A Method of Constructing an Abridged Life Table by the Combined Use of Iteration and the Cubic Spline Interpolation

Kazumasa Kobayashi
Zenji Nanjo

NUPRI Research Paper Series No. 17

March 1984

Kazumasa Kobayashi
Professor
and
Deputy Director
Nihon University
Population Research Institute

Zenji Nanjo
Professor
Fukushima Medical College
and
Affiliated Expert
Nihon University
Population Research Institute

CONTENTS

TABLES

FIGURES

ABSTRACT

This paper presents a method of constructing an abridged life table by the combined use of iteration and the cubic spline interpolation. A focus of the present study is in examining to what degree of accuracy in the life table is possible when, either population or death data, as the base material for the life table construction are given by single years of age and the remainder by 5-year age groups. It is deemed to be not always necessary to have population and death data, both by single years of age, in order to obtain the required accuracy for a complete life table. Another focus is in devising some special adjustment for resolving problems concerning end conditions in the cubic spline interpolation.

I. Introduction

If census and vital statistics tabulations provide for a population a set of data consisting of (1) the number of individuals reaching their x birthday during the year concerned (let it be year t) (E_x), (2) the number dying among them by the end of the year t at age x (${}_{\alpha}D_x$), (3) the number aged x at the beginning of the year t and dying at age x in the year t (${}_{\delta}D_x$), and (4) the number reaching their (x+1) birthday during the year t(E_{x+1}), we can easily calculate the life table probability of dying at age x for the year t by $q_x = 1 - (1 - {}_{\alpha}D_x/E_x) \cdot [1 - {}_{\delta}D_x/(E_{x+1} + {}_{\delta}D_x)]$. This is actually practised in preparing recent official complete life tables of Japan regarding ages one year and over.

The first step of making a life table without such direct data as the above is usually to obtain age-specific mortality rates of the base population. The central problem of making a life table is, then, how to reproduce exactly the age-specific mortality rates of the base population for the life table. Since the distribution of population and deaths in respective age groups is stationary in the life table and not so in the actual population, age-specific mortality rates of the life table (${}_nm_x = {}_nd_x/{}_nL_x$) is not the same as that of the base population (${}_nM_x = {}_nD_x/{}_nP_x$). Keyfitz discusses methods of constructing a life table that reproduces exactly the age-specific mortality rates which are its basis. Two methods are proposed by him. One is an iterative method based on the assumption of a sectionally stable population (Keyfitz, 1966: 19-22)[1/], and the other is one without iteration (Keyfitz, 1968: 39-42)[2/].

Most methods of life table construction developed so far, except for the youngest ages are devised to utilize such data of population and deaths as are classified either, both by single years of age or both by 5-year age groups.

In the present study we are interested in exploring methods of constructing a life table when one set of data population and deaths is available by single years of age and the remainder by n-year age groups (5>n>1). For this purpose we employ the combined use of iteration and the cubic spline interpolation. We obtain either ${}_{0.25}D_x$ or ${}_{0.25}K_x$ (population aged x to x + 0.25 at risk) by interpolation and calculate the 'expected' deaths or population at risk as follows assuming ${}_{0.25}m_x = {}_{0.25}M_x$.[3/]

$${}_{0.25}D'_x = {}_{0.25}m_x \cdot {}_{0.25}K_x$$

or

$${}_{0.25}K'_x = {}_{0.25}D_x / {}_{0.25}m_x,$$

where ${}_{0.25}m_x$ is given based on an arbitrary initial set of values for l_x.

Suppose the data of observed deaths or population at risk are available in n-year age group, that is, as ${}_nD_x$ or ${}_nK_x$, we sum up ${}_{0.25}D'_x$ to get ${}_nD'_x$ or ${}_{0.25}K'_x$ to get ${}_nK'_x$ which are comparable with observed ${}_nD_x$ or ${}_nK_x$. Then, we use the ratio ${}_{0.25}D_x/{}_{0.25}D'_x$ or ${}_{0.25}K'_x/{}_{0.25}K_x$ to correct the l_x values initially given arbitrarily. This correction can be repeated through iterative techniques until reaching some sufficient convergence. This is why we employ combined use of iteration and the cubic spline interpolation for the present study.

The method so far outlined is applicable to ages two years and over. For ages under two years we follow basically Keyfitz' idea (Keyfitz, 1968, see II of this paper). Here we also use iterative techniques, but we take a different way of obtaining the expected number of deaths based on a different curve fitting.

In the present study we examined the appropriateness of our life table method by comparing with the complete life tables for recent Japan by using the same basic data for life table construction. Methodological descriptions of life table construction and discussions of the result are given in II, III, and IV for the youngest ages, from 2 through 89 years, and ages 90 years and over, respectively. Appendix I deals with a special topic in calculating life table functions for under one year of age, with a detailed age breakdown by week and month. Appendix II presents the method of estimating midyear population.

II. The Youngest Ages

An iteration method of life table construction for the youngest ages is given by Keyfitz (1968: 240-243). It is designed to be applicable to the ages under five years. It requires an arbitrary initial set of values for the life table l_1 and l_5 as input data

(l_0 is set as 1). Upon fitting hyperbola $(ax + b)/(x + b)$ to l_0, l_1 and l_5, the values of L_0, L_1, . . . , and L_4 are calculated by integrating l_x. An iteration with the observed number of deaths is then obtained for age under one year by using formula (1) and for ages one through four years by using formula (2) as below.

$$D_0' = B_0(1 - L_0/l_0) + B_{-1}(L_0 - l_1)/l_0, \tag{1}$$

$$\begin{aligned} {}_4D_1' = {} & B_{-1}(l_1 - L_1)/l_0 + B_{-2}(L_1 - L_2)/l_0 + \cdots \\ & + B_{-5}(L_4 - l_5)/l_0. \end{aligned} \tag{2}$$

As to the expected deaths under one year of age we apply the above formula (1) with some modifications, but regarding ages one year and over, we employ another method not using births to obtain the expected number of deaths, in view of the yearly changes in child survival rates in Japan to which we intend to apply our life table methods.4/

In calculating life table functions for the youngest ages we tested ten different ways of age breakdown as shown in Table 1. When the whole range of the youngest ages which varies between under one and under five is divided into two age classes (Nos. 1 and 2), curve fitting to the three points one of which is $l_0 - 1$ is made by use of

$$l_\alpha(x) = (ax + b)/(x + b).$$

When divided into three age classes (Nos. 3 to 7), curve fitting is made to four points by use of

$$l_\beta(x) = (ax^2 + bx + c)/(x + c).$$

When divided into four age groups (Nos. 8 to 10), curve fitting to five points is made by using

$$l_\gamma(x) = (ax^3 + bx^2 + cx + d)/(x + d).$$

Findings from our study show that the fitting of the curve $l_\gamma(x)$ tends to produce the best result. The fitting of the curve $l_\beta(x)$

Table 1. Iterative Calculation of ${}_nq_x$ for the Youngest Ages with Ten Different Age Breakdowns

