WORKSHOP

ON STOP-LOSS PREMIUMS FOR THE INDIVIDUAL MODEL

BY R. KAAS AND A. E. VAN HEERWAARDEN

University of Amsterdam

AND

M. J. GOOVAERTS

K.U. Leuven and University of Amsterdam

ABSTRACT

It is shown how the upper bounds for stop-loss premiums (and approximations to tail probabilities) obtained by replacing the individual model for a portfolio of risks by the collective model can be improved upon at the cost of only slightly more computer time. The method used is simply to keep a restricted number of large risks as they are instead of approximating them by a compound Poisson distribution. In a real-life example, the relative error in the stop-loss premium is shown to be reduced drastically by keeping only 10 out of 743 risks unchanged.

KEYWORDS

Stop-loss premium; individual model.

1. INTRODUCTION

Consider the individual model for the total claims in one year on a certain portfolio containing k contracts:

(1) $$S = \sum_{i=1}^{k} X_i.$$

With this random variable S we can construct a random variable S'' having stop-loss premiums at least as large as S (see, e.g. KAAS, 1987, Th. 1.1.3), by replacing the random variable X_i in (1) by a sum Y_i of N_i independent random variables $X_i^{(j)}$ with the same distribution as X_i:

(2) $$Y_i = \sum_{j=1}^{N_i} X_i^{(j)}, \qquad S'' = \sum_{i=1}^{k} Y_i,$$

where the random variables N_i have Poisson (1) distributions and are independent of $X_i^{(j)}$. The proof of this statement goes as follows. Consider Theorem 4.3.8 of GOOVAERTS *et al.* (1984), which states that one compound distribution precedes another in stop-loss order if both the number of claims and the individual terms are correspondingly ordered. First one applies this theorem to a counting distribution degenerate on $\{1\}$ to obtain $X_i < Y_i$ for all i, where the symbol $<$ denotes stop-loss order. Next, applying the same theorem with both N_1 and N_2 degenerate on $\{k\}$, we see that indeed $S < S''$.

We assume that the random variables X_i are either equal to some amount M_i (with probability q_i), or zero, which is realistic in life-insurance applications. Without much loss of generality we assume the amounts at risk M_i to be integer-valued.

The probability distribution of S can be computed using convolution. If t denotes the highest argument of interest, this process takes a number of operations proportional to $k \cdot t$, since the random variables X_i can have only two values. The probability distribution of S'' is a compound Poisson distribution with parameter $\lambda = \Sigma q_i$ and claim-amount distribution $P[Z = z] = \Sigma_{\{i:\ M_i = z\}} q_i / \lambda$. To compute this distribution, one might either use Panjer's recursive algorithm or the fast Fourier transform technique. The number of operations involved is proportional to $t \cdot \max M_i$ for Panjer's recursion and to $b \cdot \log b$ for the FFT technique, where $b \geqslant t$ is a number so large that $P[S = b]$ is less than the required precision (see BERTRAM, 1981, or KAAS, 1987). Since both t and b will be equal to $E[S]$ plus some multiples of $\sqrt{\mathrm{Var}[S]}$, they have the same order of magnitude for large portfolios.

In KAAS (1987) it is shown by a numerical example that the stop-loss premiums of S and S'' are really quite close together. See also GERBER (1984). There remains, however, room for improvement.

2. A BETTER-FITTING, STILL TRACTABLE MODEL

Our aim is to find a random variable S' having stop-loss premiums higher than S, but not as high as S''. Moreover we want the stop-loss premiums of S' to be computable in a time longer by only a constant factor than the time needed to compute those of S'', whatever the size of the portfolio.

This is achieved by mixing the techniques mentioned in Section 1 (Panjer and convolution). Define S' as

$$S' = \sum_{i \in V} X_i + \sum_{i \notin V} Y_i, \tag{3}$$

where V is some subset of $\{1, 2, \ldots, k\}$. The policies in set V remain unchanged, the others are replaced by a compound Poisson distribution. From the additivity property of stop-loss order (see GOOVAERTS *et al.*, 1984), we see that the stop-loss premiums of S' are between those of S and S''.

