Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: - (1981)

Heft: 2

Artikel: Actuarial estimation of decremental probabilities

Autor: Seal, Hilary L.

DOI: https://doi.org/10.5169/seals-551143

HILARY L. SEAL, Apples

Actuarial Estimation of Decremental Probabilities

The actuary's estimation problems in his studies of decremental probabilities based on "lives followed through" (e.g. in a pension fund) differ from those of demographers and even medical statisticians in that complete dates of birth, disability, death, severance, etc.[1] are always available. His standard use of a "service table" based on integer ages and his calculation of "crude" probabilities of death, disability and, possibly, severance at those ages is no more than a convenience. He is well aware that what he is doing is to recognize the general relation

$$p(x, x+t) = \prod_{j=0}^{k-1} p(x+\tau_j, x+\tau_{j+1}) \tag{1}$$

where

$p(a, b) = Pr$ {an individual in active service at exact age a remains in active service until at least exact age b}

and

$$0 = \tau_0 < \tau_1 < \tau_2 < \ldots < \tau_k = t$$

but the τ's are otherwise arbitrary. In fact the actuary has changed the series $\{\tau_j\}$ into the natural numbers.

At the turn of the century German speaking actuaries, under the influence of Lexis (1877), were much concerned to check whether the relative frequencies calculated by dividing the number of deaths at a given age by the exposed to risk at that age were behaving like realizations of an underlying probability of death. Many of the results were reproduced in standard texts such as Blaschke (1906), Czuber (1910) and Landré (1911). Nowadays the equivalent techniques are known as tests of homogeneity but with few exceptions (e.g. Mises, 1964, Ch. IX) the name of Lexis no longer appears in the Anglo-Saxon textbooks.

Some thirty years ago the writer (Seal, 1949) gave a brief history of the application of Lexis's divergence coefficient to mortality data and described how it needed to be slightly modified to be used to test homogeneity in a modern setting. He concluded by applying the new techniques to six mortality experi-

[1] In what follows the decrements death and disability are frequently used instead of a longer list.

ences of the seventeenth and eighteenth centuries in which the same single age mortality rates could be evaluated under different circumstances (e.g. corresponding to specific ages at annuity purchase). The analyses of these data all showed that properly calculated relative frequencies of death were behaving as binomially distributed values prior to 1800 – and presumably since then. The actuarial use of x as an integer and $t = 1$ in (1) stemmed from long-standing recognition of calendar years as marking the progression of an individual through his life. It is intuitively obvious that if relative frequencies of mortality in calendar year spans $(x, x+1)$ are binomially distributed then so are the relative frequencies in arbitrary time spans $(x+\tau_j, x+\tau_{j+1})$. The result is that the set of relative frequencies corresponding to the time spans on the right hand side of (1) could be used as a set of consecutive binomial variates which would provide an estimate of $p(x, x+t)$ for specified, not necessarily integral, x and t. And if disabilities as well as deaths were being counted in each interval the binomial would become a trinomial, viz.

$$[\{1-{}^*p^d(x, x+t)\}+\{1-{}^*p^i(x, x+t)\}+(\{{}^*p^d(x, x+t)+{}^*p^i(x, x+t)\}-1)]^N$$

where

$$1-{}^*p^d(x, x+t)=\int_0^t \exp\left[-\int_0^s (\mu_{x+u}+\nu_{x+u})du\right]\mu_{x+s}\, ds \tag{2}$$

and

$$1-{}^*p^i(x, x+t)=\int_0^t \exp\left[-\int_0^s (\mu_{x+u}+\nu_{x+u})du\right]\nu_{x+s}\, ds \tag{3}$$

the forces μ and ν being those of mortality and disability, respectively, and the asterisk showing that the probabilities are "crude" or "dependent" (Seal, 1977). The great advantage of reducing the estimate of a survivorship or activity probability to the estimate of a binomial parameter p (or a parameter ${}^*p$ of the marginal distribution of deaths or disabilities alone) is that the relative frequency "number of successes" ÷ "number of trials" has several properties that are recognized as desirable in any statistical estimate. In N binomial trials the relative frequency of success is (a) unbiased, (b) of minimum variance, (c) minimal sufficient, (d) complete, and (e) the maximum likelihood estimate (Cox & Hinkley, 1974). Because of property (e) any transformation of $p(x, x+t)$, such as the cumulative force of decrement given by

$$\int_x^{x+t} (\mu_s+\nu_s)\, ds$$

is itself a maximum likelihood estimate (Cox & Hinkley, 1974, p. 287). However, in general maximum likelihood estimates of mortality and other decremental intensities (Elveback, 1958; Sverdrup, 1965; Hoem, 1976) only have known asymptotic probability distributions with biased means.

