JIA 79 (1953) 0221-0224

THE '*a*' AND '*b*' DISTRIBUTION CONCEPT

BY E. OLIFIERS

Consulting Actuary, Rio de Janeiro

THIS concept has been used in demographic statistics, but not in insurance statistics, to find exposed-to-risk formulae. In my paper *The Evolution of the Exposure Formulae* (1947, *T.A.S.A.* XLVIII, 76) I described the '*a*' and '*b*' Distribution Concept with special application to policies classified according to assumed ages on policy anniversaries, and by its means I was able to demonstrate the connexion between the calendar year, the policy year, and what I called the calendar policy year methods of obtaining rates of mortality and, moreover, to exhibit the formulae in their dual forms, i.e. in census form and in the more usual form suitable for an individual computation of periods of exposure by the use of cards.

The '*a*' and '*b*' concept need not be restricted to policy years. In the present note I propose to give an outline of the concept as applied to life years and by its means to link up the calendar year, life year, census, and what I call the calendar life year methods of obtaining rates of mortality. By census method is meant the method which is commonly used in forming life tables from population statistics and was also used in the A 1924–29 investigation. It will simplify the presentation to use the life year extending from one birthday to the next, which implies the use of ages last birthday for the census method. (In the A 1924–29 investigation ages nearest birthday were used; the appropriate life year to use extends from six months before to six months after a birthday.)

The following symbols all relate to persons who attained or who, if they had survived, would have attained age x in the calendar year y. Thus they are persons born in the calendar year $y-x$:

e_x^y	denotes those still under observation on their birthday in year y.
$e_{\overline{x}\rceil}^y$	denotes those still under observation on 31 December of year y.
${}_an_x^y$, ${}_a\theta_x^y$, ${}_aw_x^y$	denote those entering the experience, dying, or withdrawing from the experience respectively after their birthday anniversaries in year y.
${}_bn_{x-1}^y$, ${}_b\theta_{x-1}^y$, ${}_bw_{x-1}^y$	denote those entering the experience, dying or withdrawing from the experience respectively before their birthday anniversaries in year y.

Withdrawals are intended to be net withdrawals, i.e. withdrawals less revivals.

The suffix x or $x-1$ as the case may be denotes the age last birthday at the time of the event indicated by the particular symbol used.

It will be seen that the life year extending from the birthday in year $y-1$ to the birthday in year y is divided into an earlier part '*a*' from the birthday in year $y-1$ to 31 December of year $y-1$ and a later part '*b*' from 31 December of year $y-1$ to the birthday in year y.

The connecting link between $e_{\underline{x-1}|}^{y-1}$, e_x^y and $e_{\underline{x}|}^y$ is given by

$$e_{\underline{x-1}|}^{y-1} - {}_b\theta_{x-1}^y - {}_bw_{x-1}^y + {}_bn_{x-1}^y = e_x^y = e_{\underline{x}|}^y + {}_a\theta_x^y + {}_aw_x^y - {}_an_x^y, \tag{1}$$

in which the left-hand side expresses the prospective relation between the survivors on 31 December of year $y-1$ and those who reach their birthday in the year y, whilst the right-hand side expresses the retrospective relation between the survivors on 31 December of year y and those who reach their birthday in the year y.

The right-hand side of (1) expresses the exposed to risk in a census form on the assumption that the outgoing movement denoted by $w_x^{y,\,y+1} = {}_aw_x^y + {}_bw_x^{y+1}$ which occurs among e_x^y during the life years current on 31 December of year y is exposed to the risk of death a whole year and that the incoming movement denoted by $n_x^{y,\,y+1} = {}_an_x^y + {}_bn_x^{y+1}$ is not exposed at all to the risk of death. This assumption is not generally true so that we have to add to e_x^y the sum of the fractional periods between the exact time at which the net movements occur and the end of the life years then current, which may be denoted by

