JIA 101 (1974) 415-434

A STUDY OF THE VARIANCE OF MORTALITY RATES

BY R. H. DAW, B.Sc., F.I.A.

1. INTRODUCTION

SOME 20 to 30 years ago considerable progress was made in understanding the variance of mortality rates and, for assured lives data, the effect on it of the presence of duplicate policies, e.g. Redington and Michaelson (1940), Seal (1941, 1947, 1948 and 1949), Daw (1945 and 1951), Vajda (1945), Solomon (1948), Beard (1951), Jager (1953) and Continuous Mortality Investigation (C.M.I.) (1957). I have the impression that since that time this understanding has tended to regress. For example it is disappointing to find that the whole course of reading on the analysis of mortality data makes practically no mention of the effect of duplicate policies in increasing the variance of mortality rates. Yet this is a matter which can have a substantial effect on the results of testing a mortality table graduation as is recognized in C.M.I. (1974). It is salutory to find that G. H. Hardy (1882) was aware that the presence of duplicates increased the variance although he did not express it in quite that way.

It is therefore appropriate that the subject should have been reopened. Recently Pollard (1970) published some studies of the variance of mortality rates in Australian population and assured lives data but made no reference to previous work. Although some interesting results were produced he did not consider the nature of the variance he was studying and made little attempt to explain his results.

The present paper was started with the intention of commenting on Pollard's work and integrating it into earlier work. However, as so often happens, it developed far beyond these limits and comments on Pollard's work became largely incidental.

The Appendices to this paper describe briefly the principal methods used and also deals with technical points which have a bearing on the methods.

2. TYPES OF DEPARTURE FROM BINOMIAL VARIANCE

Consider a box containing a large number of counters, a proportion p being marked L (living) and a proportion q $(p+q = 1)$ being marked D (dead). If a large number of random samples of size N are drawn from the box (strictly with replacement after each drawing) the proportion of Ds will average q with variance pq/N, i.e. the binomial distribution will apply. Suppose also that the counters are of several different colours, the proportion of Ds being different for each colour. If random samples are drawn from such a heterogeneous population the proportion of Ds will still average q with variance pq/N. It is only if the sampling is restricted in some way related to the colours of the counters that the variance will differ from pq/N. The formulae in section 2 of Pollard (1970) give examples of various types of restriction.

Now any body of lives, even if they are all of the same age and perhaps recently selected for assurance, can, probably with a reasonable degree of success, be divided into say three groups expected to experience high, medium and low mortality. (By analogy with the counters, we have a fair idea of the colour of each counter but no knowledge whether it is marked D or L). In other words a homogeneous body of lives is only a theoretical concept and heterogeneity is the usual pattern. However just as with the counters, it may be found that heterogeneous bodies of lives behave as though they were samples drawn at random from a large (hypothetical) heterogeneous population and show binomial variance; on the other hand they may not behave in this way. The difficulty is, of course, to test how they do behave. If a variance is to be *estimated* it has to be done from the results of a number of samples of rates of mortality. It is not possible to expose the same body of lives more than once to the risk of death in a year because, apart from the increased age and the fact that some of the lives will have been removed by death, the general level of mortality varies from year to year due to weather conditions, epidemics, and such like. The exact nature of the variance estimated will therefore depend on how the samples used to estimate it have originated and we must be quite clear what variance we are attempting to estimate.

In considering this matter it will be helpful to define two types of heterogeneity, as follows:

(i) *Internal heterogeneity* will denote any features of the exposed to risk (e.g. lives subject to different chances of death) or methods of construction (e.g. duplicates) which may cause the variance of a mortality rate to depart from its theoretical binomial value, and
(ii) *External heterogeneity* will describe real differences in mortality between groups of lives (e.g. different years of experience or classes of person).

While the dividing line between internal and external heterogeneity may not always be completely clear, it is a distinction important in relation to the samples used in estimating a variance.

To illustrate the position consider Table 1 which sets out an arrangement of mortality data covering the experience of three years.

Table 1

Age	Year 1	Year 2	Year 3	Years 1 to 3 combined
x	$E_x \; \theta_x \; q_x$	$E_x \; \theta_x \; q_x$	$E_x \; \theta_x \; q_x$	$E_x \; \theta_x \; q_x$
25				
26				
27				

The experience of each main column of Table 1 at successive ages could be taken as the set of sample values from which the variance is calculated provided that some way can be found of taking account of the general increase in mortality with age i.e. the heterogeneity due to age between successive sample values. A variance estimated in this way should take account only of the effect of 'internal heterogeneity', and this applies whether the calculations are based on the figures for an individual year of experience or the combined experience of the three years of Table 1.

Possible methods of taking account of the age effect are illustrated by the assumption made in the r_x test that the third differences of the underlying rates of mortality are small (see Appendix 1), the assumption of a graduated rate of mortality appropriate to each age as in Pollard's (1970) method 3 and some of the studies in Seal (1949).

If, however, the estimate of variance is made by taking the set of sample values as the experience of a single age x in the successive individual years of Table 1 and comparing the rates of mortality of each of the three years with that of the three years combined, then each deviation will represent that of one particular year (subject to its own particular level of mortality) from the average of the three years. The resulting variance will therefore give the combined effect of both internal and external heterogeneity. Because it is known that the level of mortality fluctuates from year to year, a variance estimated in this way would be expected to exceed binomial variance. An example is given by Pollard's (1970) method 2. Pollard's method 1 attempts to remove the effect of year-to-year variation by fitting a linear trend line from which the residual variances are measured. It does not seem that such a method will eliminate much of the external heterogeneity because the year-to-year variation does not consist only of a gradual trend (hardly perceptible over a short period of years) but also contains a, usually more substantial, irregular component due to factors like epidemics and weather conditions.