To adapt relation (1) for actuarial estimation we remember that in simple binomial sampling N, the number of trials, is regarded as fixed and the random variable is measured (counted) somewhere within the span of time covered by the probability p. The intervals $(x+\tau_j,\ x+\tau_{j+1})$ in (1) are thus to be fixed by some determinate rule and the realizations of the random variables death or disability must occur *within* the interval (or, by convention, *at* the end-point of the interval). Hence we ascribe the pair of epochs $(a_j,\ b_j)$ to every individual under observation, a_j denoting his entrance epoch and b_j his putative epoch of exit from observation, although an eventual record will have to be made of his decremental epoch if this precedes b_j. The epochs a_j, b_j are ordered in time and the result is, we suppose, a consecutive series of epochs c_k $(k = 0, 1, 2, \ldots)$ any one of which denotes *one or more* of the "beginning" and "ending" dates mentioned above. If consecutive integer ages are to be made the basis of the estimation the attainment of each age by individual j will have to be given a c_k-epoch to start or end a sub-interval. The essential point to note is that random decrements (deaths and disabilities) only occur *between* two consecutive c_k-values. If data are collected in such a way that the end of each individual's exposure is randomly determined (e.g. Elveback, 1958) the only b_j-values are age-points and "loss to observation" is just another random decrement.

This conceptually straightforward method of estimation makes no assumption about the mathematical form of $p(x, x+1)$. The idea was due to Böhmer (1912) who, however, divided his year into intervals each of which contained *one* decrement *or one* accession. A rather similar procedure was suggested independently by Linder (1935) who divided the year arbitrarily into z (unequal) intervals, z being, perhaps, equal to 12. Seal (1954) disinterred the Böhmer proposal, which had by then also been made by Meier (1953), and specified the ends of the successive intervals by the *fixed* c-values described above between which several decrements could be realized. Grenander (1956) corrected Seal's statement about the variance of his estimate, and Kaplan & Meier (1958) applied Böhmer's proposal to medical statistics where b_j-values were inappropriate and "loss to observation" was another random decrement. A quarter of a century ago Böhmer-type numerical calculations for an estimate of $p(x,\ x+1)$ based on a large number of observations would have been prohibitive. Today computer calculation of the successive c-values and the counts of the intervening decrements is quite simple and leads to probability estimates with the desirable properties enumerated above.

Before pursuing the mathematical properties of the Böhmer-type estimate we mention that Hoem's (1976) review of current ideas about what he calls the "occurrence/exposure" rate and his earlier theoretical papers (1969, 1971) suggest that the assumption of constant forces of decrement throughout each integer age is now the preferred approach to estimation of mortality; Berkson & Elveback (1960) summarize the relevant mathematics very clearly. Sprague's (1879) estimates of the forces of marriage and mortality have, in fact, only to be supplemented by the distributional results mentioned in these articles.

The classic actuarial unbiased estimate of the ("crude" or "dependent") probability of death at any integer age dates back to Wittstein (1862), although the commonly accepted approximations required to produce a simple "exposed to risk" (Seal, 1977) result in an assumption that the force of mortality decreases throughout the year of age (Cantelli, 1914). Just as bad is that our b_j-values have to be assumed to come from a particular type of distribution if the usual estimate is not to diverge from the true value when the number of observations is increased to infinity (Elveback, 1958; Breslow & Crowley, 1974). These features discourage the use of standard actuarial probability estimates except when the (a_j, b_j) intervals coincide with integer ages.

