The Distribution of Aggregate Life Insurance Claims

Tom Edwalds, FSA

Munich American Reassurance Co.

ABSTRACT

This paper demonstrates the calculation of the moments of the distribution of aggregate life insurance claims from seriatim inforce data, assuming that each record represents an independent Bernoulli trial with its own known probability of a claim, q. Continuous distributions matching the first three or four moments are identified.

INTRODUCTION

With the adoption by the National Association of Insurance Commissioners of the Valuation of Life Insurance Policies Model Regulation, commonly referred to as "Regulation XXX", life insurance valuation actuaries gained the option to use their own company's product-specific mortality assumptions in the calculation of life insurance deficiency reserves. This flexibility is achieved through the application of "X factors" to the statutory valuation mortality tables. Actuaries who exercise this option must not only justify their selection of X factors before they use them, they must also assess their appropriateness in light of emerging mortality experience.

In order to assess the appropriateness of any particular mortality assumption, it is useful to understand the distribution of aggregate life insurance claims implied by this assumption. The question of the distribution of aggregate life insurance claims was most recently addressed two decades ago in the context of stop-loss reinsurance. In a paper published in TSA XXXII (1980) entitled, "The Aggregate Claims Distribution and Stop-Loss Reinsurance", Harry Panjer described a method for approximating the distribution of aggregate claims for a group life insurance contract. Panjer found that a compound Poisson process appropriately modeled the aggregate claims distribution, based on the collective risk assumption, which states that each life which leaves the group by death claim is immediately replaced by a life with identical mortality characteristics. The resulting method produces accurate and useful results, and is relatively easy to use as long as the number of distinct amounts of insurance in the group is not too great. In practice, this condition is achieved simply by rounding all face amounts to be integral multiples of some convenient unit.

In spite of its usefulness, the reliance of the Panjer method on the collective risk assumption is not intuitively appealing. Peter Kornya attempted to eschew the collective risk assumption in his paper, "Distribution of Aggregate Claims in the Individual Risk Theory Model", published in TSA XXXV (1983). The method he presented was not as computationally tractable as the Panjer method, so the Panjer method is still the one most widely used today.

This paper presents an alternative to the Panjer method for estimating the distribution of aggregate life insurance claims that does not depend on the collective risk assumption and can be used without first rounding and grouping the face amounts. This method can be used to construct hypothesis tests of mortality assumptions (both one-sided and two-sided) as well as estimate appropriate premiums for stop-loss coverage. In the method of this paper, the moments of the distribution of aggregate life insurance claims are calculated from seriatim inforce data, given a mortality assumption. The mortality assumption provides a mechanism by which each record is assigned its own known probability of a claim, q. The moments of the distribution of aggregate life insurance claims are then used to estimate the percentiles of that distribution by identifying closed-form continuous distributions with matching moments. An extra advantage of this method is that moments of the distributions of aggregate claims associated with various subsets of the inforce data can be calculated simultaneously by using database techniques.

MODEL ASSUMPTIONS AND FORMULAS

In order to use the approach described in this paper, we assume that the actuary has access to a seriatim inforce dataset that provides a complete and accurate accounting of the exposure to be studied. In particular, we assume that there is a one-to-one correspondence between the exposed lives of interest and the records in an identifiable subset of the dataset. We further assume that each record in the subset of interest contains accurate information about the birth date, gender, smoker status, and underwriting rating of the corresponding insured life and about the issue date, face amount, and amount at risk of the corresponding life insurance policy. We posit the existence of an appropriate mortality model that assigns a mortality rate to each exposed life of interest. Furthermore, we assume that any data fields necessary either to identify the exposed lives of interest or to assign a mortality rate to an exposed life are accurately coded. Finally, we assume that all of the exposures are independent.

