JSS 21 (3) (1976) 341-345

THE APPLICATION OF SEQUENTIAL ANALYSIS TO MORTALITY DATA FOR ASSURED LIVES

by

T. S. LEIGH

1. It is some years since the application of sequential analysis to actuarial work was discussed in the *Journal* and the following paragraphs illustrate how the techniques may be used to help solve a current problem.

2. The data relate to the experience of one large office during the calendar years 1970–73. It corresponds to the returns made to the Continuous Mortality Investigation relating as it does to full premium business for assured lives at duration 2 and over. All the lives have been accepted at normal rates of premium and for convenience the two forms of acceptance (Medical and Non-Medical) have been combined.

3. The office uses a valuation basis of A1949–52 although it has devised Office '70, a table suggested by its own experience. Office '70 is 80% of A1949–52 up to and including age 65 increasing by 1% thereafter to 90% at age 75 and above. The forthcoming publication of A1967–70 provided speculation as to how this would compare with Office '70 and the nature of our problem is to decide which is the more suitable.

4. Table 1 shows the values of the death rate, q_x, for decennial ages according to the various mortality bases. The values shown for A1967–70 are those relating to A1967–70 (2) i.e. based on an experience of duration 2 and over.

Table 1. Values of q_x on various mortality bases

Age x	*A1949–52*	*Office '70*	*A1967–70*
35	·00132	·00106	·00086
45	·00330	·00264	·00264
55	·01035	·00828	·00844
65	·02810	·02248	·02403
75	·07257	·06531	·06229
85	·17282	·15554	·14727
95	·34683	·31215	·30593

As can be seen from Table 1 Office '70 is heavier than A1967–70 at the younger and older ages and lighter at the middle ages. The difference between the two tables, however, is nowhere near as marked as that between them and A1949–52.

5. The theory underlying sequential analysis has been explained elsewhere and the interested reader is referred to the bibliography. The theory in this article follows that detailed by Messrs Johnson and Moore and the rudiments are briefly summarized in the following paragraphs.

6. We wish to choose between two hypotheses, H_o and H_1 and to control the probability of erroneous rejection, α and β respectively, and $P\{x|H_o\}$ and $P\{x|H_1\}$ represent the probability for the actual set of observed values (x) on the assumption that H_o and H_1 respectively are valid. We then calculate

$$R = \frac{P\{x|H_o\}}{P\{x|H_1\}}$$

and compare its value with two fixed limits A and B, being $\beta/(1-\alpha)$ and $(1-\beta)\,\alpha$ respectively. If $R \geqslant B$, H_1 is accepted and if $R \leqslant A$, H_o is accepted. If, however, $B > R > A$, further values are required and the technique is to work *backwards* from the most recent year using a cumulative total until such time as a conclusion may be drawn.

7. The following three tests will be performed:

	H_o	H_1
(i)	Accept A1949–52	Accept A1967–70
(ii)	Accept A1949–52	Accept Office '70
(iii)	Accept Office '70	Accept A1967–70

8. The observed data consist of exposed to risk (E_x^j) and deaths (d_x^j) for a number of years (j) and ages (x). Let q_x^o and q_x^1 represent the death rates according to the mortality basis underlying the hypotheses H_o and H_1 respectively. Let R_t represent the value of the statistic for t years' experience and for computational purposes we will use:

$$L_t = \log R_t$$

$$= \sum_x \left[\log\left(\frac{1-q_x^1}{1-q_x^o}\right)\right]\left(\sum_{j=1}^{t} E_x^j\right) + \sum_x \left[\log \frac{q_x^1(1-q_x^o)}{q_x^o(1-q_x^1)}\right]\left(\sum_{j=1}^{t} d_x^j\right)$$

9. The coefficients of $\sum_{j=1}^{t} E_x^j$ and $\sum_{j=1}^{t} d_x^j$ are independent of t and Tables 2 and 3 show their values at appropriate ages.

Table 2. *Coefficient of* $\sum E_x^j$

Age	*Test (i)*	*Test (ii)*	*Test (iii)*
35	·00021	·00013	·00009
45	·00030	·00030	·00000
55	·00080	·00089	−·00006
65	·00179	·00246	−·00070
75	·00478	·00331	·00135
85	·01322	·00898	·00415
95	·02636	·02244	·00381

Table 3. *Coefficient of* $\sum d_x^j$

Age	*Test (i)*	*Test (ii)*	*Test (iii)*
35	−·18630	−·09539	−·09094
45	−·09720	−·09721	·00000
55	−·08945	−·09782	·00844
65	−·06976	−·09941	·02962
75	−·07112	−·04917	−·02197
85	−·08269	−·05473	−·02976
95	−·08087	−·06821	−·01265

Table 4. *Observed results*

Age	*1973*		*1972*		*1971*		*1970*	
	E_x^1	d_x^1	E_x^2	d_x^2	E_x^3	d_x^3	E_x^4	d_x^4
35	50782	38	45698	34	50829	46	50006	49
45	30952	82	29203	96	35938	109	36035	115
55	20971	159	20543	222	25947	188	25497	224
65	5995	137	5961	170	8000	140	8316	172
75	1254	65	1232	81	1633	130	1679	97
85	407	67	428	88	537	85	671	119
95	58	20	51	15	57	21	71	20

10. Table 4 shows the observed experience for the various calendar years working backwards from 1973.

The results themselves are not directly comparable. The data for 1972 and 1973 exclude duplicate policies and thereby ensure that a life is only included once. Those for 1970 and 1971 only exclude simultaneous duplicates and duplicates for lives over age 80. However, this difference in collecting the statistics should have little bearing on the points at issue. Of more importance is the possibility of secular changes over the years and for this reason the figures in Table 4 go no further back than 1970.