No.	${}_nq_x$ to calculate	l_x to which an arbitrary initial set of values is given	Curve to fit	${}_nL_x$ to calculate by integrating l_x	Calculation of expected deaths ${}_nD'_x$
1	q_0, $4q_1$	l_1, l_5	$l_\alpha(x) = \frac{ax + b}{x + b}$	L_0, ${}_{0.25}L_{1+0.25a}$ $(a = 0,1, \ldots ,15)$	Age under one year: $D'_0 = (\sum_{i=1}^{12} B_{0,i})(l_0 - L_0)/l_0$ $+ (\sum_{i=1}^{12} B_{-1,i})(L_0 - l_1)/l_0$ ${}_{6m}D'_0 = (\sum_{i=1}^{12} B_{0,i})(l_0 - l_{6m})/l_0$ $+ \sum_{i=7}^{12} (B_{-1,i} - B_{0,i})(2\ {}_{6m}L_0 - l_1)/l_0$ ${}_{3m}D'_0 = (\sum_{i=1}^{12} B_{0,i})(l_0 - l_{3m})/l_0$ $+ \sum_{i=10}^{12} (B_{-1,i} - B_{0,i})(4\ {}_{3m}L_0 - l_1)/l_0$ ${}_{6m}D'_{6m} = (\sum_{i=7}^{12} B_{-1,i} + \sum_{i=1}^{6} B_{0,i})(l_{6m} - l_1)/l_0$ $+ \sum_{i=1}^{6} (B_{-1,i} - B_{0,i})(2\ {}_{6m}L_{6m} - l_1)/l_0$ ${}_{3m}D'_{3m} = (\sum_{i=10}^{12} B_{-1,i} + \sum_{i=1}^{9} B_{0,i})(l_{3m} - l_{6m})/l_0$ $+ \sum_{i=7}^{9} (B_{-1,i} - B_{0,i})(4\ {}_{3m}L_{3m} - l_{6m})/l_0$ Ages one through four years: ${}_{0.25}D'_x = {}_{0.25}P_x \cdot {}_{0.25}m_x$ ${}_nD'_x = \sum_{i=0}^{4n-1} ({}_{0.25}P_{x+0.25i} \cdot {}_{0.25}m_{x+0.25i})$
2	q_0, q_1	l_1, l_2		L_0, ${}_{0.25}L_{1+0.25a}$ $(a = 0,1,2,3)$	
3	q_1, $3q_2$	l_1, l_2, l_5	$l_\beta(x) = \frac{ax^2 + bx + c}{x + c}$	L_0, ${}_{0.25}L_{1+0.25a}$ $(a = 0,1, \ldots ,15)$	
4	q_1, q_2	l_1, l_2, l_3		L_0, ${}_{0.25}L_{1+0.25a}$ $(a = 0,1, \ldots ,7)$	
5	${}_{6m}q_0$, ${}_{6m}q_{6m}$, $4q_1$	l_{6m}, l_1, l_5		${}_{6m}L_0$, ${}_{6m}L_{6m}$ ${}_{0.25}L_{1+0.25a}$ $(a = 0,1, \ldots ,15)$	
6	${}_{6m}q_0$, ${}_{6m}q_{6m}$, q_1	l_{6m}, l_1, l_2		${}_{6m}L_0$, ${}_{6m}L_{6m}$ ${}_{0.25}L_{1+0.25a}$ $(a = 0,1,2,3)$	
7	${}_{3m}q_0$, ${}_{3m}q_{3m}$, ${}_{6m}q_{6m}$	l_{3m}, l_{6m}, l_1		${}_{3m}L_0$, ${}_{3m}L_{3m}$, ${}_{6m}L_{6m}$	
8	${}_{6m}q_0$, ${}_{6m}q_{6m}$, q_1, $3q_2$	l_{3m}, l_{6m}, l_1, l_2, l_5	$l_\gamma(x) = \frac{ax^3 + bx^2 + cx + d}{x + d}$	${}_{6m}L_0$, ${}_{6m}L_{6m}$, ${}_{0.25}L_{1+0.25a}$ $(a = 0,1, \ldots ,15)$	
9	${}_{3m}q_0$, ${}_{3m}q_{3m}$, ${}_{6m}q_{6m}$, $4q_1$	l_{3m}, l_{6m}, l_1, l_5		${}_{3m}L_0$, ${}_{3m}L_{3m}$, ${}_{6m}L_{6m}$, ${}_{0.25}L_{1+0.25a}$ $(a = 0,1, \ldots ,15)$	
10	${}_{3m}q_0$, ${}_{3m}q_{3m}$, ${}_{6m}q_{6m}$, q_1	l_{3m}, l_{6m}, l_1, l_2		${}_{3m}L_0$, ${}_{3m}L_{3m}$, ${}_{6m}L_{6m}$, ${}_{0.25}L_{1+0.25a}$ $(a = 0,1,2,3)$	

Iterative calculation of ${}_nq_x$: ${}_nq^*_x = (l_x - l_{x+n})({}_nD_x/{}_nD'_x)/l_x$

gives generally the second best result. The poorest outcome is obtained by the curve fitting of $l_{\alpha}(x)$.

It is also pointed out that we tend to obtain a remarkably good result if we make iteration with an age breakdown for infancy into as many as three age classifications, 0 to 3 months, 3 months to 6 months, and 6 months to one year. This tendency is evidenced by our calculation shown in Table 2 based on Japan's 1975 male life table data. Out of the six sample calculations given here the last one shows the closest result to that of our standard life table.

A whole series of steps for calculating ${}_nq_x$ for the youngest ages is described below taking No. 10 in Table 1 and (No. 6 in Table 2) as an example. In this case ${}_{3m}q_0$, ${}_{3m}q_{3m}$, ${}_{6m}q_{6m}$, and q_1 are computed.

(1) Input of data

Step 1. Input infant deaths by age of months, i.e., ${}_{3m}D_0$, ${}_{3m}D_{3m}$, ${}_{6m}D_{6m}$, and deaths at age one year, D_1, for the calendar year concerned.

Step 2. Input monthly births for the calendar year concerned, $B_{0,i}$, and the one preceding year, $B_{-1,i}$.

Step 3. Input midyear population of one year olds by 0.25-year age groups, ${}_{0.25}P_x$ (for x = 1, 1.25, 1.50, and 1.75).

Step 4. Input an arbitrary initial set of values for l_{3m}, l_{6m}, l_1 and l_2, setting $l_0 = 1$.

(2) Curve fitting to l_x values

Step 5. Fit $l_x = (ax^3 + bx^2 + cx + d)/(x + d)$ to the values l_0, l_{3m}, l_{6m}, l_1 and l_2 given by Step 4.

(3) Calculation of ${}_nL_x$

Step 6. Compute ${}_{3m}L_0$, ${}_{3m}L_{3m}$, ${}_{6m}L_{6m}$ and ${}_{0.25}L_x$ as follows:

$${}_{3m}L_0 = \int_0^{0.25} l(x)dx$$

$${}_{3m}L_{3m} = \int_{0.25}^{0.50} l(x)dx$$

$${}_{6m}L_{6m} = \int_{0.50}^{1.00} l(x)dx$$

and

Table 2. Values of Selected Life Table *Functions Under Five Years*, Males, Japan, 1975, Based on Six Different Curve-fittings

Items	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	Complete Life Table, Males, Japan, 1975**
l_x to which the curve is fitted*	l_1, l_2	l_1, l_2, l_3	l_{6m}, l_1, l_2	l_{3m}, l_{6m} l_1	l_{6m}, l_1, l_2, l_5	l_{3m}, l_{6m}, l_1, l_2	
Curve to fit	$l_\alpha(x) = \frac{ax + b}{x + b}$	$l_\beta(x) = \frac{ax^2 + bx + c}{x + c}$			$l_\gamma(x) = \frac{ax^3 + bx^2 + cx + d}{x + d}$		
l_0	100000	100000	100000	100000	100000	100000	100000
l_{3m}	99352	99252	99143	99114	99131	99111	99110
l_{6m}	99107	99056	99011	99013	99009	99008	99007
l_1	98899	98896	98893	98899	98891	98891	98890
l_2	98754	98751	98748	--	98746	98745	98745
L_0	99200	99146	99084	99181	99076	99060	99064
L_1	98813	98815	98817	--	98814	98811	98810

* l(0) is always 1 according to any of the curves fitted here.