For the jth policy record, the probability of a death claim occurring during the time period of interest is q_j, the mortality rate assigned by the mortality model. The amount of the death claim (if it occurs) is A_j, which is a constant for policy j. Now let C_j be the Bernoulli random variable representing the occurrence of a claim on policy j (i.e., $C_j = 1$ if there is a claim on policy j during the time period of interest, and $C_j = 0$ otherwise). If we let L be the aggregate claims liability arising from the exposed lives of interest over the time period of interest, we see that $L = \sum_j A_j \cdot C_j$. If we denote the means of C_j and L by $\mu(C_j)$ and $\mu(L)$, and their k^{th} central moments[1] by $\mu_k(C_j)$ and $\mu_k(L)$, respectively, it can be shown that:

[1] Recall that the k^{th} central moment of a random variable x is defined as the expected value of $(x - \mu(x))^k$.

$$\mu(C_j) = q_j$$
$$\mu_2(C_j) = q_j \cdot (1 - q_j)$$
$$\mu_3(C_j) = q_j \cdot (1 - q_j) \cdot (1 - 2 \cdot q_j)$$
$$\mu_4(C_j) = q_j \cdot (1 - q_j) \cdot (1 - 3 \cdot q_j + 3 \cdot q_j^2)$$
$$\mu_5(C_j) = q_j \cdot (1 - q_j) \cdot (1 - 2 \cdot q_j) \cdot (1 - 2 \cdot q_j + 2 \cdot q_j^2)$$
$$\mu(L) = \sum_j A_j \cdot q_j$$
$$\mu_2(L) = \sum_j A_j^2 \cdot \mu_2(C_j)$$
$$\mu_3(L) = \sum_j A_j^3 \cdot \mu_3(C_j)$$
$$\mu_4(L) = \sum_j A_j^4 \cdot \mu_4(C_j) + 3 \cdot \left[\left(\sum_j A_j^2 \cdot \mu_2(C_j) \right)^2 - \sum_j A_j^4 \cdot \mu_2^2(C_j) \right]$$
$$\mu_5(L) = \sum_j A_j^5 \cdot \mu_5(C_j) + 10 \cdot \left[\left(\sum_j A_j^2 \cdot \mu_2(C_j) \right) \cdot \left(\sum_j A_j^3 \cdot \mu_3(C_j) \right) - \sum_j A_j^5 \cdot \mu_2(C_j) \cdot \mu_3(C_j) \right]$$

Note that all of the moments of L can be calculated from sums of quantities that depend only on the amounts A_j and the moments of C_j, which in turn are dependent only on q_j. These quantities and sums can be calculated with a single pass through the data.

We know from experience that L cannot be negative and that the probability that $L = 0$ becomes vanishingly small as the number of exposures increases. Furthermore, we expect that the distribution of L will be positively skewed. We are interested in continuous distributions that might approximate the distribution of L. To investigate, we will compare the calculated moments of L with the moments of four different continuous distributions: a normal distribution, a gamma distribution, a mixture of a gamma distribution plus an exponential distribution, and a mixture of a gamma distribution plus two exponential distributions.

The normal distribution provides a baseline of comparison for the other continuous distributions. The gamma distribution and the exponential distribution are positively skewed distributions with non-negative domains. The gamma density function also has the appealing property of passing through the origin. The use of a mixture of exponential distributions plus a gamma distribution was inspired by the paper, "Toward a Unified Approach to Fitting Loss Models" by Jacques Rioux and Stuart Klugman, which appeared in ARCH volume 37 number 1 (2002).

In order to match the first two moments of the distribution of *L*, the choice of parameters for the normal distribution is obviously $\mu = \mu(L)$ and $\sigma = \sqrt{\mu_2(L)}$ for the mean and standard deviation, respectively. The gamma density function also has two parameters: the scale parameter, β, and the shape parameter, γ. The gamma density function is expressed[2] as:

$f(x) = \left(\frac{x}{\beta}\right)^{\gamma-1} \cdot \frac{e^{-x/\beta}}{\beta \cdot \Gamma(\gamma)}$. The mean is $\beta \cdot \gamma$ and the variance is $\beta^2 \cdot \gamma$, so we can match the first two moments of the distribution of *L* by choosing $\beta = \mu_2(L) / \mu(L)$ and $\gamma = \mu(L) / \beta$.

