WORKSHOP

WEIGHTED MORTALITY RATES AS EARLY WARNING SIGNALS FOR INSURANCE COMPANIES

By Leigh A. Roberts

*Institute of Statistics and Operations Research,
Victoria University, Wellington, New Zealand*

Abstract

Weighted mortality rates are commonly used in actuarial work, but the inter-relationship between the weights used and the underlying mortality rates seems not to have been widely investigated.

Calculation of the ratio of weighted mortality rates to conventional mortality rates provides a simple means for an insurance company to track changes in the underlying mortality of its portfolio over time, and acts as an early warning system for possible deterioration of underwriting results. Asymptotic distributions are found for this ratio, and for the mortality rates themselves. It is suggested that insurance companies commence to gather data for the calculation of this ratio for the insurance sector as a whole, for the main annuity and assurance classes.

Keywords

Asymptotic distribution; decremental rates; early warning signals; insurance; ratio of weighted to unweighted mortality rates; sums at risk; underlying mortality trends; weighted mortality rates.

1. Introduction

Conventional mortality rates are calculated as a ratio of the number of deaths to the number of life years lived which gave rise to those deaths, and measure the probability of dying within the next year at a given age. A variant on this theme in common use in actuarial work is the use of *weighted* mortality rates, attaching weights to each death and unit of exposure to risk: in life insurance these weights are typically sums assured or numbers of policies. For monetary weights, for instance, the numerator is expressed as dollars or pounds which have 'died', the denominator in terms of those sums which were exposed to this risk, measured in dollar or pound years.

Weighting the rate estimates in this way is a natural thing to do, in that what matters ultimately to an insurance company is the monetary amounts requiring to be paid out. Multiplying total sums at risk in an age interval, for example,

by the central weighted mortality rate yields an estimate of total payments to that group over the next year, provided that the weights remain unaltered. This point notwithstanding, the use of weighted rates is clearly somewhat dangerous if the underlying mortality rates for a given age vary within the portfolio, and the inter-relationship between the weights and the mortality rates is not understood, or changes unexpectedly over time.

The statistical heterogeneity caused by confounding mortality effects and financial factors may pose insurance companies a serious problem, because, while mortality changes relatively slowly over time, the inter-relationship between sums at risk and mortality may change suddenly and unpredictably. The 1956 Finance Act in the UK, to take an extreme example, radically changed the pattern of annuity purchases in the UK overnight: the annuitants' experience gathered by the insurance companies had thereafter to be separated into pre-1957 and post-1956 subgroups (CMIR1 (1973, p. 29)). Less extreme examples occur every year with changes to the financial environment, such as adjustments to tax rates, altering the balance of the different socio-economic groups purchasing various financial instruments, and changing the relative sizes of transactions as well. The danger is all the more acute in that such changes will tend to occur uniformly over ages, exacerbating the financial consequences.

When numbers of policies are used as weights, on the other hand, the link between the weights and the underlying mortality may not be subject to such capricious political changes; nevertheless, the relationship is not at all well understood, and the effect of duplicate policies, especially at advanced ages, can be very marked (e.g., CMIR2 (1976, p. 69); see also CMIR6 (1983, p. 45)).

Given the relatively crude understanding of the inter-relationship between weights used and the underlying mortality rates, the apparent lack of effort expended to investigate this connection is surprising, to say the least. It is not that the literature has nothing to offer on weighted averages of mortality rates. On the contrary, directly standardised demographic rates, of fertility or mortality etc., assume precisely the form of a weighted average; and ratios of weighted average mortality rates, such as the Comparative Mortality Factor (CMF) and the Standardised Mortality Ratio (SMR), are of central importance in demography (Cox (1976, p. 298)). These two ratios, moreover, possess direct analogues in economics in the Laspeyres and Paasche price indices, which in turn have an extensive literature (KITAGAWA (1964, p. 302)).

Nor is there a paucity of study on the dependence of mortality rates on many other factors, such as impairments. Both the North American Society of Actuaries and the UK actuarial bodies, amongst many other organisations, publish regular reports on impaired lives' mortality, and there is a voluminous literature. HABERMAN (1988), for example, deals with weighted averages of q/q', where q is the mortality rate of the impaired lives' group, and q' that of the population at large; and both the CMF and the SMR can also be recast as weighted averages of this ratio (Cox (1976. p. 298)).

The point is, however, that the emphasis in most mortality studies is on the comparison between two sets of mortality rates, be they for two regions, or for

impaired and normal lives, etc.; the initial difficulty arising is that of finding appropriate weights to use in taking the averages of the rates. Here we are interested instead in the effects of using different possible weights applied to the one set of underlying mortality rates. The ratio R defined below is the analogue of the Area Comparability Factor, a factor designed to correct for the varying population profiles in different regions when calculating inter-censal mortality behaviour (COX (1976, p. 305)).