It must be emphasized that in any consideration of the variance of mortality rates we must be quite clear whether we are trying to take account of internal heterogeneity alone, external heterogeneity alone or both types of heterogeneity together.

As indicated in the definition of the term, internal heterogeneity can arise from a number of factors in addition (possibly) to heterogeneity of the lives whereby every life is not necessarily subject to the same rate of mortality. Supra-binomial variance can arise, for example, from errors in the data (e.g. age mis-statements, method of finding the exposed to risk) and from the presence of duplicate policies in assured lives data. Excess variance can also arise from inaccuracies in the method of estimating the variance.

It is perhaps worth mentioning that age misstatements in population data may be of a different nature from those in assured lives. In population data the exposed to risk will usually be derived principally from a population census, the ages being those stated at the census. The ages of the deaths will be those given when the death is registered. Thus the age at the census may be discordant with the age at death and one or both could be wrong. In assured lives the date of birth will have been obtained at the time the assurance was effected and, whether right or wrong, there will be no discordance between the age in the exposed to risk and the age at death. Age errors are therefore likely to have a larger effect in population data.

The principal studies of this paper use the mortality experiences of each of the six years 1961 to 1966 of

(i) Continuous Mortality Investigation of assured lives, all classes, durations 2 and over,
(ii) Australian male population, and
(iii) Netherlands male population.

First the internal heterogeneity of these experiences is investigated, secondly the variation of the observed rates of mortality by year of experience is studied and finally an attempt is made to estimate the variance due to external heterogeneity of the six years after allowing for binominal variance and internal heterogeneity and to study how it is related to age.

3. INTERNAL HETEROGENEITY

Assured lives

The methods used to calculate mortality rates from large-scale data of assured lives are usually fairly rigorous and it may be expected that ages are accurate. Thus the method of construction is unlikely to introduce appreciable variation into the results of a study of internal heterogeneity, although Seal (1948) has shown that, for the usual form of exposed to risk formula, the theoretical distribution of deaths is not the binomial distribution, but the difference will usually be very small. The work of Redington and Michaelson (1940), Daw (1945) and Beard (1951) using the r_x test and Seal's (1949) work may be summed up by saying that no significant departures from binomial variance have been found for assured lives unless duplicate policies are present in the data.

The Australian assured lives data (1958–63) to which Pollard (1970) applies his method 3 relate to whole life with-profits policies and include duplicates which will have the effect of increasing the variance of the rates of mortality. However, Pollard can give no indication of the proportion of duplicates included in the ultimate experience. As Daw (1951) has shown, the inclusion of even quite a small proportion of duplicates may produce a substantial increase in the variance; 10% of duplicates may increase the variance to 20% to 25% above the binomial variance depending on the distribution of the duplicates. The data for individual years of duration since selection are not likely to contain many duplicates even if, unlike in the Continuous Mortality Investigation, concurrent duplicates are not excluded. I therefore consider that the presence of duplicates in the Australian data is a more likely explanation than developing heterogeneity as selection wears off of Pollard's finding of excess variance in the ultimate experience and of binomial variance in all but one of the five individual years of the select period.

The above argument is history repeating itself. When Redington and Michaelson (1940) first proposed the r_x test and applied it to the A1924–29 data, they found excess variance for durations 3 and over, but binomial variance for duration 0. They put forward the same explanation as Pollard. With the benefit of Seal's (1941) demonstration of the effects of duplicates, Daw (1945) suggested the presence of duplicates as a more likely explanation and this was accepted by Redington in the discussion of the paper.

The r_x test has now been applied to the Australian assured lives data 1958 to 1963, durations 2 and over, referred to above. Ages 22½ to 86½ gave 62 values of r_x and a value of σ_r^2 of 1·80. As indicated in Appendix 1 σ_r^2 is a measure of the average value of the ratio of the observed variance to the binomial variance. C.M.I. (1957) presents a study of the distribution of duplicates in the 1954 deaths and their effect in increasing the variance above the binomial variance; these figures, in conjunction with the above value of σ_r^2, suggest that the average proportion of duplicates in the Australian ultimate data may be of the order of 20%, provided the excess of the observed σ_r^2 of 1·80 over unity is solely due to the presence of duplicates and that the shape of their frequency distribution is similar to that in the C.M.I. data. However, this may not be a very reliable estimate bearing in mind the size of the random variation in σ_r^2 (see Table 2 and Appendix 2).

For ages 37½ to 86½ the Australian assured lives data gave $\sigma_r^2 = 1{\cdot}82$ and this may be compared with the results of Pollard's (1970) method 3 applied to the same data. For this age range the 10 values of k^2/NPQ (i.e. $\chi_4^2/4$) of Pollard's Table 5 may be averaged to give a value of $\chi_{40}^2/40$ of 2·04 which, like σ_r^2 is an estimate of the average ratio of the observed variance to binomial variance. The agreement between the two values is quite good for two very different methods, but it must be realized that Pollard's result depends on the particular graduation of the data which he used.

Turning now to the assured lives experience of the United Kingdom, the r_x test has been applied to the C.M.I. data, all classes, durations 2 and over, ages

Table 2. *C.M.I. all classes, durations 2 and over*

Year	σ_r^2
1961	2·13
1962	1·53
1963	1·79
1964	1·01
1965	1·04
1966	1·85
Average of above values	1·56

25 to 87, for each of the six years 1961 to 1966. The results are given in Table 2, each value of σ_r^2 being based on 60 values of r_x.