For a mixture of a gamma distribution plus an exponential distribution (the Mix 1 model), we consider a random variable that is the sum of a gamma distributed random variable (with parameters β and γ) and an independent exponentially distributed random variable. The density function for the exponentially distributed component is $f(x) = \lambda \cdot e^{-\lambda x}$. The mean of the mixture random variable is $1/\lambda + \beta \cdot \gamma$, the variance is $1/\lambda^2 + \beta^2 \cdot \gamma$, and the third central moment is $2/\lambda^3 + 2 \cdot \beta^3 \cdot \gamma$. We can match the first three moments of the distribution of *L* by choosing λ so that it satisfies the cubic equation:

[2] Recall that the gamma function $\Gamma(u) = \int_0^\infty z^{u-1} \cdot e^{-z} dz$; for integers, $\Gamma(n) = (n-1)!$.

$$\frac{1}{\lambda^3} - 2 \cdot \frac{\mu_2(L)}{\mu(L)} \cdot \frac{1}{\lambda^2} + \frac{\mu_3(L)}{2 \cdot \mu(L)} \cdot \frac{1}{\lambda} + \frac{[\mu_2(L)]^2}{\mu(L)} = \frac{\mu_3(L)}{2}$$

Then choose $\beta = \dfrac{\mu_2(L) - \frac{1}{\lambda^2}}{\mu(L) - \frac{1}{\lambda}}$ and $\gamma = \dfrac{\mu(L) - \frac{1}{\lambda}}{\beta}$.

For the mixture of a gamma distribution plus two exponential distributions (the Mix 2 model), we consider a random variable that is the sum of three independent random variables: one gamma distributed with parameters β and γ, one exponentially distributed with parameter λ_1, and one exponentially distributed with parameter λ_2. The mean of this mixture random variable is $\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \beta \cdot \gamma$, the variance is $\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} + \beta^2 \cdot \gamma$, the third central moment is $\frac{2}{\lambda_1^3} + \frac{2}{\lambda_2^3} + 2 \cdot \beta^3 \cdot \gamma$, and the fourth central moment is:

$$\frac{9}{\lambda_1^4} + \frac{9}{\lambda_2^4} + 6 \cdot \left(\frac{\beta^2 \cdot \gamma}{\lambda_1^2} + \frac{\beta^2 \cdot \gamma}{\lambda_2^2} + \frac{1}{\lambda_1^2 \cdot \lambda_2^2} \right) + 3 \cdot \beta^4 \cdot \gamma \cdot (\gamma + 2).$$

This author was unable to find a closed form solution for the system of equations resulting from setting these moments equal to the moments of *L*. Instead, the "Goal Seek" feature of Microsoft Excel was used to find parameters that would result in a mixture random variable whose first four central moments matched the moments of *L*.

DATA DESCRIPTION AND RESULTS

The dataset used for this analysis consisted of certain life insurance policies reinsured by Munich American Reassurance Company (MARC) at midyear 2002. This dataset was organized as an Oracle table with about six million records. SQL queries were used to match these records with mortality rates from three standard mortality tables: the SoA 75-80 Basic Table, the SoA 90-95 Basic Table, and the SoA 2001 Valuation Basic Table. The first five central moments of the Bernoulli random variable for a claim during the next quarter on each record were calculated along with the variance squared and the product of the variance and the third central moment. The amount constant for each record was the dollar amount of risk retained by MARC, scaled so that a medium sized policy would be about 1 unit. The first five moments of the aggregate claim distribution for the next quarter for each client company included in the dataset and for the entire dataset were then calculated using the formulas presented above. These moments were then copied into an Excel spreadsheet, and parameters were calculated for the continuous distributions as discussed above. For each of these continuous distributions, percentiles of the cumulative distribution function (cdf) were then calculated for a confidence interval (CI) starting two standard deviations below the mean (or at zero, if greater) and ending two standard deviations above the mean.

We present here the results for the six million policies in the total dataset, for one large client company with about 400,000 policies in force, and for one small client company with about 1,000 policies in force.