** Figures given below are specially calculated based on the q_x values before graduation.

$$_{0.25}L_x = \int_x^{x+0.25} l(a)da$$

(for x = 1, 1.25, 1.50 and 1.75).

(4) Calculation of $_{0.25}m_x$

Step 7. Compute $_{0.25}m_x$ for x = 1, 1.25, 1.50 and 1.75 as follows:

$$_{0.25}m_x = (l_x - l_{x+0.25})/_{0.25}L_x.$$

(5) Calculation of expected number of deaths $_nD'_x$

Step 8. Compute $_{3m}D'_0$, $_{3m}D'_{3m}$, $_{6m}D'_{6m}$ and D'_1 as follows:

$$_{3m}D'_0 = \sum_{i=1}^{12} B_{0,i}(l_0 - l_{3m})/l_0 + \sum_{i=10}^{12} (B_{-1,i} - B_{0,i}) \cdot (4 \cdot {}_{3m}L_0 - l_1)/l_0,$$

$$_{3m}D'_{3m} = (\sum_{i=10}^{12} B_{-1,i} + \sum_{i=1}^{9} B_{0,i})(l_{3m} - l_{6m})/l_0 + \sum_{i=7}^{9} (B_{-1,i} - B_{0,i})(4 \cdot {}_{3m}L_{3m} - l_{6m})/l_0,$$

$$_{6m}D'_{6m} = (\sum_{i=7}^{12} B_{-1,i} + \sum_{i=1}^{6} B_{0,i})(l_{6m} - l_1)/l_0 + \sum_{i=1}^{6} (B_{-1,i} - B_{0,i})(2 \cdot {}_{6m}L_{6m} - l_1)/l_0,$$

and

$$D' = \sum_{i=0}^{3} ({}_{0.25}P_{x+0.25i} \cdot {}_{0.25}m_{x+0.25i}).$$

(6) Calculation of $_nq^*_x$

Step 9. Compute $_{3m}q^*_0$, $_{3m}q^*_{3m}$, $_{6m}q^*_{6m}$, and q^*_1 as follows:

$$_{3m}q^*_0 = (1 - l_{3m}/l_0)({}_{3m}D_0/{}_{3m}D'_0),$$

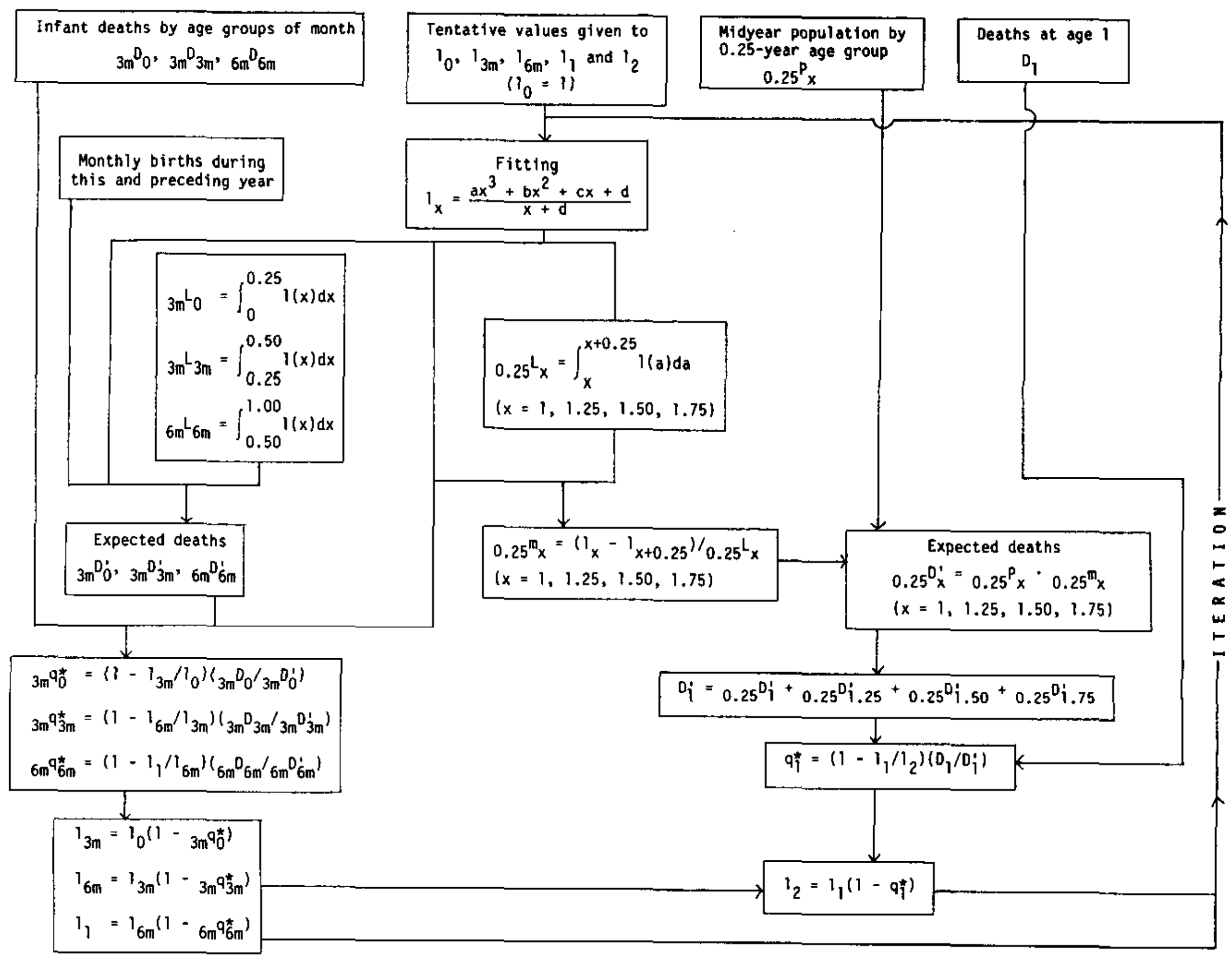
$$_{3m}q^*_{3m} = (1 - l_{6m}/l_{3m})({}_{3m}D_{3m}/{}_{3m}D'_{3m}),$$

$$_{6m}q^*_{6m} = (1 - l_1/l_{6m})({}_{6m}D_{6m}/{}_{6m}D'_{6m}),$$

and

$$q^*_1 = (1 - l_2/l_1)(D_1/D'_1).$$

Figure 1. Steps of Calculating $_nq_x$ for $0 \le x < 2$ with Three Age Divisions Under One Year Old



(7) Calculation of corrected values of l_x

Step 10. Compute corrected values of l_{3m}, l_{6m}, l_1 and l_2 as follows:

$$l_{3m} = l_0(1 - {}_{3m}q_0^*),$$

$$l_{6m} = l_{3m}(1 - {}_{3m}q_{3m}^*),$$

$$l_1 = l_{6m}(1 - {}_{6m}q_{6m}^*),$$

and

$$l_2 = l_1(1 - q_1^*).$$

(8) Iterative calculation

Step 11. Return to Step 5 with the l_x values obtained by Step 10, and repeat computation of Steps 5 to 11 until the respective ${}_nq_x^*$ values become the same as those of the preceding iteration to the sixth decimal place.

III. Ages 2 Through 89 Years

Steps of calculation in III are applicable to the case when the life table functions for the youngest ages are calculated for the first two years of life. There are altogether twelve steps (Steps 1 to 12) and among them Steps 5, 6, 7 and 9 are divided into two each, that is, a and b. The calculation taking Steps 5a, 6a, 9a and 10a will be denoted by Method I and that taking Steps 5b, 6b, 9b and 10b later by Method II (see Tables 3 and 4).