Nevertheless, the literature is not entirely silent on the dependence of mortality on the size of the insurance policy. In accord with the fact that the Society of Actuaries uses sums insured as weights for its mortality investigations, while the UK actuarial bodies generally use numbers of policies, most of this work originates from North America.

An early writer treating the behaviour of weighted rates such as those considered in this paper is CODY (1941), who finds expressions for the mean and variance of initial weighted mortality rates, and derives the ratio of the variances of the weighted and unweighted rates. More recently, KLUGMAN (1981) has contrasted the mean square errors of weighted and unweighted initial mortality rates, setting off the larger variance of the weighted rate against the bias implicit in using the unweighted rate. These theoretical developments notwithstanding, there is very little insurance data to which to apply the theory. Virtually the only publicly available data is gathered by the Society of Actuaries in its quinquennial survey of Mortality on Policies for Large Amounts (SOCACTS (1987)); and these reports are not sufficient detailed to enable one convincingly to model in detail this aspect of mortality.

As KLUGMAN (1981) has pointed out, however, it is clear from these reports that mortality is generally lighter for policies with higher sums insured. This is consistent with the PA (90) and PL (90) mortality tables, both calculated from the same life insurance data for UK pensioners relating to 1967-1970. The acronyms stand for "pensioners' amounts" and "pensioners' lives" respectively; the weights used are annual annuity payments on the one hand, and unity on the other. The PA (90) rates are substantially lower than the PL (90) rates, implying that wealthier pensioners experience better mortality than their less well-off neighbours, and the former table is used in preference to the latter simply for reasons of financial conservatism (CMIR3 (1978, p. 20)).

Mathematical methods cannot be expected to forecast mortality changes of the magnitude experienced in the UK in 1956, but one can at least expect them to track underwriting results which are changing less dramatically over time. In this paper we suggest the use of the ratio R of the weighted mortality rate to the unweighted rate for the purpose of alerting insurance companies to general alterations in portfolio characteristics, and in particular to potential deterioration in underwriting results arising from changes in the interaction between the sums at risk and the underlying mortality rates. The statistic R is very simple to estimate, and can be calculated every year for each age, or group of ages. Moreover, it is readily expressed in terms of the correlation between the weights and the mortality rates, as is seen in Section 6, and its use accords well with one's intuition.

After preliminary definitions are given in Section 2, in Section 3 we find the asymptotic distribution of R as the sample size becomes infinite, so that its estimated value can be inserted in an approximate confidence interval. A by-product of the calculations is that we also obtain the asymptotic distributions of the weighted and unweighted mortality rates, for both central and initial rates.

Simple simulations are carried out in Section 4 to obtain a rough idea of how large a sample is required before the asymptotic distributions of the ratio R and the mortality rates become acceptable approximations; a further purpose is to illustrate the use of weighted mortality rates in tracking the effect of mortality changes in a portfolio over time. Some problems of using weighted mortality rates in this manner are then addressed in Section 5.

Finally a plea is entered in Section 6 to make a start to the collection of data to keep track of the problem of the inter-relationship between the weights and the underlying mortality, for the principal annuity and assurance classes. The data is no doubt investigated more thoroughly within individual insurance companies; but as stressed above, it seems that little attention has been devoted to this matter on a sectoral level.

2. PRELIMINARY DEFINITIONS

The (sample estimate of the) central weighted mortality rate is defined as

$$(1) \qquad \widehat{M^A} = \frac{\displaystyle\sum_{1}^{N} S_j \theta_j}{\displaystyle\sum_{1}^{N} S_j (1 - \theta_j u_j)}$$

where N is the sample size; S_j is the weight, labelled as the *sum at risk*, assigned to the jth individual in the sample; the superscript A denotes *amount*, reflecting the fact that the weights will frequently be monetary amounts; and θ_j represents the fate of the jth individual, assuming the value 1 if the person dies during the year of observation, and 0 otherwise (θ, for $\theta\alpha\nu\alpha\tau o\varsigma$, is a conventional symbol for death in actuarial work). Individuals who survive are given a full year's exposure in the denominator, while deaths contribute to the exposure only while they live: the variable u_j represents the shortfall in exposure by those dying, so labelled because in practice it would frequently be taken to have a uniform distribution. All those in the sample are assumed to come under observation at the same time: there is no attempt to model entrance into the sample at varying ages or times, as is done in ROBERTS (1992b).