In the foregoing the epochs a_j, b_j, c_k were implicitely measured from an arbitrary origin. It is now more convenient to measure these epochs from exact age x, or whatever point of time is regarded as being equivalent thereto for any individual, so that we may write for the probability of a decrement $D_k^{(i)}$ of type i being equal to the specific value $d_k^{(i)}$ in the age interval $(x+c_k, x+c_{k+1})$

$$Pr\{D_k^{(i)} = d_k^{(i)} \mid N_k, {}^*p^{(i)}(x+c_k, x+c_{k+1})\} =$$

$$= \binom{N_k}{d_k^{(i)}} \{1 - {}^*p^{(i)}(x+c_k, x+c_{k+1})\}^{d_k^{(i)}} \{{}^*p^{(i)}(x+c_k, x+c_{k+1})\}^{N_k - d_k^{(i)}},$$

$$d_k^{(i)} = 0, 1, 2, \ldots, N_k$$

where

$1 - {}^*p^{(i)}(x+c_k, x+c_{k+1})$ is the "crude" or "dependent" probability of becoming a decrement of type i in the age interval $(x+c_k, x+c_{k+1})$

$$N_k = N_{k-1} + n_k - \sum_i d_{k-1}^{(i)} + f_k, \; N_{-1} \equiv 0 \qquad (4)$$

n_k is the net number of "ons" minus "offs" at the beginning of the interval and f_k is the addition to n_k to recognize the fact that some of its "offs" will have already been included in the $d^{(i)}$ of prior subintervals. If $N_k = 0$ no estimate is possible for a sub-interval commencing $x + c_k$. For simplicity we will now suppose that $i = 1$ and delete i and the asterisks from our expressions.

Consider the estimate

$$\hat{p}(x+c_k, x+c_{k+1}) = 1 - \frac{d_k}{N_k} \equiv \hat{p}_k, \; k = 0, 1, 2, \ldots \tag{5}$$

or, with $c_n = 1$

$$\hat{p}(x, x+1) = \prod_{k=0}^{n-1} \left(1 - \frac{d_k}{N_k}\right).$$

If $c_0 = 0$ and $c_2 = 1$ the estimate of

$$p_x = p(x, x+1)$$

becomes

$$\prod_{k=0}^{1} \hat{p}(x+c_k, x+c_{k+1}) = \left(1 - \frac{d_0}{N_0}\right)\left(1 - \frac{d_1}{N_1}\right)$$

and we have

$$E\left\{\prod_{k=0}^{1} \hat{p}(x+c_k, x+c_{k+1})\right\} = E\left(1 - \frac{d_0}{N_0}\right) E\left\{\left(1 - \frac{d_1}{N_1}\right) \middle| d_0\right\}$$

$$= p_0 \, E\left\{\left(1 - \frac{d_1}{N_1}\right) \middle| N_1\right\} = p_0 \, p_1 = p_x$$

This may be extended by induction to prove that the proposed estimate (5) is unbiased given N_k.

For the second moment we have

$$E\left\{\left(1 - \frac{d_0}{N_0}\right)^2 \left(1 - \frac{d_1}{N_1}\right)^2\right\} = E\left(1 - \frac{d_0}{N_0}\right)^2 E\left\{\left(1 - \frac{d_1}{N_1}\right)^2 \middle| d_0\right\}$$

$$= \left(\frac{p_0}{N_0} + \frac{N_0 - 1}{N_0} p_0^2\right)\left(\frac{p_1}{N_1} + \frac{N_1 - 1}{N_1} p_1^2\right)$$

and, inductively,

$$\mu_2'\left\{\prod_{k=0}^{n-1} \hat{p}(x+c_k, x+c_{k+1})\right\} = \prod_{k=0}^{n-1} \left(\frac{p_k}{N_k} + \frac{N_k - 1}{N_k} p_k^2\right) n = 1, 2, 3, \ldots \tag{6}$$