(1) Interpolation of l_x values at 0.25-year intervals

Step 1. Give an arbitrary initial set of values to

$$l_5, l_{10}, \ldots, \text{ and } l_{90}.$$

Step 2. Extrapolate l_{95}, l_{100}, . . . , and l_{110} by use of the following formula:

$$l_{x+5} = \exp(4 \cdot \ln l_x - 6 \cdot \ln l_{x-5} + 4 \cdot \ln l_{x-10} - \ln l_{x-15}).$$ [4/]

Table 3. Values of Selected Life Table Functions for Japanese Females, 1980: *Method I, Method II and Official Complete Life Table*

x	Method I				Method II				Official Complete Life Table				
	l_x	$_nq_x$	$_nL_x$	$\mathring{e}_x$	l_x	$_nq_x$	$_nL_x$	$\mathring{e}_x$	l'_x	l_x	$_nq_x$	$_nL_x$	$\mathring{e}_x$
2	99245	0.001257	297527	77.36	99245	0.001259	297526	77.36	99245	99254	0.001360	297526	77.35
5	99121	0.001054	495309	74.46	99120	0.001053	495309	74.46	99121	99119	0.001039	495307	74.46
10	99016	0.000701	494918	69.54	99016	0.000705	494916	69.54	99016	99016	0.000697	494912	69.53
15	98947	0.001351	494428	64.58	98946	0.001341	494428	64.58	98948	98947	0.001364	494435	64.58
20	98813	0.001849	493627	59.67	98814	0.001858	493629	59.67	98814	98812	0.001842	493614	59.66
25	98630	0.002393	492586	54.77	98630	0.002396	492583	54.77	98631	98630	0.002403	492580	54.77
30	98394	0.003121	491243	49.90	98394	0.003115	491242	49.90	98394	98393	0.003110	491246	49.90
35	98087	0.004464	489411	45.05	98087	0.004464	489411	45.05	98089	98087	0.004476	489408	45.04
40	97649	0.006732	486722	40.24	97649	0.006729	486722	40.24	97647	97648	0.006708	486720	40.23
45	96992	0.010572	482587	35.49	96992	0.010563	482591	35.49	96990	96993	0.010619	482593	35.49
50	95967	0.016258	476180	30.84	95968	0.016278	476182	30.84	95965	95963	0.016215	476156	30.84
55	94406	0.024325	466699	26.31	94406	0.024339	466690	26.31	94405	94407	0.024320	466686	26.30
60	92110	0.039136	452213	21.90	92108	0.039123	452208	21.90	92118	92111	0.039181	452219	21.89
65	88505	0.065846	429260	17.68	88504	0.065876	429247	17.68	88497	88502	0.065671	429232	17.68
70	82677	0.120814	390504	13.73	82674	0.120698	390514	13.73	82679	82690	0.121018	390551	13.73
75	72689	0.215479	326920	10.25	72695	0.215594	326936	10.25	72676	72683	0.216185	326812	10.24
80	57026	0.370065	233642	7.33	57023	0.370208	233599	7.33	56994	56970	0.368937	233615	7.33
85	35923	0.555070	127393	5.13	35912	0.554652	127390	5.13	35916	35949	0.555676	127473	5.12
90	15983	1.000000	56926	3.56	15994	1.000000	56926	3.56	15968	15973	1.000000	56668	3.55

Table 4. Values of Selected Life Table Functions for Japanese Males, 1980 and Japanese Males and Females, 1975: Method I, Method II and Official Complete Life Tables

	Method I		Method II		Official Complete Life Table		
x	l_x	e_x	l_x	e_x	l'_x	l_x	e_x
			Males, 1980				
2	99064	72.04	99064	72.04	99065	99070	72.03
10	98712	64.28	98712	64.28	98713	98714	64.28
20	98243	54.56	98245	54.56	98245	98245	54.56
30	97380	45.00	97382	44.99	97382	97386	45.00
40	96103	35.52	96105	35.52	96102	96100	35.52
50	92816	26.57	92815	26.57	92817	92822	26.57
60	85699	18.31	85700	18.31	85716	85699	18.31
70	69839	11.18	69839	11.18	69842	69852	11.18
80	37844	6.07	37847	6.07	37814	37803	6.08
90	7123	3.15	7116	3.16	7176	7117	3.17
			Males, 1975				
2	98745	70.63	98745	70.64	98745	98752	70.63
10	98295	62.94	98296	62.94	98295	98301	62.94
20	97721	53.27	97724	53.27	97722	97731	53.27
30	96692	43.78	96695	43.79	96697	96700	43.78
40	95132	34.41	95135	34.41	95132	95139	34.41
50	91477	25.56	91479	25.57	91479	91487	25.56
60	83779	17.39	83779	17.39	83774	83798	17.38
70	65999	10.53	66020	10.54	66006	66007	10.53
80	33169	5.72	33195	5.72	33159	33198	5.70
90	5383	3.12	5404	3.11	5382	5357	3.05
			Females, 1975				
2	99009	75.65	99009	75.65	99009	99019	75.65
10	98709	67.87	98710	67.87	98709	98716	67.87
20	98450	58.03	98451	58.04	98450	98454	58.04
30	97873	48.34	97874	48.35	97873	97879	48.35
40	96942	38.76	96942	38.76	96938	96947	38.76
50	94916	29.46	94919	29.46	94912	94922	29.46
60	90298	20.68	90305	20.68	90288	90303	20.68
70	79126	12.78	79133	12.78	79127	79122	12.78
80	50711	6.76	50704	6.76	50704	50705	6.76
90	12073	3.38	12133	3.36	12053	12001	3.39

Step 3. Interpolate l_x values at every 0.25-year interval between ages 2 and 90 years based on the values of l_1 and l_2 already fixed and those of l_5 through l_{110} given by Steps 1 and 2 above, by use of the cubic spline functions.

(2) Input of data

Step 4. Input l_x values for $x = 2, 2.25, \ldots, 89.75$, and 90 provided by Step 3.

Step 5a. Input midyear population by 0.25-year age group, ${}_{0.25}P_x$, for $x = 2, 2.25, \ldots$, and 89.75.

Step 5b. Input midyear population by 5-year age group, ${}_5P_x$, for $x = 5, 10, \ldots$, and 85, and ${}_3P_2$.

Step 6a. Input deaths during the year concerned by 5-year age group, ${}_5D_x$, for $x = 5, 10, \ldots$, and 85, and ${}_3D_2$.

Step 6b. Input deaths during the year concerned by 0.25-year age group, ${}_{0.25}D_x$, for $x = 2, 2.25, \ldots$, and 89.75.

(3) Calculation of ${}_{0.25}L_x$ and ${}_{0.25}m_x$

Step 7. Compute ${}_{0.25}L_x$ values for $x = 2, 2.25, \ldots$, and 89.75 by integration as follows:

$$L = \int_x^{x+0.25} l(a)da.$$

Step 8. Compute ${}_{0.25}m_x$ values for $x = 2, 2.25, \ldots$, and 89.75 by the following formula:

$${}_{0.25}m_x = (l_x - l_{x+0.25})/{}_{0.25}L_x.$$

(4) Calculation of expected deaths ${}_nD'_x$

Step 9a. Compute ${}_3D'_2$ and ${}_5D'_x$ as follows:

$${}_3D'_2 = \sum_{x=2}^{4} \sum_{n=0}^{3} {}_{0.25}m_{x+0.25n} \cdot {}_{0.25}P_{x+0.25n}$$

and

$${}_5D'_x = \sum_{x=5i}^{5i+4} \sum_{n=0}^{3} {}_{0.25}m_{x+0.25n} \cdot {}_{0.25}P_{x+0.25n}$$

for $i = 1, 2, \ldots$, and 17.