The initial weighted mortality rate is similar, save that all individuals are given a full year's exposure to risk regardless of whether they live or die; i.e.,

we set the variable $u = 0$ in (1):

$$(2) \qquad \widehat{Q^A} = \frac{\sum_1^N S_j \theta_j}{\sum_1^N S_j}$$

The unweighted rates are labelled as *life* rates, and are obtained by setting $S = 1$ in (1) and (2):

$$\widehat{M^L} = \frac{\sum_1^N \theta_j}{\sum_1^N (1 - \theta_j u_j)} \; ; \qquad \widehat{Q^L} = \frac{\sum_1^N \theta_j}{N}$$

The population parameters corresponding to these latter sample estimates are customarily labelled m and q, whence the notation used above.

For the ratios of weighted to unweighted rates we may use either the central or initial rates:

$$\widehat{R^M} = \frac{\widehat{M^A}}{\widehat{M^L}} \; ; \qquad \widehat{R^Q} = \frac{\widehat{Q^A}}{\widehat{Q^L}}$$

3. ASYMPTOTIC DISTRIBUTIONS

All individuals are assumed to be mutually independent: i.e., (S_j, θ_j, u_j) are mutually independent random vectors, for $j = 1, \ldots, N$; and we further assume that u_j is independent of both S_j and θ_j, for each j. The variates S_j and θ_j are *not* assumed independent: their possible dependence is the problem addressed in this paper.

We shall obtain the asymptotic distribution of $\widehat{R^M}$ as the sample size $N \to \infty$, from which the asymptotic distributions of $\widehat{M^A}$, $\widehat{M^L}$, $\widehat{R^Q}$, $\widehat{Q^A}$ and $\widehat{Q^L}$ may all be immediately inferred. Nevertheless, the general derivation is not so easy to follow, and we illustrate the procedure briefly by first finding the asymptotic distribution of $\widehat{M^A}$.

3.1. Asymptotic distribution of $\widehat{M^A}$

Dividing the numerator and denominator of $\widehat{M^A}$ by the (deterministic) sample size N, and letting $N \to \infty$, both quantities tend to probability limits, say ξ_1 and ξ_2 respectively:

$$T_1 = \sum_1^N S_j \theta_j / N \to \xi_1 = E(S\theta), \quad \text{and}$$

$$T_2 = \sum_1^N S_j (1 - \theta_j u_j) / N \to \xi_2 = E(S) - E(u) E(S\theta);$$

whence $\widehat{M^A} = \dfrac{T_1}{T_2} \to M^A = \dfrac{\xi_1}{\xi_2}$ in probability, as $N \to \infty$.

Each of the numerator and denominator, suitably centred and normalised, possesses a limiting normal distribution by the Central Limit Theorem; and so will the mortality rate $\widehat{M^A}$, an intuitive argument for which is as follows.

$$\widehat{M^A} = \frac{\xi_1(1+\varepsilon_1)}{\xi_2(1+\varepsilon_2)} = \frac{\xi_1}{\xi_2}\,[1+\varepsilon_1-\varepsilon_2+O\,(\varepsilon_1^2+\varepsilon_2^2)],$$

where $\varepsilon_j \to 0$ in probability, as $N \to \infty$, for $j = 1, 2$; and the expansion is valid as long as $|\varepsilon_2| < 1$. Neglecting the remainder term, and denoting the covariance matrix of $(\xi_1\varepsilon_1, \xi_2\varepsilon_2)'$ by Σ_0/N, the asymptotic distribution of $\sqrt{N}\,(\widehat{M^A} - M^A)$ is $\mathcal{N}\,(0, \sigma_0^2)$, with

(3)
$$\frac{\sigma_0^2}{(M^A)^2} = \left(\frac{1}{\xi_1}\ -\ \frac{1}{\xi_2}\right) \Sigma_0 \begin{pmatrix} 1/\xi_1 \\ -1/\xi_2 \end{pmatrix}$$

$$= \frac{\sigma_{11}}{\xi_1^2} - \frac{2\sigma_{12}}{\xi_1\xi_2} + \frac{\sigma_{22}}{\xi_2^2}$$

where $\sigma_{11} = E(S^2\theta) - \xi_1^2;\ \ \sigma_{12} = E(S^2\theta)\,[1 - E(u)] - \xi_1\xi_2;$

and $\sigma_{22} = E(S^2) - E(S^2\theta)\,[2\,E(u) - E(u^2)] - \xi_2^2.$

σ_{11}/N is the variance of T_1, σ_{22}/N the variance of T_2, and σ_{12}/N their covariance. The formal justification for our conclusion follows from a standard functional central limit theorem (RAO (1973), p. 387).

Instead of further developing this formula, we shall obtain the analogous result for the statistic $\widehat{R^M}$, from which the asymptotic distributions of the other statistics, including $\widehat{M^A}$ itself, are easily derived.

