Insurance Liabilities

Introduction

Insurance companies assume risks of other entities (individuals and companies) in return for payments of premium. While premium is normally paid upfront, before the inception of an insurance contract, loss payments by the insurance company will not take place until after an insured event occurs and is reported. Even when the insured event is reported, the exact magnitude of the loss payments due to the insured ("ultimate loss") may not be known for quite some time. The determination of this amount (known as "adjustment process") may be lengthy and involve litigation. As a result, the loss payments attributable to any given policy may occur over a period of time even after the coverage period ends, potentially in the course of many years (e.g., over a lifetime of an insured individual)—a process known as "loss development." Thus, loss payments become a long-term liability for the insurance company and give rise to uncertain future cash flows.

Insurance companies are required to establish reserves, i.e., money set aside, for the future loss payments. The reserves consist of case reserves, as set by the case adjusters, bulk reserves—these judgmental adjustments to case reserves are established by actuaries on an aggregate basis, and Incurred But Not Reported (IBNR) reserves—the latter being an actuarial allowance for yet unknown events. The sum of already paid losses and reserves for future payments is referred to as incurred losses.

The insurance risk modeling in ERM requires analysis and simulation of cash flows stemming from claim payments. From the accounting point of view, insurance liabilities are represented by the incurred losses, while the risk management perspective requires analysis of the uncertainty in the actual cash flows which are represented by paid losses. A change in the value of the incurred losses reflects a change in the internal estimate of the unpaid loss. A measure of variability of such estimates would not be representative of the intrinsic fluctuations in the paid losses, and, therefore, would not provide a measure of the risk of the associated cash flows. Also, the current practice of reserve setting is not based on statistically sound analysis of claim data and is subject to varying company policies. As a result, in ERM we concentrate on the development and fluctuations of the paid losses. The incurred loss data may be used as a supplement in order to estimate the development beyond the period covered by the paid loss data on Schedule P (10 years).

Loss Triangles and Link Ratios

Historical loss data are conveniently organized in the form of triangles.

Table 2. Cumulative Paid Losses

Accident Year	Development Year 0	1	2	3	4	5	6	7	8	9
1981	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
1982	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	
1983	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466		
1984	5,655	11,555	15,766	21,266	23,425	26,083	27,067			
1985	1,092	9,565	15,836	22,169	25,955	26,180				
1986	1,513	6,445	11,702	12,935	15,852					
1987	557	4,020	10,946	12,314						
1988	1,351	6,947	13,112							
1989	3,133	5,395								
1990	2,063									

The actuarial technique routinely used to analyze loss triangles is the "chain ladder link-ratios" method. In this approach, one finds average ratios of cumulative payments in each pair of adjacent development years. For example, the average ratio of payments in column 1 to same accident year payments in column 0 (averaging could be straight, dollar weighted, "excluding high-low," etc.) would give the development factor between "ages" 0 and 1. From a statistical point of view, this technique is based on the assumption that a cumulative payment in one development year is a predictor of the incremental payment in the next year, which means, in particular, that incremental payments of the same accident year are not independent. On the contrary, statistical loss data analysis in the literature [14] and Seabury's own research show that for most real loss development arrays, the incremental payments are independent random variables, and, therefore, standard development factor (link-ratio) techniques are inappropriate.

This finding has important implications for both reserve setting (estimates of the mean) and risk analysis (estimates of uncertainty). First of all, it can be proven that the "chain ladder link-ratios" method produces upward-biased reserve estimates [18]. The consequence for the reserve uncertainty is rather subtle: under the assumptions behind the "chain ladder" approach, the probability distribution of the incremental payments would necessarily widen as a function of the development "age" (similarly to how the volatility of a common stock scales with time). The later development years would then make a bigger contribution to the reserve risk than their contribution to the reserve itself. On the other hand, under the assumption of statistically independent incremental payments, there is no inherent widening of the distribution with time—although the volatility may still increase late in the development because of the low number of claims remaining open.

Statistical Modeling Framework

It is clear from the above discussion that the natural way to analyze and model paid loss data is on the incremental rather than cumulative basis.

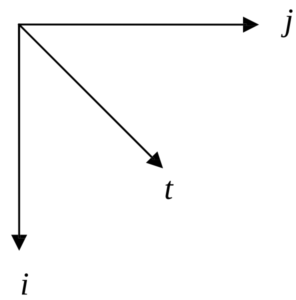
Table 3. Incremental Paid Losses

		Development Year									
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Accident Year	**1981**	5,012	3,257	2,638	898	1,734	2,642	1,828	599	54	172
	1982	106	4,179	1,111	5,270	3,116	1,817	(103)	673	535	
	1983	3,410	5,582	4,881	2,268	2,594	3,479	649	603		
	1984	5,655	5,900	4,211	5,500	2,159	2,658	984			
	1985	1,092	8,473	6,271	6,333	3,786	225				
	1986	1,513	4,932	5,257	1,233	2,917					
	1987	557	3,463	6,926	1,368						
	1988	1,351	5,596	6,165							
	1989	3,133	2,262								
	1990	2,063									

The incremental loss development array is subject to multiple trends acting in different directions: the natural loss development within each accident year (horizontal direction in Table 3), the change in exposure from one accident year to another (vertical direction), and inflation—since calendar years are represented by diagonals in Table 3, this last type of trend acts from one diagonal to another. Seabury ERM introduces a modeling framework for the incremental paid loss data that is able to capture all these trends.

