OPTIMIZING THE TERM OF AN INVESTIGATION INTO DECREMENTAL RATES

BY G. C. TAYLOR, B.A., F.I.A.

THE PROBLEM

One of the earliest problems faced by actuaries and one which has persisted to modern times is that of designing and carrying out an investigation into a certain set of decremental rates, especially mortality rates. Many investigations have been conducted with a view to the derivation of a table of mortality rates, some of them of crucial importance to the British life insurance industry. Such tables as the OM, the A1924/29 and the A1949/52 tables have been so widely known and used that mention of them is scarcely necessary. However, despite the importance of such investigations, little detailed attention has been given to methods for determining a suitable term of the experience. The OM table was derived from data covering a term of 30 years, A1924/29 from a 6-year term and A1949/52 from a 4-year term. Were these terms reasonable or not?

The reason that the term of such an investigation is problematical is well-known in a general way. The problem is that, historically, mortality varies with time, and so any experience incorporates, almost inevitably, a certain amount of heterogeneity from this cause. Thus the term of the investigation should not be too long. On the other hand, it is clear that the results will not be improved by decreasing the term of the investigation indefinitely, for as the amount of data becomes smaller so do the variances of one’s estimates of the mortality rates increase, making the estimates less reliable. Actuarial text-books are always careful to issue a warning on this point. For example, Benjamin and Haycocks (1970, p. 177) say: ‘Care is usually taken, therefore, to keep the period of investigation as short as is compatible with a reasonable volume of data in order to minimize the heterogeneity due to time changes.’ The same problem is also dealt with briefly by Anderson and Dow (1952, pp. 6–7). However, such statements are usually qualitative, merely noting that the term of the investigation must not be too short, nor too long. It is left to one’s intuition to strike the required balance.

The reports on the various mortality investigations which have been made in the past do not go any deeper than this. For example, choosing two investigations at random, the report on The American–Canadian Mortality Investigation, 1900–15 (J.I.A. 1917, 51, 260) simply states without further justification that ‘it was decided to take the experience during the years 1900 to 1915’, and similarly the report on the mortality of life annuitants for 1900–20 by Elderton and Oakley (1923, p. 43) simply states ‘it has been arranged to investigate the mortality which occurred in each of three periods: I, 1900–1907; II, 1907–1914; III, 1914–1920’.

The problem under consideration can be summarized as follows. Given that
the object of a mortality investigation is to estimate a set of current mortality rates, it is acknowledged that

(i) Increasing the term of the investigation increases the bias of one's estimators by introducing further heterogeneity to the data.
(ii) Decreasing the term of the investigation increases the unreliability of one's estimators in the sense of increasing their variances.

What term, then, produces the best attainable balance between the two opposing forces?

It is the purpose of this paper to provide an answer to this question.

Hitherto, this discussion has been predominantly in terms of mortality investigations. However, the methods developed here will have wider application than that. An investigation to determine mortality rates is only a special case of investigations into decremental rates in general. Thus, the techniques of this paper apply equally well to investigations into ill-health retirement rates in a pension fund, monthly rates of falling pregnant, failure rates of bus engines, in fact in any investigation at all into decremental rates.

**OPTIMIZING THE TERM IN RESPECT OF A SINGLE MORTALITY RATE**

In this and subsequent sections the practice of setting the discussion in the context of an investigation to determine rates of mortality will be continued in order that ideas be fixed. Nevertheless, the wider applicability of the techniques, as mentioned in the last paragraph of the previous section, is implied throughout.

Although most investigations will be concerned with estimating rates for a number of ages, it is instructive to consider initially an investigation into mortality at one particular age. This approach will illustrate the essence of the technique which is then extended to the more complicated situation of the next section.

Suppose that the mortality rate prevailing at present is $q_x$. Suppose also that during the investigation $n_x$ lives are observed from age $x$ to age $(x + 1)$ (fractional exposures are ignored here), the probability of death for the $i$th life being $q_{xi}$. Then the usual (ungraduated) estimate of $q_x$ is

$$
\hat{q}_x = \frac{1}{n} \sum_{i=1}^{n} \hat{q}_{xi},
$$

where

$$
\hat{q}_{xi} = 0 \text{ if the } i\text{th} \text{ life survives};
= 1 \text{ if the } i\text{th} \text{ life dies};
\hat{q}_{xi} \sim \text{Bin}(1, q_{xi}).
$$

The mean square deviation of the estimate $\hat{q}_x$ from the parameter $q_x$ is
Now, if it is assumed that the mortality experiences of the various lives are independent, then equation (1) gives

\[ \text{Var } \hat{q}_x = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var } \hat{q}_{xi} \]
\[ = \frac{1}{n^2} \sum_{i=1}^{n} q_{xi} (1 - q_{xi}) \] (by relation (2))
\[ = \frac{1}{n} V_{x(n)}, \text{ say.} \]

If \( E_{x(n)} \) is written for \( \delta(\hat{q}_x) \), equation (3) becomes

\[ y_{x(n)} = \frac{1}{n} V_{x(n)} + (E_{x(n)} - q_x)^2. \] (4)