The distribution of S can be determined by first computing the probability vector of the second sum of (3), either by Panjer's recursion or by FFT, and subsequently adding the terms of the first sum using convolution. By taking the number of elements $|V|$ in V small enough, the $|V| \cdot t$ extra operations this requires do not add substantially to the number of operations used in the first step. If we use Panjer's recursion, we may take $|V|$ to be of order $\max M_i$. With FFT we may (asymptotically) leave about $\log b$ policies unchanged. For large k, the standard deviation of S becomes negligible compared to $E[S]$, so we may replace log b above by log k.

Having decided upon how many elements to include in V, the question remains which policies we should leave unchanged. We consider two criteria.

Denote by $\pi_Z(t)$ the stop-loss premium with retention t of an arbitrary random variable Z. From GERBER (1984, Th. 3, and property (1)) it follows that the so-called stop-loss distance (the maximum over all retentions of the absolute difference in stop-loss premium) between S and S' satisfies

(4) $$\max_{t \in V} | \pi_{S'}(t) - \pi_S(t) | \leqslant \tfrac{1}{2} \sum_{i \notin V} q_i^2 M_i.$$

Note that $\pi_{S'}(t) \geqslant \pi_S(t)$ for all t. To minimize this upper bound the policies with maximal $q_i^2 M_i$ should be included in V.

Another option is to minimize the total error made in the stop-loss premium over all integer retentions instead of only its maximum. This can be done as follows. First, it can be shown that the means of S, S' and S'' are equal. For any non-negative random variable Z we have

(5) $$\int_0^\infty \int_t^\infty (z-t)\; dF_Z(z)\; dt = \int_0^\infty \int_0^z (z-t)\; dt\; dF_Z(z)$$

$$= \int_0^\infty \tfrac{1}{2} z^2 \; dF_Z(z) = \tfrac{1}{2} E[Z^2]1.$$

From $E[S] = E[S']$ and (5) we deduce directly for the intergral over all retentions of the difference in stop-loss premiums for S' and S:

(6) $$\int_0^\infty \{\pi_{S'}(t) - \pi_S(t)\}\; dt = \tfrac{1}{2}\{\mathrm{Var}[S'] - \mathrm{Var}[S]\}.$$

Also, for an arbitrary random variable Z with values only in $\{0, 1, \dots\}$ we have by the piecewise linearity of $\pi_Z(t)$ for non-integral values of t,

(7) $$\int_0^\infty \pi_Z(t)\; dt = \sum_{i=0}^\infty \tfrac{1}{2}\{\pi_Z(i) + \pi_Z(i+1)\} = \tfrac{1}{2}\pi_Z(0) + \sum_{i=1}^\infty \pi_Z(i).$$

As $\pi_Z(0) = E[Z]$ for non-negative random variables Z, for the left-hand side of (6) we have

(8) $$\int_0^\infty \{\pi_{S'}(t) - \pi_S(t)\}\; dt = \sum_{i=0}^\infty \{\pi_{S'}(i) - \pi_S(i)\}.$$

Combining (8) and (6) we see that the total (absolute) error in the stop-loss premiums is minimized making the variance of S' as small as possible. The variance of S equals

(9) $$\mathrm{Var}[S] = \sum_{i=1}^k q_i(1-q_i)M_i^2.$$

For the variance of S'' we have

$$\text{(10)} \qquad \operatorname{Var}[S''] = \sum_{i=1}^{k} q_i M_i^2 = \operatorname{Var}[S] + \sum_{i=1}^{k} q_i^2 M_i^2,$$

and then of course the variance of S' can be written as

$$\text{(11)} \qquad \operatorname{Var}[S'] = \sum_{i \in V} q_i(1-q_i)M_i^2 + \sum_{i \notin V} q_i M_i^2 = \operatorname{Var}[S''] - \sum_{i \in V} q_i^2 M_i^2.$$

To minimize the variance of S', the policies to be included in V are those with the highest contributions to the variance of S''. This means that $q_i^2 M_i^2$ should be maximal, or equivalently $q_i M_i$ should be maximal. So, according to this second criterion, only the policies with the lowest risk premiums should be replaced by a compound Poisson distribution.

Intuitively it is more appealing to distinguish the policies on their risk premium than on their value of $q_i^2 M_i$. Also, since (4) gives only an upper bound to the maximum of the error, we can expect the error in the tails to be less using the second criterion.

