3

Life tables and selection

3.1 Summary

In this chapter we define a life table. For a life table tabulated at integer ages only, we show, using fractional age assumptions, how to calculate survival probabilities for all ages and durations.

We discuss some features of national life tables from Australia, England & Wales and the United States.

We then consider life tables appropriate to individuals who have purchased particular types of life insurance policy and discuss why the survival probabilities differ from those in the corresponding national life table. We consider the effect of 'selection' of lives for insurance policies, for example through medical underwriting. We define a select survival model and we derive some formulae for such a model.

We discuss briefly how mortality rates change over time, and illustrate one way to allow for mortality trends in a survival model.

3.2 Life tables

Given a survival model, with survival probabilities ${}_tp_x$, we can construct the **life table** for the model from some initial age x_0 to a maximum age ω. We define a function $\{l_x\}$ for $x_0 \le x \le \omega$ as follows. Let l_{x_0} be an arbitrary positive number (called the **radix** of the table) and, for $0 \le t \le \omega - x_0$, define

$$l_{x_0+t} = l_{x_0}\,{}_tp_{x_0}.$$

From this definition we see that for $x_0 \le x \le x+t \le \omega$,

$$\begin{aligned} l_{x+t} &= l_{x_0}\,{}_{x+t-x_0}p_{x_0} \\ &= l_{x_0}\,{}_{x-x_0}p_{x_0}\,{}_tp_x \\ &= l_x\,{}_tp_x, \end{aligned}$$

so that

$$\boxed{{}_tp_x = l_{x+t}/l_x.} \tag{3.1}$$

For any $x \geq x_0$, we can interpret l_{x+t} as the expected number of survivors at age $x+t$ from l_x independent individuals aged x. This interpretation is more natural if l_x is an integer, and follows because the number of survivors to age $x+t$ is a random variable with a binomial distribution with parameters l_x and ${}_tp_x$. That is, suppose we have l_x independent lives aged x, and each life has a probability ${}_tp_x$ of surviving to age $x+t$. Then the number of survivors to age $x+t$ is a binomial random variable, $\mathbf{L}_t$, say, with parameters l_x and ${}_tp_x$. The expected value of the number of survivors is then

$$E[\mathbf{L}_t] = l_x\,{}_tp_x = l_{x+t}.$$

We always use the table in the form l_y/l_x which is why the radix of the table is arbitrary – it would make no difference to the survival model if all the l_x values were multiplied by 100, for example.

From (3.1) we can use the l_x function to calculate survival probabilities. We can also calculate mortality probabilities. For example,

$$q_{30} = 1 - \frac{l_{31}}{l_{30}} = \frac{l_{30} - l_{31}}{l_{30}} \tag{3.2}$$

and

$${}_{15|30}q_{40} = {}_{15}p_{40}\;{}_{30}q_{55} = \frac{l_{55}}{l_{40}}\left(1 - \frac{l_{85}}{l_{55}}\right) = \frac{l_{55} - l_{85}}{l_{40}}. \tag{3.3}$$

In principle, a life table is defined for all x from the initial age, x_0, to the limiting age, ω. In practice, it is very common for a life table to be presented, and in some cases even defined, at integer ages only. In this form, the life table is a useful way of summarizing a lifetime distribution since, with a single column of numbers, it allows us to calculate probabilities of surviving or dying over integer numbers of years starting from an integer age.

It is usual for a life table, tabulated at integer ages, to show the values of d_x, where

$$\boxed{d_x = l_x - l_{x+1},} \tag{3.4}$$

in addition to l_x, as these are used to compute q_x. From (3.4) we have

$$d_x = l_x\left(1 - \frac{l_{x+1}}{l_x}\right) = l_x(1 - p_x) = l_x\,q_x.$$

Table 3.1 *Extract from a life table.*

x	l_x	d_x
30	10 000.00	34.78
31	9 965.22	38.10
32	9 927.12	41.76
33	9 885.35	45.81
34	9 839.55	50.26
35	9 789.29	55.17
36	9 734.12	60.56
37	9 673.56	66.49
38	9 607.07	72.99
39	9 534.08	80.11

We can also arrive at this relationship if we interpret d_x as the expected number of deaths in the year of age x to $x+1$ from a group of l_x lives aged exactly x, so that, using the binomial distribution again

$$\boxed{d_x = l_x \, q_x \, .} \tag{3.5}$$

Example 3.1 Table 3.1 gives an extract from a life table. Calculate

(a) l_{40},

(b) ${}_{10}p_{30}$,

(c) q_{35},

(d) ${}_5q_{30}$, and

(e) the probability that a life currently aged exactly 30 dies between ages 35 and 36.

Solution 3.1 (a) From equation (3.4),

$$l_{40} = l_{39} - d_{39} = 9\,453.97.$$

(b) From equation (3.1),

$${}_{10}p_{30} = \frac{l_{40}}{l_{30}} = \frac{9\,453.97}{10\,000} = 0.94540.$$

(c) From equation (3.5),

$$q_{35} = \frac{d_{35}}{l_{35}} = \frac{55.17}{9\,789.29} = 0.00564.$$

(d) Following equation (3.2),

$$_5q_{30} = \frac{l_{30} - l_{35}}{l_{30}} = 0.02107.$$

(e) This probability is $_{5|}q_{30}$. Following equation (3.3),

$$_{5|}q_{30} = \frac{l_{35} - l_{36}}{l_{30}} = \frac{d_{35}}{l_{30}} = 0.00552.$$

□

3.3 Fractional age assumptions

A life table $\{l_x\}_{x \ge x_0}$ provides exactly the same information as the corresponding survival distribution, S_{x_0}. However, a life table tabulated at integer ages only does not contain all the information in the corresponding survival model, since values of l_x at integer ages x are not sufficient to be able to calculate probabilities involving non-integer ages, such as $_{0.75}p_{30.5}$. Given values of l_x at integer ages only, we need an additional assumption or some further information to calculate probabilities for non-integer ages or durations. Specifically, we need to make some assumption about the probability distribution for the future lifetime random variable between integer ages.

We use the term **fractional age assumption** to describe such an assumption. It may be specified in terms of the force of mortality function or the survival or mortality probabilities.

In this section we assume that a life table is specified at integer ages only and we describe the two most useful fractional age assumptions.

3.3.1 Uniform distribution of deaths

The uniform distribution of deaths (UDD) assumption is the most common fractional age assumption. It can be formulated in two different, but equivalent, ways as follows.

UDD1

For integer x, and for $0 \le s < 1$, assume that

$$\boxed{_sq_x = s\,q_x\,.} \tag{3.6}$$

UDD2

Recall from Chapter 2 that K_x is the integer part of T_x, and define a new random variable R_x such that

$$T_x = K_x + R_x.$$

The UDD2 assumption is that, for integer x, $R_x \sim \mathrm{U}(0,1)$, and R_x is independent of K_x.

The equivalence of these two assumptions is demonstrated as follows. First, assume that UDD1 is true. Then for integer x, and for $0 \le s < 1$,

$$\begin{aligned}
\Pr[R_x \le s] &= \sum_{k=0}^{\infty} \Pr[R_x \le s \text{ and } K_x = k] \\
&= \sum_{k=0}^{\infty} \Pr[k \le T_x \le k+s] \\
&= \sum_{k=0}^{\infty} {}_kp_x \; {}_sq_{x+k} \\
&= \sum_{k=0}^{\infty} {}_kp_x \; s\,(q_{x+k}) \quad \text{using UDD1} \\
&= s \sum_{k=0}^{\infty} {}_kp_x \; q_{x+k} \\
&= s \sum_{k=0}^{\infty} \Pr[K_x = k] \\
&= s.
\end{aligned}$$

This proves that $R_x \sim \mathrm{U}(0,1)$. To prove the independence of R_x and K_x, note that

$$\begin{aligned}
\Pr[R_x \le s \text{ and } K_x = k] &= \Pr[k \le T_x \le k+s] \\
&= {}_kp_x \; {}_sq_{x+k} \\
&= s \; {}_kp_x \; q_{x+k} \\
&= \Pr[R_x \le s]\Pr[K_x = k]
\end{aligned}$$

since $R_x \sim \mathrm{U}(0,1)$. This proves that UDD1 implies UDD2.

To prove the reverse implication, assume that UDD2 is true. Then for integer x, and for $0 \le s < 1$,

$$\begin{aligned}
{}_sq_x &= \Pr[T_x \le s] \\
&= \Pr[K_x = 0 \text{ and } R_x \le s] \\
&= \Pr[R_x \le s]\;\Pr[K_x = 0]
\end{aligned}$$

as K_x and R_x are assumed independent. Thus,

$$ {}_sq_x = s\,q_x \,. \tag{3.7}$$

Formulation UDD2 explains why this assumption is called the Uniform Distribution of Deaths, but in practical applications of this assumption, formulation UDD1 is the more useful of the two.