If the principle is now adopted that the smaller the mean square deviation from \( q_x \) of the estimate of \( q_x \), the better is that estimate, then the best set of lives to use for the mortality investigation is the set for which \( y_{x(n)} \) is a minimum. Therefore, allotting \( i = 1 \) to the most recently observed life, \( i = 2 \) to the second most recently observed life, etc., one seeks a value of \( n \) such that

\[ y_{x(n)} - y_{x(n-1)} \leq 0 \text{ and } y_{x(n+1)} - y_{x(n)} > 0. \]

From equation (4),

\[ y_{x(n)} - y_{x(n-1)} = \frac{1}{n} V_{x(n)} + (E_{x(n)} - q_x)^2 - \frac{1}{n-1} V_{x(n-1)} - (E_{x(n-1)} - q_x)^2 \]
\[ = (E_{x(n)} - E_{x(n-1)}) [(E_{x(n)} - E_{x(n-1)}) + 2(E_{x(n-1)} - q_x)] \]
\[ + \left( \frac{1}{n} V_{x(n)} - \frac{1}{n-1} V_{x(n-1)} \right). \] (5)

But

\[ E_{x(n)} = \delta(\hat{q}_x) = \delta \left[ \frac{1}{n} \sum_{i=1}^{n} \hat{q}_{xi} \right] = \frac{1}{n} \sum_{i=1}^{n} q_{xi}. \]

Therefore,

\[ E_{x(n)} - E_{x(n-1)} = \frac{1}{n} (q_{xn} - E_{x(n-1)}), \] (6)

and

\[ \frac{1}{n} V_{x(n)} - \frac{1}{n-1} V_{x(n-1)} = \frac{1}{n^2} \left[ q_{xn}(1 - q_{xn}) - \frac{2n - 1}{n-1} V_{x(n-1)} \right]. \] (7)
Substituting equations (6) and (7) in (5), and noting that \((2n-1)/(n-1) \approx 2\) for large \(n\):

\[
y_{x(n)} - y_{x(n-1)} = \frac{1}{n^2}[(q_{xn} - E_{x(n-1)}) (q_{xn} - E_{x(n-1)} + 2n(E_{x(n-1)} - q_x)) + q_{xn}(1 - q_{xn}) - 2V_{x(n-1)}].
\]

(8)

Setting \(y_{x(n)} - y_{x(n-1)} = 0\) yields

\[
n = \frac{(2V_{x(n-1)} - q_{xn}(1 - q_{xn})) - (q_{xn} - E_{x(n-1)})^2}{2(q_{xn} - E_{x(n-1)})(E_{x(n-1)} - q_x)}.
\]

(9)

Usually, in practice, equation (9) will define a unique \(n\) and then the optimal term for the investigation will be that which provides for the inclusion of the most recent \(n\) lives. If there is no \(n\) satisfying (9), then we would conclude that an indefinite enlargement of the investigation data would be beneficial to the estimate of \(q_x\). However, this circumstance would not arise in practice. If equation (9) has several solutions for \(n\), then it would be necessary to consider which of them minimizes \(y_{x(n)}\).

**OPTIMIZING THE TERM IN RESPECT OF THE WHOLE MORTALITY TABLE**

It is now proposed to consider the more usual investigation which aims to estimate the mortality rate at each of a large number of ages. It is clear from equation (9) that the optimal term for one particular age will not, in general, be the optimum for other ages. Therefore, it is necessary to consider the optimization of the term in the respect of the whole mortality table as a separate problem. Consider an investigation over all ages between \(\alpha\) and \(\beta\) with \(n_x\) lives observed at age \(x\). Corresponding to equation (1) is

\[
\hat{q}_x = \frac{1}{n_x} \sum_{i=1}^{n_x} \hat{q}_{xi}.
\]

One possible approach to the present problem is to form a weighted average of the mortality rates and minimize the mean square deviation of the observed from the 'expected' value of this average. Thus, define

\[
q = \sum_{x=\alpha}^{\beta} w_x \hat{q}_x,
\]

(10)

and

\[
q = \sum_{x=\alpha}^{\beta} w_x q_x,
\]

(11)

where

\[
\sum_{x=\alpha}^{\beta} w_x = 1.
\]
The problem is now to minimize (by appropriate choice of \( n = \sum \alpha_n \)) the function

\[ Y_n = \delta \{(q - q)^2\}. \]

Corresponding to equation (3),

\[ Y_n = \text{Var} \ q + [\delta(q) - q]^2, \]

which, upon application of equations (10) and (11), becomes

\[
Y_n = \sum_{x=\alpha}^{\beta} w_x^2 \text{Var} \ q_x + \left[ \sum_{x=\alpha}^{\beta} w_x (\delta(q_x) - q_x) \right]^2
= \sum_{x=\alpha}^{\beta} \frac{w_x^2}{n_x} V_{x(n_x)} + \left[ \sum_{x=\alpha}^{\beta} w_x (E_{x(n_x)} - q_x) \right]^2. \tag{12}
\]

The determination of the optimal \( n \) in the present case must differ slightly from that of the previous section. In that section, \( Y_x(n) \) would (in practical examples) initially decrease with increasing \( n \) and, after a certain value of \( n \), would increase. However, the effect on \( Y_n \) of increasing \( n \) by unity will, in the critical range of \( n \), be very much dependent upon the age of the additional life included. Hence \( Y_n \) will not be divisible into two monotonic sections, but will fluctuate frequently between positive and negative gradients in the critical range of \( n \). Of course, this would not be of any special importance if one were prepared to compute \( Y_n \) for each \( n \) and simply pick out the minimum \( Y_n \). However, it does cause difficulties in an analytical approach to the problem.