3. NUMERICAL RESULTS

We tested the procedure outlined above on a real-life portfolio of widow/orphan pensions with a stop-loss coverage for each year's losses. The data consist of the capital lost in case of death and the mortality rate of the insured. To make exact computation of the probability function feasible we rounded the capitals to integers, after having applied a scaling factor. The resulting portfolio (after removal of capitals equal to zero) had the following characteristics:

number of policies, k	743
range of the risk capitals	$\{1, \ldots, 50\}$
expected number of deaths, $\sum q_i$	1.71
mean claim	3.18
expected value of S	5.44
variance of S	45.41

Table 1 gives the error in stop-loss premium made by applying the collective model. Two sizes $|V|$ were investigated: $|V| = 10$ and $|V| = 20$. Inclusion of policies in these sets V was done according to both criteria.

The exact values of the stop-loss premium, again expressed as percentages of $E[S]$, are given in Table 2, together with the relative errors of the approximating models. Again we see that the collective model S'' is a good approximation to the individual model, but, especially in the tails, the extra effort caused by using S' is certainly worthwhile. As expected, the errors for small retentions ($< E[S] + \sqrt{\operatorname{Var}[S]}$, say), were smaller using the first criterion, but for bigger retentions errors resulted almost as large as those for S''.

TABLE 1

ERROR BOUNDS AND ACTUALLY OBSERVED MAXIMAL ERROR (OVER ALL RETENTIONS) IN STOP-LOSS PREMIUMS FOR S'' AND S', AS A PERCENTAGE OF $E[S]$

	S''	$S', \|V\| = 10$	$S', \|V\| = 20$	$S', \|V\| = 10$	$S', \|V\| = 20$
		First criterion		Second criterion	
Total error	1.06	0.78	0.62	0.68	0.54
Error bound	0.18	0.15	0.14	0.16	0.14
Maximal error	0.046	0.037	0.031	0.041	0.035

TABLE 2

STOP-LOSS PREMIUMS FOR THE INDIVIDUAL MODEL S AS A PERCENTAGE OF $E[S]$; RELATIVE ERRORS (*100%) FOR THE COLLECTIVE MODEL S'' AND THE MIXED MODELS S' WITH $|V| = 10$ AND $|V| = 20$

Retention	S	S''	$S', \|V\| = 10$	$S', \|V\| = 20$	$S', \|V\| = 10$	$S', \|V\| = 20$
			First criterion		Second criterion	
0	100.000	0.00	0.00	0.00	0.00	0.00
3	61.476	0.05	0.04	0.04	0.05	0.04
6	38.345	0.12	0.10	0.08	0.11	0.09
9	24.846	0.18	0.14	0.11	0.15	0.12
12	16.775	0.23	0.16	0.13	0.17	0.13
15	11.744	0.27	0.18	0.15	0.18	0.14
18	8.414	0.30	0.20	0.17	0.19	0.14
21	6.062	0.34	0.22	0.19	0.19	0.14
24	4.427	0.39	0.24	0.21	0.20	0.15
27	3.246	0.45	0.27	0.23	0.21	0.15
30	2.379	0.53	0.30	0.25	0.22	0.16
33	1.727	0.64	0.36	0.29	0.24	0.17
36	1.245	0.76	0.44	0.34	0.26	0.19
39	0.887	0.91	0.56	0.41	0.29	0.20
42	0.612	1.11	0.73	0.50	0.32	0.22
45	0.420	1.37	0.96	0.63	0.37	0.25
48	0.280	1.75	1.31	0.82	0.45	0.28
51	0.173	2.41	1.92	1.18	0.58	0.34
54	0.108	3.29	2.77	1.66	0.76	0.43
57	0.068	4.41	3.88	2.26	0.96	0.51
60	0.044	5.77	5.24	2.97	1.16	0.57
63	0.029	7.38	6.84	3.75	1.33	0.61
66	0.019	9.21	8.64	4.59	1.45	0.59
69	0.013	11.46	10.89	5.67	1.52	0.58
72	0.009	14.20	13.53	6.98	1.47	0.57
75	0.006	17.42	16.77	8.86	1.42	0.20
78	0.004	21.11	20.31	11.28	1.40	0.61
81	0.003	24.78	23.92	14.64	1.39	0.64
84	0.002	28.64	27.81	19.20	1.46	0.67
87	0.001	34.37	32.84	24.66	1.61	0.76
90	0.001	42.79	41.48	33.17	1.82	0.91

4. SOME FURTHER COMMENTS

As pointed out by B. SUNDT at the XX ASTIN-Colloquium, Scheveningen (1987) the idea of separating large and small risks in the computations need not be restricted to the case of random variables X_k of purely life-insurance type. In the next more complicated case, for instance, of including both death risk M_k with probability q_k and disability risk Q_k with probability i_k on policy k, we may approximate S by the compound Poisson distribution S'' with parameter $\lambda = \sum q_k + \sum i_k$ and claim amount distribution $P[Z = z] = (\sum_{\{k:\, M_k = z\}} q_k + \sum_{\{k:\, Q_k = z\}} i_k)/\lambda$. Note, however, that the considerations of how many policies to leave unchanged depend on the fact that convolution with a two-valued risk takes only $O(t)$ steps.

