THE STANDARD DEVIATION OF THE PRESENT VALUE OF A LIFE ANNUITY AS ESTIMATED FROM A GRADUATED MORTALITY TABLE

Stuart Klugman
University of Iowa

1. Introduction

All net premium calculation methods are the end result of an estimation procedure. In the particular case of an immediate life annuity, the goal is to estimate the value of $E[a_{\overline{T}|}]$ where T is a random variable equal to the number of full years lived by an annuitant from issue to death. If the one-year survival probabilities are known for the population of annuitants, then $E[a_{\overline{T}|}] = a_{\overline{1}|}\,{}_{1|}q_x + a_{\overline{2}|}\,{}_{2|}q_x + \cdots = a_x$ for policies issued at age x. Since the survival probabilities are unknown, estimates of their values must be used. The resulting annuity value is thus an estimator of a_x and the standard deviation of this estimator provides a measure of its accuracy. This paper will consider the calculation of this standard deviation.

2. Definitions and qualifications

It will be assumed that the estimates of p_x have been obtained from a graduated mortality table. The study which produced these values had n_y independent observations at age y

and the observations at different ages are also assumed to be independent At each age y there were θ_y observed deaths and a crude survival rate of $\hat{p}_y = 1 - \theta_y/n_y$. The crude rates themselves provide an estimator $\hat{a}_x = v\hat{p}_x + v^2\,{}_2\hat{p}_x + \cdots$ of a_x.

Most graduation methods involve a linear transformation of the crude rates. If $\hat{\underset{\sim}{p}}$ is a column vector of the crude rates then the graduated rates will be $\tilde{\underset{\sim}{p}} = \underset{\sim}{a} + B\hat{\underset{\sim}{p}}$ where $\underset{\sim}{a}$ and B are a vector and matrix of known values, respectively. This scheme includes graduation by interpolation, moving-weighted-average, Whittaker-Henderson and Bayesian approaches. The post-graduation estimator is $\tilde{a}_x = v\tilde{p}_x + v^2\,{}_2\tilde{p}_x + \cdots$.

There are two matters that will not be addressed by this paper. One is the lack of independence between observations in most mortality studies. It should be noted that the method proposed in the next section will still work if the estimates $\hat{p}_y$ and $\hat{p}_{y+1}$ are correlated, but the correlation would have to be known. The second concerns other sources of error. This paper considers only sampling error. There is also a possibility of error due to sampling from the wrong population. This is most always the case as the "true" a_x actually depends on mortality rates which apply at some future time while the sampling must be from a current or past population.

3. The Standard Deviation

The standard deviation of $\hat{a}_x$ was obtained by Tinner (JIA LVI pp. 301-308) in 1925. He found $E[\hat{a}_x] = a_x$ and

$$\mathrm{Var}[\hat{a}_x] = \sum_{t=1}^{\infty} v^{2t} {}_tp_x^2(1+2a_{x+t})[(1+\alpha_x)(1+\alpha_{x+1})\cdots(1+\alpha_{x+t-1})-1]$$

where $\alpha_y = q_y/n_y p_y$. If all $n_y \to \infty$ then clearly $\mathrm{Var}[\hat{a}_x] \to 0$ (provided $\mathrm{Var}[\hat{a}_x] < \infty$ for some set of sample sizes) and so $\hat{a}_x$ is a consistent estimator of a_x. Although $\mathrm{Var}[\hat{a}_x]$ cannot be directly calculated, a reasonable approximation can be obtained by using $\hat{p}_y$ in place of p_y or by obtaining the p_y from a standard table.

It is unusual for the crude rates to be used in the evaluation of actuarial functions. A linear graduation procedure will produce smoother and more representative survival rates. It is believed that the graduated rates provide better estimates of the population rates and so $\tilde{a}_x$ should be a better estimator of a_x. This should be reflected by a reduced variance. As the graduated estimators of p_y are no longer independent, calculation of $\mathrm{Var}[\tilde{a}_x]$ will be a more complex problem.

Instead of $\mathrm{Var}[\tilde{a}_x]$, $E[\tilde{a}_x^2]$ will be obtained.

$$E[\tilde{a}_x^2] = E[(\sum_{t=1}^{\infty} v^t {}_t\tilde{p}_x)^2] = E[\sum_{t=1}^{\infty}\sum_{r=1}^{\infty} v^{t+r} {}_t\tilde{p}_x \, {}_r\tilde{p}_x]$$

$$= E[\sum_{t=1}^{\infty} v^{2t} {}_t\tilde{p}_x^2 + 2\sum_{t=2}^{\infty}\sum_{r=1}^{t-1} v^{t+r} {}_t\tilde{p}_x \, {}_r\tilde{p}_x]$$

$$= \sum_{t=1}^{\infty} v^{2t} E[{}_t\tilde{p}_x^2] + 2\sum_{t=2}^{\infty}\sum_{r=1}^{t-1} v^{t+r} E[{}_r\tilde{p}_x^2 \, {}_{t-r}\tilde{p}_{x+r}].$$