One way of resolving this difficulty is to regard the age of each additional life included in the investigation as a random variable with probability \( k_x \) of being \( x \). The problem is then to seek an integer \( n \) such that

\[ \delta_x(Y_n - Y_{n-1})_x \leq 0 \text{ and } \delta_x(Y_{n+1} - Y_{n})_x > 0, \]

where \((Y_n - Y_{n-1})_x\) is the difference between \( Y_n \) and \( Y_{n-1} \) when the \( n \)th life included in the investigation has age \( x \). According to the Law of Large Numbers, the actual minimum \( Y_n \) would be expected to occur in this vicinity if \( n \) is fairly large.

By equation (12)

\[
(Y_n - Y_{n-1})_x = w_x^2 \left[ \frac{1}{n_x} V_{x(n_x)} - \frac{1}{n_x-1} V_{x(n_x-1)} \right] + \left[ \sum_{x=\alpha}^{\beta} w_x (E_{x(n_x)} - q_x) \right]^2
- \left[ \sum_{x=\alpha}^{\beta} w_x (E_{x(n_x)} - q_x) + w_x (E_{x(n_x-1)} - q_x) \right]^2
= w_x^2 \left[ \frac{1}{n_x} V_{x(n_x)} - \frac{1}{n_x-1} V_{x(n_x-1)} \right]
+ w_x (E_{x(n_x)} - E_{x(n_x-1)}) \left[ 2 \sum_{x=\alpha}^{\beta} w_x (E_{x(n_x)} - q_x) - w_x (E_{x(n_x)} - E_{x(n_x-1)}) \right]. \tag{13}
\]
Equations (6) and (7) can now be applied to (13) to give (approximately)

\[
(Y_n - Y_{n-1})_x = \frac{w^2_x}{n^2_x} \left[ q_{xn_x}(1 - q_{xn_x}) - 2V_{x(n_x-1)} \right]
\]

\[
+ \frac{w^2_x}{n^2_x} \left[ q_{xn_x} - E_{x(n_x-1)} \right] \left[ 2 \sum_{z=\alpha}^{\beta} w_z(E_{z(n_x)} - q_z) - \frac{w_x}{n_x} (q_{xn_x} - E_{x(n_x-1)}) \right]
\]

\[
= \frac{w^2_x}{n^2_x} \left[ q_{xn_x}(1 - q_{xn_x}) - 2V_{x(n_x-1)} \right] - (q_{xn_x} - E_{x(n_x-1)})^2 \right]
\]

\[
+ 2 \frac{w_x}{n_x} \left[ q_{xn_x} - E_{x(n_x-1)} \right] \sum_{z=\alpha}^{\beta} w_z(E_{z(n_x)} - q_z).
\]

Hence

\[
\mathbf{E}_x(Y_n - Y_{n-1})_x = \sum_{x=\alpha}^{\beta} \frac{k_x w_x^2}{n^2_x} \left[ q_{xn_x}(1 - q_{xn_x}) - 2V_{x(n_x-1)} - (q_{xn_x} - E_{x(n_x-1)})^2 \right]
\]

\[
+ 2 \left[ \sum_{x=\alpha}^{\beta} k_x w_x(q_{xn_x} - E_{x(n_x-1)}) \right] \left[ \sum_{x=\alpha}^{\beta} w_x(E_{x(n_x)} - q_x) \right]. \tag{14}
\]

If the weights \(w_x\) are now chosen as \(w_x = n_x/n\) (i.e. \(\hat{q}\) and \(q\) are crude death rates) and \(k_x\) is taken as \(n_x/n\) also, then equation (14) becomes

\[
\mathbf{E}_x(Y_n - Y_{n-1})_x = \frac{1}{n^2} \sum_{x=\alpha}^{\beta} k_x \left[ q_{xn_x}(1 - q_{xn_x}) - 2V_{x(n_x-1)} - (q_{xn_x} - E_{x(n_x-1)})^2 \right]
\]

\[
+ 2 \left[ \sum_{x=\alpha}^{\beta} k_x(q_{xn_x} - E_{x(n_x-1)}) \right] \left[ \sum_{x=\alpha}^{\beta} k_x(E_{x(n_x)} - q_x) \right].
\]

Setting this equal to zero gives

\[
n = \frac{\sum_{x=\alpha}^{\beta} k_x \left[ 2V_{x(n_x-1)} + (q_{xn_x} - E_{x(n_x-1)})^2 - q_{xn_x}(1 - q_{xn_x}) \right]}{2 \left[ \sum_{x=\alpha}^{\beta} k_x(q_{xn_x} - E_{x(n_x-1)}) \right] \left[ \sum_{x=\alpha}^{\beta} k_x(E_{x(n_x)} - q_x) \right]}. \tag{15}
\]

