1
Introduction

The leaders and governments of most civilizations have collected statistical information about their peoples. The Sumerians, five thousand years ago, for example, enumerated their population for taxation purposes. Later, the Romans based conscription on enumeration of the population. The Christian Church has over many centuries compiled, in the form of parish registers, a huge amount of demographic material.[1] Today, the quantity of data available about current populations, even in some of the less-developed countries, is enormous.

It was only in the seventeenth century, however, that men became interested in studying the numbers of their fellow human-beings from the purely scientific point of view. The Englishman, John Graunt (1620–74) was probably the first of these, and his work was thorough and of a high standard (J. Graunt, 1662[2]; I. Sutherland, 1963). He produced the first life table and studied the population of London in some detail. Many others were to follow his example.

We are primarily concerned with mathematical models for human populations. Leonhard Euler produced such a model as early as 1767[3],and the population analysis of Thomas Malthus (1798[2]) could

[1] Several groups are currently studying the data available from parish registers. Those at the Institute of Genetics, Pavia, Italy are interested in genetic data obtainable from parish records in the Parma Valley, Italy (L. L. Cavalli-Sforza, 1958). The Cambridge Group for the History of Population and Social Structure are interested mainly in demographic results from pre-industrial populations in England (E. A. Wrigley, 1966; E. A. Wrigley and R. S. Schofield, 1968). Some of the better-known work has been done in France (P. Goubert, 1960).

[2] These references are extremely interesting (and sometimes amusing) reading.

[3] 'Recherches générales sur la mortalité et la multiplication du genre humain'. *Histoire de l'academie royale des sciences et belles-lettres, Année 1760*, pp. 144–64.' Preussische Akademie der Wissenschaften zu Berlin. The date of publication is 1767. See also I. Todhunter (1865). *A History of the Mathematical Theory of Probability*, pp. 240–1. (Reprinted in 1965 by Chelsea Publishing Company, New York.)

be described as the formulation of a mathematical model. Modern population theory, however, might be said to have begun with the deterministic theories of A. J. Lotka (1907) and F. R. Sharpe and A. J. Lotka (1911).

An account of the life table and its use is given in chapter 2. The next eight chapters are devoted to various mathematical models which have been proposed for analysing the growth of human populations. These models vary considerably in their complexity and in their mathematical treatment.

2
The life table

2.1 *Introduction*

Consider a large number n_0 of lives all aged exactly 0 years. After x years, some of these lives will have died, and there will be, say, n_x lives remaining, all aged x. It is clear that n_x forms monotonic non-increasing sequence of integers with increasing x. If n_x is large, the probability of survival from age x to age $x+t$, denoted by $_tp_x$, will be approximately equal to n_{x+t}/n_x. That is,

$$_tp_x \doteqdot n_{x+t}/n_x. \qquad (2.1.1)$$

This is the type of argument first used by demographers to develop the important theoretical concept of the *life table*. The first life table, compiled by John Graunt in 1662, was mentioned in chapter 1. This is commonly regarded as his masterpiece. The following quotation comes from his book:

> '9. Whereas we have found, that of 100 quick Conceptions about 36 of them die before they be six years old, and that perhaps but one surviveth 76, we, having seven Decads between six and 76, we sought six mean proportional numbers between 64, the remainder, living at six years, and the one, which survives 76, and finde, that the numbers following are practically near enough to the truth; for men do not die in exact Proportions, nor in Fractions: from whence arises this Table following.

Viz. of 100 there dies within the first six years	36
The next ten years, or Decad	24
The second Decad	15
The third Decad	09
The fourth	6
The next	4
The next	3
The next	2
The next	1

'10. From whence it follows, that of the said 100 conceived there remains alive at six years end 64.

At Sixteen years end	40
At Twenty six	25
At Tirty six	16
At Fourty six	10
At Fifty six	6
At Sixty six	3
At Seventy six	1
At Eighty	0'

In spite of its humble origin, the modern life table is a fine theoretical tool. A continuous, well-behaved, monotonic decreasing function l_x is defined such that the probability of survival from exact age x to exact age $x+t$, ${}_tp_x$ is equal to l_{x+t}/l_x. Some arbitrary value like 10,000 or 100,000 is usually assigned to l_0, and this value is called the *radix*.