3.2. Asymptotic distribution of $\widehat{R^M}$

As above, we define the following random variables and probability limits.

$$T_3 = \sum_1^N \theta_j/N \to \xi_3 = E(\theta), \quad T_4 = \sum_1^N (1-\theta_j u_j)/N \to \xi_4 = 1 - E(u)\,E(\theta),$$

so that $\widehat{R^M} = \dfrac{T_1/T_2}{T_3/T_4} \to R^M = \dfrac{\xi_1\xi_4}{\xi_2\xi_3}$ in probability, as $N \to \infty$.

In analogy with the previous case, the limiting distribution of $\sqrt{N}\,(\widehat{R^M} - R^M)$ is $\mathcal{N}\,(0, (R^M)^2\,\zeta'\Sigma\zeta)$, where $\zeta' = (\xi_1^{-1}, -\xi_2^{-1}, -\xi_3^{-1}, \xi_4^{-1})$, and Σ contains

elements $\sigma_{ij} = N \times \text{cov}(T_i, T_j)$. The leading 2×2 submatrix of Σ is the covariance matrix Σ_0 previously employed in (3).

Defining

$$\mu = \frac{E(S\theta)}{E(S)\,E(\theta)}\;; \qquad v = \frac{E(S^2\theta)}{E(S)\,E(S\theta)}\;;$$

$$\alpha = E(u)\,E(\theta); \qquad \text{and} \qquad \beta = E(u^2) - 2\,E(u),$$

and making the innocuous assumption that the mortality rate q_j for the jth individual is bounded away from unity, so that $q_j < q < 1$ for all j and some q, the asymptotic mean and variance of R^M are given by

$$\text{asmean}\,(\widehat{R^M}) = R^M = \mu\left[1 + \frac{\alpha\,(\mu - 1)}{1 - \alpha\mu}\right] = \mu + O(q)$$

$$N \times \frac{\text{asvar}\,(\widehat{R^M})}{(R^M)^2} = \frac{1}{E(\theta)}\left[\frac{v}{\mu} - 1\right] - 2\,\frac{1 - E(u)}{1 - \alpha\mu}\,(v - \mu)$$

$$+ \frac{E(S^2) + \beta E(S^2\theta)}{[E(S)]^2\,[1 - \alpha\mu]^2} + \frac{1 + \beta E(\theta)}{[1 - \alpha]^2} - 2\,\frac{1 + \mu\beta E(\theta)}{[1 - \alpha\mu]\,[1 - \alpha]}$$

$$= \frac{1}{E(\theta)}\left[\frac{v}{\mu} - 1\right] + O(1) = \frac{E(S^2\theta)}{[E(S\theta)]^2} - \frac{1}{E(\theta)} + O(1)$$

$$= \frac{1}{E(\theta)}\left[\frac{v}{\mu} - 1\right] - 2[1 - E(u)]\,(v - \mu) + \frac{V(S)}{[E(S)]^2} + O(q)$$

in an obvious, if rather cumbersome, notation. Details are given in ROBERTS (1992a).

That $v/\mu \geq 1$ is easily shown from Schwarz' inequality; the inequality is strict unless S reduces to a constant (disallowing the pathological case in which $E(\theta) = 1$). The factor $(v - \mu)$ appearing in the expression for the variance is therefore positive unless all the sums at risk are equal.

3.3. Asymptotic distributions for the remaining statistics

From the asymptotic distribution for $\widehat{R^M}$, the limiting distribution of $\widehat{M^A}$ is found by setting $T_3 = T_4 = 1$, causing the corresponding terms in the covariance matrix to vanish, while that of $\widehat{M^L}$ is found by setting $T_1 = T_2 = 1$. The limiting distributions of $\widehat{R^Q}$, $\widehat{Q^A}$ and $\widehat{Q^L}$ are then deduced by setting the variable $u = 0$.

For completeness we gather the remaing results together below.

$$\text{asymptotic mean} \qquad N \times \frac{\text{asymptotic variance}}{(\text{asymptotic mean})^2}$$

$\widehat{M^A}$
$$\frac{\mu E(\theta)}{1-\alpha\mu}$$
$$\frac{E(S^2\theta)}{[E(S\theta)]^2} - 2v\frac{1-E(u)}{1-\alpha\mu} + \frac{E(S^2)+\beta E(S^2\theta)}{[E(S)]^2\,[1-\alpha\mu]^2}$$
$$= \frac{E(S^2\theta)}{[E(S\theta)]^2} + O(1)$$

$\widehat{M^L}$
$$\frac{E(\theta)}{1-\alpha}$$
$$\frac{1}{E(\theta)} - 2\frac{1-E(u)}{1-\alpha} + \frac{1+\beta E(\theta)}{[1-\alpha]^2} = \frac{1}{E(\theta)} + O(1)$$