Equation (15) gives an expression for the required value of \(n\) in terms of \(q_x\), the mortality rate at the end of the investigation, \(q_{xn_x}\), the mortality rate at the start of the investigation, \(E_{x(n_x)}\) and \(E_{x(n_x-1)}\), average mortality rates over the whole of the investigation, and \(V_{x(n_x-1)}\), which is a multiple of the variance of the mortality rate taken over the whole investigation. If any prior information on these rates is available, it may be used to calculate the optimal \(n\).
THE CASE OF MORTALITY RATES VARYING LINEARLY WITH TIME

It is extremely unlikely that reliable prior information on all of the terms included in (15) will be available. Nevertheless, if the term of the investigation is not going to turn out to be too large, it might be reasonable to assume that mortality rates change linearly with time.

Consider a life office which has $n_x$ lives in force at age $x$ each year and whose mortality rate at age $x$ has, in recent years, been decreasing linearly at the rate of $\delta_x$ per annum. Then

$$q_{xi} = q_x + \frac{i}{v_x} \delta_x.$$ 

If $n_x = v_x t$, then $q_{xnx} = q_x + t \delta_x$. Also, if $n_x$ and $v_x$ are fairly large, then

$$E_x(n_x - 1) \approx E_x(n_x) = q_x + \frac{1}{2} t \delta_x,$$

and

$$V_x(n_x - 1) \approx V_x(n_x) = \frac{1}{n_x} \sum_{i=1}^{n_x} \left( q_x + \frac{i}{n_x} t \delta_x \right) (1 - q_x - \frac{i}{n_x} t \delta_x) \approx q_x (1 - q_x) + \frac{1}{2} (1 - 2 q_x) t \delta_x - \frac{1}{2} t^2 \delta_x^2.$$

Then equation (15) becomes

$$n \approx \frac{\sum_{x=a}^{b} q_x (1 - q_x) + \frac{7}{12} \delta_x^2 t^2}{2 \left[ \sum_{x=a}^{b} k_x \delta_x t \right]^2} = \frac{2 \sum k_x q_x (1 - q_x)}{t^2 \left[ \sum k_x \delta_x \right]^2} + \frac{7}{6} \frac{\sum k_x \delta_x^2}{\left[ \sum k_x \delta_x \right]^2}. \quad (16)$$

The final term of equation (16) will be of the order of 7/6 which is insignificant if $n$ is large. Therefore, approximately,

$$n = \frac{2 \sum k_x q_x (1 - q_x)}{t^2 \left[ \sum k_x \delta_x \right]^2}. \quad (17)$$

But $n = \Sigma n_x = t \Sigma v_x = N t$, say, where $N = \Sigma v_x$ = total number of lives in force at any time. Therefore, equation (17) becomes

$$t^3 = \frac{2 \sum k_x q_x (1 - q_x)}{N \left[ \sum k_x \delta_x \right]^2},$$

that is

$$t = \left\{ \frac{2 \sum k_x q_x (1 - q_x)}{N \left[ \sum k_x \delta_x \right]^2} \right\}^{\frac{1}{3}}. \quad (18)$$

The value of $t$ given by equation (18) is the optimal term of the investigation in years.
A NUMERICAL EXAMPLE

One would usually be ill-advised to expect equation (18) to provide an accurate value of $t$, for there is bound to be a certain unreliability in prior information about the $q_x$'s and $\delta_x$'s. Nevertheless, this equation is useful for order-of-magnitude calculations, e.g. should the term of investigation be 1 year or 10 years?

It is interesting to consider the investigation of 1958–63 mortality of Assured Lives in Australia (durations 2 and over). In the discussion of the report on the investigation (Transactions of the Institute of Actuaries of Australia and New Zealand, 1968), G. C. Lane remarked that the 1953–58 mortality followed very closely that of the A1949/52 ultimate table with the age rated down 1 year, and that the 1958–63 mortality followed very closely that of the same table with age rated down $1\frac{1}{2}$ years. If this is accepted $q_x$ may be taken as

$$q_x = q^{A49/52}_{x-1}$$

and $\delta_x$ as

$$\delta_x = \frac{1}{2}(q^{A49/52}_{x-1} - q^{A49/52}_{x-1\frac{1}{2}}).$$

Then, with the data of Table 7 of the report referred to above, Table 1 can be obtained. The totals of columns (5) and (7), in conjunction with equation (18) yield

$$t = \left[\frac{2 \times 35626.4}{0.5 \times (354.49)^2}\right]^{\frac{1}{3}}$$

whence $t = 1.4$ approximately.

Thus the optimal term for an investigation aiming to calculate 1963 mortality by means of data from years prior to this is about $1\frac{1}{2}$ years when the size of the exposure is as shown in Table 1.