TABLE 1
Distribution of Aggregate Claims for Total Dataset Using SoA 75-80 Basic Table

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			4589.73	203.56	0.33	3.34	4182.61	4996.84
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	4589.73	sigma	203.56	0.00	3.00	0.0228	0.9772
		lambda	beta	gamma				
Gamma			9.03	508.39	0.09	3.01	0.0203	0.9749
Mix 1		0.0095	6.75	664.01	0.33	3.44	0.0156	0.9700
Mix 2		0.0118	6.13	721.60	0.33	3.36	0.0149	0.9695
		0.0118						

TABLE 2
Distribution of Aggregate Claims for Total Dataset Using SoA 90-95 Basic Table

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			3186.61	166.57	0.40	3.50	2853.48	3519.75
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	3186.61	sigma	166.57	0.00	3.00	0.0228	0.9772
		lambda	beta	gamma				
Gamma			8.71	365.99	0.10	3.02	0.0199	0.9745
Mix 1		0.0108	6.19	499.66	0.40	3.58	0.0142	0.9688
Mix 2		0.0118	5.60	543.46	0.40	3.50	0.0137	0.9684
		0.0169						

TABLE 3
Distribution of Aggregate Claims for Total Dataset Using SoA 2001 VBT

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			2525.83	151.09	0.45	3.64	2223.65	2828.00
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	2525.83	sigma	151.09	0.00	3.00	0.0228	0.9772
		lambda	beta	gamma				
Gamma			9.04	279.48	0.12	3.02	0.0194	0.9741
Mix 1		0.0114	6.20	393.46	0.45	3.69	0.0132	0.9679
Mix 2		0.0118	5.70	420.56	0.45	3.64	0.0128	0.9676
		0.0221						

TABLE 4
Distribution of Aggregate Claims for Large Client Using SoA 75-80 Basic Table

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			151.26	23.00	0.62	3.85	105.27	197.26
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	151.26	sigma	23.00	0.00	3.00	0.0228	0.9772
		lambda	beta	gamma				
Gamma			3.50	43.26	0.30	3.14	0.0140	0.9697
Mix 1		0.0689	2.33	58.70	0.62	3.99	0.0087	0.9649
Mix 2		0.0848	1.95	65.31	0.62	3.86	0.0079	0.9641
		0.0844						

TABLE 5
Distribution of Aggregate Claims for Large Client Using SoA 90-95 Basic Table

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			103.85	19.55	0.78	4.28	64.74	142.95
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	103.85	sigma	19.55	0.00	3.00	0.0228	0.9772
		lambda	beta	gamma				
Gamma			3.68	28.21	0.38	3.21	0.0119	0.9681
Mix 1		0.0739	2.21	40.94	0.78	4.42	0.0059	0.9627
Mix 2		0.0798	1.79	46.12	0.78	4.28	0.0051	0.9619
		0.1135						

TABLE 6
Distribution of Aggregate Claims for Large Client Using SoA 2001 VBT

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			81.61	17.51	0.88	4.62	46.59	116.63
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	81.61	sigma	17.51	0.00	3.00	0.0228	0.9772
		lambda	beta	gamma				
Gamma			3.76	21.72	0.43	3.28	0.0103	0.9670
Mix 1		0.0787	2.11	32.72	0.88	4.71	0.0043	0.9614
Mix 2		0.0810	1.78	35.28	0.88	4.62	0.0038	0.9610
		0.1530						

TABLE 7

Distribution of Aggregate Claims for Small Client Using SoA 75-80 Basic Table

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			2.02	1.80	2.58	16.19	0.00	5.61
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	2.02	sigma	1.80	0.00	3.00	0.1307	0.9772
		lambda	beta	gamma				
Gamma			1.60	1.26	1.78	7.76	0.0000	0.9511
Mix 1		4.2337	1.78	1.00	1.95	8.80	0.0000	0.9506
Mix 2		13.5996	1.79	1.00	1.97	8.89	0.0000	0.9504
		6.4390						

TABLE 8

Distribution of Aggregate Claims for Small Client Using SoA 90-95 Basic Table

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			1.55	1.56	2.95	20.51	0.00	4.68
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	1.55	sigma	1.56	0.00	3.00	0.1605	0.9772
		lambda	beta	gamma				
Gamma			1.58	0.99	2.02	9.09	0.0000	0.9502
Mix 1		0.6445	49.07	0.00	2.95	103.05	0.0000	0.9479
Mix 2		10.0583	3.91	0.07	2.95	20.51	0.0000	0.9464
		0.8365						