The values in Table 2 appear at first sight to show substantial variation. As indicated in Appendix 2 statistical significance should be considered using σ_r, not σ_r^2. The average of the values of σ_r for the C.M.I. experience (see Table A2 of Appendix 2) is 1·24 and taking the standard deviation of the observed values as $\frac{1 \cdot 5\ \sigma_r}{\sqrt{(2n)}} = \cdot 17$ (where $n = 60$ and σ_r is taken as 1·24), the range for $\pm$ two standard deviations is ·90 to 1·58 which covers the range of the observed values, i.e. 1·00 to 1·46. The variation is therefore no more than might be expected for an average value of σ_r of 1·24.

Values of σ_r^2 for the C.M.I. data have also been calculated for ten-year age-groups using the individual values of r_x for all the six years of experience. These values are given in Table 3, each being based on 60 values of r_x.

Table 3. *C.M.I. all classes, durations 2 and over*

Age group	σ_r^2
25–34	1·15
35–44	1·72
45–54	1·75
55–64	1·80
65–74	1·35
75–84	1·56

For the C.M.I. data 1961–66, the overall value of σ_r^2 of 1·56 (Table 2) is not very different from the overall variance ratio of 1·45 (C.M.I. 1957) which measures only the effect of duplicates in the C.M.I. deaths in 1954. Also the values of σ_r^2 (Table 3) for age groups are in reasonable agreement with the figures in C.M.I. (1957) but with lower values at the young ages and a rather higher level at the middle ages. These differences are probably mere random variations and the figures give no reason to think that the internal heterogeneity in the C.M.I. data is due to factors other than the presence of duplicates.

An attempt was made to check how far the value of σ_r^2 measured the increase in variance due to the presence of duplicates by comparing the figure of 1·45 given in C.M.I. (1957) with the value of σ_r^2 calculated from the corresponding C.M.I. data for 1954; the resulting value of σ_r^2 was 1·47. In view of the size of the random variation of σ_r^2 so close an agreement must be fortuitous, but again there is no reason to think that factors other than the presence of duplicates are operating.

Population data

Population data used in the study of mortality would normally be expected to be less accurate than those for assured lives; age misstatements would occur to a greater extent and have a larger effect, and the population figures (whether census or estimates) would be subject to some error. Also the figure taken for the exposed risk may be only an approximation to the true figure which should be related to the deaths. The values of σ_r^2 obtained from the r_x test applied to population data will be inflated by reason of any inaccuracies of the data or the method of deriving rates of mortality.

Daw (1945) applied the r_x test to the data used to construct the English Life Table No. 10 (males) which was based on the deaths in the three years 1930–32. Over the adult age range the exposed to risk at age x was taken as three times the population enumerated at that age at the census on 26/27 April 1931 and this was related to the deaths at age x in the three years. Unfortunately in calculating the values of r_x the census population was used instead of three times that population. The value of 1·39 for σ_r given by Daw (1945) therefore needs to be multiplied by $\sqrt{3}$, giving the corrected value of σ_r as 2·41 (or σ_r^2 as 5·80).

In constructing E.L.T. No. 12 from the census population of 23/24 April 1961 and deaths in the three years 1960 to 1962 a more accurate estimate was made of the exposed to risk to be related to the deaths than the comparatively crude estimate of three times the census population used in E.L.T. No. 10. This improvement of method is reflected in the lower value of σ_r^2 of 3·66 obtained for E.L.T. No. 12 (Males) ages 20 to 89, as compared with 5·80 for E.L.T. No. 10 (Males).

The Australian Life Tables 1960–62 (Commonwealth Actuary, 1965) are based on the census of 30 June 1961 and deaths in the three years 1960–62. In obtaining the exposed to risk the census population at each age was adjusted for net migration into and out of Australia during the three years but no attempt was made to adjust for other differences between the correct exposed to risk and three times the adjusted census population. The exposed to risk figures used are not published in the official report but were supplied to me by the Commonwealth Actuary. The value of σ_r^2 for males aged 25 to 88 calculated from these exposed to risk figures and the three years' deaths was 2·65. Because this figure is lower than that for E.L.T. No. 12, it does not necessarily follow that the Australian data are more accurate than the E.L.T. data. Errors in the data may not be proportional to the exposed to risk and broadly the exposed to risk of E.L.T. No. 12 is about 5 times that of the Australian data.

Pollard (1970), (1971) applies his method 2 to Australian population data for the 6 years 1961 to 1966, so these data have been studied in this paper. The annual Australian Demography Bulletins give for each age the recorded number of deaths and an estimate of the population at 30 June based on the last available census. The r_x test has been applied to the male figures for ages 25–84 in each of the 6 years taking the mid-year population estimate as the exposed to risk; Table 5 gives the resulting values of σ_r^2.

The figures in Table 4 are rather surprising. As would be expected the lowest values of σ_r^2 are those of 1961 and 1966, these being census years for which the population figures will be most accurate. For the other years it might have been

Table 4. *Australian male population data*

Year	σ_r^2
1961	·83
1962	4·08
1963	2·35
1964	3·16
1965	1·94
1966	1·60
Average of above values	2·33

expected that the farther away from the previous census year the less accurate would be the population estimates; in fact the highest σ_r^2 is that of 1962, only one year away from the census, and the general trend thereafter is downwards, the lowest value of an inter-census year being that of 1965 which is four years away from the previous census. It appears from the 1965 Demography Bulletin that the population estimates for that year must have been made before the census of 1966 was taken.