In fact, the term of the investigation was 5 years, which is significantly greater than optimal. Moreover, as will be shown in the next section, there are other factors, not yet considered, which can reduce the optimal term further. Of course, the whole analysis has been performed under the assumptions of linearity of mortality rates over time, and constancy of numbers in force. These assumptions may be somewhat inaccurate, but nevertheless it does seem clear that the optimal term (in accordance with equation (15)) is significantly less than 5 years.

It is interesting to calculate the size of the experience for which a 5-year term would have been optimal. If a second experience were identical to the one just considered except scaled down in size by a factor of $\rho$, then the optimal term would be $(2.8/\rho)^{\frac{1}{3}}$. For this to be equal to 5, then $\rho$ must be approximately $24\%$. Thus, an office contributing $24\%$ of the Assured Lives in Australia experience (a fairly small office) and typical in its mortality experience would have found an investigation term of 5 years appropriate.

One of the most striking observations that can be made of Table 1 is the great
Table 1

<table>
<thead>
<tr>
<th>Age group</th>
<th>Central age $x$</th>
<th>Exposed to risk $5 \times Nk_x$</th>
<th>$q_x$</th>
<th>$q_x(1-q_x)$</th>
<th>$\delta_x$</th>
<th>$(3) \times (6)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9½–11½</td>
<td>11</td>
<td>6253</td>
<td>-00111</td>
<td>6-94</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>12½–16½</td>
<td>15</td>
<td>32662</td>
<td>-00111</td>
<td>36-25</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>17½–21½</td>
<td>20</td>
<td>68521</td>
<td>-00111</td>
<td>76-06</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>22½–26½</td>
<td>25</td>
<td>118092</td>
<td>-00112</td>
<td>132-26</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>27½–31½</td>
<td>30</td>
<td>177932</td>
<td>-00115</td>
<td>204-62</td>
<td>-000001</td>
<td>-1-779</td>
</tr>
<tr>
<td>32½–36½</td>
<td>35</td>
<td>252750</td>
<td>-00125</td>
<td>315-94</td>
<td>-000004</td>
<td>1-0110</td>
</tr>
<tr>
<td>37½–41½</td>
<td>40</td>
<td>299262</td>
<td>-00165</td>
<td>493-78</td>
<td>-00013</td>
<td>3-8904</td>
</tr>
<tr>
<td>42½–46½</td>
<td>45</td>
<td>348003</td>
<td>-00276</td>
<td>957-01</td>
<td>-000033</td>
<td>11-4841</td>
</tr>
<tr>
<td>47½–51½</td>
<td>50</td>
<td>375381</td>
<td>-00504</td>
<td>1880-66</td>
<td>-000060</td>
<td>22-5229</td>
</tr>
<tr>
<td>52½–56½</td>
<td>55</td>
<td>344340</td>
<td>-00884</td>
<td>3016-42</td>
<td>-000094</td>
<td>32-3680</td>
</tr>
<tr>
<td>57½–61½</td>
<td>60</td>
<td>273203</td>
<td>-1483</td>
<td>3991-50</td>
<td>-000149</td>
<td>40-7072</td>
</tr>
<tr>
<td>62½–66½</td>
<td>65</td>
<td>194971</td>
<td>-2431</td>
<td>4624-71</td>
<td>-000237</td>
<td>46-2081</td>
</tr>
<tr>
<td>67½–71½</td>
<td>70</td>
<td>146738</td>
<td>-3942</td>
<td>5556-97</td>
<td>-000377</td>
<td>55-3202</td>
</tr>
<tr>
<td>72½–76½</td>
<td>75</td>
<td>98104</td>
<td>-6322</td>
<td>5809-72</td>
<td>-000588</td>
<td>57-6852</td>
</tr>
<tr>
<td>77½–81½</td>
<td>80</td>
<td>50920</td>
<td>-9971</td>
<td>4571-09</td>
<td>-000886</td>
<td>45-1151</td>
</tr>
<tr>
<td>82½–86½</td>
<td>85</td>
<td>20459</td>
<td>-15312</td>
<td>2652-92</td>
<td>-001261</td>
<td>25-7988</td>
</tr>
<tr>
<td>87½–91½</td>
<td>90</td>
<td>6235</td>
<td>-22613</td>
<td>1091-13</td>
<td>-001655</td>
<td>10-3189</td>
</tr>
<tr>
<td>92½–96½</td>
<td>95</td>
<td>949</td>
<td>-31709</td>
<td>205-50</td>
<td>-001958</td>
<td>1-8581</td>
</tr>
<tr>
<td>97½–101½</td>
<td>100</td>
<td>12</td>
<td>-41818</td>
<td>2-92</td>
<td>-002045</td>
<td>-0245</td>
</tr>
</tbody>
</table>

$\sum_{x} 35626-4 = 5 \sum_{x} Nk_x q_x(1-q_x)$

$\sum_{x} 354-4904 = 5 \sum_{x} Nk_x \delta_x$
variation in optimal terms for the individual ages. The optimal term for age \( x \) alone is obtained by replacing \( k_x \) by 1 and \( N \) by \( Nk_x \) in (18) to give
\[
t = \left[ \frac{2q_x(1-q_x)}{Nk_x\delta_x^2} \right]^{1/3}.
\]
The figures in Table 1 show that for low ages and high ages \( t \) is quite large, while for the middle ages it is relatively small. For example, at age 30, \( t = (10 \times 204.62/0.03)^{1/3} = 40.9 \). As age increases, \( t \) decreases until reaching a minimum at age 75 of \( (10 \times 5809.72/3327.56)^{1/3} = 2.6 \). As age continues to increase, \( t \) then increases again until, at age 95, \( t = (10 \times 205.50/3.45)^{1/3} = 8.4 \). These rather large disparities suggest that it may be worthwhile to consider separate investigations of mortality at the high and/or low ages.