Step 9b. Compute ${}_3P'_2$ and ${}_5P'_x$ as follows:

$$ {}_3P'_2 = \sum_{x=2}^{4} \sum_{n=0}^{3} {}_{0.25}D'_{x+0.25n} / {}_{0.25}m_{x+0.25n}, $$

and

$$ {}_5P'_x = \sum_{x=5i}^{5i+4} \sum_{n=0}^{3} {}_{0.25}D'_{x+0.25n} / {}_{0.25}m_{x+0.25n} $$

for i = 1, 2, . . . , and 17.

(5) Calculation of ${}_nq^*_x$

Step 10a. Compute ${}_3q^*_2$ and ${}_5q^*_x$ as follows:

$$ {}_3q^*_2 = (1 - l_5/l_2)\,{}_3D_2/{}_3D'_2, $$

and

$$ {}_5q^*_x = (1 - l_{x+5}/l_x)\,{}_5D_x/{}_5D'_x $$

for x = 5, 10, . . . , and 85.

Step 10b. Compute ${}_3q^*_2$ and ${}_5q^*_x$ as follows:

$$ {}_3q^*_2 = (1 - l_5/l_2)\,{}_3P'_2/{}_3P_2, $$

and

$$ {}_5q^*_x = (1 - l_{x+5}/l_x)\,{}_5P'_x/{}_5P_x $$

for x = 5, 10, . . . , and 85.

(6) Calculation of corrected values of l_x

Step 11. Compute corrected values of l_5 and l_{5n} for n = 2 to 18 as follows:

$$ l_5 = l_2(1 - {}_3q^*_2), $$

and

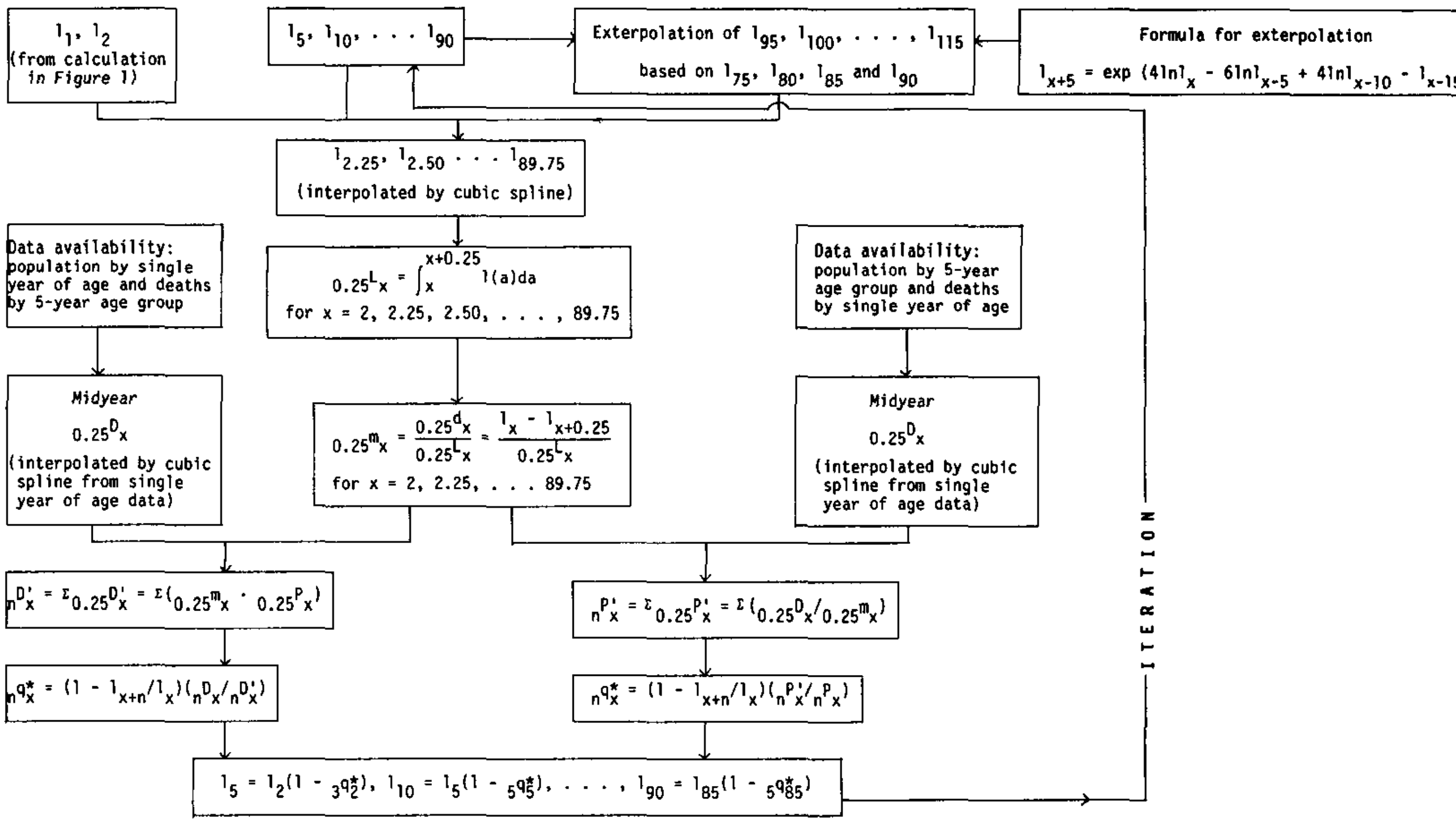
$$ l_{5n} = l_{5n-5}(1 - {}_5q^*_{5n-5}) $$

for n = 2 to 18.

(7) Repetition of calculation

Step 12. Return to Step 2 with the l_x values obtained by Step 11, and repeat computation of Steps 2 to 12 until the respective ${}_nq^*_x$ values become the same as those of the preceding iteration to the sixth decimal place.[5/]

Figure 2. Steps of Calculating $_nq_x$ for $2 \leq x < 90$



Following the steps described above we constructed life tables of Japanese (both male and female) for 1975 and 1980. Some of the results are given in Tables 3 and 4 comparing them with the official complete life tables.

Table 3 is the 1975 life table for females. Columns (1) to (4) show the results by Method I and columns (5) to (8) show those by Method II. Columns (9) to (12) give the corresponding data from the official complete life table. The figures by Methods I and II are extremely close to each other and also sufficiently close to the corresponding complete life table figures. As to the values of $\mathring{e}_x$, in particular, we find almost no difference between our calculation and the complete life table. Table 4 is for 1980 males and 1975 both males and females. The general tendency is quite similar in this table too.

IV. Ages 90 years and over

Steps of calculation are described below. We use regula falsi as an iterative method for obtaining a solution of equation $f(x) = 0$.

(1) Input of data

Step 1. Input the values of l_{80}, l_{85} and l_{90} which were obtained previously.

Step 2. Survivors at age 95 years be denoted by l'_{95} as a variable. Give a tentative value to l'_{95} which is denoted by $l^{(1)}$.

(2) Extrapolation of l_x

Step 3. Fit a Makeham curve $l_{x+t} = l_x\, S^t\, g^{c^t - 1}$ to l_{80}, l_{85}, l_{90} and l'_{95}.

Step 4. Interpolate l_x between l_{90} and l_{104} at 0.25-yeaer interval (i.e., calculate $l_{90+0.25(i-1)}$ for $i = 1, 2, \ldots,$ 57).