The variance is a minimum amomg variances of unbiased estimates of p_x based on the set $\{a_i, b_j\}$ (Grenander, 1956).
Similarly

$$\mu_3'\left\{\prod_{k=0}^{n-1}\hat{p}(x+c_k,\, x+c_{k+1})\right\}=\prod_{k=0}^{n-1}\left\{\frac{p_k}{N_k^2}+3\,\frac{N_k-1}{N_k^2}\,p_k^2+\frac{(N_k-1)^{(2)}}{N_k^2}\,p_k^3\right\} \quad (7)$$

and

$$\mu_4'\left\{\prod_{k=0}^{n-1}\hat{p}(x+c_k,\, x+c_{k+1})\right\}=$$

$$=\prod_{k=0}^{n-1}\left\{\frac{p_k}{N_k^3}+7\,\frac{N_k-1}{N_k^3}\,p_k^2+6\,\frac{(N_k-1)^{(2)}}{N_k^3}\,p_k^3+\frac{(N_k-1)^{(3)}}{N_k^3}\,p_k^4\right\} \quad (8)$$

(Johnson & Kotz, 1969, p. 51).
Estimates of (6), (7) and (8) can be obtained from (5). The distribution of $\hat{p}(x, x+1)$ tends to normality with mean p_x and variance obtained from (6). Estimates of these parameters are calculated in the usual manner and a Pearson curve could be fitted by using estimates of the first four moments about the estimated mean.

Illustration

Estimation at two consecutive ages, 40 and 41, will serve to illustrate the procedure. It is supposed that 1000 individuals are "beginners" at age 40 and the ends of fractional intervals at which one individual was "existing" are noted as c_j ($j=1, 2, \ldots$) in the following table, the final unit value at each age denoting the end of the integer age interval. No new "beginners" are supposed to be added at age 41. The "crude" probabilities of death and disability for each of the intervals (needed to simulate the numbers of deaths and disabilities in the interval) were obtained from (2) and (3), respectively, where

$$\mu_x = 0.0000260\times 10^{0{,}043x}$$

$$\nu_x = 0.0000932\times 10^{0{,}043x}$$

The c-values of these two Gompertz forces were chosen to be equal so that (2) would result in

$$\frac{bc^x\, e^H}{H \ln c}\,\{e^{-H}-e^{-H\,c^t}\} \quad \text{where } H=(b+B)\,c^x/\ln c$$

$b=0.0000260 \qquad B=0.0000932$

and (3) would be B/b times this. The two $*q$'s for each interval were then used to calculate the first few terms of the appropriate binomial with N given by (4) and a pair of pseudorandom numbers from a uniform distribution were compared with the cumulation of the two sets of binomial terms to obtain the numbers of deaths and disabilities, respectively, in the interval. These are also shown in the table.

Age 40	$j=0$		1		2		3		4				
c_j	0		0.151		0.808		1.0		–				
No. exposed	1000		999		994		992						
Deaths		0		0		0		–					
Disabilities		0		4		1		–					
$10^3 \cdot {}^*\hat{q}^d_{40}$										0.0	±0.0	(1.4)	0.0
$10^3 \cdot {}^*\hat{q}^i_{40}$										4.0	±2.2	(5.1)	5.0
Age 41													
c_j	0		0.384		0.534		0.781		1.0				
No. exposed	992		986		982		980		978				
Deaths		3		0		1		0					
Disabilities		2		3		0		1					
$10^3 \cdot {}^*\hat{q}^d_{41}$										5.0	±2.0	(1.6)	4.0
$10^3 \cdot {}^*\hat{q}^i_{41}$										6.1	±2.5	(5.7)	6.1

The estimated crude probabilities (multiplied by 10^3) calculated from (5) and their standard errors calculated from (6) are shown on the right with the true values in parentheses. For comparison what Sprague (1879) would call the annual death (disability) rate among those who do not become disabled (die) in the year is shown at the extreme right of the table, e.g. $\hat{\hat{q}}^d_{41} = 4/\{992-(4/2)\}$ $= .00404$.

Prof. Hilary L. Seal
La Mottaz
1143 Apples

References

Berkson, J. & Elveback, L. (1960): Competing exponential risks, with particular reference to the study of smoking and lung cancer. *J. Amer. Statist. Assn. 55*, 415–428.

Blaschke, E. (1906): *Vorlesungen über Mathematische Statistik.* Teubner, Leipzig.

Böhmer, P.E. (1912): Theorie der unabhängigen Wahrscheinlichkeiten. *Rapp. 7e Congr. Internat. Actu. 2*, 327–346.

Breslow, N. & Crowley, J. (1974): A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist. 2*, 437–453.