As might be expected the values of σ_r^2 for the census years (1961 and 1966) are less than that for the 1960–62 life tables because the census population at the middle of the year is a better measure of the exposed to risk of that year than is three times this population of the three-year period centered on the census date, even when adjustment is made for migration. For 1961 and 1966 the values of σ_r corresponding to those of σ_r^2 in Table 4 both differ from unity by less than twice the standard deviation, indicating no significant departure from binomial variance (see Appendix 2) and those for the other years are all outside this range.

As has been shown above for assured lives, apart from the effect of duplicates, there does not seem to be much evidence of departures from binomial variance. It therefore appeared reasonable to consider that the high values of σ_r^2 found for the population data so far examined were due to inaccuracies in the method of constructing the tables and to age misstatements. It was therefore surprising to find that Jager (1953), when applying the r_x test to the Netherlands population data obtained values of σ_r^2 surprisingly close to unity. For males in 1947 to 1949 his values are 1·08, ·85 and 1·04 respectively, and for females in 1949, 1·25. This indicated that a more detailed study should be made of the Netherlands data.

A pamphlet entitled 'The Netherlands Central Bureau of Statistics, organisation function and activities', issued in 1969 by the Director General of Statistics, describes the methods used to study mortality. Censuses are normally taken every 10 years but each municipality keeps a population register; this is a card index of all persons resident in the municipality and date of birth and sex are among the information recorded. This register is constantly kept up to date, movements into and out of the municipality being recorded. From these records annual mortality tables are prepared for the whole country. The figures for the years 1961 to 1966 are given in Netherlands Central Bureau of Statistics (1967 and 1972). The following sets of published figures are relevant to the present study:

(i) Population on 31 December of each year at age x last birthday (denoted by P_x^t where t is the year)
(ii) Deaths during the year at age x last birthday at beginning of year of death (denoted by θ_x^t)
(iii) The rates of mortality for each year obtained by the formula

$$q_{x+\frac{1}{2}}^t = \frac{\theta_x^t}{\frac{1}{2}(P_x^{t-1} + P_{x+1}^t + \theta_x^t)} \qquad (1)$$

Owing to the method of collecting the data any age errors will correspond in both population and deaths and will therefore be of the assured lives type referred to in section 2.

The r_x test has been applied to the data for male lives aged 25 to 87 in the six years 1961 to 1966 and the results are given in Table 5.

Table 5. *Netherlands male population data*

Year	σ_r^2
1961	·65
1962	1·62
1963	·90
1964	1·10
1965	·72
1966	·66
Average of above values	·94

Although the values in Table 5 are rather more variable than those of Jager (1953), none of the corresponding σ_r's differ from unity by more than twice the standard deviation (see Appendix 2). There is therefore no evidence of internal heterogeneity in the Netherlands population data.

The Netherlands data demonstrate that population data can be produced and used to calculate mortality rates which exhibit no significant departure from binomial variance. It therefore seems reasonable to attribute the high values of σ_r^2 in the population data of England and Wales and of Australia to errors in the method of obtaining the crude values of q_x and to age misstatements.

The variation of σ_r^2 according to age will now be considered. The values of σ_r^2 for the population data examined above have been calculated for ten-year age-groups. The figures for the three population life tables dealt with above are given in Table 6 together with the values for the whole age range which have already been quoted in the text.

Table 6. *Values of σ_r^2 for male Life Tables*

Age group	E.L.T. No. 10	E.L.T. No. 12	Australia 1960–62
25–34	3·16	·59	·50
35–44	2·05	2·21	1·12
45–54	6·35	1·63	1·67
55–64	8·13	1·34	4·67
65–74	6·08	9·02	2·67
75–84	9·96	7·54	5·35
All ages	5·80	3·66	2·65

Although the values of σ_r^2 for ten-year age-groups are based on only ten individual values of r_x and may therefore be expected to show substantial variations, it seems to be a feature of Table 6 that the larger values occur at the higher ages.

Table 7 gives the values of σ_r^2 for ten-year age-groups for the Australian and Netherlands population data for the six years 1961 to 1966 calculated as described above for the C.M.I. data; all but one of the figures are based on 60 values of r_x.

Table 7. *Values of σ_r^2, 1961 to 1966*

Age group	Australia	Netherlands
25–34	·96	1·16
35–44	1·16	·82
45–54	1·55	·77
55–64	2·25	·72
65–74	4·15	1·01
75–84	4·48*	1·16

* For ages 75–81 only

The Australian figures in Table 7 show a rather similar age pattern to those for the Australian Life Table 1960–62 (Table 6). The Netherlands values of σ_r^2 are fairly constant and all the corresponding values of σ_r are within two standard deviations of unity.

The Netherlands data show that the tendency for σ_r^2 to increase with age is not necessarily a feature of population data. Possible explanations of the high values of σ_r^2 for the three life tables in Table 6 and the Australian population data in Table 7 may therefore be that the errors in the method of obtaining the crude values of q_x are larger at the higher ages or that age misstatements and discordances are more frequent at these ages.

As has already been mentioned the Australian figures for 1961 and 1966 are more accurate than those for the intervening years. It is therefore of interest to examine the age trend of σ_r^2 separately for the census years 1961 and 1966 (combined) and 1962 to 1965; the figures are given in Table 8 and, apart from the oldest age-group, those for 1961 and 1966 are each based on 20 values of r_x and those for 1962 - 1965 on 40 values of r_x.