(3) Calculation of expected population aged 90 years and over

Step 5. Calculate ${}_{0.25}L_x$ for $x = 90, 90.25, \ldots,$ 103.75 as follows:

$${}_{0.25}L_x = \frac{13}{96}\,(l_x + l_{x+0.25}) - \frac{1}{96}\,(l_{x-0.25} + l_{x+0.5}).$$

Step 6. Calculate ${}_{0.25}m_x$ for $x = 90, 90.25, . . . , 103.75$ by

$$_{0.25}m_x = \frac{l_x - l_{x+0.25}}{_{0.25}L_x}.$$

Step 7. Interpolate D_x at 0.25-year interval between D_{90} and D_{104} by the cubic spline functions (to obtain ${}_{0.25}D_x$).

Step 8. Calculate the expected population age 90 years and over ${}_{\infty}P'_{90}$ by

$$_{\infty}P'_{90} = \sum_{i=1}^{57} [{}_{0.25}D_{90+0.25(i-1)}/{}_{0.25}m_{90+0.25(i-1)}]$$

(4) Correction of l'_{95}

Step 9. Compute

$$y = f(l'_{95}) = {}_{\infty}P'_{90} - {}_{\infty}P_{90}.$$

Step 10. $y = f(l'_{95}) = f_1$, say.

Step 11. Give another tentative value to l'_{95} which is denoted by $l^{(2)}$ $(l^{(1)} < l^{(2)})$.

Step 12. Repeat computation of Steps 3 to 9 to get new value of y.

$$y = f(l'_{95}) = f_2, \text{ say.}$$

Step 13. Compute

$$f_1 \cdot f_2$$

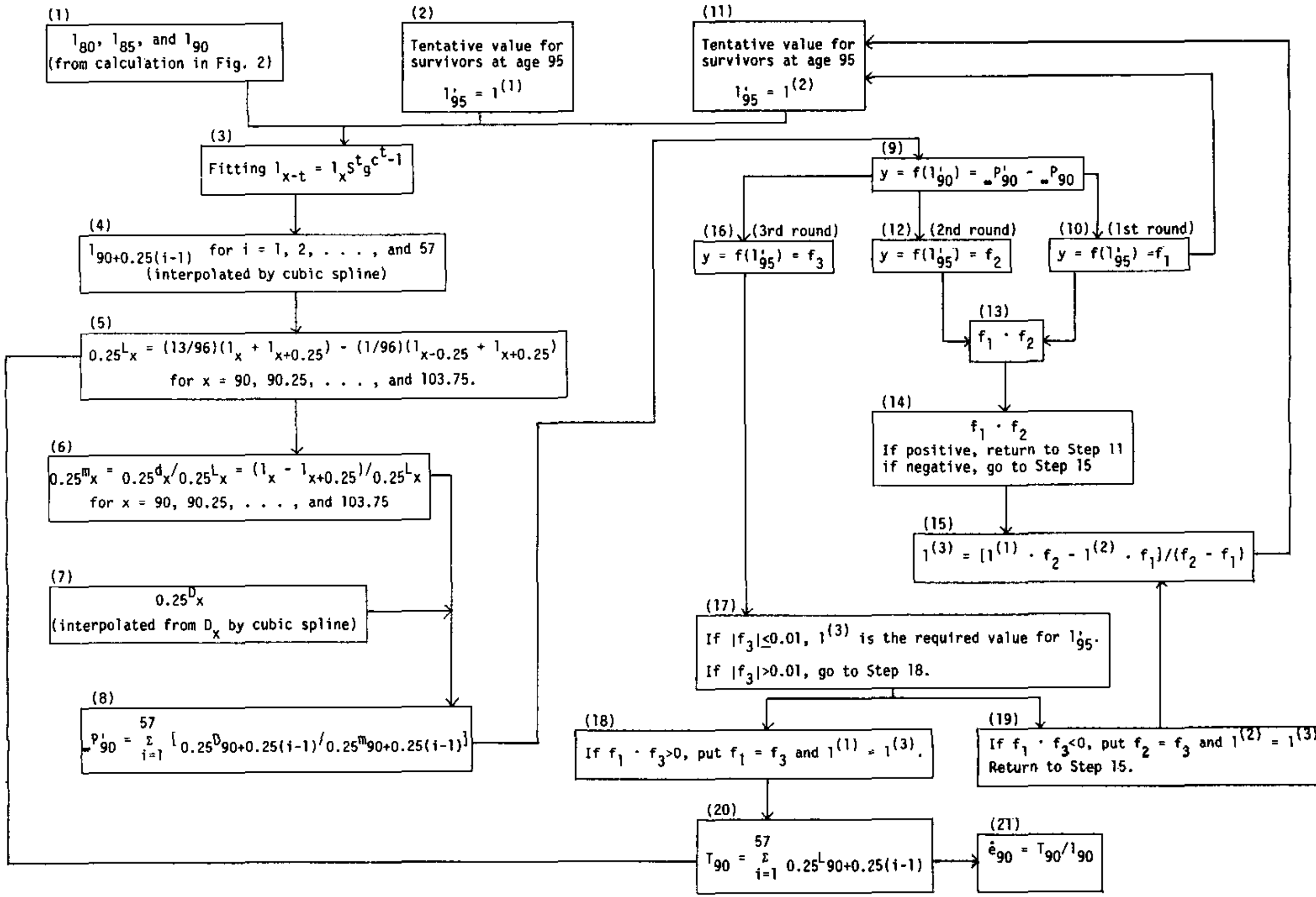
Step 14. If the product $f_1 \cdot f_2$ is positive, try Steps 3 to 13 until the product is found to be negative. When the value of $f_1 \cdot f_2$ is negative, go to Step 15.

Step 15. Compute

$$l^{(3)} = [l^{(1)} \cdot f_2 - l^{(2)} \cdot f_1]/(f_2 - f_1).$$

Step 16. Return to Step 3 with the value of $l^{(3)}$ (for l'_{95}) and

Figure 3. Steps of Calculating $_{\infty}L_{90}$



compute down to Step 9 which produces

$$f_3 = f(l'_{95}) = f(l^{(3)}).$$

Step 17. If $|f_3| \leq 0.01$, $l^{(3)}$ is the required value for l'_{95}. If $|f_3| > 0.01$, go to Step 18.

Step 18. If $f_1 \;\; f_3 > 0$, put $f_1 = f_3$ and $l^{(1)} = l^{(3)}$.

Step 19. If $f_1 \;\; f_3 < 0$, put $f_2 = f_3$ and $l^{(2)} = l^{(3)}$. Return to Step 15.

(5) Calculation of T_{90} and $\mathring{e}_{90}$

Step 20. Compute T_{90} by

$$T_{90} = \sum_{i=1}^{57} 0.25 L_{90+0.25(i-1)}.$$

Step 21. Compute $\mathring{e}_{90}$ by

$$\mathring{e}_{90} = T_{90}/l_{90}.$$

V. Conclusion

There are two major findings from the present study.

1. If either population or death data is available by single years of age and the other by 5-year age groups, the life table constructed based on these produces values quite close to those of the life table constructed based on population and death data, both of which are classified by single years of age.
2. Of the three fractional functions fitted to l_x for the youngest ages, $l_Y(x)$ (See II) which is fitted to five points is found to have a sufficiently good fitting.

For the convenience of users, practical procedures of life table construction are described in details, step by step. Although our method is devised to be adaptable to the present situation of data availability for the whole of Japan, it will also be applicable to other populations after appropriate modifications.

Appendix I. Calculation of Probabilities of Dying under One Year by Weeks and Months of Age

(1) Input of data

Step 1. Input numbers of monthly births from April of the preceding year (t-1) to December of the year concerned (t).