$$f(n_x^{y,\,y+1} - w_x^{y,\,y+1}).$$

The following formula is obtained for the life year exposed to risk E_x^y

$$E_x^y = e_x^y + f(n_x^{y,\,y+1} - w_x^{y,\,y+1}) = e_{\underline{x}|}^y + {}_a\theta_x^y + {}_aw_x^y - {}_an_x^y + f(n_x^{y,\,y+1} - w_x^{y,\,y+1})$$
$$\doteqdot e_{\underline{x}|}^y + {}_a\theta_x^y, \tag{2}$$

in which it is assumed that

$${}_an_x^y - {}_aw_x^y \doteqdot f(n_x^{y,\,y+1} - w_x^{y,\,y+1}).$$

The calendar year, life year, census and calendar life year mortality rates may now be written in their census form, i.e. in a form which would enable data to be collected by censuses of business in force on 31 December of each year and schedules of deaths for each year.

Calendar year rate:

$$q_{x-1}^y = \frac{{}_b\theta_{x-1}^y + {}_a\theta_x^y}{E_{\underline{x-1}|}^y} = \frac{{}_b\theta_{x-1}^y + {}_a\theta_x^y}{\frac{1}{2}(e_{\underline{x-1}|}^{y-1} + e_{\underline{x}|}^y + {}_b\theta_{x-1}^y + {}_a\theta_x^y)}. \tag{3}$$

Life year rate:

$$q_x^{y,\,y+1} = \frac{{}_a\theta_x^y + {}_b\theta_x^{y+1}}{E_x^y} = \frac{{}_a\theta_x^y + {}_b\theta_x^{y+1}}{e_{\underline{x}|}^y + {}_a\theta_x^y}. \tag{4}$$

Census rate:

$$q_x^y = \frac{m_x}{1 + \frac{1}{2}m_x} = \frac{{}_b\theta_x^y + {}_a\theta_x^y}{\frac{1}{2}(e_{\underline{x}|}^{y-1} + e_{\underline{x}|}^y + {}_b\theta_x^y + {}_a\theta_x^y)}. \tag{5}$$

Calendar life year rate:

$$q_x'^y = \frac{{}_b\theta_x^y + {}_a\theta_x^y}{\frac{1}{2}(E_x^{y-1} + E_x^y)} = \frac{{}_b\theta_x^y + {}_a\theta_x^y}{\frac{1}{2}(e_{\underline{x}|}^{y-1} + {}_a\theta_x^{y-1} + e_{\underline{x}|}^y + {}_a\theta_x^y)}. \tag{6}$$

Although for the sake of comparison each of the above four mortality rates is expressed in terms of a and b symbols it is only necessary to tabulate a and b deaths separately for the life year rate (4) and the calendar life year rate (6). For the calendar year rate (3) deaths may be tabulated according to year of birth $y-x$ and for the census rate (5) according to age last birthday x at death.

The factor $\frac{1}{2}$ in the denominator of the calendar year rate (3) involves the assumption of an even distribution of the net movement occurring in the year y

as may be seen by adding $e^{y-1}_{\underline{x-1}|} + {}_b\theta^y_{x-1} - {}_aw^y_x + {}_an^y_x$ to both sides of (1) and dividing by two. The alternative expression $e^{y-1}_{\underline{x-1}|} + \tfrac{1}{2}({}_bn^y_{x-1} + {}_an^y_x - {}_bw^y_{x-1} - {}_aw^y_x)$ for the exposed to risk is then obtained.

The calendar life year rate (6) is found from the observation of lives born in two successive years. ${}_b\theta^y_x$ in the numerator denotes the deaths in the later part b of the life year starting on the birthday in year $y-1$ and thus arises out of persons born in the year $y-x-1$. The appropriate exposed to risk is E^{y-1}_x. ${}_a\theta^y_x$ in the numerator denotes the deaths in the earlier part a of the life year starting on the birthday in year y and thus arises out of persons born in the year $y-x$. The appropriate exposed to risk is E^y_x. The factor $\tfrac{1}{2}$ in the exposed to risk $\tfrac{1}{2}(E^{y-1}_x + E^y_x)$ expresses the assumption of an equal distribution of the living over each of the calendar years $y-1$ and y.