An immediate consequence is that

$$\boxed{l_{x+s} = l_x - s\,d_x} \tag{3.8}$$

for $0 \le s < 1$. This follows because

$$_sq_x = 1 - \frac{l_{x+s}}{l_x}$$

and substituting $s\,q_x$ for $_sq_x$ gives

$$s\,\frac{d_x}{l_x} = \frac{l_x - l_{x+s}}{l_x}.$$

Hence

$$l_{x+s} = l_x - s\,d_x$$

for $0 \le s \le 1$. Thus, we assume that l_{x+s} is a linearly decreasing function of s.

Differentiating equation (3.6) with respect to s, we obtain

$$\frac{d}{ds}\,{_sq_x} = q_x, \quad 0 \le s < 1$$

and we know that the left-hand side is the probability density function for T_x at s, because we are differentiating the distribution function. The probability density function for T_x at s is $_sp_x\,\mu_{x+s}$ so that under UDD

$$\boxed{q_x = {_sp_x}\,\mu_{x+s}} \tag{3.9}$$

for $0 \le s < 1$.

The left-hand side does not depend on s, which means that the density function is a constant for $0 \le s < 1$, which also follows from the uniform distribution assumption for R_x.

Since q_x is constant with respect to s, and $_sp_x$ is a decreasing function of s, we can see that μ_{x+s} is an increasing function of s, which is appropriate for ages of interest to insurers. However, if we apply the approximation over successive ages, we obtain a discontinuous function for the force of mortality, with discontinuities occurring at integer ages, as illustrated in Example 3.4. Although this is undesirable, it is not a serious drawback.

Example 3.2 Given that $p_{40} = 0.999473$, calculate $_{0.4}q_{40.2}$ under the assumption of a uniform distribution of deaths.

Solution 3.2 We note that the fundamental result in equation (3.7), that for fractions of a year s, $_sq_x = s\,q_x$, requires x to be an integer. We can manipulate the required probability ${}_{0.4}q_{40.2}$ to involve only probabilities from integer ages as follows

$$\begin{aligned}{}_{0.4}q_{40.2} &= 1 - {}_{0.4}p_{40.2} = 1 - \frac{l_{40.6}}{l_{40.2}} \\ &= 1 - \frac{{}_{0.6}p_{40}}{{}_{0.2}p_{40}} = 1 - \frac{1-0.6q_{40}}{1-0.2q_{40}} \\ &= 2.108 \times 10^{-4}.\end{aligned}$$

□

Example 3.3 Use the life table in Example 3.1 above, with the UDD assumption, to calculate (a) ${}_{1.7}q_{33}$ and (b) ${}_{1.7}q_{33.5}$.

Solution 3.3 (a) We note first that

$${}_{1.7}q_{33} = 1 - {}_{1.7}p_{33} = 1 - (p_{33})\,({}_{0.7}p_{34}).$$

We can calculate p_{33} directly from the life table as $l_{34}/l_{33} = 0.995367$ and ${}_{0.7}p_{34} = 1 - 0.7\,q_{34} = 0.996424$ under UDD, so that ${}_{1.7}q_{33} = 0.008192$.

(b) To calculate ${}_{1.7}q_{33.5}$ using UDD, we express this as

$$\begin{aligned}{}_{1.7}q_{33.5} &= 1 - {}_{1.7}p_{33.5} \\ &= 1 - \frac{l_{35.2}}{l_{33.5}} \\ &= 1 - \frac{l_{35} - 0.2d_{35}}{l_{33} - 0.5d_{33}} \\ &= 0.008537.\end{aligned}$$

□

Example 3.4 Under the assumption of a uniform distribution of deaths, calculate $\lim_{t\to 1^-} \mu_{40+t}$ using $p_{40} = 0.999473$, and calculate $\lim_{t\to 0^+} \mu_{41+t}$ using $p_{41} = 0.999429$.

Solution 3.4 From formula (3.9), we have $\mu_{x+t} = q_x/{}_tp_x$ for $0 \le t < 1$. Setting $x = 40$ yields

$$\lim_{t\to 1^-} \mu_{40+t} = q_{40}/p_{40} = 5.273 \times 10^{-4},$$

while setting $x = 41$ yields

$$\lim_{t\to 0^+} \mu_{41+t} = q_{41} = 5.71 \times 10^{-4}.$$

□

Example 3.5 Given that $q_{70} = 0.010413$ and $q_{71} = 0.011670$, calculate ${}_{0.7}q_{70.6}$ assuming a uniform distribution of deaths.

Solution 3.5 As deaths are assumed to be uniformly distributed between ages 70 and 71 and ages 71 and 72, we first write the probability as

$${}_{0.7}q_{70.6} = {}_{0.4}q_{70.6} + (1 - {}_{0.4}q_{70.6})\,{}_{0.3}q_{71}.$$

Following the same arguments as in Solution 3.3, we obtain

$$_{0.4}q_{70.6} = 1 - \frac{1 - q_{70}}{1 - 0.6q_{70}} = 4.191 \times 10^{-3},$$

and as $_{0.3}q_{71} = 0.3q_{71} = 3.501 \times 10^{-3}$, we obtain $_{0.7}q_{70.6} = 7.678 \times 10^{-3}$. □

3.3.2 Constant force of mortality

A second fractional age assumption is that the force of mortality is constant between integer ages. Thus, for integer x and $0 \le s < 1$, we assume that μ_{x+s} does not depend on s, and we denote it μ_x^*. We can obtain the value of μ_x^* by using the fact that

$$p_x = \exp\left\{-\int_0^1 \mu_{x+s} ds\right\}.$$

Hence the assumption that $\mu_{x+s} = \mu_x^*$ for $0 \le s < 1$ gives $p_x = e^{-\mu_x^*}$ or $\mu_x^* = -\log p_x$. Further, under the assumption of a constant force of mortality, for $0 \le s < 1$ we obtain

$$_sp_x = \exp\left\{-\int_0^s \mu_x^* \, du\right\} = e^{-\mu_x^* s} = (p_x)^s.$$

Similarly, for $t, s > 0$ and $t + s < 1$,

$$_sp_{x+t} = \exp\left\{-\int_0^s \mu_x^* \, du\right\} = (p_x)^s.$$

Thus, under the constant force assumption, the probability of surviving for a period of $s < 1$ years from age $x + t$ is independent of t provided that $s + t < 1$.

The assumption of a constant force of mortality between integer ages leads to a step function for the force of mortality over successive years of age, whereas we would expect the force of mortality to increase smoothly. However, if the true force of mortality increases slowly over the year of age, the constant force of mortality assumption is reasonable.

Example 3.6 Given that $p_{40} = 0.999473$, calculate $_{0.4}q_{40.2}$ under the assumption of a constant force of mortality.

Solution 3.6 We have $_{0.4}q_{40.2} = 1 - {}_{0.4}p_{40.2} = 1 - (p_{40})^{0.4} = 2.108 \times 10^{-4}$. □

Example 3.7 Given that $q_{70} = 0.010413$ and $q_{71} = 0.011670$, calculate $_{0.7}q_{70.6}$ under the assumption of a constant force of mortality.

Solution 3.7 As in Solution 3.5 we write

$$_{0.7}q_{70.6} = {}_{0.4}q_{70.6} + (1 - {}_{0.4}q_{70.6})\, {}_{0.3}q_{71},$$

where $_{0.4}q_{70.6} = 1 - (p_{70})^{0.4} = 4.178 \times 10^{-3}$ and $_{0.3}q_{71} = 1 - (p_{71})^{0.3} = 3.515 \times 10^{-3}$, giving $_{0.7}q_{70.6} = 7.679 \times 10^{-3}$. □

Note that in Examples 3.2 and 3.5 and in Examples 3.6 and 3.7 we have used two different methods to solve the same problems, and the solutions agree to five decimal places. It is generally true that the assumptions of a uniform distribution of deaths and a constant force of mortality produce very similar solutions to problems. The reason for this can be seen from the following approximations. Under the constant force of mortality assumption

$$q_x = 1 - e^{-\mu^*} \approx \mu^*$$

provided that μ^* is small, and for $0 < t < 1$,

$$_tq_x = 1 - e^{-\mu^* t} \approx \mu^* t.$$

In other words, the approximation to $_tq_x$ is t times the approximation to q_x, which is what we obtain under the uniform distribution of deaths assumption.

3.4 National life tables

Life tables based on the mortality experience of the whole population of a country are regularly produced for many countries in the world. Separate life tables are usually produced for males and for females and possibly for some other groups of individuals, for example on the basis of smoking habits.

