AN APPROACH TO THE ANALYSIS OF CLAIMS EXPERIENCE IN MOTOR LIABILITY EXCESS OF LOSS REINSURANCE

H. G. VERBEEK
Netherlands

Introduction

In this paper motor liability insurance is considered from the viewpoint of an excess of loss reinsurer.

The reinsurer pays portions of claim amounts in excess of a contractually agreed limit, further referred to as excess point.

Aspects of practical interest are the rating of new contracts and cost-projections on claims occurred in past years which have not yet been reported.

What causes a major difficulty to a rating approach is the timelag between the occurrence of a claim and its settlement.

During this timelag, which may extend to ten years or even longer, the size of a claim will usually increase considerably. This is due to such factors as inflation and more victim oriented legal procedures.

As a result of this phenomenon, a claim, initially not involving the reinsurer, may confront him many years later, when its size overtakes the excess point. Hence, each risk year produces a generation of claims stretching far into the future.

The problem facing the reinsurer trying to compute premiums is that he must look, say ten years into the future, on the basis of incomplete data.

The table below is a schematic presentation of the time history of a contract that started k years ago.

The experience is assumed to be expressed in numbers of claims exceeding a constant excess point.

Successive rows show the generations of excess claims down to the last completed risk year k.

In the pages that follow we will attempt to project expected numbers of excess claims into the future, using information as schematized in this table.

Table I

year occurred	year reported				
	1	2		$k-1$	k
1	n_{11}	n_{12}	————————	$n_{1,\,k-1}$	n_{1k}
2	n_{21}	n_{22}	————————	$n_{2,\,k-1}$	O
				O	O
k	n_{k1}	O		^	O

We will restrict our analysis to expected numbers and ignore the size aspect. If this succeeds, the step from expected numbers to expected claims cost does not present substantial difficulties.

The mean amounts of excess claims for fixed excess points empirically show a rather stable pattern, which is strictly true in the Pareto case.

Hence, to find estimates of the net claims cost we can multiply projected numbers by conservative estimates of mean excess claims.

General outline and assumptions used

In this section we shall state all the assumptions which we are going to use. These assumptions permit us to formulate a multiplicative table of parameters corresponding to the realisations shown in table I. Certain relations are then seen to exist between the parameters.

We shall make use of these relations to fit the entire collection of realisations in table I to the elements of the table of parameters. We shall do this by employing the Maximum likelihood principle, thus finding a number of equations solvable by a recursive procedure.

The solution takes the form of two different sets of parameters which are the basis for our projection.

We now list our assumptions:

1. The claim counting variable in the basic risk business is a Poisson distribution. There may be reasons to prefer a variable having a fluctuating basic probability structure. We have

abstained from such model in view of the generally small num-
bers of observations available for analysis.

2. Claim events that take place in a certain risk year have a
probability r_1 of being reported in the same year, a probability r_2
of being reported one year later and so on. All claims occurring
in a risk year will be assumed to have been reported at the end of
the k-th year. Hence the successive probabilities add up to unity
if summed over the k years.

3. To each of the k years covering our claims history we shall assign
a conditional claim size distribution function. This distribution
generates the size of any claim reported in that particular year
regardless of the risk year in which the generation started.
Hence, no size is attached to the claim event before it is reported.
We shall denote the sequence of distributions by $F_m(x)$ where the
index m runs from 1, the oldest year, to k, the year preceding the
present.

The analytical structure of these distributions will not be
specified.

No explicit assumption is needed in the context in which they are
used.

4. We shall suppose that the respective generations of claim
realisations have been suitably pre-manipulated so as to permit
them to be treated as coming from identical basic business.

Accordingly we assume that the k generations have been pro-
duced by Poisson variables all having the same parameter α.

5. No correlation between the r's and the claim size variables will be
assumed.

Multiplicative model of parameters

*The assumed Poisson distribution for the number of claim events
in the basic risks implies that all individual elements in the triangular
table I are also realisations of Poisson variables.*

This follows from the fact that the counting variable of an indi-
vidual element can be obtained by compounding the basic Poisson
variable twice in succession with a binomial distribution. Com-
pounding Poisson variables with binomial distributions any number
of times will reproduce Poisson variables with a change in the
parameter.