Another question raised at this colloquium by W. S. MEIJER is how these computational techniques relate to the techniques of stochastic simulation widely used in practice. A rule of thumb is to use simulation only as a last resort and avoid it whenever the model considered admits more exact procedures. Indeed, suppose we estimate the stop-loss premium at a certain retention t by n pseudo-random draws from the distribution of S, or rather of $Z = (S-t)_+$. To achieve the same relative accuracy of 1% as in Table 1, we have to take n so large that

$$\frac{1}{\sqrt{n}} \frac{\sqrt{\mathrm{Var}[Z]}}{E[Z]} \approx 0.01. \tag{12}$$

For $t = 12 \approx E[S] + \sqrt{\mathrm{Var}[S]}$ and the above portfolio we obtained a value of $n \approx 185{,}000$, and for $t = 18 \approx E[S] + 2\sqrt{\mathrm{Var}[S]}$ we even have to generate some 400,000 replications of the portfolio, which amounts to about 300 million random drawings.

A procedure for obtaining quick first estimates we can recommend is to use the familiar Gamma approximation to the distribution of S (see, e.g. BOWERS *et al.*, 1987). The Gamma distribution admits an analytical expression for the associated stop-loss premiums. We found that the relative error in the stop-loss premium did not exceed 7.2% for retentions in the range $(0, E[S] + 3\sqrt{\mathrm{Var}[S]})$. The NP-approximation might be used too.

Instead of using convolution to obtain the exact distribution of the individual model one might use the recursive algorithm described in DE PRIL (1986). As was pointed out by KUON *et al.* (1987), his algorithm is very time-consuming. Worth mentioning too is the algorithm of KORNYA (1983) to approximate the distribution of S with a controllable error bound.

It should be noted that our method works best for portfolios of small to intermediate size, when the risk capitals are of the same order of magnitude as the retention. For really large portfolios, the error reduction is less spectacular, though still significant.

We think one of the main advantages of our approach is that it can be explained easily not only to clients and managers, but also to those members of actuarial departments who have not sufficiently kept up with the latest developments in risk theory.

REFERENCES

BERTRAM, J. (1981) Numerische Berechnung von Gesamtschadenverteilungen: *Blätter der deutsche Gesellschaft für Versicherungs-Mathematiker* XV-2, 175–194.

BOWERS, N. L., GERBER, H. U., HICKMAN, J. C., JONES, D. A. and NESBITT, C. J. (1987) *Actuarial Mathematics*. Society of Actuaries, Itasca, Ill.

DE PRIL, N. (1986) On the exact computation of the aggregate claims distribution in the individual life model. *ASTIN-Bulletin* **16**, 109–112.

GERBER, H. U. (1984) Error bounds for the compound Poisson approximation. *Insurance: Mathematics and Economics* **3**, 191–194.

GOOVAERTS, M. J., HAEZENDONCK, J. and DE VYLDER, F. (1984) *Insurance Premiums*. North-Holland, Amsterdam.

KAAS, R. (1987) *Bounds and Approximations for Some Risk Theoretical Quantities*. University of Amsterdam.

KORNYA, P. S. (1983) Distribution of aggregate claims in the individual risk theory model. *Society of Actuaries: Transactions* **35**, 823–858.

KUON, S., REICH, A. and REIMERS, L. (1987) Panjer vs Kornya vs De Pril: A comparison from a practical point of view. *XX ASTIN-Colloquium, Scheveningen.*

R. KAAS, A. E. VAN HEERWAARDEN, M. J. GOOVAERTS
Institute for Actuarial Science and Econometrics, Jodenbreestraat 23, NL-1011 NH, Amsterdam.

https://doi.org/10.2143/AST.18.1.2014963 Published online by Cambridge University Press