An interesting point emerges from the use of equation (18). Since the system of weights \( w_x \) will vary between life offices, so will the optimal term vary. Thus an investigation of the combined mortality experience of a number of life offices must be performed over a term which, inevitably, is unsuitable for some of the offices. Often, the weights will not vary sufficiently between offices to render this point of any practical consequence, but it is interesting, in principle, to remark that the term of the investigation will be dependent upon which institutions are to use the derived mortality rates.

**POSSIBLE MODIFICATIONS OF THE OPTIMAL TERM FORMULA**

**Sums at risk**

The basic optimization equation (18) was derived on the assumption that mean square deviation of the actual crude death rate from the 'expected' should be minimized. From the point of view of a life office, this is probably not the most natural investigation, since it disregards the financial importance of the various ages. A more acceptable weighted average mortality rate might be
\[
\frac{\sum_{x=\alpha}^{\beta} n_x v_x q_x}{\sum_{x=\alpha}^{\beta} n_x v_x} = \frac{\sum_{x=\alpha}^{\beta} n_x v_x q_x}{\sum_{x=\alpha}^{\beta} n_x v_x},
\]
where \( v_x \) is the mean sum at risk per life aged \( x \). That is, in equations (10) and (11), the weights \( w_x \) are taken as
\[
w_x = n_x v_x.
\]
These weights may be normalized if desired. Making this substitution in equation (14) yields
\[
\xi_x(Y_n - Y_{n-1}) = \sum_{x=\alpha}^{\beta} k_x v_x^2 \left[ q_{xn_x} (1-q_{xn_x}) - 2V_x (n_x-1) - (q_{xn_x} - E_x(n_x-1))^2 \right] \\
+ 2n \left[ \sum_{x=\alpha}^{\beta} k_x v_x (q_{xn_x} - E_x(n_x-1)) \right] \left[ \sum_{x=\alpha}^{\beta} k_x v_x (E_x(n_x) - q_x) \right],
\]
where \( n_x \) is the number of lives aged \( x \).
whence equation (15) is replaced by

\[
\frac{\sum_{x=\alpha}^{\beta} k_x v_x^2 \left[ 2V_{x(n_x-1)} + (q_{xn_x} - E_{x(n_x-1)})^2 - q_{xn_x}(1 - q_{xn_x}) \right]}{2 \left[ \sum_{x=\alpha}^{\beta} k_x v_x (q_{xn_x} - E_{x(n_x-1)}) \right]} \left[ \sum_{x=\alpha}^{\beta} k_x (E_{x(n_x)} - q_x) \right]^{1/2},
\]

(15a)

and equation (18) by

\[
t = \left\{ \frac{2 \sum_{x=\alpha}^{\beta} k_x v_x q_x (1 - q_x)}{N \left[ \sum_{x=\alpha}^{\beta} k_x v_x \delta_x \right]^2} \right\}^{1/2}.
\]

(18a)

Application of (18a) is illustrated in Table 2, from which it is found that

\[
t = \left[ \frac{2 \times 3963721}{1.5 \times (2225.42)^2} \right]^{1/2} = 2 \text{ years, approximately.}
\]

This is larger than the term of 1 1/2 years produced by equation (18), the increase resulting from the fact that the inclusion of the terms \(v_x\) in (18a) increases the weight of the younger membership for whom optimal values of \(t\) are larger.

**Graduation**

The effect of graduation on the foregoing reasoning is a most important factor which has not yet been considered. The whole idea of the paper is, as stated in the first section, to examine how the decrease in variance of decrement rates achieved by lengthening the term of the investigation can be 'traded off' against the corresponding increase in bias. Now, if the crude decrement rates are graduated, then the variance of each rate will be lessened by the error-reducing factor of the graduation, but the bias will be unchanged. This suggests a smaller value of \(t\) than implied by equations (18) and (18a). In fact, if the error-reducing index of the graduation is \(e_x\) at age \(x\), then equation (3) becomes

\[
y_{x(n)} = e_x^2 \text{ Var } \hat{q}_x + [\delta(\hat{q}_x) - q_x]^2.
\]

(3a)

Following the working subsequent to equation (3), one finds that equation (3a) leads to equations (15) and (18) being replaced by

\[
n = \frac{\sum_{x=\alpha}^{\beta} k_x \left[ e_x^2 [2V_{x(n_x-1)} - q_{xn_x}(1 - q_{xn_x})] + (q_{xn_x} - E_{x(n_x-1)})^2 \right]}{2 \left[ \sum_{x=\alpha}^{\beta} k_x (q_{xn_x} - E_{x(n_x-1)}) \right]} \left[ \sum_{x=\alpha}^{\beta} k_x (E_{x(n_x)} - q_x) \right]^{1/2},
\]