Let us now find the parameters of the entire collection of Poisson variables corresponding to the elements in table I.

The following relation exists between the parameter α in the basic business and that of the variable counting the claims exceeding the excess point x_0

$$\lambda_1 = \alpha[1 - F_1(x_0)].$$

This refers to the oldest risk year. A claim has a probability r_1 of being reported in its year of occurrence. Hence the Poisson parameter corresponding to the element n_{11} in table I, is $r_1\lambda_1$.

The excess claims in the second risk year have the parameter

$$\lambda_2 = \alpha[1 - F_2(x_0)].$$

Hence the parameter $r_1\lambda_2$ corresponds to the element n_{21}.

In view of assumption 3 in the previous section, the claim size variables in the second year of the first generation and the first year of the second generation are identical.

It follows that n_{12} has the parameter $r_2\lambda_2$.

Repeating this argumentation k times we can write down the entire table of parameters corresponding to the realisations in table I:

Table II

year occurred	year reported				
	1	2	————————	k—1	k
1	$r_1\lambda_1$	$r_2\lambda_2$	————————	$r_{k-1}\lambda_{k-1}$	$r_k\lambda_k$
2	$r_1\lambda_2$	$r_2\lambda_3$	————————	$r_{k-1}\lambda_k$	O
			$r_j\lambda_j$	O	O
				O	O
k—1	$r_1\lambda_{k-1}$	$r_2\lambda_k$	O	O	O
k	$r_1\lambda_k$	O	O	O	O

We should note that in successive columns the r's carry the same index number, while in successive diagonals the lambdas have the same index.

Estimation of the parameters

In this section we shall obtain a recursive computational proce-dure for finding ML estimates of the two sequences r_m and λ_m.

We define:

$$P[n_{11} \ .. \ n_{1k}, \ .., \ n_{k1} \mid r_1 \ ..r_k, \ \lambda_1 \ ..\lambda_k] =$$

$$= \prod_{i=1}^{k} \ \prod_{j=1}^{k-i+1} \ (r_j\lambda_{i+j-1})^{n_{ij}} \ exp - r_j\lambda_{i+j-1} \mid n_{ij}!$$

We further define the loglikelihood function

$$L = ln \ P[n_{11}..n_{1k}, \ ..n_{k1} \mid r_1..r_k, \lambda_1..\lambda_k]$$

Hence

$$L = - \ \sum_{i=1}^{k} \ \sum_{j=1}^{k-i+1} \ r_j\lambda_{i+j-1} + \sum_{i=1}^{k} \ \sum_{j=1}^{k-i+1} \ n_{ij} \ ln \ r_j\lambda_{i+j-1} -$$

$$- \ \sum_{i=1}^{k} \ \sum_{j=1}^{k-i+1} \ ln \ n_{ij}!$$

Setting partial derivatives with respect to the k lambdas and k r's to zero and ensuring the latter to sum to unity by introducing the multiplier s, we find these $2 \ k$ equations:

$$\frac{\partial L}{\partial r_1} = - \lambda_1 - \lambda_2 - ... - \lambda_k + \frac{n_{11}}{r_1} + \frac{n_{21}}{r_1} +$$

$$+ ... + \frac{n_{k-1.1}}{r_1} + \frac{n_{k1}}{r_1} + s = 0$$

$$\frac{\partial L}{\partial r_2} = - \lambda_2 - ... - \lambda_k + \frac{n_{12}}{r_2} + \frac{n_{22}}{r_2} + ... + \frac{n_{k-1.2}}{r_2} + s = 0$$

. . . .

$$\frac{\partial L}{\partial r_k} = - \lambda_k + \frac{n_{1k}}{r_k} + s = 0$$

$$\frac{\partial L}{\partial \lambda_1} = - r_1 + \frac{n_{11}}{\lambda_1} = 0$$

$$\frac{\partial L}{\partial \lambda_2} = - r_2 - r_1 + \frac{n_{21}}{\lambda_2} + \frac{n_{12}}{\lambda_2} = 0$$

. . . .

$$\frac{\partial L}{\partial \lambda_k} = - r_k - ... - r_2 - r_1 + \frac{n_{k1}}{\lambda_k} + \frac{n_{k-1.2}}{\lambda_k} + ... + \frac{n_{1k}}{\lambda k} = 0$$

Multiplying the above equations respectively by r_1, r_2 ... r_k λ_1, λ_2 ... λ_k and adding up the resulting set of equations, we find that the multiplier s equals zero.