Table 3.2 shows values of $q_x \times 10^5$, where q_x is the probability of dying within one year, for selected ages x, separately for males and females, for the populations of Australia, England & Wales and the United States. These tables are constructed using records of deaths in a particular year, or a small number of consecutive years, and estimates of the population in the middle of that period. The relevant years are indicated in the column headings for each of the three life tables in Table 3.2. Data at the oldest ages are notoriously unreliable. For this reason, the United States Life Tables do not show values of q_x for ages 100 and higher.

For all three national life tables and for both males and females, the values of q_x follow exactly the same pattern as a function of age, x. Figure 3.1

Table 3.2 *Values of* $q_x \times 10^5$ *from some national life tables.*

	Australian Life Tables 2000–02		English Life Table 15 1990–92		US Life Tables 2002	
x	Males	Females	Males	Females	Males	Females
0	567	466	814	632	764	627
1	44	43	62	55	53	42
2	31	19	38	30	37	28
10	13	8	18	13	18	13
20	96	36	84	31	139	45
30	119	45	91	43	141	63
40	159	88	172	107	266	149
50	315	202	464	294	570	319
60	848	510	1 392	830	1 210	758
70	2 337	1 308	3 930	2 190	2 922	1 899
80	6 399	4 036	9 616	5 961	7 028	4 930
90	15 934	12 579	20 465	15 550	16 805	13 328
100	24 479	23 863	38 705	32 489	–	–

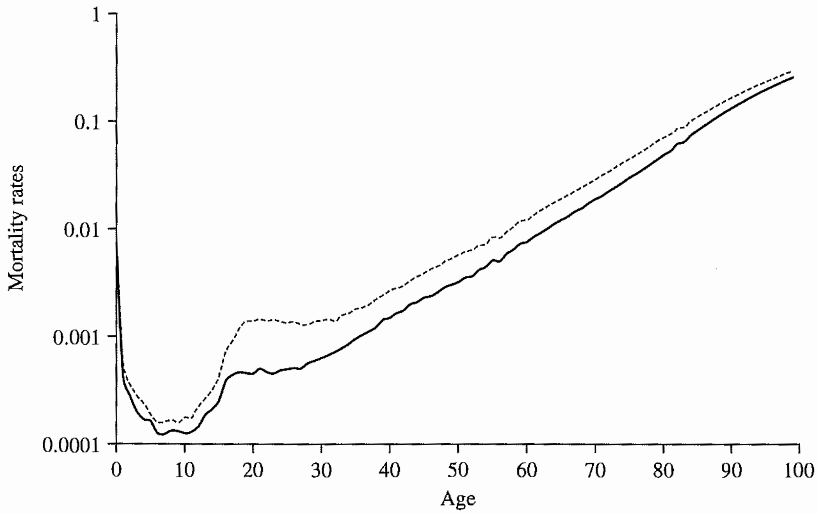

Figure 3.1 US 2002 mortality rates, male (dotted) and female (solid).

shows the US 2002 mortality rates for males and females; the graphs for England & Wales and for Australia are similar. (Note that we have plotted these on a logarithmic scale in order to highlight the main features. Also, although the information plotted consists of values of q_x for $x = 0, 1, \ldots, 99$, we have plotted a continuous line as this gives a clearer representation.) We note the following points from Table 3.2 and Figure 3.1.

- The value of q_0 is relatively high. Mortality rates immediately following birth, *perinatal mortality*, are high due to complications arising from the later stages of pregnancy and from the birth process itself. The value of q_x does not reach this level again until about age 55. This can be seen from Figure 3.1.
- The rate of mortality is much lower after the first year, less than 10% of its level in the first year, and declines until around age 10.
- In Figure 3.1 we see that the pattern of male and female mortality in the late teenage years diverges significantly, with a steeper incline in male mortality. Not only is this feature of mortality for young adult males common for different populations around the world, it is also a feature of historical populations in countries such as the UK where mortality data have been collected for some time. It is sometimes called the accident hump, as many of the deaths causing the 'hump' are accidental.
- Mortality rates increase from age 10, with the accident hump creating a relatively large increase between ages 10 and 20 for males, a more modest increase from ages 20 to 40, and then steady increases from age 40.
- For each age, all six values of q_x are broadly comparable, with, for each country, the rate for a female almost always less than the rate for a male of the same age. The one exception is the Australian Life Table, where q_{100} is slightly higher for a female than for a male. According to the Australian Government Actuary, Australian mortality data indicate that males are subject to lower mortality rates than females at very high ages, although there is some uncertainty as to where the cross-over occurs due to small amounts of data at very old ages.
- The Gompertz model introduced in Chapter 2 is relatively simple, in that it requires only two parameters and has a force of mortality with a simple functional form, $\mu_x = Bc^x$. We stated in Chapter 2 that this model does not provide a good fit across all ages. We can see from Figure 3.1 that the model cannot fit the perinatal mortality, nor the accident hump. However, the mortality rates at later ages are rather better behaved, and the Gompertz model often proves useful over older age ranges. Figure 3.2 shows the older ages US 2002 Males mortality rate curve, along with a Gompertz curve fitted to the US 2002 Table mortality rates. The Gompertz curve provides a pretty close fit – which is a particularly impressive feat, considering that Gompertz proposed the model in 1825.

A final point about Table 3.2 is that we have compared three national life tables using values of the probability of dying within one year, q_x, rather than the force of mortality, μ_x. This is because values of μ_x are not published for any ages for the US Life Tables. Also, values of μ_x are not published for age 0 for

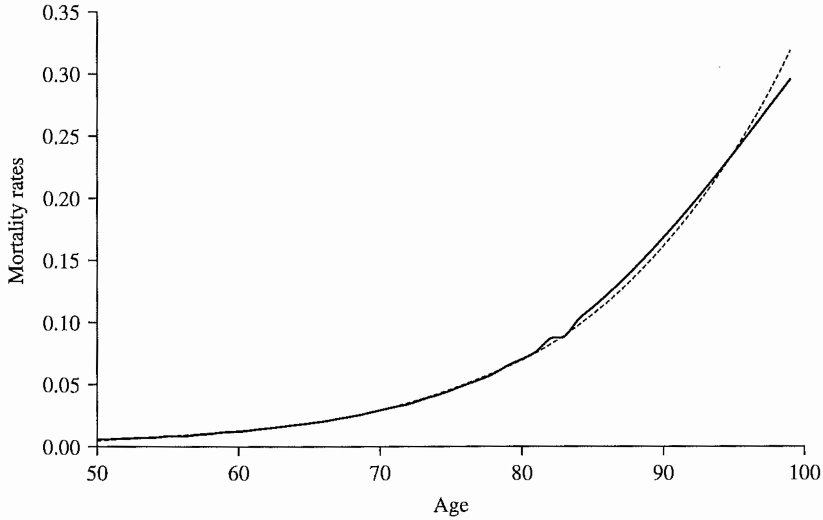

Figure 3.2 US 2002 male mortality rates (solid), with fitted Gompertz mortality rates (dotted).

the other two life tables – there are technical difficulties in the estimation of μ_x within a year in which the force of mortality is changing rapidly, as it does between ages 0 and 1.

3.5 Survival models for life insurance policyholders

Suppose we have to choose a survival model appropriate for a man, currently aged 50 and living in the UK, who has just purchased a 10-year term insurance policy. We could use a national life table, such as English Life Table 15, so that, for example, we could assume that the probability this man dies before age 51 is 0.00464, as shown in Table 3.2. However, in the UK, as in some other countries with well-developed life insurance markets, the mortality experience of people who purchase life insurance policies tends to be different from the population as a whole. The mortality of different types of life insurance policyholders is investigated separately, and life tables appropriate for these groups are published.

Table 3.3 shows values of the force of mortality ($\times 10^5$) at two-year intervals from age 50 to age 60 taken from English Life Table 15, Males (ELTM 15), and from a life table prepared from data relating to term insurance policyholders in the UK in 1999–2002 and which assumes the policyholders *purchased their policies at age 50*. This second set of values comes from Table A14 of a 2006 working paper of the Continuous Mortality Investigation in the UK. Hereafter

Table 3.3 *Values of the force of mortality* $\times 10^5$.

x	ELTM 15	CMI A14
50	440	78
52	549	152
54	679	240
56	845	360
58	1057	454
60	1323	573

we refer to this working paper as CMI, and further details are given at the end of this chapter. The values of the force of mortality for ELTM 15 correspond to the values of q_x shown in Table 3.2.

The striking feature of Table 3.3 is the difference between the two sets of values. The values from the CMI table are very much lower than those from ELTM 15, by a factor of more than 5 at age 50 and by a factor of more than 2 at age 60. There are at least three reasons for this difference.