Table 8. *Values of σ_r^2 for Australia*

Age group	1961 and 1966	1962–1965
25–34	·99	·94
35–44	·85	1·32
45–54	·30	2·18
55–64	1·65	2·55
65–74	2·11	5·16
75–81	1·46	5·98
Average of values in Table 4	1·22	2·88

The figures for 1961 and 1966 are lower than for 1962–65 in all but the youngest age-group. There is a pronounced age trend in 1962–65 and although 1961 and 1966 do not show a regular age trend the larger values of σ_r^2 occur at the older ages. However for no age-group is the corresponding value of σ_r for 1961 and 1966 more than twice its standard deviation from unity. The substantial difference between the two sets of figures might be thought to indicate that age misstatements cannot play a large part in the heterogeneity of the Australian data, and that errors in the estimated populations are largely responsible for the high values of σ_r^2 at the higher ages for 1962–65. Alternatively it might be that the age misstatements in population and deaths tend to match better in the census years than in the intervening years.

The results of this section of the paper can be summed up as follows:

(i) No evidence has been found that internal heterogeneity among lives results in the variance of mortality rates differing from the binomial variance; this applies to investigations based on assured lives and on population data.

(ii) Departures from binomial variance are due to features of or errors in the data or the way in which they are used. Examples are the presence of duplicates in assured lives data, and the use of approximations to the correct exposed to risk in population mortality investigations.

(iii) While a study of the mortality process needs the elimination of any variance due to features or errors of the data (e.g. duplicates), there are situations where it is the variance shown by the data which needs to be studied. For example in testing the graduation of a mortality table the variance taken into account should be the increased variance due to the presence of duplicates. In considering excess loss assurance of death claims (e.g. the payment of all death claims in excess of a specified number among a company's employees) it is necessary to take account of both binomial variance and the variance of the year-to-year levels of mortality (see Section 4) in determining the distribution of deaths.

It is interesting to find from the example of the Netherlands that population data can be produced which appear to be as accurate as assured lives data but without the complication of the presence of duplicates. While as an actuary I applaud the production of accurate figures, as a citizen I would not wish to advocate that the Netherlands methods should be adopted in this country.

4. VARIANCE OF MORTALITY RATES ACCORDING TO YEAR OF EXPERIENCE

In this section the variance of observed mortality rates over successive years will be considered, i.e. the variance across the columns of Table 1. An attempt will then be made to estimate the variance due to external heterogeneity caused by variation in the general level of mortality from year to year.

The χ^2 test (see Appendix 3) has been applied by Seal (1941) and Daw (1945) to study the variance of the mortality of U.K. assured lives over the six years 1924–29 at a few individual ages; they found little evidence of departure from binomial variance. Solomon (1948) in a carefully designed investigation compared the mortality of U.K. assured lives in the three periods of exposure 1924–28, 1929–33, 1934–38 for the age-group 46–55; he found a statistically significant difference between the mortality of each quinquennium.

Now it is known that mortality varies from year to year due to factors such as a cold winter or an epidemic. Thus whether or not the difference between the variance by years of experience and the binomial variance is found to be statistically significant would seem to depend on the size of the exposed to risk studied; if the exposed to risk is large enough statistical significance must be found. I have therefore come to the conclusion that the investigations just mentioned and similar investigations are largely misconceived. It is not usually (but may occasionally be) a meaningful question to ask whether there is a statistically significant difference between the rates of mortality observed in successive years or periods of experience—we know that such differences exist. Also two of the three investigations mentioned made a separate statistical test for each age and so considerably cut down the size of the exposed to risk studied by each test. If any such test is to be made it should, if possible, be over the whole age range, e.g. by adding the values of χ^2 for the individual ages.

The next step will be to study the variance of the mortality rates over the six years 1961 to 1966 for the three sets of data considered in section 3, i.e. a study across the columns of Table 1. The purpose is not to determine whether this variance differs significantly from binomial variance (which has been criticized above), but to use the results in an attempt to estimate the variance due to the external heterogeneity arising from years of experience.

The χ^2 test has been applied at each individual age to the figures of exposed to risk and deaths for the six years (see Appendix 3). The χ^2 calculated for each age will have 5 degrees of freedom (denoted by χ_5^2) since each is based on figures for six separate years. Column (2) of Table 9 shows the total of the values of χ_5^2 for ten year-age-groups, i.e. values of χ_{50}^2. The total of the χ_5^2s over the whole age range will be a χ_{300}^2 (χ_{285}^2 for Australia). Now χ_f^2/f is an estimate of the ratio of the actual variance of the rates of mortality over the six years to the binomial variance. The actual variance will include all the factors included in σ_r^2 (except any variance due to variations in the internal heterogeneity between successive ages) and will also include the variance between successive years. Thus $(\Sigma\chi_5^2)/f$ will be expected to be greater than the corresponding value of σ_r^2 and this can be seen from Table 9 to be the case for most age-groups.

If

(i) the binomial variance adjusted for any internal heterogeneity (i.e. that measured by σ_r^2), and

(ii) the variance due to external heterogeneity between years of experience (which will be denoted by var Y)

can be regarded as independent and additive, the difference between $(\Sigma\chi_5^2)/f$ and σ_r^2 gives a measure of var Y expressed as a proportion of the binomial variance. Var Y is an estimate of the variance of the general levels of mortality applicable to the six years 1961 to 1966 and, for the ten-year age-groups used in Table 9, is given by the formula

$$\text{var } Y = \left(\frac{\Sigma\chi_5^2}{50} - \sigma_r^2\right)\frac{\bar{p}\,\bar{q}}{E} \tag{2}$$

where $\frac{\bar{p}\bar{q}}{E}$ is an estimate of the average binomial variance. The values of $\bar{q}$ have been calculated from the total exposed to risk and deaths for the ten-year age-group and E has been taken as the average exposed to risk for the ten ages and the six years of experience, i.e. as one-sixtieth of the total exposed to risk for the age-group.