Although l_x is continuous and is defined for all values of x greater than zero, it is an empirical function and seldom (if ever) has an explicit mathematical form. It is usually tabulated for integral values of x, and intermediary values (if required) are found by interpolation. The difference $l_x - l_{x+1}$ represents the number of deaths aged x last birthday out of l_x lives who attain exact age x, and it is denoted by d_x.

The life table for Australian males (1961) is given on pages 175–177.

2.2 *Mortality computations*

Using the life table, it is possible to compute various probabilities involving mortality. Consider, for example, a life now aged 30. What is the probability that this life will die between exact age 40 and exact age 50?

According to our definition of l_x, the probability that a life aged 30 will survive to exact age 40 is l_{40}/l_{30}, and the probability that a life aged 40 will die before exact age 50 is $1-(l_{50}/l_{40})$. The mortality prospects of an individual at different ages are assumed independent, and we conclude that the life aged 30 will die between ages 40 and 50 with probability $(l_{40}/l_{30})(1-l_{50}/l_{40})$. This probability simplifies to $(l_{40}-l_{50})/l_{30}$, and for the Australian life table (males) 1961 it is equal to $(92{,}859-88{,}473)/94{,}726 = 0.046$.

2.3 *The force of mortality*

An example of the use of the life table function l_x for mortality computations is given in section 2.2. Consider now a life aged x. What is the probability that this life will die between exact age $x+t$ and exact age $x+t+\mathrm{d}t$? An argument similar to that given in section 2.2 indicates that the probability is $(l_{x+t}-l_{x+t+\mathrm{d}t})/l_x$. The function l_x is well-behaved and $l_{x+t+\mathrm{d}t}$ may be expanded in a Taylor series about the point $x+t$. It follows that

$$\frac{l_{x+t}-l_{x+t+\mathrm{d}t}}{l_x} = \frac{l_{x+t}}{l_x}\left\{-\frac{1}{l_{x+t}}\frac{\mathrm{d}}{\mathrm{d}t}(l_{x+t})\,\mathrm{d}t\right\}+o(\mathrm{d}t)$$

$$= {}_tp_x\mu_{x+t}\,\mathrm{d}t+o(\mathrm{d}t), \qquad (2.3.1)$$

where

$$\mu_x = -\frac{1}{l_x}\frac{\mathrm{d}}{\mathrm{d}x}l_x = -\frac{\mathrm{d}}{\mathrm{d}x}\log l_x. \qquad (2.3.2)$$

Formula (2.3.1) should be noted: ${}_tp_x$ is the probability that a life survives from age x to age $x+t$, and $\mu_{x+t}\mathrm{d}t$ is the probability that a life aged $x+t$ will die during the time element $\mathrm{d}t$. The mortality function μ_x, defined by equation (2.3.2), is usually referred to as the *force of mortality at age x*. It is of considerable theoretical importance.

2.4 *Numerical evaluation of the force of mortality*

Numerical values of μ_x are often required. The function l_x is usually an empirical function, so numerical differentiation is necessary to determine μ_x. The following formula is frequently employed by the compilers of life tables, and it is accurate provided l_x is a polynomial of fourth degree in the vicinity of x:

$$\mu_x = \frac{8(l_{x-1}-l_{x+1})-(l_{x-2}-l_{x+2})}{12l_x}. \qquad (2.4.1)$$

It may be derived by expanding l_{x-2}, l_{x-1}, l_{x+1} and l_{x+2} in Taylor series about the point x and eliminating the terms involving second, third and fourth derivatives of l_x.

It is not possible to compute μ_0 and μ_1 using this formula. The force of mortality μ_x is changing very rapidly in the age range 0 to 2, so any formula using values of l_x for integral x only is unlikely to give reliable results. The problem of computing μ_x for $0 \leqslant x \leqslant 2$ is extremely complex, and in the Australian life table (males) for 1961, for example, μ_x is only given for ages greater than two years. Approximate values can be obtained by fitting a hyperbola to the l_x function in this age range.