Also, $E[\tilde{a}_x] = \sum_{t=1}^{\infty} v^t E[{}_t\tilde{p}_x]$. Using the exact distributions would make evaluation of the required moments extremely difficult. To avoid some of the problem, a lognormal approximation will be used. Assume $\ell n(\tilde{\underset{\sim}{p}}) \sim N(\underset{\sim}{\nu}, \Delta)$. The values of $\underset{\sim}{\nu}$ and Δ will be obtained by the method of moments. From the lognormal distribution, $E[\tilde{p}_y] = \exp(\nu_y + \delta_{yy}/2) = \mu_y$ where δ_{xy} represents an element of Δ and $Cov[\tilde{p}_y, \tilde{p}_z] = \exp(\nu_y + \nu_z + \delta_{yy}/2 + \delta_{zz}/2 + \delta_{yz}) - \mu_y \mu_z = \gamma_{yz}$. At the same time $E[\hat{p}_y] = p_y$, $Var[\hat{p}_y] = p_y(1-p_y)/n_y = \sigma_{yy}$ and $Cov(\hat{p}_y, \hat{p}_z) = 0 = \sigma_{yz}$ for $y \neq z$. Let Σ be the matrix containing the σ_{yz}. Since $\tilde{\underset{\sim}{p}} = \underset{\sim}{a} + B\hat{\underset{\sim}{p}}$, $E[\tilde{\underset{\sim}{p}}] = \underset{\sim}{a} + B\underset{\sim}{p} = \underset{\sim}{\mu}$ and $Var[\tilde{\underset{\sim}{p}}] = B\Sigma B' = \Gamma$ with γ_{yz} being the elements of Γ. With $\underset{\sim}{\mu}$ and Γ known it is possible to solve for $\underset{\sim}{\nu}$ and Δ:

$$\delta_{yz} = \ell n(\gamma_{yz}/\mu_y \mu_z + 1)$$
$$\nu_y = \ell n\ \mu_y - \delta_{yy}/2.$$

For calculating the moments, note that all the needed moments are of the form $E[\tilde{p}_x^{r_x}\ \tilde{p}_{x+1}^{r_{x+1}} \cdots]$ with $r_y = 0, 1$ or 2 and $r_y \geq r_{y+1}$. The expectation can be written as

$$E[\prod_{y=x}^{\infty} \tilde{p}_y^{r_y}] = E[\exp(\ell n \prod_{y=x}^{\infty} \tilde{p}_y^{r_y})] = E[\exp(\sum_{y=x}^{\infty} r_y\ \ell n\, \tilde{p}_y]$$

$$= E[\exp(\underset{\sim}{r}'\ \ell n\,(\tilde{\underset{\sim}{p}}))] = M_{\ell n(\tilde{\underset{\sim}{p}})}(\underset{\sim}{r})$$

$$= \exp(\underset{\sim}{r}'\underset{\sim}{\nu} + \underset{\sim}{r}'\Delta\,\underset{\sim}{r}/2).$$

$M(\underset{\sim}{r})$ is the moment generating function of the random variable $\ell n(\tilde{\underset{\sim}{p}})$.

4. An Example

The ideas in the preceding section will be illustrated with a small example. Consider the crude rates for terminations due to death or recovery on group long-term disability cases (TSA 1976 Reports, pp. 137-139) for ages 40-49 (Table 1). The graduated rates were obtained by a Whittaker-Henderson Type B formula with second differences, a smoothing coefficent of 5000, and exposures as weights.

Table 1

duration (y)	n_y	θ_y	$\hat{p}_y$	$\tilde{p}_y$	$\sigma_{yy}(\times 10^5)$
0	6625	1311	.8021	.7946	2.396
1	4088	823	.7987	.8218	3.933
2	2498	349	.8603	.8591	4.811
3	1724	135	.9217	.8975	4.186
4	1131	58	.9487	.9286	4.327
5	764	46	.9398	.9524	7.405
6	484	26	.9463	.9737	10.499

For the crude rates and $v = .95$, $\hat{a} = 3.2316$ and St Dev($\hat{a}$) = .0345 using $\hat{p}_y$ for p_y. If $\tilde{p}_y$ is used for p_y then St Dev($\hat{a}$) = .0352. From the Whittaker-Henderson graduation, the relevant matrices may be found in Tables 2-4 and an illustration of the calculation of $\tilde{a}$ and $E[\tilde{a}]$ (using $\tilde{p}$ for p) is presented in Table 5. The two values are nearly identical as their ratio is similar to exp[sum of above diagonal elements of Δ] which turns out to be very close to 1.