(a) The data on which the two life tables are based relate to different calendar years; 1990–92 in the case of ELTM 15 and 1999–2002 in the case of CMI. Mortality rates in the UK, as in many other countries, have been decreasing for some years so we might expect rates based on more recent data to be lower (see Section 3.11 for more discussion of mortality trends). However, this explains only a small part of the differences in Table 3.3. An interim life table for England & Wales, based on male population data from 2002–2004, gives $\mu_{50} = 391 \times 10^{-5}$ and $\mu_{60} = 1008 \times 10^{-5}$. Clearly, mortality in England & Wales has improved over the 12-year period, but not to the extent that it matches the CMI values shown in Table 3.3. Other explanations for the differences in Table 3.3 are needed.

(b) A major reason for the difference between the values in Table 3.3 is that ELTM 15 is a life table based on the *whole male population* of England & Wales, whereas CMI Table A14 is based on the experience of males who are *term insurance policyholders*. Within any large group, there are likely to be variations in mortality rates between subgroups. This is true in the case of the population of England and Wales, where social class, defined in terms of occupation, has a significant effect on mortality. Put simply, the better your job, and hence the wealthier you are likely to be, the lower your mortality rates. Given that people who purchase term insurance policies are likely to be among the better paid people in the population, we have an explanation for a large part of the difference between the values in Table 3.3.

(c) The third reason, which is the most significant, arises from the selection process which policyholders must complete before the insurer will issue the insurance policy. The selection, or underwriting process ensures that people who purchase life insurance cover are healthy at the time of purchase, so the CMI figures apply to lives who were all healthy at age 50, when the insurance was purchased. The ELT tables, on the other hand, are based on data from both healthy and unhealthy lives. This is an example of **selection**, and we discuss it in more detail in the following section.

3.6 Life insurance underwriting

The values of the force of mortality in Table 3.3 are based on data for males who purchased term insurance at age 50. CMI Table A14 gives values for different ages at the purchase of the policy ranging from 17 to 90. Values for ages at purchase 50, 52, 54 and 56 are shown in Table 3.4.

There are two significant features of the values in Table 3.4, which can be seen by considering the rows of values for ages 56 and 62.

(a) Consider the row of values for age 56. Each of the four values in this row is the force of mortality at age 56 based on data from the UK over the period 1999–2002 for males who are term insurance policyholders. The only difference is that they purchased their policies at different ages. The more recently the policy was purchased, the lower the force of mortality. For

Table 3.4 *Values of the force of mortality* $\times 10^5$ *from CMI Table A14.*

	Age at purchase of policy			
x	50	52	54	56
50	78	—	—	—
52	152	94	—	—
54	240	186	113	—
56	360	295	227	136
58	454	454	364	278
60	573	573	573	448
62	725	725	725	725
64	917	917	917	917
66	1159	1159	1159	1159

example, for a male who purchased his policy at age 56, the value is 0.00136, whereas for someone of the same age who purchased his policy at age 50, the value is 0.00360.

(b) Now consider the row of values for age 62. These values, all equal to 0.00725, do not depend on whether the policy was purchased at age 50, 52, 54 or 56.

These features are due to life insurance underwriting, which we described in Chapter 1. Recall that the life insurance underwriting process evaluates medical and lifestyle information to assess whether the policyholder is in normal health.

The important point for this discussion is that the mortality rates in the CMI tables are based on individuals accepted for insurance at normal premium rates, that is, individuals who have passed the required health checks. This means, for example, that a man aged 50 who has just purchased a term insurance at the normal premium rate is known to be in good health (assuming the health checks are effective) and so is likely to be much healthier, and hence have a lower mortality rate, than a man of age 50 picked randomly from the population. When this man reaches age 56, we can no longer be certain he is in good health – all we know is that he was in good health six years ago. Hence, his mortality rate at age 56 is higher than that of a man of the same age who has just passed the health checks and been permitted to buy a term insurance policy at normal rates. This explains the differences between the values of the force of mortality at age 56 in Table 3.4.

The effect of passing the health checks at issue eventually wears off, so that at age 62, the force of mortality does not depend on whether the policy was purchased at age 50, 52, 54 or 56. This is point (b) above. However, note that these rates, 0.00725, are still much lower than μ_{62} $(=0.01664)$ from ELTM 15. This is because people who buy term life insurance in the UK tend to have lower mortality than the general population. In fact the population is made up of many heterogeneous lives, and the effect of initial selection is only one area where actuaries have tried to manage the heterogeneity. In the US, there has been a lot of activity recently developing tables for 'preferred lives', who are assumed to be even healthier than the standard insured population. These preferred lives tend to be from higher socio-economic groups. Mortality and wealth are closely linked.

3.7 Select and ultimate survival models

A feature of the survival models studied in Chapter 2 is that probabilities of future survival depend only on the individual's current age. For example, for a given survival model and a given term t, ${}_tp_x$, the probability that an individual currently aged x will survive to age $x + t$, depends only on the current age x. Such survival models are called **aggregate survival models**, because lives are all aggregated together.

The difference between an aggregate survival model and the survival model for term insurance policyholders discussed in Section 3.6 is that in the latter case, probabilities of future survival depend not only on current age but also on how long ago the individual entered the group of policyholders, i.e. when the policy was purchased.

This leads us to the following definition. The mortality of a group of individuals is described by a **select and ultimate survival model**, usually shortened to **select survival model**, if the following statements are true.

(a) Future survival probabilities for an individual in the group depend on the individual's current age *and* on the age at which the individual joined the group.
(b) There is a positive number (generally an integer), which we denote by d, such that if an individual joined the group more than d years ago, future survival probabilities depend only on current age. The initial selection effect is assumed to have worn off after d years.

We use the following terminology for a select survival model. An individual who enters the group at, say, age x, is said to be **selected**, or just **select**, at age x. The period d after which the age at selection has no effect on future survival probabilities is called the **select period** for the model. The mortality that applies to lives after the select period is complete is called the **ultimate** mortality, so that the complete model comprises a select period followed by the ultimate period.

Going back to the term insurance policyholders in Section 3.6, we can identify the 'group' as male term insurance policyholders in the UK. A select survival model is appropriate in this case because passing the health checks at age x indicates that the individual is in good health and so has lower mortality rates than someone of the same age who passed these checks some years ago. There are indications in Table 3.4 that the select period, d, for this group is less than or equal to six years. See point (b) in Section 3.6. In fact, the select period is five years for this particular model. Select periods typically range from one year to 15 years for life insurance mortality models.

For the term insurance policyholders in Section 3.6, being selected at age x meant that the mortality rate for the individual was lower than that of a term insurance policyholder of the same age who had been selected some years earlier. Selection can occur in many different ways and does not always lead to lower mortality rates, as Example 3.8 shows.

Example 3.8 Consider men who need to undergo surgery because they are suffering from a particular disease. The surgery is complicated and there is a probability of only 50% that they will survive for a year following surgery. If they do survive for a year, then they are fully cured and their future mortality follows the Australian Life Tables 2000–02, Males, from which you are given the following values:

$$l_{60} = 89\,777, \quad l_{61} = 89\,015, \quad l_{70} = 77\,946.$$

Calculate

(a) the probability that a man aged 60 who is just about to have surgery will be alive at age 70,
(b) the probability that a man aged 60 who had surgery at age 59 will be alive at age 70, and
(c) the probability that a man aged 60 who had surgery at age 58 will be alive at age 70.

Solution 3.8 In this example, the 'group' is all men who have had the operation. Being selected at age x means having surgery at age x. The select period of the survival model for this group is one year, since if they survive for one year after being 'selected', their future mortality depends only on their current age.

(a) The probability of surviving to age 61 is 0.5. Given that he survives to age 61, the probability of surviving to age 70 is

$$l_{70}/l_{61} = 77\,946/89\,015 = 0.8757.$$

Hence, the probability that this individual survives from age 60 to age 70 is

$$0.5 \times 0.8757 = 0.4378.$$

(b) Since this individual has already survived for one year following surgery, his mortality follows the Australian Life Tables 2000–02, Males. Hence, his probability of surviving to age 70 is

$$l_{70}/l_{60} = 77\,946/89\,777 = 0.8682.$$

(c) Since this individual's surgery was more than one year ago, his future mortality is exactly the same, probabilistically, as the individual in part (b). Hence, his probability of surviving to age 70 is 0.8682. □

Selection is not a feature of national life tables since, ignoring immigration, an individual can enter the population only at age zero. It is an important feature of many survival models based on data from, and hence appropriate to, life insurance policyholders. We can see from Tables 3.3 and 3.4 that its effect on the force of mortality can be considerable. For these reasons, select survival models are important in life insurance mathematics.

The select period may be different for different survival models. For CMI Table A14, which relates to term insurance policyholders, it is five years, as noted above; for CMI Table A2, which relates to whole life and endowment policyholders, the select period is two years.

In the next section we introduce notation and develop some formulae for select survival models.