2.5 *Other life table functions*

The life table functions l_x, μ_x, and ${}_tp_x$ are defined above. Other life table functions are defined in terms of l_x, and we shall have occasion to use some of them in later chapters:

$$d_x = l_x - l_{x+1} \quad \text{(deaths)},$$

$$ {}_np_x = l_{x+n}/l_x,$$

$$p_x = {}_1p_x,$$

$$ {}_nq_x = 1 - {}_np_x,$$

$$q_x = 1 - p_x \quad \text{(mortality rate)},$$

$$L_x = \int_0^1 l_{x+t}\,dt,$$

$$m_x = d_x/L_x \quad \text{(central mortality rate)}.$$

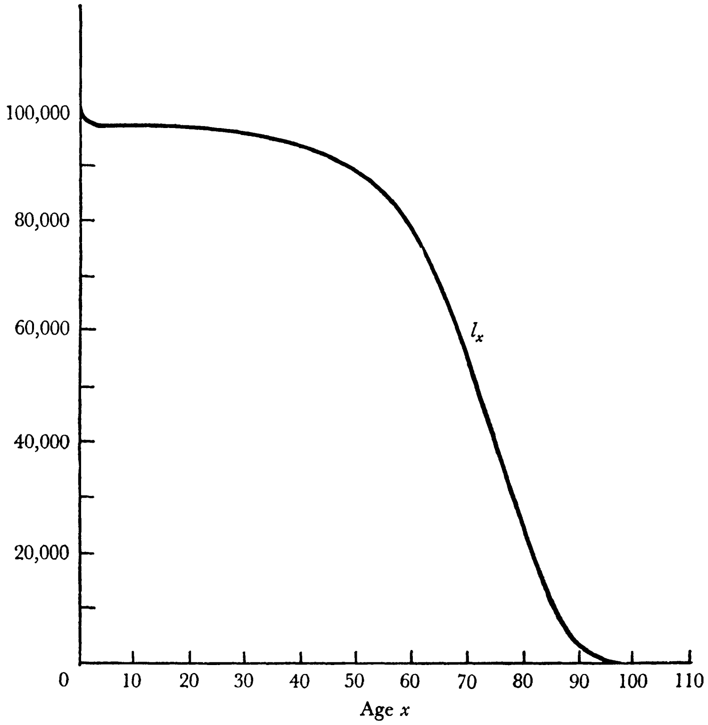

Figure 2.6.1. Graph of l_x (Australian life table (males) 1961).

2.6 *The graphs of certain life table functions*

The graphs of l_x, μ_x, and q_x for Australian Males (1961) are given in figures 2.6.1, 2.6.2 and 2.6.3 respectively. The following comments about μ_x help in the understanding of these curves:

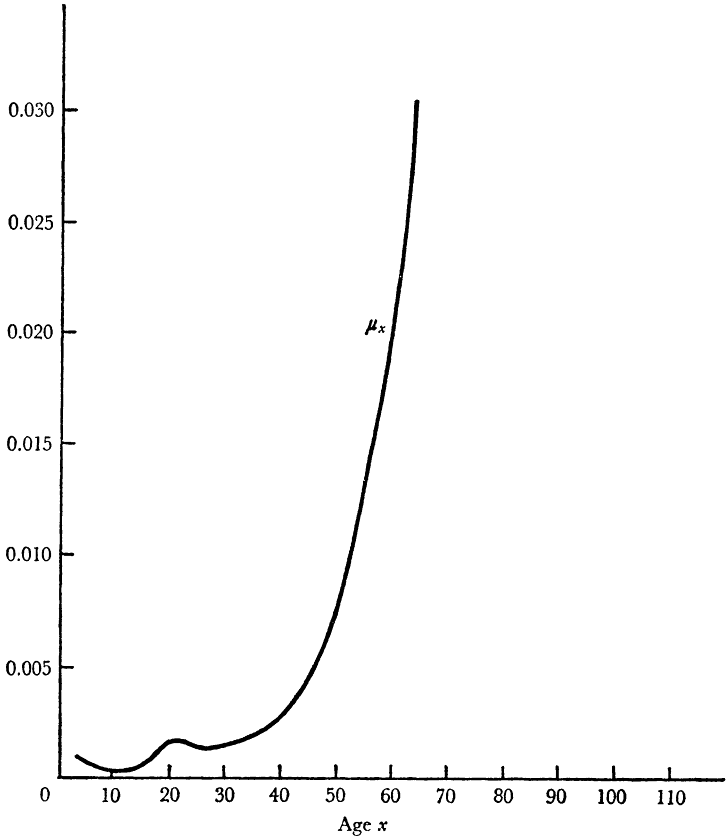

Figure 2.6.2. Graph of μ_x (Australian life table (males) 1961).

(i) μ_0 is large.
(ii) It drops rapidly in the first two years of life.
(iii) It is a minimum near 11 years of age.
(iv) It rises *very* gradually to reach a maximum in the early twenties.
(v) It then drops a little before continuing its gradual ascent.

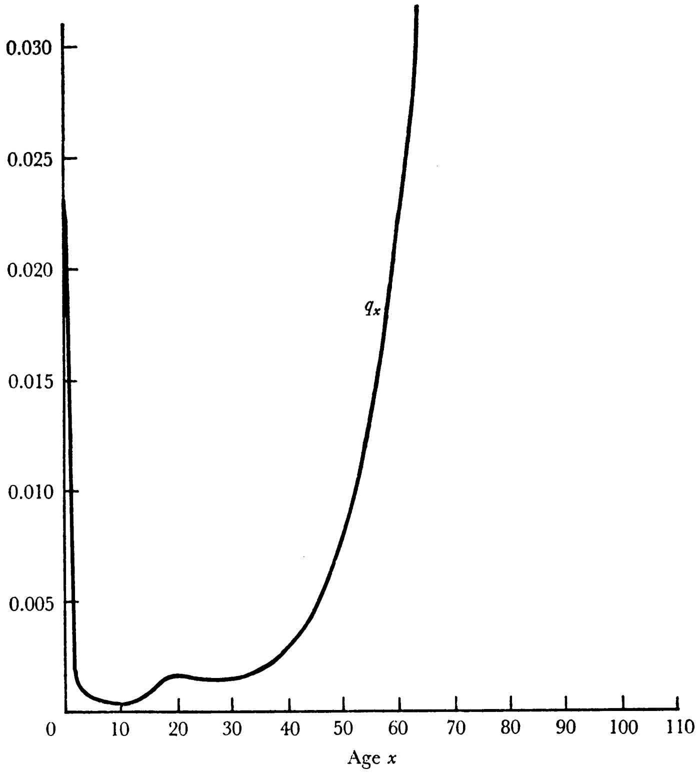

Figure 2.6.3. Graph of q_x (Australian life table (males) 1961).

The maximum value in the early twenties is caused by motor accidents, etc.

(vi) After age 50, μ_x starts to rise more rapidly.

The q_x curve is of course very similar to the μ_x curve.

2.7 *The Euler–Maclaurin expansion*

The Euler–Maclaurin expansion connects the integral of a function with the sum of a series of equally spaced ordinates. To prove the formula, we make use of the differential operator D, and the finite difference operators Δ and E. These three operators are defined as follows:

$$\left.\begin{aligned} \mathrm{D}U_x &= \frac{\mathrm{d}}{\mathrm{d}x}U_x, \\ \Delta U_x &= U_{x+1} - U_x, \\ \mathrm{E}U_x &= U_{x+1}. \end{aligned}\right\} \qquad (2.7.1)$$

Relationships exist between the operators, and the following two are important in the present context:

$$\left.\begin{aligned} E &\equiv 1+\Delta; \\ E &\equiv e^{D}. \end{aligned}\right\} \qquad (2.7.2)$$

The first relationship is very easy to prove, and the second may be derived by considering the Taylor series expansion:

$$\begin{aligned} EU_x &= U_{x+1} \\ &= U_x + U_x' + \frac{1}{2!}U_x'' + \frac{1}{3!}U_x''' + \dots \\ &= \left(1 + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \dots\right) U_x \\ &= e^{D} U_x. \end{aligned}$$