Step 2. Input values of l_{1w}, l_{2w}, l_{3w}, l_{4w}, l_{2m}, l_{3m}, l_{6m}, l_{9m}, and l_1 which are given arbitrarily. The value of l_0 is set as 1.

(2) Calculation of $_nL_x$

Step 3. Fit cubic spline curve to the above l_x values and by integrating them calculate $_{1w}L_0$, $_{1w}L_{1w}$, $_{1w}L_{2w}$, $_{1w}L_{3w}$, $_{2m-4w}L_{4w}$, $_{1m}L_{2m}$, $_{3m}L_{3m}$, $_{3m}L_{6m}$ and $_{3m}L_{9m}$.

(3) Calculation of expected numbers of deaths

Step 4. B_0 = number of births occurring in January to December of year t.

B_1 = number of births between December 25 of year t-1 and December 24 of year t.

B_2 = number of births between December 18 of year t-1 and December 17 of year t.

B_3 = number of births between December 11 of year t-1 and December 10 of year t.

B_4 = number of births between December 4 of year t-1 and December 3 of year t.

B_5 = number of births between November of year t-1 and October of year t.

B_6 = number of births in October of year t-1 to September of year t.

B_7 = number of births in July of year t-1 to June of year t.

B_8 = number of births in April of year t-1 to March of year t.

B_9 = number of births in January to December of year t.

B_{10} = number of births in December of year t-1.

B_{11} = number of births in December of year t.

Step 5. Compute B_1, B_2, B_3, and B_4 as follows:

$$B_1 = B_9 + (7/31)(B_{10} - B_{11}),$$

$$B_2 = B_9 + (14/31)(B_{10} - B_{11}),$$

$$B_3 = B_9 + (21/31)(B_{10} - B_{11}),$$

and

$$B_4 = B_9 + (28/31)(B_{10} - B_{11}).$$

Step 6. Put

$$W_1 = W_2 = W_3 = W_4 = 365/7$$

$$W_5 = 365/(365/6 - 28)$$

$$W_6 = 12$$

and

$$W_7 = W_8 = W_9 = 4$$

Step 7. Compute ${}_nD'_x$ as follows:

$${}_{1w}D'_0 = B_0(l_0 - W_1 \cdot {}_{1w}L_0)/l_0$$

$$+ B_1(W_1 \cdot {}_{1w}L_0 - l_{1w})/l_0,$$

$${}_{1w}D'_{1w} = B_1(l_{1w} - W_2 \cdot {}_{1w}L_{1w})/l_0$$

$$+ B_2(W_2 \cdot {}_{1w}L_{1w} - l_{2w})/l_0,$$

$${}_{1w}D'_{2w} = B_2(l_{2w} - W_3 \cdot {}_{1w}L_{2w})/l_0$$

$$+ B_3(W_3 \cdot {}_{1w}L_{2w} - l_{3w})/l_0,$$

$${}_{1w}D'_{3w} = B_3(l_{3w} - W_4 \cdot {}_{1w}L_{3w})/l_0$$

$$+ B_4(W_4 \cdot {}_{1w}L_{3w} - l_{4w})/l_0,$$

$${}_{2m-4w}D'_{4w} = B_4(l_{4w} - W_5 \cdot {}_{2m-4w}L_{4w})/l_0$$

$$+ B_5(W_5 \cdot {}_{2m-4w}L_{4w} - l_{2m})/l_0,$$

$$_{1m}D'_{2m} = B_5(l_{2m} - W_6 \cdot {}_{1m}L_{2m})/l_0$$

$$+ B_6(W_6 \cdot {}_{1m}L_{2m} - l_{3m})/l_0,$$

$$_{3m}D'_{3m} = B_6(l_{3m} - W_7 \cdot {}_{3m}L_{3m})/l_0$$

$$+ B_7(W_7 \cdot {}_{3m}L_{3m} - l_{6m})/l_0,$$

$$_{3m}D'_{6m} = B_7(l_{6m} - W_8 \cdot {}_{3m}L_{6m})/l_0$$

$$+ B_8(W_8 \cdot {}_{3m}L_{6m} - l_{9m})/l_0,$$

and

$$_{3m}D'_{9m} = B_8(l_{9m} - W_9 \cdot {}_{3m}L_{9m})/l_0$$

$$+ B_9(W_9 \cdot {}_{3m}L_{9m} - l_1)/l_0.$$

(4) Calculation of ${}_nq^*_x$

Step 8. Compute ${}_nq^*_x$ as follows:

$$_{1w}q^*_0 = (1 - l_{1w}/l_0)({}_{1w}D_0/{}_{1w}D'_0),$$

$$_{1w}q^*_{1w} = (1 - l_{2w}/l_{1w})({}_{1w}D_{1w}/{}_{1w}D'_{1w}),$$

$$_{1w}q^*_{2w} = (1 - l_{3w}/l_{2w})({}_{1w}D_{2w}/{}_{1w}D'_{2w}),$$

$$_{1w}q^*_{3w} = (1 - l_{4w}/l_{3w})({}_{1w}D_{3w}/{}_{1w}D'_{3w}),$$

$$_{2m-4w}q^*_{4w} = (1 - l_{2m}/l_{4w})({}_{2m-4w}D_{4w}/{}_{2m-4w}D'_{4w}),$$

$$_{1m}q^*_{2m} = (1 - l_{3m}/l_{2m})({}_{1m}D_{2m}/{}_{1m}D'_{2m}),$$

$$_{3m}q^*_{3m} = (1 - l_{6m}/l_{3m})({}_{3m}D_{6m}/{}_{3m}D'_{6m}),$$

$$_{3m}q^*_{6m} = (1 - l_{9m}/l_{6m})({}_{3m}D_{6m}/{}_{3m}D'_{6m}),$$

and

$$_{3m}q^*_{9m} = (1 - l_1/l_{9m})({}_{3m}D_{9m}/{}_{3m}D'_{9m}).$$

(5) Correction of l_x

Step 9. Compute corrected values of l_{1w}, l_{2w}, . . . , and l_1 as follows:

$$l_{1w} = l_0(1 - {}_{1w}q_0^*),$$

$$l_{2w} = l_{1w}(1 - {}_{1w}q_{1w}^*),$$

$$l_{3w} = l_{2w}(1 - {}_{1w}q_{2w}^*),$$

$$l_{4w} = l_{3w}(1 - {}_{1w}q_{3w}^*),$$

$$l_{2m} = l_{4w}(1 - {}_{2m-4w}q_{4w}^*),$$

$$l_{3m} = l_{2m}(1 - {}_{1m}q_{2m}^*),$$

$$l_{6m} = l_{3m}(1 - {}_{3m}q_{3m}^*),$$

$$l_{9m} = l_{6m}(1 - {}_{3m}q_{6m}^*),$$

and

$$l_1 = l_{9m}(1 - {}_{3m}q_{9m}^*).$$

(6) Iterative calculation

Step 10. Return to Step 3 with the l_x values obtained by Step 9 and repeat computation of Steps 3 to 10 until the difference between new and old values of l_x become smaller than 0.01, respectively.

The following table shows l_x and ${}_nq_x$ values for Japanese males, 1975, calculated on the basis of the above method in comparison with the official complete life table. Coincidence of our calculation to the complete life table is almost perfect.

x	l_x		${}_nq_x$ x 100,000	
	Our calculation	Official complete life table	Our calculation	Official complete life table
0	100000.0	100000	613.0	613
1w	99387.0	99387	89.9	90
2w	99297.6	99298	44.2	44
3w	99253.8	99254	25.6	26
4w	99228.4	99228	76.5	77
2m	99152.5	99153	42.6	43
3m	99110.2	99110	103.6	104
6m	99007.5	99008	66.9	--
9m	98941.2	--	51.8	--
1	98890.0	98890	--	--

Appendix II. Estimation of Midyear Population

The population census of Japan is taken regularly as of October first, every fifth year. Intercensal estimates of population by sex and single years of age are, accordingly, made also as of October first of each intercensal year. For constructing a life table for a calendar year, therefore, we need to estimate midyear population by sex and age for the denominator from the enumerated or estimated population as of October first. Below are described steps of this estimation.