Table 2

B matrix for the Whittaker-Henderson graduation

.8356	.2050	.0175	-.0196	-.0188	-.0126	-.0072
.3322	.3948	.1913	.0772	.0197	-.0035	-.0117
.0465	.3131	.3419	.1998	.0831	.0223	-.0067
-.0752	.1830	.2895	.3113	.1801	.0843	.0271
-.1101	.0711	.1834	.2746	.2782	.1908	.1119
-.1091	-.0187	.0729	.1901	.2825	.3213	.2610
-.0986	-.0989	-.0344	.0964	.2615	.4120	.4621

Table 3

Γ matrix for variance of graduated rates ($\Gamma = B\Sigma B'$)
(all elements are $\times 10^6$)

18.447	9.929	3.497	-.287	-2.189	-3.156	-3.827
9.929	10.818	9.098	6.012	2.854	-.037	-2.778
3.497	9.098	11.543	10.303	7.304	3.800	.208
-.287	6.012	10.303	11.548	10.520	8.503	6.164
-2.189	2.854	7.304	10.520	12.624	14.071	15.186
-3.156	- 037	3.800	8.503	14.071	20.319	26.640
-3.827	-2.778	.208	6.164	15.186	26.640	39.014

Table 4

Parameters for the lognormal random variable $\underset{\sim}{\nu}$ vector

-.2300	-.1962	-.1519	-.1082	-.0742	-.0488	-.0268

Δ matrix
(all elements are x 10^6)

29.219	15.204	5.123	-.403	-2.967	-4.171	-4.947
15.204	16.017	12.886	8.151	3.741	-.047	-3.473
5.123	12.886	15.640	13.364	9.157	4.645	.249
-.403	8.151	13.364	14.339	12.626	9.949	7.054
-2.967	3.741	9.157	12.626	14.643	15.913	16.799
-4.171	-.047	4.645	9.949	15.913	22.402	28.731
-4.947	-3.473	.249	7.054	16.799	28.731	41.159

Table 5

Calculation of $\tilde{a}$ and $E[\tilde{a}]$

duration (y)	v^y	${}_y\tilde{p}$	$v^y\,{}_y\tilde{p}$	$E[{}_y\tilde{p}]$	$v^y E[{}_y\tilde{p}]$
1	.95	.7946	.7549	.7946	.7548
2	.9025	.6530	.5893	.6530	.5894
3	.8574	.5610	.4810	.5610	.4810
4	.8145	.5035	.4101	.5035	.4101
5	.7738	.4675	.3617	.4675	.3618
6	.7351	.4453	.3273	.4453	.3273
7	.6983	.4336	.3028	.4336	.3028
			$\tilde{a}$ = 3.2272		$E[\tilde{a}]$ = 3.2273

The standard deviation of $\tilde{a}$ turns out to be .0345. This shows that, at least in this example, the graduation had very little effect on the variance. It is much more sensitive to the choice of values used for p_y than it is to the graduation induced dependence. Of course, no general conclusions can be made from the analysis of an isolated example such as the above.

5. Conclusions

Calculation of the standard deviation of an estimator of a_x can serve at least four useful purposes. It first provides a measure of the success of a graduation procedure. Secondly, it can be used prior to a mortality study to determine adequate sample sizes. A third use would be for setting margins on mortality tables. Instead of analyzing the rates themselves, it may be more appropriate to deal with the actuarial functions that are actually used. A final use was demonstrated by Jenkins and Lew in creating the 1949 Annuity Tables (TSA I, p. 442). The final table was intended to be used for all types of annuities. To determine if this table would be useful for a specific type of annuity, $\hat{a}^*_x$ based on a study only of that type of annuitant was obtained and then its difference from $\hat{a}_x$ in units of $Var[\hat{a}_x]$ was measured where $\hat{a}_x$ was based on the complete experience table If this distance was small (e g., less than 2) then the table would be acceptable for that type of annuity. For example, experience under immediate refund annuities was worse at a level of 7 standard deviations at issue age 60. The standard deviations were calculated using Steffensen's approximation which is always conservative.

Although computation of the standard deviation by the above method is not difficult for a computer, it would be convenient to have an approximation which gives quick but reasonably accurate results. It would also be helpful to know if the level of dependence introduced by the graduation is significant enough to warrant the use of this method. Experience with larger studies than that used in this paper may provide some enlightenment.

6. Acknowledgements

I want to thank Dick Ziock for suggesting this topic and Jim Hickman for pointing out the reference to Jenkins and Lew. The idea of using the lognormal approximation was suggested by Harry Panjer.