$\widehat{R^Q}$
$$\mu$$
$$(v-\mu)\left[\frac{E(S)}{E(S\theta)} - 2\right] + \frac{V(S)}{[E(S)]^2}$$
$$= \frac{E(S^2\theta)}{[E(S\theta)]^2} - \frac{1}{E(\theta)} + O(1)$$

$\widehat{Q^A}$
$$\mu E(\theta)$$
$$\frac{E(S^2\theta)}{[E(S\theta)]^2} - 2v + \frac{E(S^2)}{[E(S)]^2} = \frac{E(S^2\theta)}{[E(S\theta)]^2} + O(1)$$

$\widehat{Q^L}$
$$E(\theta)$$
$$\frac{1}{E(\theta)} - 1$$

When deaths are uniformly distributed over the time period considered, i.e. when u has the uniform distribution between 0 and 1, the expressions given involving the central rates simplify substantially. In that case, $\alpha = E(\theta)/2$, $\beta = -2/3$ and $E(u) = .5$.

Regardless of the sums at risk, when all individuals are subject to the same survival curve, say with central rate m_0 and initial rate q_0, then $E(\widehat{M^L}) = M^L = m_0$, $E(\widehat{Q^L}) = Q^L = q_0$. The relationship between the two rates is given by $m_0 = q_0/[1 - E(u)q_0]$: see, i.a., ELANDT-JOHNSON and JOHNSON (1980).

The asymptotic bias arising from the use of the weighted rates to estimate the unweighted rates, or vice versa, is seen to vanish when $\mu = 1$; i.e., when S and θ are uncorrelated. In this case, it is to be expected that the variance is larger for the weighted rate, and this is verified for the initial rates below when S and θ are independent.

We note that the expressions given for the means and variances of the central and initial rates agree as far as the highest order terms, as they also do for $\widehat{R^M}$ and $\widehat{R^Q}$.

The asymptotic distribution of $\widehat{Q^A}$ is found in KLUGMAN (1981), the variance agreeing with his formula for Var $(\hat{q}_G)$. The numerical dominance of the first

term $E(S^2\theta)[E(S\theta)]^{-2}$ in the variance is evident from the succeeding calculations in Klugman's paper. The ratio of the variances of $\widehat{Q^A}$ and $\widehat{Q^L}$ is $E(S^2\theta)E(\theta)[E(S\theta)]^{-2} = v/\mu$, neglecting lower order terms, which reduces to $E(S^2) \times [E(S)]^{-2}$ when S and θ are independent, in agreement with formulae in CODY (1941, p. 71) and KLUGMAN (1981, Section II). The asymptotic variance of the weighted rate estimate exceeds that of the unweighted estimate unless all sums at risk are equal, because $v > \mu$, as noted above.

The variance of $\widehat{Q^L}$ can be written as the "binomial" variance $Q(1-Q)/N$, with $Q = E(\theta)$, conforming with POLLARD (1970, eqn. (8)) (in the notation of that paper, n_i and q_i are both deterministic). Pollard actually deals with the number of deaths, but this is precisely a multiple of the unweighted initial mortality rate.

Finally suppose that the jth group has sum at risk S_j; that the individuals in that group have a common mortality rate of q_j; and that the probability of an individual selected at random from the portfolio belonging to the jth group is π_j, where $\Sigma \pi_j = 1$. Then

$$Q^A = \frac{E(S\theta)}{E(S)} = \frac{\Sigma S_j \pi_j q_j}{\Sigma S_j \pi_j}$$

Thus Q^A is the weighted mean of the group mortality rates, with weights the sum at risk times the probability of the policyholder belonging to that category. There are comparable results for the central rates. Armed with the further definitions $d_j = [1 - E(u)q_j]/[1 - E(u)q_0]$, where the 0th group has been chosen arbitrarily as a reference group, we have that

$$M^A = \frac{E(S\theta)}{E(S) - E(u)E(S\theta)} = \frac{\Sigma S_j \pi_j d_j m_j}{\Sigma S_j \pi_j d_j}$$

The asymptotic mean is again a weighted sum of the central rates for each group; the weights, while slightly more complicated, differ in value but little from the weights used for the initial rates, the parameters d_j being close to unity.

4. SIMULATION RESULTS

The calculation of mortality rates was simulated in order firstly to obtain a rough idea of how large a sample needs to be before the asymptotically valid normal distribution becomes a reasonable approximation to the finite sample distribution. A second purpose was to investigate how effective the calculation of the ratio of amounts' and lives' mortality rates is at tracking the behaviour of mortality within the portfolio over time.