It may be helpful to mention that the statistical model envisaged by formula (2) is the Lexian modification of binomial sampling, where the variance considered is that of k sets of n repeated trials in which the value of q is constant for each of the n trials within a set but varies from set to set. The variance of the k values of q which apply to the k sets corresponds to var Y in which k is 6 and 'sets' are years of experience (see e.g. Aitken (1939), p. 52).

A value of χ_5^2 for age x is based on the q_xs at age x only, but r_x involves the values of q_x, q_{x+1}, q_{x+2} and q_{x+3}. Thus σ_r^2 for the ten-year age-group x to $x+9$ will utilize values of $q_x, q_{x+1} \dots, q_{x+12}$ but with lower weights to the three values at each end of the range than at the intermediate values. $(\Sigma\chi_5^2)/50$ and σ_r^2 are therefore not truly comparable, since they cannot each relate with equal weight to mortality at the same ages. However, the agreement will be improved for ten-year age-groups if σ_r^2 for ages x to $x+9$ is compared with $(\Sigma\chi_5^2)/50$ for ages $x+1$ to $x+10$. For this reason the values of $(\Sigma\chi_5^2)/f$ for age-groups 26–35, 36–45, etc., in Table 9 are compared with σ_r^2 for age-groups 25–34, 35–44, etc., which have already been given in Tables 3 and 7.

In some of the age-groups of Table 9 $(\Sigma\chi_5^2)/f$ is less than σ_r^2, which by formula (2) would give the anomalous result of a negative var Y; the values of var Y have been omitted for these age-groups. A few such discrepancies are to

Table 9. *Variance between years of experience 1961 to 1966*

Age group x	$\Sigma\chi_5^2$	$\frac{\Sigma\chi_5^2}{f}$	σ_r^2	10 var Y (by formula (2))	$\frac{\text{var } Y}{\bar{q}^2}$
(1)	(2)	(3)	(4)	(5)	(6)
Continuous Mortality Investigation					
26–35	73	1·46	1·15	·03	·00488
36–45	101	2·02	1·72	·05	·00169
46–55	136	2·72	1·75	·46	·00172
56–65	159	3·17	1·80	3·23	·00159
66–75	60	1·21	1·35	negative	
76–85	101	2·03	1·56	140·48	·00144
All ages	630	2·10	1·56	·59	·00144
Australian male population					
26–35	58	1·15	·96		
36–45	54	1·09	1·16	negative	
46–55	68	1·36	1·55	negative	
56–65	120	2·39	2·25		
66–75	218	4·36	4·15		
76–82	101	2·90	4·48	negative	
All ages	619	2·17	2·33	negative	
Netherlands male population					
26–35	47	·95	1·16	negative	
36–45	44	·88	·82	·02	·00036
46–55	68	1·36	·77	·65	·00135
56–65	93	1·86	·72	4·16	·00119
66–75	134	2·68	1·01	21·63	·00116
76–85	173	3·46	1·16	157·14	·00150
All ages	559	1·86	·94	2·50	·00125

be expected when two variance ratios obtained by such different methods as are $(\Sigma\chi_5^2)/f$ and σ_r^2 are compared. The fact that only one age-group in the C.M.I. and in the Netherlands figures show this discrepancy is a further indication of the basic accuracy and consistency of these data. In the Australian population data three of the six age-groups and the all-ages group have $(\Sigma\chi_5^2)/f$ less than σ_r^2. In these data a substantial part of the variance appears to be due to errors in the population estimates and these will no doubt affect $(\Sigma\chi_5^2)/f$ and σ_r^2 in very different ways. In fact it is probably expecting too much of the Australian population estimates for inter-census years to use them as exposed to risk figures for a mortality study; they were so used because Pollard (1970) applied two of his methods to these data.

Table 9 gives in column (5) the values of var Y calculated by formula (2). These figures are of comparatively little interest in themselves as they represent the experience of only six years which may or may not be typical of the general run of years. If there is any trend in the mortality rates over the years this will contribute to and be included in the estimates of var Y. A study of the three sets of data used does not indicate any consistent trend and gives the impression that at the most only a very small part of var Y can be due to a trend in the mortality rates.

Jager (1953) mentions a number of Dutch investigations which regard the variation in mortality rates over a series of years as consisting of two parts, the binomial variance and the year-to-year variance; the present study makes a similar assumption but takes the random variance as being proportional to the binomial variance and thus allows for features of the data such as the presence of duplicates. Jager (1953) also mentions a similar investigation which allows for a linear trend over the period of years studied. All these works appear to be published in the Dutch language but a summary of them is to be found in Derksen (1948).

Of more interest than the actual values of var Y given in Table 9 is their age trend. The values increase rapidly with increasing age as might perhaps have been expected from a study of the graphs in the long series of notes in *J.I.A.* on 'The recent trend of mortality in Great Britain' and for certain specific causes of death, particularly bronchitis, in Daw (1954).

The next question to be considered is the nature of the age increase. Column (6) of Table 9 gives the values for each age group of (var $Y)/\bar{q}^2$; apart from the youngest ages these values are remarkably constant both within and between each set of data, considering the way in which they have been derived.