3.8 Notation and formulae for select survival models

A select survival model represents an extension of the ultimate survival model studied in Chapter 2. In Chapter 2, survival probabilities depended only on the current age of the individual. For a select survival model, probabilities of survival depend on current age and (within the select period) age at selection, i.e. age at joining the group. However, the survival model for those individuals all selected at the same age, say x, depends only on their current age and so fits the assumptions of Chapter 2. This means that, provided we fix and specify the age at selection, we can adapt the notation and formulae developed in Chapter 2 to a select survival model. This leads to the following definitions:

${}_tp_{[x]+s} =$ Pr[a life currently aged $x+s$ who was select at age x survives to age $x+s+t$],

${}_tq_{[x]+s} =$ Pr[a life currently aged $x+s$ who was select at age x dies before age $x+s+t$],

$\mu_{[x]+s}$ is the force of mortality at age $x+s$ for an individual who was select at age x,

$$\mu_{[x]+s} = \lim_{h\to 0^+}\left(\frac{1-{}_hp_{[x]+s}}{h}\right).$$

From these definitions we can derive the following formula

$${}_tp_{[x]+s} = \exp\left\{-\int_0^t \mu_{[x]+s+u}\,du\right\}.$$

This formula is derived precisely as in Chapter 2. It is only the notation which has changed.

For a select survival model with a select period d and for $t \geq d$, that is, for durations at or beyond the select period, the values of $\mu_{[x-t]+t}$, ${}_sp_{[x-t]+t}$ and ${}_{u|s}q_{[x-t]+t}$ do not depend on t, they depend only on the current age x. So, for $t \geq d$ we drop the more detailed notation, $\mu_{[x-t]+t}$, ${}_sp_{[x-t]+t}$ and ${}_{u|s}q_{[x-t]+t}$, and write μ_x, ${}_sp_x$ and ${}_{u|s}q_x$. For values of $t < d$, we refer to, for example, $\mu_{[x-t]+t}$ as being in the **select** part of the survival model and for $t \geq d$ we refer to $\mu_{[x-t]+t}$ ($\equiv \mu_x$) as being in the **ultimate** part of the survival model.

3.9 Select life tables

For an ultimate survival model, as discussed in Chapter 2, the life table $\{l_x\}$ is useful since it can be used to calculate probabilities such as ${}_{t|u}q_x$ for non-negative values of t, u and x. We can construct a **select life table** in a similar way but we need the table to reflect duration as well as age, during the select period. Suppose we wish to construct this table for a select survival model for ages at selection from, say, x_0 (≥ 0). Let d denote the select period, assumed to be an integer number of years.

The construction in this section is for a select life table specified at all ages and not just at integer ages. However, select life tables are usually presented at integer ages only, as is the case for ultimate life tables.

First we consider the survival probabilities of those individuals who were selected at least d years ago and hence are now subject to the ultimate part of the model. The minimum age of these people is $x_0 + d$. For these people, future survival probabilities depend only on their current age and so, as in Chapter 2, we can construct an ultimate life table, $\{l_y\}$, for them from which we can calculate probabilities of surviving to any future age.

Let l_{x_0+d} be an arbitrary positive number. For $y \geq x_0 + d$ we define

$$l_y = {}_{(y-x_0-d)}p_{x_0+d}\, l_{x_0+d}. \tag{3.10}$$

Note that ${}_{(y-x_0-d)}p_{x_0+d} = {}_{(y-x_0-d)}p_{[x_0]+d}$, because d years after selection at age x_0, the probability of future survival depends only on the current age, $x_0 + d$. From this definition we can show that for $y > x \geq x_0 + d$

$$l_y = {}_{y-x}p_x\, l_x. \tag{3.11}$$

This follows because

$$\begin{aligned} l_y &= \left({}_{(y-x_0-d)}p_{x_0+d}\right) l_{x_0+d} \\ &= \left({}_{y-x}p_{[x_0]+x-x_0}\right) \left({}_{(x-x_0-d)}p_{[x_0]+d}\right) l_{x_0+d} \\ &= \left({}_{y-x}p_x\right) \left({}_{(x-x_0-d)}p_{x_0+d}\right) l_{x_0+d} \\ &= {}_{y-x}p_x\, l_x. \end{aligned}$$

This shows that within the ultimate part of the model we can interpret l_y as the expected number of survivors to age y out of l_x lives currently aged x ($< y$), who were select at least d years ago.

Formula (3.10) defines the life table within the ultimate part of the model. Next, we need to define the life table within the select period. We do this for a life select at age x by 'working backwards' from the value of l_{x+d}. For $x \geq x_0$ and for $0 \leq t \leq d$, we define

$$l_{[x]+t} = \frac{l_{x+d}}{{}_{d-t}p_{[x]+t}} \tag{3.12}$$

which means that if we had $l_{[x]+t}$ lives aged $x+t$, selected t years ago, then the expected number of survivors to age $x+d$ is l_{x+d}. This defines the select part of the life table.

Example 3.9 For $y \geq x+d > x+s > x+t \geq x \geq x_0$, show that

$$ {}_{y-x-t}p_{[x]+t} = \frac{l_y}{l_{[x]+t}} \tag{3.13} $$

and

$$ {}_{s-t}p_{[x]+t} = \frac{l_{[x]+s}}{l_{[x]+t}}. \tag{3.14} $$

Solution 3.9 First,

$$ \begin{aligned} {}_{y-x-t}p_{[x]+t} &= {}_{y-x-d}p_{[x]+d} \times {}_{d-t}p_{[x]+t} \\ &= {}_{y-x-d}p_{x+d} \times {}_{d-t}p_{[x]+t} \\ &= \frac{l_y}{l_{x+d}} \frac{l_{x+d}}{l_{[x]+t}} \\ &= \frac{l_y}{l_{[x]+t}}, \end{aligned} $$

which proves (3.13). Second,

$$ \begin{aligned} {}_{s-t}p_{[x]+t} &= \frac{{}_{d-t}p_{[x]+t}}{{}_{d-s}p_{[x]+s}} \\ &= \frac{l_{x+d}}{l_{[x]+t}} \frac{l_{[x]+s}}{l_{x+d}} \\ &= \frac{l_{[x]+s}}{l_{[x]+t}}, \end{aligned} $$

which proves (3.14). □

This example, together with formula (3.11), shows that our construction preserves the interpretation of the ls as expected numbers of survivors within both the ultimate and the select parts of the model. For example, suppose we have $l_{[x]+t}$ individuals currently aged $x+t$ who were select at age x. Then, since ${}_{y-x-t}p_{[x]+t}$ is the probability that any one of them survives to age y, we can see from formula (3.13) that l_y is the expected number of survivors to age y. For $0 \leq t \leq s \leq d$, formula (3.14) shows that $l_{[x]+s}$ can be interpreted as the expected number of survivors to age $x+s$ out of $l_{[x]+t}$ lives currently aged $x+t$ who were select at age x.

Example 3.10 Write an expression for ${}_{2|6}q_{[30]+2}$ in terms of $l_{[x]+t}$ and l_y for appropriate x, t and y, assuming a select period of five years.

Solution 3.10 Note that ${}_{2|6}q_{[30]+2}$ is the probability that a life currently aged 32, who was select at age 30, will die between ages 34 and 40. We can write this probability as the product of the probabilities of the following events:

- a life aged 32, who was select at age 30, will survive to age 34, and,
- a life aged 34, who was select at age 30, will die before age 40.

Hence,

$$\begin{aligned} {}_{2|6}q_{[30]+2} &= {}_2p_{[30]+2}\ {}_6q_{[30]+4} \\ &= \frac{l_{[30]+4}}{l_{[30]+2}}\left(1-\frac{l_{[30]+10}}{l_{[30]+4}}\right) \\ &= \frac{l_{[30]+4}-l_{40}}{l_{[30]+2}}. \end{aligned}$$

Note that $l_{[30]+10} \equiv l_{40}$ since 10 years is longer than the select period for this survival model. □

Table 3.5 *Extract from US Life Tables, 2002.*

x	l_x
70	80 556
71	79 026
72	77 410
73	75 666
74	73 802
75	71 800

Example 3.11 A select survival model has a select period of three years. Its ultimate mortality is equivalent to the US Life Tables, 2002, Females. Some l_x values for this table are shown in Table 3.5.

You are given that for all ages $x \geq 65$,

$$p_{[x]} = 0.999, \quad p_{[x-1]+1} = 0.998, \quad p_{[x-2]+2} = 0.997.$$

Calculate the probability that a woman currently aged 70 will survive to age 75 given that

(a) she was select at age 67,
(b) she was select at age 68,
(c) she was select at age 69, and
(d) she is select at age 70.