To derive the Euler–Maclaurin expansion, we consider the sum

$$\begin{aligned} U_0 + U_1 + U_2 + \dots + U_{n-1} &= (E^0 + E^1 + E^2 + \dots + E^{n-1})\, U_0 \\ &= (E^n - 1)\,(E-1)^{-1} U_0 \\ &= (E^n - 1)\,(e^{D}-1)^{-1} U_0 \\ &= (E^n - 1)\left(\frac{1}{D} - \frac{1}{2} + \frac{1}{12}D + \dots\right) U_0. \end{aligned}$$

That is $$\sum_{0}^{n-1} U_i = \int_0^n U_x \, dx - \tfrac{1}{2}(U_n - U_0) + \tfrac{1}{12}(U_n' - U_0') + \dots \qquad (2.7.3)$$

and this is the Euler–Maclaurin expansion.

2.8 *The expectation of life*

The *curtate expectation of life* is the average number of *whole* years lived after age x by a life who attains age x. It is denoted by e_x. Clearly,

$$\begin{aligned} e_x &= \sum_{n=0}^{\infty} n \frac{d_{x+n}}{l_x} \\ &= \frac{1}{l_x}\{(l_{x+1} - l_{x+2}) + 2(l_{x+2} - l_{x+3}) + 3(l_{x+3} - l_{x+4}) + \dots\} \\ &= \frac{1}{l_x}\sum_{n=1}^{\infty} l_{x+n} \\ &= \frac{1}{l_x}\sum_{n=0}^{\infty} l_{x+n} - 1. \end{aligned} \qquad (2.8.1)$$

The *complete expectation of life* is the average number of years of life lived after age x by a life who attains age x. It is denoted by $\mathring{e}_x$. The probability that a life aged x will die at age $x+t$ is given by equation (2.3.1), and it follows that the expectation of life at age x is given by

$$\begin{aligned}\mathring{e}_x &= \frac{1}{l_x}\int_0^\infty t l_{x+t}\mu_{x+t}\,\mathrm{d}t \\ &= -\frac{1}{l_x}\int_0^\infty t\frac{\mathrm{d}}{\mathrm{d}t}(l_{x+t})\,\mathrm{d}t \\ &= \frac{1}{l_x}\int_0^\infty l_{x+t}\,\mathrm{d}t \quad \text{(integration by parts).} \qquad (2.8.2)\end{aligned}$$

By general reasoning, it seems plausible that $\mathring{e}_x$ should be greater than e_x by about half a year. The Euler–Maclaurin expansion does in fact provide a relationship between the two expectations. When applied to formula (2.8.1),

$$\begin{aligned}e_x + 1 &= \frac{1}{l_x}\int_0^\infty l_{x+t}\,\mathrm{d}t + \frac{1}{l_x}(\tfrac{1}{2}l_x) - \frac{1}{12}\frac{1}{l_x}(l_x') + \dots \\ &= \mathring{e}_x + \tfrac{1}{2} + \tfrac{1}{12}\mu_x + \dots .\end{aligned}$$

Hence

$$\mathring{e}_x \doteqdot e_x + \tfrac{1}{2} - \tfrac{1}{12}\mu_x. \qquad (2.8.3)$$

2.9 *The uniform distribution of deaths*

The assumption is frequently made that deaths between age x and age $x+1$ are uniformly distributed over the year of age. This assumption is usually very accurate for human populations (except perhaps at the very young ages), and it is a very convenient one for numerical work. It is equivalent to assuming that l_{x+t} is linear in the range $0 \leqslant t \leqslant 1$. One can deduce that

$$L_x = \tfrac{1}{2}(l_x + l_{x+1}), \qquad (2.9.1)$$

$$m_x = 2q_x/(2-q_x), \qquad (2.9.2)$$

$$q_x = 2m_x/(2+m_x), \qquad (2.9.3)$$

and

$${}_tq_x = tq_x \quad (0 \leqslant t \leqslant 1), \qquad (2.9.4)$$

under this assumption.