(15b)

and
### Table 2

<table>
<thead>
<tr>
<th>Age group</th>
<th>Central age</th>
<th>Exposed to risk</th>
<th>$v_x$</th>
<th>$q_x$</th>
<th>$(3) \times (4)^2 \times q_x(1-q_x) \left( = \frac{1}{5} \left( q_x^{A49/52} - q_x^{A49/52} \right) \right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9½ – 11½</td>
<td>11</td>
<td>6253</td>
<td>50:0</td>
<td>0.0111</td>
<td>17350</td>
</tr>
<tr>
<td>12½ – 16½</td>
<td>15</td>
<td>32662</td>
<td>50:0</td>
<td>0.0111</td>
<td>90625</td>
</tr>
<tr>
<td>17½ – 21½</td>
<td>20</td>
<td>68521</td>
<td>53:0</td>
<td>0.0111</td>
<td>213653</td>
</tr>
<tr>
<td>22½ – 26½</td>
<td>25</td>
<td>118092</td>
<td>50:0</td>
<td>0.0112</td>
<td>330650</td>
</tr>
<tr>
<td>27½ – 31½</td>
<td>30</td>
<td>177932</td>
<td>45:0</td>
<td>0.0115</td>
<td>414356</td>
</tr>
<tr>
<td>32½ – 36½</td>
<td>35</td>
<td>252750</td>
<td>40:0</td>
<td>0.0125</td>
<td>505504</td>
</tr>
<tr>
<td>37½ – 41½</td>
<td>40</td>
<td>299262</td>
<td>30:0</td>
<td>0.0165</td>
<td>444402</td>
</tr>
<tr>
<td>42½ – 46½</td>
<td>45</td>
<td>348003</td>
<td>22:0</td>
<td>0.0276</td>
<td>463193</td>
</tr>
<tr>
<td>47½ – 51½</td>
<td>50</td>
<td>375381</td>
<td>16:0</td>
<td>0.0504</td>
<td>481449</td>
</tr>
<tr>
<td>52½ – 56½</td>
<td>55</td>
<td>344340</td>
<td>12:5</td>
<td>0.0884</td>
<td>471316</td>
</tr>
<tr>
<td>57½ – 61½</td>
<td>60</td>
<td>273203</td>
<td>8:2</td>
<td>0.1483</td>
<td>268388</td>
</tr>
<tr>
<td>62½ – 66½</td>
<td>65</td>
<td>194971</td>
<td>5:6</td>
<td>0.2431</td>
<td>145031</td>
</tr>
<tr>
<td>67½ – 71½</td>
<td>70</td>
<td>146738</td>
<td>3:5</td>
<td>0.3942</td>
<td>68073</td>
</tr>
<tr>
<td>72½ – 76½</td>
<td>75</td>
<td>98104</td>
<td>2:3</td>
<td>0.6322</td>
<td>30733</td>
</tr>
<tr>
<td>77½ – 81½</td>
<td>80</td>
<td>50920</td>
<td>1:6</td>
<td>0.9971</td>
<td>11702</td>
</tr>
<tr>
<td>82½ – 86½</td>
<td>85</td>
<td>20459</td>
<td>1:4</td>
<td>1.5312</td>
<td>5200</td>
</tr>
<tr>
<td>87½ – 91½</td>
<td>90</td>
<td>6235</td>
<td>1:3</td>
<td>2.2613</td>
<td>1844</td>
</tr>
<tr>
<td>92½ – 96½</td>
<td>95</td>
<td>949</td>
<td>1:1</td>
<td>3.1709</td>
<td>249</td>
</tr>
<tr>
<td>97½ – 101½</td>
<td>100</td>
<td>12</td>
<td>1:0</td>
<td>4.1818</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
\frac{3963721}{5} = 5 \sum N k_x v_x q_x^2 (1 - q_x)
\]

\[
\frac{2225420}{5} = 5 \sum N k_x v_x \delta x
\]
an Investigation into Decremental Rates

\[ t = \left\{ \frac{2 \sum_{x} e_{x}^{2} k_{x} q_{x}(1 - q_{x})^{\frac{1}{3}}}{N \left[ \sum_{x} k_{x} \delta_{x} \right]^{2}} \right\} \quad (18b) \]

respectively. If it is reasonable to take \( e_{x} = e \) (constant) for all \( x \), then (18b) becomes

\[ t = e^{\frac{2}{3}} \left\{ \frac{2 \sum_{x} k_{x} q_{x}(1 - q_{x})}{N \left[ \sum_{x} k_{x} \delta_{x} \right]^{2}} \right\}. \quad (18c) \]

Similarly, a graduation of error-reducing power \( e \) at all ages modifies equation (18a) to

\[ t = e^{\frac{2}{3}} \left\{ \frac{2 \sum_{x} k_{x} q_{x}^{2}(1 - q_{x})}{N \left[ \sum_{x} k_{x} q_{x} \delta_{x} \right]^{2}} \right\}. \quad (18d) \]

In other words, such a graduation multiplies the optimal term of the investigation by a factor of \( e^{\frac{2}{3}} \).