Solution 3.11 (a) Since the woman was select three years ago and the select period for this model is three years, she is now subject to the ultimate part of the survival model. Hence the probability she survives to age 75 is l_{75}/l_{70}, where the ls are taken from US Life Tables, 2002, Females. The required probability is

$$_5p_{70} = 71\,800/80\,556 = 0.8913.$$

(b) We have

$$_5p_{[68]+2} = \frac{l_{[68]+2+5}}{l_{[68]+2}} = \frac{l_{75}}{l_{[68]+2}} = \frac{71\,800}{l_{[68]+2}}.$$

We calculate $l_{[68]+2}$ by noting that

$$l_{[68]+2} \times p_{[68]+2} = l_{[68]+3} = l_{71} = 79\,026.$$

We are given that $p_{[68]+2} = 0.997$. Hence, $l_{[68]+2} = 79\,264$ and so

$$_5p_{[68]+2} = 0.9058.$$

(c) We have

$$_5p_{[69]+1} = \frac{l_{[69]+1+5}}{l_{[69]+1}} = \frac{l_{75}}{l_{[69]+1}} = \frac{71\,800}{l_{[69]+1}}.$$

We calculate $l_{[69]+1}$ by noting that

$$l_{[69]+1} \times p_{[69]+1} \times p_{[69]+2} = l_{[69]+3} = l_{72} = 77\,410.$$

We are given that $p_{[69]+1} = 0.998$ and $p_{[69]+2} = 0.997$. Hence, $l_{[69]+1} = 77\,799$ and so

$$_5p_{[69]+1} = 0.9229.$$

(d) We have

$$_5p_{[70]} = \frac{l_{[70]+5}}{l_{[70]}} = \frac{l_{75}}{l_{[70]}} = \frac{71\,800}{l_{[70]}}.$$

Proceeding as in (b) and (c),

$$l_{[70]} \times p_{[70]} \times p_{[70]+1} \times p_{[70]+2} = l_{[70]+3} = l_{73} = 75\,666,$$

giving

$$l_{[70]} = 75\,666/(0.997 \times 0.998 \times 0.999) = 76\,122.$$

Hence

$$_5p_{[70]} = 0.9432.$$

□

Table 3.6 *CMI Table A5: male non-smokers who have whole life or endowment policies.*

Age, x	Duration 0 $q_{[x]}$	Duration 1 $q_{[x-1]+1}$	Duration 2+ q_x
60	0.003469	0.004539	0.004760
61	0.003856	0.005059	0.005351
62	0.004291	0.005644	0.006021
63	0.004779	0.006304	0.006781
⋮	⋮	⋮	⋮
70	0.010519	0.014068	0.015786
71	0.011858	0.015868	0.017832
72	0.013401	0.017931	0.020145
73	0.015184	0.020302	0.022759
74	0.017253	0.023034	0.025712
75	0.019664	0.026196	0.029048

Example 3.12 CMI Table A5 is based on UK data from 1999 to 2002 for male non-smokers who are whole life or endowment insurance policyholders. It has a select period of two years. An extract from this table, showing values of $q_{[x-t]+t}$, is given in Table 3.6. Use this survival model to calculate the following probabilities:

(a) ${}_4p_{[70]}$,

(b) ${}_3q_{[60]+1}$, and

(c) ${}_{2|}q_{73}$.

Solution 3.12 Note that CMI Table A5 gives values of $q_{[x-t]+t}$ for $t = 0$ and $t = 1$ and also for $t \geq 2$. Since the select period is two years $q_{[x-t]+t} \equiv q_x$ for $t \geq 2$. Note also that each row of the table relates to a man *currently* aged x, where x is given in the first column. Select life tables, tabulated at integer ages, can be set out in different ways – for example, each row could relate to a fixed age at selection – so care needs to be taken when using such tables.

(a) We calculate ${}_4p_{[70]}$ as

$$\begin{aligned}
{}_4p_{[70]} &= p_{[70]}\, p_{[70]+1}\, p_{[70]+2}\, p_{[70]+3} \\
&= p_{[70]}\, p_{[70]+1}\, p_{72}\, p_{73} \\
&= (1 - q_{[70]})\,(1 - q_{[70]+1})\,(1 - q_{72})\,(1 - q_{73}) \\
&= 0.989481 \times 0.984132 \times 0.979855 \times 0.977241 \\
&= 0.932447.
\end{aligned}$$

(b) We calculate ${}_3q_{[60]+1}$ as

$$\begin{aligned}{}_3q_{[60]+1} &= q_{[60]+1} + p_{[60]+1}\, q_{62} + p_{[60]+1}\, p_{62}\, q_{63}\\ &= q_{[60]+1} + (1 - q_{[60]+1})\, q_{62} + (1 - q_{[60]+1})\,(1 - q_{62})\, q_{63}\\ &= 0.005059 + 0.994941 \times 0.006021\\ &\quad + 0.994941 \times 0.993979 \times 0.006781\\ &= 0.017756.\end{aligned}$$

(c) We calculate ${}_{2|}q_{73}$ as

$$\begin{aligned}{}_{2|}q_{73} &= {}_2p_{73}\, q_{75}\\ &= (1 - q_{73})\,(1 - q_{74})\, q_{75}\\ &= 0.977241 \times 0.974288 \times 0.029048\\ &= 0.027657.\end{aligned}$$

□

Example 3.13 A select survival model has a two-year select period and is specified as follows. The ultimate part of the model follows Makeham's law, so that

$$\mu_x = A + Bc^x$$

where $A = 0.00022$, $B = 2.7 \times 10^{-6}$ and $c = 1.124$. The select part of the model is such that for $0 \le s \le 2$,

$$\mu_{[x]+s} = 0.9^{2-s}\,\mu_{x+s}.$$

Starting with $l_{20} = 100\,000$, calculate values of

(a) l_x for $x = 21, 22, \ldots, 82$,
(b) $l_{[x]+1}$ for $x = 20, 21, \ldots, 80$, and,
(c) $l_{[x]}$ for $x = 20, 21, \ldots, 80$.

Solution 3.13 First, note that

$${}_tp_x = \exp\left\{-At - \frac{B}{\log c}c^x(c^t - 1)\right\}$$

and for $0 \le t \le 2$,

$$\begin{aligned}{}_tp_{[x]} &= \exp\left\{-\int_0^t \mu_{[x]+s}ds\right\}\\ &= \exp\left\{0.9^{2-t}\left(\frac{1-0.9^t}{\log(0.9)}A + \frac{c^t - 0.9^t}{\log(0.9/c)}Bc^x\right)\right\}.\end{aligned} \qquad (3.15)$$

(a) Values of l_x can be calculated recursively from

$$l_x = p_{x-1}l_{x-1} \quad \text{for} \quad x = 21, 22, \ldots, 82.$$

Table 3.7 *Select life table with a two-year select period, Example 3.13.*

x	$l_{[x]}$	$l_{[x]+1}$	l_{x+2}	$x+2$	x	$l_{[x]}$	$l_{[x]+1}$	l_{x+2}	$x+2$
			100 000.00	20	50	98 552.51	98 450.67	98 326.19	52
			99 975.04	21	51	98 430.98	98 318.95	98 181.77	53
20	99 995.08	99 973.75	99 949.71	22	52	98 297.24	98 173.79	98 022.38	54
21	99 970.04	99 948.40	99 923.98	23	53	98 149.81	98 013.56	97 846.20	55
22	99 944.63	99 922.65	99 897.79	24	54	97 987.03	97 836.44	97 651.21	56
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
47	98 856.38	98 778.94	98 684.88	49	79	77 465.70	75 531.88	73 186.31	81
48	98 764.09	98 679.44	98 576.37	50	80	75 153.97	73 050.22	70 507.19	82
49	98 663.15	98 570.40	98 457.24	51					

(b) Values of $l_{[x]+1}$ can be calculated from

$$l_{[x]+1} = l_{x+2}/p_{[x]+1} \quad \text{for} \quad x = 20, 21, \ldots, 80.$$

(c) Values of $l_{[x]}$ can be calculated from

$$l_{[x]} = l_{x+2}/{}_2p_{[x]} \quad \text{for} \quad x = 20, 21, \ldots, 80.$$

Sample values are shown in Table 3.7. The full table up to age 100 is given in Table D.1 in Appendix D. □

This model is used extensively throughout this book for examples and exercises. We call it the **Standard Select Survival Model** in future chapters.

The ultimate part of the model, which is a Makeham model with $A = 0.00022$, $B = 2.7 \times 10^{-6}$ and $c = 1.124$, is also used in many examples and exercises where a select model is not required. We call this the **Standard Ultimate Survival Model**.

3.10 Some comments on heterogeneity in mortality

We noted in Section 3.5 the significant difference between the mortality of the population as a whole, and the mortality of insured lives. It is worth noting, further, that there is also considerable variability when we look at the mortality experience of different groups of insurance company customers and pension plan members. Of course, male and female mortality differs significantly, in shape and level. Actuaries will generally use separate survival models for men and women when this does not breach discrimination laws. Smoker and non-smoker mortality differences are very important in whole life and term insurance; smoker mortality is substantially higher at all ages for both sexes, and separate smoker/non-smoker mortality tables are in common use.