The near constancy of (var $Y)/\bar{q}^2$ can be expressed alternatively by saying that the variance of the mortality rates from year to year is proportional to the square of the rate of mortality. However the study of a short period of six years for assured lives in the United Kingdom and for Netherlands male population data gives little ground for a general assertion that the year-to-year variance of mortality rates is always proportional to the square of the mortality rate; that needs a much wider network of calculations on accurate mortality data for other periods of years and other countries. A further question not covered by any of the calculations in this paper is whether a similar relation holds for female lives or whether the biological difference between the sexes gives rise to a different relationship.

5. CONCLUSION

The results of this paper may be summed up by saying that no evidence has been found that internal heterogeneity of lives results in the variance of mortality rates departing from binomial variance; all the departures which have been found are man-made (e.g. duplicate policies). I think it is true to say that mortality data and the methods of studying them are not sufficiently accurate to detect possible small real departures from binomial variance, but nevertheless that if any such departures do exist they must be small.

For two sets of data it is found that the variance due to changing levels of mortality from year to year (i.e. var Y) is proportional to the square of the rate of mortality. This gives rise to the question of whether the relationship is peculiar to these experiences or is a more general phenomenon; to answer this further numerical work is needed.

While this paper started out as a study of the variance of mortality rates it has turned out to be of perhaps more relevance to methods of constructing mortality tables. Left to themselves mortality rates appear to show a variance little, if at all, different from binomial variance, it is usually the methods of construction or errors in the data which produce supra-binomial variance. The results may be helpful in indicating where methods of construction may be improved. The paper also shows the importance of considering the results of studies of the variance of mortality rates in relation to the methods used to produce the rates and the nature of the variation which is being measured.

6. ACKNOWLEDGMENT

I would like to give my grateful thanks to Mr. F. M. Redington for his comments and advice. He unerringly picked out from an earlier draft all the points on which I had vague doubts and, as a result, I like to think that the paper now makes a more positive contribution to the subject.

APPENDIX 1

THE r_x TEST

The r_x test was devised by Redington and Michaelson (1940) and later independently by Jager (1953). It is designed to test whether the variance of mortality rates follows the binomial distribution or whether internal heterogeneity is present using that term as it is defined in Section 2 of this paper. The third differences of graduated rates of mortality are usually small so that the third differences of the ungraduated q_xs can be regarded as mainly composed of errors, random or otherwise. On the assumption of binomial variance the standard deviation of q_x is $\sqrt{\left(\frac{p_x q_x}{E_x}\right)}$ and the standard deviation of $\Delta^3 q_x$ ($\sigma(\Delta^3 q_x$, say) is given by

$$\sigma(\Delta^3 q_x) = \sqrt{\left\{\frac{p_{x+3}\,q_{x+3}}{E_{x+3}} + 9\,\frac{p_{x+2}\,q_{x+2}}{E_{x+2}} + 9\,\frac{p_{x+1}\,q_{x+1}}{E_{x+1}} + \frac{p_x q_x}{E_x}\right\}} \qquad \text{(A1)}$$

and the statistic r_x by

$$r_x = \Delta^3 q_x / \sigma(\Delta^3 q_x) \qquad \text{(A2)}$$

The values of r_x calculated for each age from the mortality data should be very nearly normally distributed with mean zero and variance of unity if the binomial distribution is applicable to the mortality data. If the variance differs from binomial variance the observed variance of r_x (say σ_r^2) will differ from unity. Thus σ_r^2 gives a measure of the average value of the ratio of the observed variance to the binomial variance.

As described by Redington and Michaelson (1940) and as applied by Daw (1945), $\sigma(\Delta^3 q_x)$, the denominator of r_x was calculated using graduated rates of mortality. However, Jager (1953) used the ungraduated rates; this has the advantage of making the r_x test completely independent of any graduation. In his reply to the discussion of his paper Daw (1945) gave an example showing that the use of graduated or ungraduated mortality rates in calculating $\sigma(\Delta^3 q_x)$ had only a negligible effect on the value of σ_r^2 and a few more such comparisons have been made in the course of the present study. The values of σ_r^2 for the Australian male population data given in Table 4 were calculated using the ungraduated rates and three of these have been recalculated using the graduated q_xs of the Australian male Life Table 1960–62; the comparison is set out in Table A1.

These figures show that the alternative methods have only a small effect on σ_r^2 in spite of the fact that the graduated rates used are not strictly applicable to any of the years of experience and are probably least appropriate for 1966.

All of the values of σ_r^2 given in this paper except those for the Australian assured lives and Life Table and E.L.T. No. 10 have been calculated using the ungraduated q_xs and so are quite independent of any graduation.

Table A1. *Values of σ_r^2 from Australian male population data calculating σ ($\Delta^3 q_x$) from:*

Year	Ungraduated q_x of the year (see Table 4)	Graduated q_x of Australian Male Life Tables 1960–62
1962	4·08	4·11
1963	2·35	2·39
1966	1·60	1·68

APPENDIX 2

STANDARD DEVIATION OF σ_r

Daw (1945) gives in Appendix 3 the standard deviation of σ_r as $1{\cdot}5\ \sigma_r/\surd(2n)$ where n is the number of values of r_x on which σ_r is based (i.e. three less than the range of ages used to calculate σ_r). The tables in the paper give only the values of σ_r^2 but the sampling distribution of σ_r is likely to diverge less from the normal distribution than that of σ_r^2. Table A2 has therefore been included showing the all-ages values of σ_r corresponding to those of σ_r^2 given in the paper, in order that some idea may be obtained of the statistical significance of any of these values.