Cantelli, F.P. (1914): Genesi e costruzione delle tavole di mutualità. *Boll. notiz. sul credito e sulla previdenza* 3–4 (included in the author's *Alcune Memorie Matematiche*, Giuffrè, Milan [1958]).

Cox, D.R. & Hinkley, D.V. (1974): *Theoretical Statistics.* Chapman & Hall, London.

Czuber, E. (1910): *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung* Bd. 2, Teubner, Leipzig.

Elveback, L. (1958): Estimation of survivorship in chronic disease: The "actuarial" method. *J. Amer. Statist. Assn. 53*, 420–440.

Grenander, U. (1956): On the theory of mortality measurement. *Skand. Aktuar. – Tidskr. 39*, 70–96, 125–153.

Hoem, J.M. (1969): The sampling distribution of an estimator arising in connection with the truncated exponential distribution. *Ann. Math. Statist. 40*, 702–703.

– (1971): Point estimation of forces of transition in demographic models. *J. Roy Statist. Soc. B, 33*, 275–289.

– (1976): The statistical theory of demographic rates: A review of current developments. *Scand. J. Statist. 3*, 169–185.

Johnson, N.L. & Kotz, S. (1969): *Distributions in Statistics: Discrete Distributions.* Houghton Mifflin, Boston.

Kaplan, E.L. & Meier, P. (1958): Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assn. 53*, 457–481.

Kendall, M.G. & Stuart, A. (1979): *The Advanced Theory of Statistics, Vol. 2*, Griffin, London.

Landre, C.L. (1911): *Mathematisch-technische Kapitel zur Lebensversicherung.* Fischer, Jena.

Lexis, W. (1877): *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft.* Wagner, Freiburg i. Br.

Linder, A. (1935): Über die Berechnung der Wahrscheinlichkeiten unabhängiger Ordnungen aus den Beobachtungszahlen. *Mitt. Verein. Schweiz. Vers.-Math. 30*, 35–52.

Meier, P. (1953): Nonparametric estimation of survivorship (Abstract). *Ann. Math. Statist. 24*, 688.

Mises, R. von (1964): *Mathematical Theory of Probability and Statistics.* Academic Press, New York.

Seal, H.L. (1949): Mortality data and the binomial probability law. *Skand. Aktuar.-Tidskr. 32*, 188–216.

– (1954): The estimation of mortality and other decremental probabilities. *Skand. Aktuar.-Tidskr. 37*, 137–162.

– (1977): Studies in the History of Probability and Statistics. XXXV Multiple decrements or competing risks. *Biometrika 64*, 429–439.

Sprague, T.B. (1879): On the construction of a combined marriage and mortality table from observations made as to the rates of marriage and mortality among any body of men; etc. *J. Inst. Actu. 21*, 406–452.

Sverdrup, E. (1965): Estimates and test procedures in connection with stochastic models for deaths, recoveries and transfers between states of health. *Skand. Aktuar.-Tidskr. 48*, 184–211.

Wittstein, T. (1862): Die Mortalität in Gesellschaften mit successiv eintretenden und ausscheidenden Mitgliedern. *Arch. Math. u. Physik 39*, 67–92.

Summary

It is suggested that the standard "exposed to risk" method of estimating mortality rates, depending on the assumption that mortality decreases throughout the year of age, be replaced by a method closely analogous to that proposed by Böhmer (1912) and, independently, by Linder (1935).

Zusammenfassung

Der Autor schlägt vor, die zur Schätzung der Sterblichkeit verwendete Standardmethode der «Anzahl Personen unter Risiko», welche auf der Hypothese basiert, dass die Sterblichkeit im Laufe des Jahres abnimmt, zu ersetzen durch ein Verfahren, welches ähnlich ist einer Methode vorgeschlagen von Böhmer (1912) und, unabhängig, von Linder (1935).

Résumé

L'article suggère de remplacer, lors de l'estimation des taux de mortalité, la méthode standard du «nombre de personnes exposées au risque», dépendant de l'hypothèse d'une mortalité décroissante en cours d'année, par une méthode très voisine de celle qui a été proposée par Böhmer (1912) et, de façon indépendante, par Linder (1935).