In addition, insurers will generally use product-specific mortality tables for different types of contracts. Individuals who purchase immediate or deferred annuities may have different mortality from those purchasing term insurance. Insurance is sometimes purchased under group contracts, for example by an employer to provide death-in-service insurance for employees. The mortality experience from these contracts will generally be different from the experience of policyholders holding individual contracts. The mortality experience of pension plan members may differ from the experience of lives who purchase individual pension policies from an insurance company. Interestingly, the differences in mortality experience between these groups will depend significantly on country. Studies of mortality have shown, though, that the following principles apply quite generally.

◇ Wealthier lives experience lighter mortality overall than less wealthy lives.
◇ There will be some impact on the mortality experience from self-selection; an individual will only purchase an annuity if he or she is confident of living long enough to benefit. An individual who has some reason to anticipate heavier mortality is more likely to purchase term insurance. While underwriting can identify some selective factors, there may be other information that cannot be gleaned from the underwriting process (at least not without excessive cost). So those buying term insurance might be expected to have slightly heavier mortality than those buying whole life insurance, and those buying annuities might be expected to have lighter mortality.
◇ The more rigorous the underwriting, the lighter the resulting mortality experience. For group insurance, there will be minimal underwriting. Each person hired by the employer will be covered by the insurance policy almost immediately; the insurer does not get to accept or reject the additional employee, and will rarely be given information sufficient for underwriting decisions. However, the employee must be healthy enough to be hired, which gives some selection information.

All of these factors may be confounded by tax or legislative systems that encourage or require certain types of contracts. In the UK, it is very common for retirement savings proceeds to be converted to life annuities. In other countries, including the USA, this is much less common. Consequently, the type of person who buys an annuity in the USA might be quite a different (and more self-select) customer than the typical individual buying an annuity in the UK.

3.11 Mortality trends

A challenge in developing and using survival models is that survival probabilities are not constant over time. Commonly, mortality experience gets lighter over time. In most countries, for the period of reliable records, each generation, on average, lives longer than the previous generation. This can be explained by advances in health care and by improved standards of living. Of course, there are exceptions, such as mortality shocks from war or from disease, or declining life expectancy in countries where access to health care worsens, often because of civil upheaval. The changes in mortality over time are sometimes separated into three components: trend, shock and idiosyncratic. The trend describes the gradual reduction in mortality rates over time. The shock describes a short-term jump in mortality rates from war or pandemic disease. The idiosyncratic risk describes year to year random variation that does not come from trend or shock, though it is often difficult to distinguish these changes.

While the shock and idiosyncratic risks are inherently unpredictable, we can often identify trends in mortality by examining mortality patterns over a number of years. We can then allow for mortality improvement by using a survival model which depends on both age and calendar year. A common model for projecting mortality is to assume that mortality rates at each age are decreasing annually by a constant factor, which depends on the age and sex of the individual. That is, suppose $q(x, Y)$ denotes the mortality rate for a life aged x in year Y, so that $q(x, 0)$ denotes the mortality rate at age x for a baseline year, $Y = 0$. Then, the estimated one-year mortality probability for a life aged x at time $Y = s$ is

$$q(x, s) = q(x, 0)\, r_x^s \quad \text{where} \quad 0 < r_x \leq 1.$$

The r_x terms are called mortality **reduction factors**, and typical values are in the range 0.95 to 1, where the higher values (implying less reduction) tend to apply at older ages. Using $r_x = 1$ for the oldest ages reflects the fact that, although many people are living longer than previous generations, there is little or no increase in the maximum age attained; the change is that a greater proportion of lives survive to older ages. In practice, the reduction factors are applied for integer values of s.

Figure 3.3 shows reduction factors for females based on mortality in Australia in the 25 years prior to the production of Australian Life Tables 2000–02. This shows the greatest reduction in mortality rates has occurred at the youngest ages, that mortality rates have not fallen greatly from mid-teens to late thirties, and that as age increases from around age 60, reduction factors are increasing.

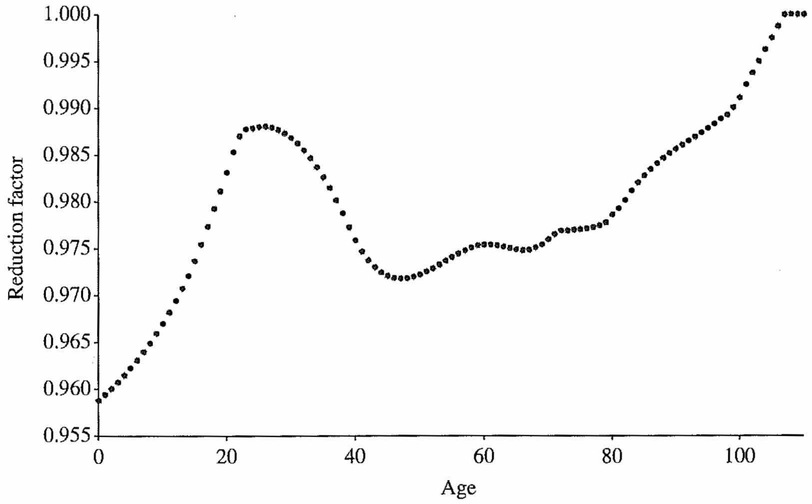

Figure 3.3 Reduction factors, r_x, based on Australian female mortality.

Given a baseline survival model, with mortality rates $q(x,0) = q_x$, say, and a set of age-based reduction factors, r_x, we can calculate survival probabilities from the baseline year, ${}_tp(x,0)$, say, as

$$\begin{aligned} {}_tp(x,0) &= p(x,0)\,p(x+1,1)\ldots p(x+t-1,t-1) \\ &= (1-q_x)\,(1-q_{x+1}\,r_{x+1})\left(1-q_{x+2}\,r_{x+2}^2\right)\ldots\left(1-q_{x+t-1}\,r_{x+t-1}^{t-1}\right). \end{aligned} \tag{3.16}$$

Some survival models developed for actuarial applications implicitly contain some allowance for mortality improvement. When selecting a survival model to use for valuation and risk management, it is important to verify the projection assumptions.

The use of reduction factors allows for predictable improvements in life expectancy. However, if the improvements are underestimated, then mortality experience will be lighter than expected, leading to losses on annuity and pension contracts. This risk, called longevity risk, is of great recent interest, as mortality rates have declined in many countries at a much faster rate than anticipated. As a result, there has been increased interest in stochastic mortality models, where the force of mortality in future years follows a stochastic process which incorporates both predictable and random changes in longevity, as well as pandemic-type shock effects.

Table 3.8 shows the effect of reduction factors on the calculation of expectation of life. In this table we show values of e_x under two scenarios. The

Table 3.8 *Values of e_x without and with reduction factors.*

x	Scenario 1 e_x	Scenario 2 e_x	x	Scenario 1 e_x	Scenario 2 e_x
0	82.36	94.97	50	34.01	38.80
10	72.86	84.18	60	24.94	28.06
20	63.00	72.84	70	16.57	18.26
30	53.22	61.47	80	9.48	10.15
40	43.51	50.06	90	4.83	5.03

first scenario is that no reduction factors apply to the female mortality rates of Australian Life Tables 2000–02, and the second scenario is that the reduction factors shown in Figure 3.3 apply, with survival probabilities calculated according to formula (3.16).

The values in Table 3.8 show that the application of reduction factors to mortality rates can have a significant effect on expected future lifetime, particularly at younger ages. However, the values in this table should be treated with caution. The key underlying assumption in the calculations is that mortality rates will continue to reduce in the future, and this assumption is questionable. Nevertheless, the table does illustrate the basic fact that allowing for mortality improvement may have a significant effect on expectation of life.

3.12 Notes and further reading

The mortality rates in Section 3.4 are drawn from the following sources:

- Australian Life Tables 2000–02 were produced by the Australian Government Actuary (2004).
- English Life Table 15 was prepared by the UK Government Actuary and published by the Office for National Statistics (1997).
- US Life Tables 2002 were prepared in the Division of Vital Statistics of the National Center for Health Statistics in the US – see Arias (2004).

The Continuous Mortality Investigation in the UK has been ongoing for many years. Findings on mortality and morbidity experience of UK policyholders are published via a series of formal reports and working papers. In this chapter we have drawn on CMI (2006).

In Section 3.5 we noted that there can be considerable variability in the mortality experience of different groups in a national population. Coleman and Salt (1992) give a very good account of this variability in the UK population.

The paper by Gompertz (1825), who was the Actuary of the Alliance Insurance Company of London, introduced the force of mortality concept.

See, for example, Lee and Carter (1992), Li *et al.* (2010) or Cairns *et al.* (2009) for more detailed information about stochastic mortality models.

3.13 Exercises

Exercise 3.1 Sketch the following as functions of age x for a typical (human) population, and comment on the major features.