Since at advanced ages, where there is no new business, the deaths appearing in the numerator of the calendar year rate (3) and the life year rate (4) must necessarily arise out of the in force appearing in the respective denominators, the calendar year rate (3) and the life year rate (4) are superior to the census rate (5) and the calendar life year rate (6) in which that absolute correspondence between numerator and denominator does not exist. Thus the mortality rates obtained from formulae (5) and (6) may, but those obtained from formulae (3) and (4) cannot, exceed unity at the older ages. It is worth pointing out, however, that if q_x is obtained from $1-e^{-m_x}$(see Spurgeon, *Life Contingencies*, p. 16) instead of from $m_x/(1+\tfrac{1}{2}m_x)$ as in the census rate (5) the rate of mortality cannot exceed unity no matter what value is found for the central rate of mortality $({}_b\theta^y_x + {}_a\theta^y_x)/\tfrac{1}{2}(e^{y-1}_{\underline{x}|} + e^y_{\underline{x}|})$.

By dividing the deaths in the two mortality rates (3) and (5) according to a and b deaths it may be seen that if the suffix $x-1$ of ${}_b\theta^y_{x-1}$ of the numerator and of $\tfrac{1}{2}(e^{y-1}_{\underline{x-1}|} + {}_b\theta^y_{x-1})$ of the denominator of the calendar year rate (3) is advanced one year to x the calendar year rate (3) becomes the census rate (5). It may also be seen by the a and b concept that the calendar life year rate (6) is most akin to the life year rate (4), for if this rate is written

$$\frac{{}_b\theta^{y+1}_x + {}_a\theta^y_x}{\tfrac{1}{2}(e^y_{\underline{x}|} + {}_a\theta^y_x + e^y_{\underline{x}|} + {}_a\theta^y_x)},$$

and if y of ${}_b\theta^{y+1}_x$ of the numerator and of the first $\tfrac{1}{2}(e^y_{\underline{x}|} + {}_a\theta^y_x)$ of the denominator is brought back one year to $y-1$, the life year rate (4) becomes the calendar life year rate (6). If observations are taken over a number of years the overlap between the calendar life year rate (6) and the theoretically more accurate life year rate (4) becomes greater and greater. On the other hand, the census rate (5) becomes no nearer to the calendar year rate (3) if observations are taken over a number of years. It may also be seen that the only difference between the census rate (5) and the calendar life year rate (6) lies in one term of the denominator, i.e. ${}_a\theta^{y-1}_x$ having been substituted for ${}_b\theta^y_x$ in order to give a year of exposure to each death occurring in the life year current on 31 December of year $y-1$ and in this difference lies the superiority of the calendar life year rate (6).

My interest in mortality exposed to risk formulae began in 1945 because the Brazilian State Reinsurance Institute was authorized under a Brazilian Statutory provision to demand that life assurance companies should supply an individual card for each policy for the purpose of conducting a continuous mortality investigation on the lines of the early voluntary investigations of the British Life Offices. The saving of work which would follow dispensing with cards and supplying the information by a census of policies in force at 31 December each year made me wish to follow the example of the British Life Offices in their A1924–29 investigation, but modifying the formulae to suit the conditions in which the Brazilian continuous mortality investigation had to proceed. My first paper on the '*a*' and '*b*' Distribution Concept was presented to the Brazilian Institute of Actuaries under the title *Uma Simbologia Racional das Formulas dos Expostos ao Risco*, and was followed two years later by the paper *The Evolution of the Exposure Formulae* to which reference has been made earlier in this note. The arguments were sufficiently convincing to induce the Brazilian State Reinsurance Institute to change as from 31 December 1949 the process of collecting mortality statistics from individual cards to a census of the business in force at 31 December of each year.