(1) Input of data

Step 1. Input population by single years of age (for ages 1 through 90 years of a particular sex) as of October first of the year concerned.

Step 2. Input coefficient of adjustment for age-unknowns of population.

Step 3. Input deaths by single years of age (for ages 1 through 90 years) during the calendar year concerned.

Step 4. Input coefficients of adjustment for age-unknowns and underregistration of deaths.

(2) Adjustment of population and deaths

Step 5. Compute adjusted population by single years of age as of October first of the year concerned using the input data of Steps 1 and 2.

Step 6. Compute adjusted deaths by single years of age during the calendar year concerned using the input data of Steps 3 and 4.

(3) Interpolation of population and deaths at 0.25-year interval

Step 7. Compute cumulated population up to age x for x = 1, 2, . . . , and 90 from Step 5.

Step 8. Compute cumulated deaths up to age x for x = 1, 2, . . . , and 90 from Step 6.

Step 9. From Step 7 compute cumulated population up to age x at 0.25-year interval for x = 1, 1.25, 1.50, . . . , and 90.25 by the cubic spline interpolation.

Step 10. From Step 8 compute cumulative deaths up to age x at 0.25-year interval for x = 1, 1.25, 1.50, . . . , and 90.25 by the cubic spline interpolation.

Step 11. Compute one-fourth of the deaths of Step 10 to obtain those for the three months from July to September.

Step 12. Interpolated the deaths of Step 11 at 0.125-year interval by use of the cubic spline functions.

Step 13. Compute deaths by 0.25-year age groups ${}_{0.25}D_{x+0.125}$ for x = 1, 1.25, 1.50, . . . , and 90.25.

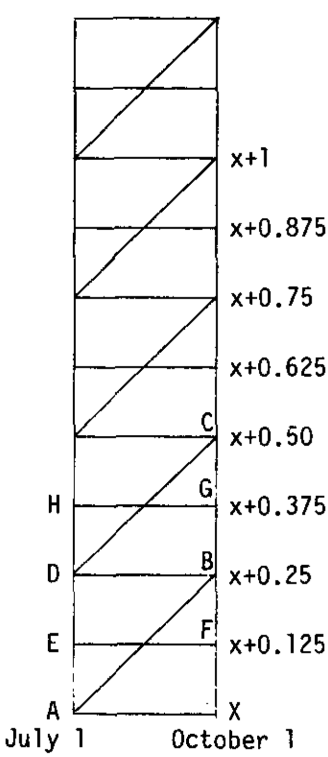
Step 14. Compute population by 0.25-year age groups ${}_{0.25}\hat{P}_{x+0.25}$ for x = 1, 1.25, 1.50, . . . , and 90.

(4) Calculation of midyear population

Step 15. Compute midyear population by 0.25-year age group ${}_{0.25}P_x$ as follows:

$$ {}_{0.25}P_x = {}_{0.25}\hat{P}_{x+0.25} + {}_{0.25}D_{x+0.25} $$

for x = 1, 1.25, . . . , and 89.75. (see the Lexis diagram shown below).



Assuming

Deaths(ABCD) = Deaths(EFGH),

Population(AD) is estimated by

Population(AD) = Population(BC) + Deaths(ABCD)
= Population(BC) + Deaths(EFGH).

Therefore,

$$ {}_{0.25}P_x = {}_{0.25}\hat{P}_{x+0.25} + {}_{0.25}D_{x+0.125}. $$

Notes

1/ Through the iterative computation ${}_5M'_x$ is calculated by

$$ {}_5M'_x = (l_x - \exp(-5r) \cdot l_{x+5}) / $$

$$ [(65/24)(l_x + \exp(-5r) \cdot l_{x+5}) $$

$$ - (5/24)(\exp(-10r) \cdot l_{x+10} + \exp(5r) \cdot l_{x-5})] - r, $$

and correction of ${}_nq_x$ by ${}_nq^*_x = {}_nq_x({}_nM_x/{}_nM'_x)$ is repeated.

2/ The following formula is used to obtain ${}_nq_x$:

$$ {}_nq_x = 1 - \exp[-n\,{}_nM_x + n/(48 \cdot {}_nP_x)({}_nP_{x+n} - {}_nP_{x-n}) $$

$$ \cdot ({}_nM_{x+n} - {}_nM_{x-n})]. $$

This was first given in Keyfitz and Frauenthal (1975: 889-899).

3/ The principal characteristic of our life table method is in assuming that the life table age-specific death rate ${}_nm_x$ is the same as the age-specific death rate of the observed population ${}_nM_x$ when the age interval n is taken small enough to be one-quarter of a year. Keyfitz describes that ${}_{0.2}m_x$ may be assumed to be the same as ${}_{0.2}M_x$ in his discussion on repeated application of matrix for life table construction (Keyfitz, 1968: 234). Our assumption is just its modification. This assumption makes the life table construction quite simple. The small age interval of one-quarter of a year is obtained conveniently by the cubic spline interpolation. *An advantage of the cubic spline is in its flexibility of* obtaining interpolation intervals of any length needed. This is not similar to the case of the osculatory interpolation by multipliers of Sprague, Greville, Beers and so on which is restricted to a quadratic interpolation only. One more advantage is that the curve fitting of the cubic spline functions permits direct calculation of areas under the curve by integration which is especially useful for obtaining life table stationary population $({}_{0.25}L_x)$

by integrating life table survivors (l_x).

4/ This procedure is specially taken to obtain somewhat more effective results of interpolation for ages under 90 years than those calculated without using base data extending beyond age 90, in view of the fact that the usual three choices of the end condition for the cubic spline interpolation do not always guarantee satisfactory returns.

5/ Iterative calculation will be completed, if the difference between new and old values of l_{90} becomes smaller than 0.01. In fact, however, the whole age range under 90 years is found sometimes too wide to converge. In such a case, it is suggested to divide the whole age range into three parts and carry out iteration blockwise one after another. Concretely speaking, first of all, values of l_3 to l_{45} are determined by carrying out iteration until the difference between new and old values of l_{45} becomes smaller than 0.01. Secondly, interpolation and iteration are carried out for the second part (l_x values for ages under 45 are fixed now) until the difference between new and old values of l_{75} becomes smaller than 0.01. After the values of l_x for the same part (under age 75) have been determined, interpolation and iteration for the third part of the age range are carried out (l_x values for ages under 75 are fixed now) until the difference between new and old values of l_{90} becomes smaller than 0.01.

References

Ahlberg, J.H., E.N. Nilson, and J.L. Walsh. 1967. The Theory of Splines and Their Applications. New York. Academic Press.

Ichida, Kozo and Fujiichi Yoshimoto. 1979. Spline Kansu to sono oyo (Spline functions and their applications). Tokyo. Kyoiku Shuppan.

Keyfitz, Nathan. 1966. "A Life Table that Agrees with the Data." Journal of American Statistical Association, Vol. 69, pp. 305-311.

_______________. 1968. "A Life Table that Agrees with the Data, II." Journal of American Statistical Association, Vol. 63, pp. 1253-1267.

_______________. 1968b. Introduction to the Mathematics of Population. Readings. Addison-Wesley Publishing Company.

_______________ and Wilhelm Flieger. 1971. Population, Facts and Methods of Demography. San Francisco. W.H. Freeman and Company.