The sample investigated was considered to consist of two groups, with equal numbers of individuals in each group. These numbers were assumed stationary through time, so that deaths are replaced by new lives in the investigation. Mortality rates were initially identical for the two groups, as were the sums at

risk; over the 10 year sampling period, however, while the characteristics of the. first group remained unchanged, the sum at risk of the second group increased by 5% p.a.; and the initial unweighted rates for the second group were successively assumed to increase, then to decrease, by 5% p.a. For each scenario, 500 simulations were carried out.

Situations sampled ranged over mortality rates of .005, .01 and .05 at the beginning of the simulated period, and for sample sizes between 100 and 100,000 life years per annum.

For each run, the intervals bounded by the 2.5 and the 97.5 percentiles (95% highest density regions, or HDRs) were found for both the empirical sample distribution and the approximating normal distribution; when these intervals were close to coincident, the asymptotic distribution was taken to be a good approximation to the finite sample distribution. Using this criterion, we conclude that the limiting normal distribution is an adequate approximation for all of the mortality rates as soon as the expected numbers of deaths reaches about 5; for the ratio R, whether using initial or central rates, this stage is not reached until expected deaths reaches about 10. The larger sample size required for R is probably due to the fact that a default value must be assigned to R when there are no deaths (for which purpose unity was chosen here): there is little possibility of approaching the limiting distribution unless there is a negligible chance of no deaths arising.

Once the expected number of deaths reaches about 20, the limiting normal distribution would seem to be a good approximation for all statistics, in the sense that the two HDRs almost overlapped. Some typical results are shown below, for an initial mortality rate of 5%, and with mortality increasing for half the population at 5% p.a. We note that exact moments of the mortality rates for small sample sizes can be calculated from formulae given in ROBERTS (1992a); but this is impractical for other than very small samples.

As regards the second purpose of the simulation, viz. that related to keeping track of changes in the portfolio over time, one wants the ratio R to differ significantly from unity for moderate sample sizes, when the weighted and unweighted rates are behaving differently from one another. The criterion chosen for deciding when R differs from unity is that the asymptotic mean of R should differ from unity by more than 4 (asymptotic) standard deviations. When this criterion is satisfied, there is at least a 97.5% probability that the 95% confidence interval surrounding the point estimate of R will not contain the value unity; thus the conventional two sided statistical test of whether $R = 1$ will be rejected at the 5% level, with probability at least .975.

To see this, recall that $E(\widehat{R^Q}) = \mu$; let $\mathrm{Var}(\widehat{R^Q}) = \sigma^2$, and suppose that $\mu - 1 > 4\sigma$. Using the normal approximation, $\mathrm{Prob}(\widehat{R^Q} > \mu - 2\sigma) \approx .975$. On the assumption that the standard error estimated from the sample is close to the true standard deviation, the width of the 95% confidence interval encasing $\widehat{R^Q}$ will be close to 4σ, and $\mathrm{Prob}(\widehat{R^Q} - 2\sigma > 1) > .975$.

The length of time, and/or the sample size, required to ascertain that R is diverging from unity is surprisingly large. The heavy unbroken lines in Figure 1 map the asymptotic 95% HDR, viz. $\mu \pm 2\sigma$, over time. For a sample size of
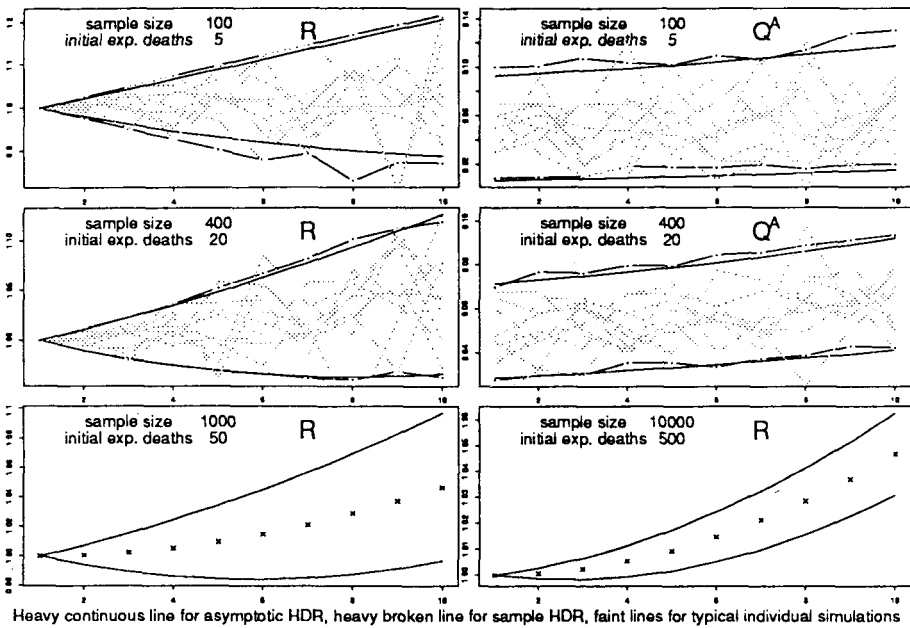
Heavy continuous line for asymptotic HDR, heavy broken line for sample HDR, faint lines for typical individual simulations

FIGURE 1. Simulated values of $\widehat{R^Q}$ and $\widehat{Q^A}$.