We now introduce the notation:

$$v_j = \sum_{i=1}^{k+1-j} n_{ij}$$

$$d_j = \sum_{i=1}^{j} n_{i \cdot j - i + 1}$$

denoting respectively the sums of successive columns and diagonals of the observed claims in table I.

After some manipulation and reordering we can write the equations as follows:

$$\lambda_k = d_k \qquad (1)$$
$$r_k \lambda_k = v_k \qquad (2)$$
$$\lambda_{k-1} - r_k \lambda_{k-1} = d_{k-1} \qquad (3)$$
$$r_{k-1} \lambda_{k-1} + r_{k-1} \lambda_k = v_{k-1} \qquad (4)$$
$$\cdots \cdots \cdots \cdots \qquad \vdots$$
$$\lambda_1 - \lambda r_2 - \ldots r_{k-1} \lambda_1 - r_k \lambda_1 = d_1 \qquad (2k-1)$$
$$r_1 \lambda_1 + r_1 \lambda_2 + \ldots r_1 \lambda_{k-1} + r_1 \lambda_k = v_1 \qquad (2k)$$

This way of writing immediately suggests a simple recursive procedure for obtaining its solution. As d and v on the right hand side are known quantities, we note that λ_k is already solved and equals d_k as equation 1 states.

Proceeding in this manner we find numerial values for the two sequences r_m and λ_m m, assuming the integers 1 through k.

Projection of expected numbers of claims

The two sequences extracted from the data in table I by means of the procedure derived in the previous section are the basic ingredients for this projection.

For this purpose we use the sequence of the r's in unaltered form.

As regards the lambda sequence we observe that it will reflect the claims inflation process acting on the claim size variable in the k prior years.

Hence their values generally show a steeply increasing trend and

for projection purposes we shall need k new values obtained by extrapolating k yearly periods into the future.

Empirical data suggest that extrapolation by means of exponential curves may be realistic.

We shall not discuss this aspect, but just assume that we have accomplished finding k further lambdas.

We are then in a position to extend table II as follows:

Table III

year occurred	year reported					
	1	2	——	$k-2$	$k-1$	k
1	$r_1\lambda_1$	$r_2\lambda_2$	——	$r_{k-2}\lambda_{k-2}$	$r_{k-1}\lambda_{k-1}$	$r_k\lambda_k$
2	$r_1\lambda_2$	$r_2\lambda_3$	——	$r_{k-2}\lambda_{k-1}$	$r_{k-1}\lambda_k$	$(r_k\lambda_{k+1})$
3	$r_1\lambda_3$	$r_2\lambda_4$	——	$r_{k-2}\lambda_k$	$(r_{k-1}\lambda_{k+1})$	$(r_k\lambda_{k+2})$
k	$r_1\lambda_k$	$(r_2\lambda_{k+1})$	——	$(r_{k-2}\lambda_{2k-3})$	$(r_{k-1}\lambda_{2k-2})$	$(r_k\lambda_{2k-1})$
$k+1$	$(r_1\lambda_{k+1})$	$(r_2\lambda_{k+2})$	——	$(r_{k-2}\lambda_{2k-2})$	$(r_{k-1}\lambda_{2k-1})$	$(r_k\lambda_{2k})$

The extrapolated parameters, referring to future years, are the ones shown between brackets.

Thus we have extrapolated the successive generations of expected claims and also projected the expected claims for the next risk year, as shown in the bottom row of this table. By adding the elements of this row we find the expected number of claims generated by the next risk year $k+1$.

Expected "run-off" on prior risk years is obtained by summing the bracketed elements in the higher rows.

Conclusion

The approach outlined in the preceding pages hinges on the extrapolation of a time series k years into the future, where k might be a number in the order of 10.

14

Without wishing to minimise the pretentiousness of such an aim, we would like to stress that it is inherent in any rating approach to the type of business we have discussed, although it may not always show up explicitly.

In our view, a method such as presented here has the advantage of making explicit the quantities that have to be used for extrapolation.