The data have been presented in decennial age-groupings and are assumed to be concentrated at the mid-points of the various groupings although ages 35 and 95 respectively have been chosen for the 'up to age 40' and 'over age 90' groupings respectively.

12. The tests are performed assuming 90% confidence limits α and β are both made equal to 0·05 and

$$-\log_{10}A = \log_{10}B = 1{\cdot}279$$

13. For Test (i) the first year's experience gives

$$\begin{aligned} L_1 &= (0{\cdot}00021\times 50{,}782)+(0{\cdot}00030\times 30{,}952)+\ldots \\ &+ (-0{\cdot}18630\times 38)+(-0{\cdot}09720\times 82)+\ldots \\ &= 60{\cdot}368-50{\cdot}617 \\ &= 9{\cdot}751 \end{aligned}$$

As $9{\cdot}751>1{\cdot}279$, the hypothesis H_1 'Accept A1967–70' is accepted.

14. For Test (ii) the first year's experience gives

$$\begin{aligned} L_1 &= (0{\cdot}00013\times 50{,}782)+(0{\cdot}00030\times 30{,}952)+\ldots \\ &+ (-0{\cdot}09539\times 38)+(-0{\cdot}09721\times 82)+\ldots \\ &= 58{\cdot}408-48{\cdot}995 \\ &= 9{\cdot}413 \end{aligned}$$

As $9{\cdot}413>1{\cdot}279$ the hypothesis H_1 'accept Office '70' is accepted.

15. For Test (iii) the first year's experience gives

$$\begin{aligned} L_1 &= (0{\cdot}00009\times 50{,}782)+(0{\cdot}00000\times 30{,}952)+\ldots \\ &+ (-0{\cdot}09094\times 38)+(0{\cdot}00000\times 82)+\ldots \\ &= 2{\cdot}719-1{\cdot}731 \\ &= 0{\cdot}988 \end{aligned}$$

This result does not enable us to reach a decision and we, therefore, take the second year's experience which gives

$$\begin{aligned} &(0{\cdot}00009\times 45{,}698)+(0{\cdot}0000\times 29{,}203)+\ldots \\ &+ (-0{\cdot}09094\times 34)+(0{\cdot}00000\times 96)+\ldots \\ &= 2{\cdot}340-0{\cdot}772 \\ &= 1{\cdot}568 \end{aligned}$$

leading to

$$\begin{aligned} L_2 &= 0{\cdot}988+1{\cdot}568 \\ &= 2{\cdot}556 \end{aligned}$$

As $2{\cdot}556>1{\cdot}279$ the hypothesis H_1 'Accept A1967–70' is accepted.

16. Whilst there can be little doubt about the validity of the conclusions drawn from Tests (i) and (ii), there is the probability of the wrong hypothesis being accepted for Test (iii). Nevertheless this

probability is not very high as L_2 is also acceptable at the 95% confidence level ($2{\cdot}556 > 1{\cdot}591$).

17. However, if we continued to work backwards, $-2{\cdot}032$ and $-0{\cdot}922$ would be obtained for 1971 and 1970 respectively leading to $L_3 = 0{\cdot}524$ and $L_4 = -0{\cdot}398$. Both these values are inconclusive but it is somewhat disconcerting to realize that had the test been performed with 1971 as the first year the hypothesis H_o, 'Accept Office '70', would have been accepted.

18. Of course, there is not a vast difference between Office '70 and A1967–70 and it is worth noting that by using the techniques of sequential analysis a decision may be reached on the basis of two years' experience. Nevertheless any conclusions drawn must be tempered by the approximations surrounding the age-groupings, the treatment of duplicate policies, and by the comments made in the previous paragraphs concerning the possibility of a wrong hypothesis being accepted.

19. The above result should help the office to reach a decision regarding the use of A1967–70 and a close watch can be kept on mortality statistics in subsequent years using the above techniques.

REFERENCES

Armitage, P. (1947) 'Some sequential tests of Student's hypothesis.' *J.R. Statist. Soc.* Series B, **9**, 250–63.

Benjamin, S. (1960) 'Théorie statistique des sondages.' *XVIth International Actuarial Congress*, Brussels, thème B2, 244–7.

Johnson, N.L. and Moore, P.G. (1957) 'Applications of sequential analysis to mortality data.' *J.S.S.* **14**, 84 and **15**, 90.

Joint Mortality Investigation Committee (1974) 'Considerations affecting the preparation of standard tables of mortality.' *J.I.A.* **101**, 133.