10,000 life years p.a., 7 years is needed for the mean to exceed unity by four standard deviations; for a sample size of 1,000 p.a., the criterion is nowhere near being satisfied even after the 10 years shown, and in fact 18 years are required. Even for a sample size of 100,000 p.a., the criterion is not met until 3 years have elapsed.

For just two groups of identical, independent individuals, N times the variance of $\hat{R}$ reduces to

$$\frac{\mu^2}{E(\theta)}\left[\frac{v}{\mu} - 1\right] = \frac{bc(a-1)^2}{\pi_0^3 q_0 (1+ac)^2 (1+bc)^3} \,,$$

where $a = S_1/S_0$; $b = q_1/q_0$; $c = \pi_1/\pi_0$; and only the highest order terms are retained (so that it is immaterial whether we use $\widehat{R^M}$ or $\widehat{R^Q}$). The criterion which we are using for assessing when R departs from 1 becomes

$$|\mu - 1| = \left|\frac{1+abc}{\pi_0(1+ac)(1+bc)} - 1\right| > \frac{4}{\sqrt{N q_0}}\frac{|a-1|\sqrt{bc}}{\pi_0^{3/2}(1+ac)(1+bc)^{3/2}}$$

Now set $c = 1$, $\pi_0 = 1/2$. The criterion reduces to

$$\left|\frac{(1-a)(1-b)}{(1+a)(1+b)}\right| > \frac{4 \times 2^{3/2}}{\sqrt{N q_0}}\frac{|a-1|\sqrt{b}}{(1+a)(1+b)^{3/2}}\,; \quad \text{or}$$

$$N q_0 > \frac{128\, b}{(1+b)(1-b)^2}$$

Note that when $b = 1$, the requisite sample size is infinite: we cannot distinguish between two groups with equal mortality rates. When $b = 0$, the sample size required becomes zero because $\widehat{R^Q}$ reduces to a constant, distinct from unity unless $a = 1$.

5. PRACTICAL CONSIDERATIONS

Despite the simplicity of the results in this paper, several problems may arise in their application to tracking the inter-relationship between mortality and weights over time.

In view of our simulation results, one could first admit that for smaller companies, or for the less popular classes of insurance, the expected number of deaths in one portfolio over a year may be well short of the 10 or 20 that seem to be needed for the limiting normal distribution to be valid. The second point arising concerns the large sample size necessary effectively to track the behaviour of mortality over time, even for only broad subgroupings of the exposure. On both counts, amalgamation of experience, either over time or over companies, may be necessary.

Even assuming that companies are willing to undertake the calculation of weighted rates over time, however, and that there is sufficient data for this to be worthwhile, there remain technical problems to be overcome. Of primary importance is the bias introduced by selective withdrawals in the measurement of mortality experience. This source of possible bias assumes some importance given the very high level of early withdrawals experienced by some classes of life insurance: in Australia, for example, 15% of ordinary life insurance policies lapse within the first year, and some 50% have surrendered within about 6 years (ISC (1992, p. 96)). The use of central mortality rates will decrease the potential bias, but the formulae derived above are far simpler for initial rates. Even if such problems are satisfactorily resolved within the individual company, moreover, there is still the problem at the sectoral level of obtaining data from the various companies on a comparable basis.

It is theoretically straightforward to derive the asymptotic distribution of rates similar to those used above, allowing for other decrements besides death (such rates are set out in ROBERTS (1992b), although the asymptotic distributions are not given there); but it is to be expected that there will be some correlation between sums at risk and lapse rates, and the accurate estimation of lapse and mortality rates for many different bands of sums at risk may present difficulties.

There is also the question of what weight to choose as the "sum at risk". For the purposes of the insurance sector as a whole, this may best be chosen as the sum assured plus accrued bonus, or annual annuity payment in force. Within an individual company, however, the difference between the sum assured and the reserve may be a more sensible choice for insurance policies for sufficiently large portfolios. BATTEN (1978, ch. 7) considers several possibilities for the weighting variable.

Finally, as KLUGMAN (1981) points out, the reason that weights used in practice are generally sums insured or numbers of policies is that of practicality: the information is easily available. The possibility arises then of using such a ratio as R, with both the numerator and denominator containing *weighted* mortality rates. The weights used in the numerator could remain sums insured, or other appropriate monetary weights; those in the denominator could be numbers of policies. The analysis in this paper would still be valid: the variable θ would now refer to the death of a policy, and the various terms like $E(S\theta)$ would need to be interpreted as expected values over a population of policies.