(a) μ_x,

(b) l_x, and

(c) d_x.

Exercise 3.2 You are given the following life table extract.

Age, x	l_x
52	89 948
53	89 089
54	88 176
55	87 208
56	86 181
57	85 093
58	83 940
59	82 719
60	81 429

Calculate

(a) ${}_{0.2}q_{52.4}$ assuming UDD (fractional age assumption),

(b) ${}_{0.2}q_{52.4}$ assuming constant force of mortality (fractional age assumption),

(c) ${}_{5.7}p_{52.4}$ assuming UDD,

(d) ${}_{5.7}p_{52.4}$ assuming constant force of mortality,

(e) ${}_{3.2|2.5}q_{52.4}$ assuming UDD, and

(f) ${}_{3.2|2.5}q_{52.4}$ assuming constant force of mortality.

Exercise 3.3 Table 3.9 is an extract from a (hypothetical) select life table with a select period of two years. Note carefully the layout – each row relates to a fixed age at selection.

Table 3.9 *Extract from a (hypothetical) select life table.*

x	$l_{[x]}$	$l_{[x]+1}$	l_{x+2}	$x+2$
75	15 930	15 668	15 286	77
76	15 508	15 224	14 816	78
77	15 050	14 744	14 310	79
⋮	⋮	⋮	⋮	⋮
80			12 576	82
81			11 928	83
82			11 250	84
83			10 542	85
84			9 812	86
85			9 064	87

Table 3.10 *Mortality rates for female non-smokers with term insurance.*

Age, x	Duration 0 $q_{[x]}$	Duration 1 $q_{[x-1]+1}$	Duration 2 $q_{[x-2]+2}$	Duration 3 $q_{[x-3]+3}$	Duration 4 $q_{[x-4]+4}$	Duration 5+ q_x
69	0.003974	0.004979	0.005984	0.006989	0.007994	0.009458
70	0.004285	0.005411	0.006537	0.007663	0.008790	0.010599
71	0.004704	0.005967	0.007229	0.008491	0.009754	0.011880
72	0.005236	0.006651	0.008066	0.009481	0.010896	0.013318
73	0.005870	0.007456	0.009043	0.010629	0.012216	0.014931
74	0.006582	0.008361	0.010140	0.011919	0.013698	0.016742
75	0.007381	0.009376	0.011370	0.013365	0.015360	0.018774
76	0.008277	0.010514	0.012751	0.014988	0.017225	0.021053
77	0.009281	0.011790	0.014299	0.016807	0.019316	0.023609

Use this table to calculate

(a) the probability that a life currently aged 75 who has just been selected will survive to age 85,

(b) the probability that a life currently aged 76 who was selected one year ago will die between ages 85 and 87, and

(c) ${}_{4|2}q_{[77]+1}$.

Exercise 3.4 CMI Table A23 is based on UK data from 1999 to 2002 for female non-smokers who are term insurance policyholders. It has a select period of five years. An extract from this table, showing values of $q_{[x-t]+t}$, is given in Table 3.10.

Use this survival model to calculate

(a) ${}_2p_{[72]}$,
(b) ${}_3q_{[73]+2}$,
(c) ${}_{1|}q_{[65]+4}$, and
(d) ${}_7p_{[70]}$.

Table 3.11 *Mortality rates for female smokers with term insurance.*

Age x	Duration 0 $q_{[x]}$	Duration 1 $q_{[x-1]+1}$	Duration 2 $q_{[x-2]+2}$	Duration 3 $q_{[x-3]+3}$	Duration 4 $q_{[x-4]+4}$	Duration 5+ q_x
70	0.010373	0.013099	0.015826	0.018552	0.021279	0.026019
71	0.011298	0.014330	0.017362	0.020393	0.023425	0.028932
72	0.012458	0.015825	0.019192	0.022559	0.025926	0.032133
73	0.013818	0.017553	0.021288	0.025023	0.028758	0.035643
74	0.015308	0.019446	0.023584	0.027721	0.031859	0.039486
75	0.016937	0.021514	0.026092	0.030670	0.035248	0.043686
76	0.018714	0.023772	0.028830	0.033888	0.038946	0.048270
77	0.020649	0.026230	0.031812	0.037393	0.042974	0.053262

Exercise 3.5 CMI Table A21 is based on UK data from 1999 to 2002 for female smokers who are term insurance policyholders. It has a select period of five years. An extract from this table, showing values of $q_{[x-t]+t}$, is given in Table 3.11. Calculate

(a) ${}_7p_{[70]}$,
(b) ${}_{1|2}q_{[70]+2}$, and
(c) ${}_{3.8}q_{[70]+0.2}$ assuming UDD.

Exercise 3.6 A select survival model has a select period of three years. Calculate ${}_3p_{53}$, given that

$$q_{[50]} = 0.01601, \quad {}_2p_{[50]} = 0.96411,$$
$$ {}_{2|}q_{[50]} = 0.02410, \quad {}_{2|3}q_{[50]+1} = 0.09272.$$

Exercise 3.7 When posted overseas to country A at age x, the employees of a large company are subject to a force of mortality such that, at exact duration t years after arrival overseas ($t = 0, 1, 2, 3, 4$),

$$q^A_{[x]+t} = (6 - t)q_{x+t}$$

where q_{x+t} is on the basis of US Life Tables, 2002, Females. For those who have lived in country A for at least five years the force of mortality at each age

Table 3.12 *An extract from the United States Life Tables, 2002, Females.*

Age, x	l_x
30	98 424
31	98 362
32	98 296
33	98 225
34	98 148
35	98 064
⋮	⋮
40	97 500

is 50% greater than that of US Life Tables, 2002, Females, at the same age. Some l_x values for this table are shown in Table 3.12.
Calculate the probability that an employee posted to country A at age 30 will survive to age 40 if she remains in that country.

Exercise 3.8 A special survival model has a select period of three years. Functions for this model are denoted by an asterisk, *. Functions without an asterisk are taken from the Canada Life Tables 2000–02, Males. You are given that, for all values of x,

$$p^*_{[x]} = {}_4p_{x-5}; \quad p^*_{[x]+1} = {}_3p_{x-1}; \quad p^*_{[x]+2} = {}_2p_{x+2}; \quad p^*_x = p_{x+1}.$$

A life table, tabulated at integer ages, is constructed on the basis of the special survival model and the value of l^*_{25} is taken as 98 363 (i.e. l_{26} for Canada Life Tables 2000–02, Males). Some l_x values for this table are shown in Table 3.13.

(a) Construct the $l^*_{[x]}$, $l^*_{[x]+1}$, $l^*_{[x]+2}$, and l^*_{x+3} columns for $x = 20, 21, 22$.
(b) Calculate ${}_{2|38}q^*_{[21]+1}$, ${}_{40}p^*_{[22]}$, ${}_{40}p^*_{[21]+1}$, ${}_{40}p^*_{[20]+2}$, and ${}_{40}p^*_{22}$.

Exercise 3.9 (a) Show that a constant force of mortality between integer ages implies that the distribution of R_x, the fractional part of the future life time, conditional on $K_x = k$, has the following truncated exponential distribution for integer x, for $0 \le s < 1$ and for $k = 0, 1, \ldots$

$$\Pr[R_x \le s \mid K_x = k] = \frac{1 - \exp\{-\mu^*_{x+k}s\}}{1 - \exp\{-\mu^*_{x+k}\}} \tag{3.17}$$

where $\mu^*_{x+k} = -\log p_{x+k}$.

Table 3.13 *Canada Life Tables 2000–02, Males.*

Age, x	l_x
15	99 180
16	99 135
17	99 079
18	99 014
19	98 942
20	98 866
21	98 785
22	98 700
23	98 615
24	98 529
25	98 444
26	98 363
⋮	⋮
62	87 503
63	86 455
64	85 313
65	84 074

(b) Show that if formula (3.17) holds for $k = 0, 1, 2, \ldots$ then the force of mortality is constant between integer ages.

Exercise 3.10 Verify formula (3.15).

Answers to selected exercises

3.2 (a) 0.001917
(b) 0.001917
(c) 0.935422
(d) 0.935423
(e) 0.030957
(f) 0.030950

3.3 (a) 0.66177
(b) 0.09433
(c) 0.08993

3.4 (a) 0.987347
(b) 0.044998
(c) 0.010514
(d) 0.920271

3.5 (a) 0.821929

(b) 0.055008

(c) 0.065276

3.6 0.90294

3.7 0.977497

3.8 (a) The values are as follows:

x	$l^*_{[x]}$	$l^*_{[x]+1}$	$l^*_{[x]+2}$	l_{x+3}
20	99 180	98 942	98 700	98 529
21	99 135	98 866	98 615	98 444
22	99 079	98 785	98 529	98 363

(b) 0.121265, 0.872587, 0.874466, 0.875937, 0.876692.