For example, most acceptable summation graduation formulae have an error-reducing index in the vicinity of \( 0.4 \) (see Pollard 1970, 1971). Hence the graduation reduces the optimal term of the investigation by a factor of \( (0.4)^{\frac{2}{3}} \), i.e. about \( 0.5 \). If such a graduation had been applied to the data of the 1958–63 Assured Lives in Australia experience, then equations (18c) and (18d) would give optimal terms of about \( 0.75 \) and 1 respectively.

SOME GENERAL COMMENTS

Quite apart from being useful for numerical work, the family of equations (18), (18a)–(18d) indicate some general properties of the optimal term. These may be summarized as follows:

(i) The optimal term is largely determined by three factors (see equation (18)): the size of the experience (as measured by the number of lives in force at any particular time); the level of mortality; the rate at which mortality has been changing over recent times.

(ii) For given level and rate of change of mortality, the optimal term is inversely proportional to the cube root of the size of the experience.

(iii) For given size of experience and rate of change of mortality, the optimal term increases as mortality increases (assuming that mortality rates are considerably less than \( \frac{1}{2} \) at most ages).

(iv) For given size of experience and level of mortality the optimal term decreases as the rate of change of mortality increases.

It is appropriate that some comment be devoted to the assumptions underlying the main equation (18). There are two such assumptions which deserve scrutiny. They are that mortality rates have been changing linearly in the recent past and that the uniform rates of change can be measured.

The first assumption may sometimes be thought to be invalid. For example,
the term of the A1949/52 mortality investigation was determined mainly by factors extraneous to the present paper for the very reason that the first assumption did not seem valid. The report on this investigation (J.I.A. 1956, 82, 3) states:

Over the period 1940–47, the data in the C.M.I. were abnormal in so far as they included deaths due to the Second World War. Some war deaths were recorded in 1946 and 1947 because claims under policies on the lives of deceased prisoners-of-war were still being reported in those years, and consequently 1948 was the first year for which the data represented normal civilian mortality. In 1947, however, mortality in Britain among the general population was heavy, whereas in 1948 it was exceptionally light; and there is reason to presume that these two features were interdependent in so far as the adverse conditions in the early months of 1947 may have precipitated many deaths which otherwise might not have occurred until the following year. Since 1947 had to be excluded it seemed wise to exclude 1948 as well and to commence with the experience of 1949. At the time when the Committee began to consider the construction of the new table, the experience of 1952 had just become available and the period 1949–52 commended itself as a suitable basis. The years 1949 and 1950 had a very similar experience intermediate between 1947 and 1948. In 1951, mortality at the middle and older ages was heavy on account of an influenza epidemic. The year 1952 was lighter than 1949–50, but not so light as 1948. Epidemics and 'heavy years' occur from time to time, and one heavy year combined with one fairly light and two moderate years seemed a reasonable and prudent basis for tables designed for use in the conduct of life assurance.

Thus assumption of linear change in mortality rates was certainly held to be invalid in this case, and, consequently, the foremost aim in fixing a term for the investigation was to ensure that the correct mixture of light, heavy and medium years was taken in order that the mortality experience should be 'average'.

Even if this assumption is tenable, there may be some difficulty in measuring the rate at which mortality has been changing at each age. Since the purpose of a mortality investigation is to obtain information about mortality, it is clear that any prior information will usually be of a sketchy nature. Therefore, our values of $\delta_x$ in equations (18) and (18a)–(18d) may be rather unreliable. Against this, however, must be set the argument that in such a situation an informed and intelligent calculation of the optimal term with the attendant reservations kept in mind is better than an arbitrary decision.

Moreover, once an investigation has been carried out (whatever term is used), a great deal more information about the $\delta_x$'s comes to hand. If this extra information reveals a serious error in the term adopted and a consequent serious distortion of the correct schedule of mortality rates, then it may be used to produce a much more reliable calculation of the optimal term, whereupon a new investigation may be performed.

In any event, whether numerical calculations can be made or not, the above equations do provide some useful qualitative information, as noted at the beginning of this section. Perhaps the most useful observation is (in informal terms) that the smaller is a life office (say), the less particular should it be about the data which it includes in an investigation of its own experience. A similar sentiment has previously been expressed in connexion with sickness experiences (Taylor 1971, pp. 74–5):
When every piece of data in the investigation provides an unbiased estimate of the rate which is being sought, our purposes are best served by the inclusion of all available data and the consequent reduction of the variance of our estimate. However, in the case of a sickness experience distorted by financial immaturity, not all the data provide unbiased estimates of the stable sickness rates. Thus, by the inclusion of all data a reduction in variance of the estimate is achieved at the expense of biasing the estimate. Therefore, whether or not the data of the unstable durations of membership should be used depends upon the size of the experience. In a large established society, it would probably be wise to exclude all data relating to unstable durations. On the other hand, all data should be included in the experience of a small society.

REFERENCES

(Typescripts of these two papers are held in Institute of Actuaries Library.)