6. CONCLUSION

Referring to Section 3.3, recall that the moments for the M and Q rates and their ratios agree as far as terms of the highest order, and the following remarks will hold true whether one works with central or initial rates. The weights are assumed to be monetary weights, although comparable statements could be made for general weights.

Consider the ratio $\widehat{R^Q}$, the asymptotic mean of which is:

$$R^Q = \mu = \frac{E(S\theta)}{E(S)\,E(\theta)} = 1 + \frac{\text{cov}\,(S,\,\theta)}{E(S)\,E(\theta)} = 1 + \text{constant} \times \text{corr}\,(S,\,\theta).$$

We first note that the quantity μ is dimensionless, so that inflation affecting all members of the portfolio equally should not alter its value, as long as the underlying mortality rates in the portfolio remain unaltered (or the changes in mortality act uniformly on the whole portfolio). The second point to note is the ease with which μ is estimated, or rather $\widehat{R^Q}$ calculated: e.g., $\widehat{E(S\theta)} = \Sigma\, S_j/N$, where the summation is taken over deaths.

A third aspect is the appeal to our intuition engendered by the quantity μ, which is simply related to the correlation between S and θ as shown: the constant is the product of the coefficients of variation (standard deviation divided by the mean) of the two variates. A complete lack of correlation, i.e. no differential mortality for wealthier individuals in the portfolio, means that μ will remain as unity regardless of what happens separately to sums at risk and mortality.

The insurance company will wish to estimate both the ratio μ and its components $Q^A = E(S\theta)/E(S)$ and $Q^L = E(\theta)$, for various classes of business and for different age groups, and to track these quantities over time. The expected sum to be paid out is the exposure to risk times the former quantity, while the second factor follows changes in the underlying mortality over time. Estimates should be encased in a confidence interval, standard errors being estimated from the above expressions for the asymptotic variance. Insurance companies should also commence to calculate these quantities on a sectoral level, to oversee changes in the insurance market as a whole.

REFERENCES

BATTEN, R. W. (1978) *Mortality Table Construction.* Prentice-Hall.
CMIR1 (1973) *Continuous Mortality Investigation Report No. 1.* Continuous Mortality Investigation Bureau. Institute of Actuaries, Faculty of Actuaries.
CMIR2 (1976) *Continuous Mortality Investigation Report No. 2.* Continuous Mortality Investigation Bureau. Institute of Actuaries, Faculty of Actuaries.
CMIR3 (1978) *Continuous Mortality Investigation Report No. 3.* Continuous Mortality Investigation Bureau. Institute of Actuaries, Faculty of Actuaries.
CMIR6 (1983) *Continuous Mortality Investigation Report No. 6.* Continuous Mortality Investigation Bureau. Institute of Actuaries, Faculty of Actuaries.
CODY, D. C. (1941) Actuarial note: the standard deviation in the rate of mortality by amounts. *Transactions, Actuarial Society of America* **42**, (1) 69–73.
COX, P. R. (1976) *Demography.* 5th edition. Cambridge University Press.
ELANDT-JOHNSON, R. C. and JOHNSON, N. L. (1980) *Survival Models and Data Analysis.* Wiley.
HABERMAN, S. (1988) Measuring relative mortality experience. *Jnl. of the Institute of Actuaries* **115**, (II) 271–298.
ISC (1992) *Insurance and Superannuation Commission Annual Report 1991-1992.* Australian Government Publication.
KITAGAWA, E. M. (1964) Standardised comparisons in population research. *Demography* **1**, 296–315.
KLUGMAN, S. A. (1981) On the variance and mean squared error of decrement estimators. *Trans., Soc. of Actuaries* **33**, 301–311.
POLLARD, A. H. (1970) Random mortality fluctuations and the binomial hypothesis. *Jnl. of the Institute of Actuaries* **96**, (II) 251–264.
RAO, C. R. (1973) *Linear statistical inference and its applications.* Second edition. Wiley.
ROBERTS, L. A. (1992a) Exact Moments and Asymptotic Distributions of Weighted Mortality Rates. Technical Report No. 23, Institute of Statistics and Operations Research, Victoria University, Wellington.
ROBERTS, L. A. (1992b) On ratios of random variables and generalised mortality rates. *Jnl. Applied Probability* **29**, (2) 268–279.
SOCACTS (1987) 1985, 1986 and 1987 reports of mortality, morbidity and other experience. *Trans., Soc. of Actuaries*, 35–69.

LEIGH ROBERTS
Institute of Statistics and Operations Research,
Victoria University of Wellington,
P.O. Box 600, Wellington, New Zealand.