When considering whether a value of σ_r is significantly different from unity (i.e. binomial variance) the value of σ_r used in the above formula for its standard

Table A2. *Values of σ_r and its standard deviation*

	n	1961	1962	1963	1964	1965	1966	S.D. of σ_r $= 1{\cdot}5/\surd(2n)$
Experiences of 1961–66								
C.M.I.	60	1·46	1·24	1·34	1·00	1·02	1·36	·14
Australian males	57	·91	2·02	1·53	1·78	1·39	1·27	·14
Netherlands males	60	·80	1·27	·95	1·05	·85	·81	·14

		n	σ_r	
England and Wales:	E.L.T. No. 10, males	70	2·41	·13
	E.L.T. No. 12, males	67	1·91	·13
Australia:	Assured lives 1958–63	62	1·34	·13
	Life Table 1960–62, males	61	1·63	·14

deviation should be unity (not the observed σ_r), since this is part of the hypothesis being considered. Assuming that σ_r is normally distributed, a value outside the range $1 \pm 2\,(1\cdot5)/\surd(2n)$ will be taken as statistically significant at the 5% level; Table A2 gives the values of $1\cdot5/\surd(2n)$.

APPENDIX 3

THE χ^2 TEST

The χ^2 test is described by Seal (1941); it is used in this paper to study the variance of observed mortality rates over successive years, i.e. the variance across the columns of Table 1. When six years of experience are involved the criterion to be calculated for each individual age is

$$\chi_5{}^2 = \sum_{t=1}^{6} \frac{(\theta_x{}^t - E_x{}^t \bar{q}_x)^2}{E_x{}^t \bar{p}_x \bar{q}_x} \tag{A3}$$

where t (= 1, 2, . . . 6) represents the year of experience,

$$\bar{q}_x = \sum_t \theta_x{}^t / \sum_t E_x{}^t \tag{A4}$$

and χ_5^2 is approximately distributed as χ^2 with 5 degrees of freedom.

The χ^2 test is similar to Pollard's (1970) method 2 but would seem to have a better theoretical basis as it does not involve a rating up or down of the observed deaths as does Pollard's method.

It is sometimes found that the proportions of the exposed to risk for an age-group which are applicable to the individual ages of the group show a distinct time trend. If the χ^2 test is applied to an age-group over a period of years of experience where this time trend exists, the value of χ^2 for the age-group will be artificially inflated because the average age of the age-group will be changing over the period. For this reason the χ^2 tests made in this paper have been for individual ages and not for age-groups. Pollard's (1970) method 2 can suffer from the objection just described.

REFERENCES

AITKEN, A. C. Statistical Mathematics. London: Oliver and Boyd, 1939.

BEARD, R. E. Some notes on graduation. *J.I.A.* 1951, **77**, 382.

COMMONWEALTH ACTUARY. Australian Life Tables 1960–62. Commonwealth Bureau of census and statistics, Canberra, 1965.

Continuous investigation into the mortality of assured lives. Memorandum on a special enquiry into the distribution of duplicate policies. *J.I.A.* 1957, **83**, 34.

C.M.I. Considerations affecting the preparation of standard tables of mortality. *J.I.A.* 1974, **101**, 133.

DAW, R. H. On the validity of statistical tests of the graduation of a mortality table. *J.I.A.* 1945, **72**, 174.

DAW, R. H. Duplicate policies in mortality data. *J.I.A.* 1951, **77**, 261.

DAW, R. H. Some statistical aspects of mortality from degenerative heart disease. *J.I.A.* 1954, **80**, 69.

DERKSEN, J. B. D. The calculation of mortality rates in the construction of life tables. A mathematical statistical study. *Population Studies* 1948, **1**, 457.

HARDY, G. H. On the mortality observed amongst the various classes of bonus policies in the British Empire Mutual Assurance Company. *J.I.A.* 1882, **23**, 1.

JAGER, J. DE. Stochastical investigations on mortality tables. *Het Verzekerings-Archief, Actuarieel Bijvoegsel* 1953, **30**, 76.

NETHERLANDS CENTRAL BUREAU OF STATISTICS. Life tables for the Netherlands 1961–1965. (1967)

NETHERLANDS CENTRAL BUREAU OF STATISTICS. Life tables for the Netherlands 1966–1970. (1972).

POLLARD, A. H. Random mortality fluctuations and the binomial hypothesis. *J.I.A.* 1970, **96**, 251.

POLLARD, A. H. Correspondence. *J.I.A.* 1971, **98**, 91.

REDINGTON, F. M. and MICHAELSON, R. L. An aspect of the *a priori* probability theory of mortality. *Trans. Twelfth Int. Cong. Act., Lucerne*, 1940, **I**, 225.

SEAL, H. L. Tests of a mortality table graduation. *J.I.A.* 1941, **71**, 5.

SEAL, H. L. A probability distribution of deaths at age *x* when policies are counted instead of lives. *Skand. Actuar.* 1947, **30**, 18.

SEAL, H. L. The probability of decrements from a population. A study in discrete random processes. *Skand. Actuar.* 1948, **31**, 14.

SEAL, H. L. Mortality data and the binomial law. *Skand. Actuar.* 1949, **32**, 188.

SOLOMON, L. The analysis of heterogeneous mortality data. *J.I.A.* 1948, **74**, 94.

VAJDA, S. The analysis of variance of mortality rates. *J.I.A.* 1945, **72**, 240.