TABLE 9

Distribution of Aggregate Claims for Small Client Using SoA 2001 VBT

			Mean	Std Dev	Skewness	Kurtosis	CI start	CI end
Calculated Statistics			1.19	1.34	3.55	28.75	0.00	3.88
Distribution		**Parameters**			Skewness	Kurtosis	cdf CI start	cdf CI end
Normal	mu	1.19	sigma	1.34	0.00	3.00	0.1882	0.9772
		lambda	beta	gamma				
Gamma			1.52	0.78	2.26	10.67	0.0000	0.9496
Mix 1		1.0418	3.86	0.06	3.55	28.86	0.0000	0.9582
Mix 2		888.5242	3.83	0.06	3.55	28.75	0.0000	0.8927
		1.0482						

DISCUSSION

Looking at Tables 1, 2, and 3 above, we see that for the six million policies in the total dataset the confidence intervals by mortality table are disjoint, demonstrating the significant differences in the overall mortality levels of the three tables used in this study. The confidence interval using the 75-80 table ranges from 91% to 109% of the mean; using the 90-95 table the range is 90% to 110%, and using the 2001 VBT the range is 88% to 112%. The lower the expected claims, the wider the confidence interval, relatively speaking. Each confidence interval includes about 95.5% of the distribution for all of the continuous distributions fitted to the moments calculated using that mortality table. While the normal distribution estimates that the probability that aggregate claims will exceed the confidence interval is equal to the probability that they will not even reach the confidence interval, the mixture models show that in fact the probability of claims exceeding the range is at least twice the probability of an unusually favorable outcome. For example, in Table 1 the "cdf CI end" for the normal model is 0.9772, so the estimated probability of aggregate claims exceeding the confidence interval is 1 – 0.9772 = 0.0228, which is the same as the "cdf CI start", the estimated probability that aggregate claims will not even reach the confidence interval. But in that same table, the "cdf CI end" for the Mix 2 model is 0.9695, leaving an estimated probability of aggregate claims exceeding the confidence interval equal to 1 – 0.9695 = 0.0305, which is more than twice the "cdf CI start" value of 0.0149.

Comparing the percentiles of the confidence interval between the Mix 1 and Mix 2 models, we see that the shift was fairly small. It appears that for this group of policies, the Mix 1 model provides a satisfactory fit for estimating the percentiles of the aggregate claims distribution.

For the large client, we see from Tables 4, 5, and 6 that the confidence intervals are quite wide and overlapping. Using the 2001 VBT, the confidence interval ranges from 57% to 143% of the mean. This shows that it is difficult to draw conclusions about the mortality assumption even for a block of 400,000 policies using only one quarter of experience. The mixture models again show that the probability of claims exceeding the range is significantly higher than the probability of an unusually favorable outcome, up to ten times as great when using the 2001 VBT. Again, the shift in the percentiles of the confidence interval was fairly small between the Mix 1 and Mix 2 models, suggesting that the Mix 1 model again provides a satisfactory fit.

Tables 7, 8, and 9 illustrate the well-known fact that the experience of a small block over a short period is unlikely to reveal much information. For the small client, the confidence intervals are one-sided and very wide. Using the 2001 VBT, the range is up to 326% of the mean. This case also tested the limits of this technique. While it was simple to calculate the moments of the aggregate claims distribution for this small client simultaneously with the total dataset and all clients in the dataset, matching these moments with continuous distributions proved to be quite problematic. While the equations for choosing the parameters of the Mix 1 model are guaranteed to have real solutions, the solutions are not necessarily positive, as the parameters must be. Also, even when all of the calculated parameters are positive, the values may be too extreme for calculating percentiles of the gamma distribution using Excel. For the mixture models for the small client using the 90-95 Table and the 2001 VBT, it was necessary to use numerical integration techniques to estimate the percentile values of the confidence interval.

CONCLUSION

The technique outlined in this paper appears to provide a useful method for constructing hypothesis tests of the mortality assumption for sufficiently large blocks of life insurance (apparently about a hundred thousand life years of exposure is sufficiently large). For such a block, it appears to be satisfactory to approximate the aggregate claims distribution with a mixture of a gamma distribution plus an exponential distribution with parameters chosen so that the first three moments match. For small blocks of business, this technique does not appear to be satisfactory. More experience with this technique in practice will help define the conditions for which it is most appropriate.