The model parameterizes the trends in each of the three directions—development years, accident years, and payment/calendar years. In what follows, the development years are denoted by *j*, $j = 0,1,2,...,s-1$; accident years by *i*, $i = 1,2,...,s$; and payment years by *t*, $t = 1,2,...,s$.

Figure 3. Trends in Loss Data



The payment year variable t can be expressed as $t = i + j$. This relationship implies that both development and accident year trends are projected onto payment year trends.

In its most general form, the model can be written as:

$$P_{i,j} = \exp\left(\alpha_i + \sum_{k=1}^{j} \gamma_k + \sum_{t=1}^{i+j} \iota_t\right) + \varepsilon_{i,j} \qquad (3.2.14)$$

Here, $P_{i,j}$ is the incremental payment amount for accident year i and development year j—the payment takes place in payment year $i + j$; $\varepsilon_{i,j}$ is zero-mean random error (not necessarily normally distributed).

The parameters α_i in the accident year direction determine the level from year to year; often the level (after adjusting for exposures) shows little change over many years, requiring only a few parameters. The parameters γ_k in the development year direction represent the trend from one development year to the next. This trend is often linear (on the log scale) across many of the later development years, often requiring only one parameter to describe the tail of the data. The parameters ι_t in the payment year direction describe the trend from payment year to payment year. If the original data are inflation adjusted, the payment year parameters represent superimposed (social) inflation, which may be stable for many years. This is determined in the analysis. Consequently, the (optimal) identified model for a particular loss development array is likely to be parsimonious. This allows us to have a clearer picture of what is happening in the incremental loss process.

The distribution of variables $P_{i,j}$ in the model (3.2.14) is determined by the distribution of the random term $\varepsilon_{i,j}$. It is well known that insurance losses exhibit long-tail distributions. Accordingly, the natural candidates would be such distributions as lognormal and fat-tail power-law family. Note that the random term in (3.2.14) is additive, so that the incremental payment might become negative. This is actually advantageous, because the real insurance data, due to the practices of salvage and subrogation, do exhibit this kind of anomaly (as evident from the data in Table 3).

Implementation and Calibration

The model defined by Eq. (3.2.14) is nonlinear and cannot be reduced to a linear regression. Accordingly, the parameters in (3.2.14) should be obtained through a nonlinear regression analysis (general nonlinear minimization). In the current ERM version, we accomplish a less ambitious goal and approximate Eq. (3.2.14) by a linear model on a log-scale:

$$y(i, j) \equiv \ln(P_{i,j}) = \alpha_i + \sum_{k=1}^{j} \gamma_k + \sum_{t=1}^{i+j} \iota_t + \varepsilon_{i,j}. \qquad (3.2.15)$$

Note that Equation (3.2.15) does not allow negative incremental payments; therefore, we are forced to drop negative data points from the historical data.

The original equation (3.2.14) has an inherent advantage missing from (3.2.15): any scheme used for minimization the random terms in (3.2.14) would be dominated by the large dollar amounts (usually, early development years), precisely the ones that contribute most into the risk of the future payments. In the log-scale model (3.2.15), if one were to perform an ordinary least-squares regression (OLS), the parameters could be driven by the large variations in the small dollar amounts of the old development age. Therefore, our approximation of (3.2.14) by (3.2.15) makes it necessary to perform a weighted least-squares regression in (3.2.15), the weights being the payments $P_{i,j}$ themselves.

The random term in (3.2.15) is assumed to be normal, so that the incremental payments follow a lognormal distribution. We do not assume, however, that the error terms $\varepsilon_{i,j}$ come from a single distribution. Instead, we model $\varepsilon_{i,j}$ as $N(0, \sigma(P_{i,j}))$, where the standard deviation $\sigma(P_{i,j})$ becomes a function of the incremental payment $P_{i,j}$. The chosen scheme of weighted least-square regression implies that the variances σ_j^2 are inversely proportional to $P_{i,j}$, the proportionality factor being determined by the regression. We, however, make an additional step and assume that $\sigma(P_{i,j})$ is a general non-increasing function of $P_{i,j}$. This assumption is motivated by our research that shows that the payment volatility (on log-scale) remains practically constant over a wide range of payment magnitude, but once the payments drop below a certain threshold, the volatility begins to increase. Therefore, having done the weighted regression as discussed above, we perform an additional non-parametric fit of the squared residuals with a monotonous (more specifically, non-increasing) function of $P_{i,j}$; this least-square fit is weighted, $P_{i,j}$ being the weights again. The resulting function provides the required estimate of variance $\sigma^2(P_{i,j})$:

$$\begin{gathered} \varepsilon_{i,j}^2 = \sigma^2(P_{i,j}) + \upsilon_{i,j}, \\ \sigma^2(P) \le \sigma^2(P') \text{ for } P > P'. \end{gathered} \tag{3.2.16}$$

In Eq. (3.2.16), $\upsilon_{i,j}$ is an error term; if $\varepsilon_{i,j}^2$ happens to be non-increasing as a function of $P_{i,j}$, then all $\upsilon_{i,j}$ would be equal to 0.

Theoretically speaking, Eq. (3.2.15) is overparameterized by one parameter and exhibits "perfect" multicolinearity: it remains invariant under the transformation

$$\iota_i = \iota_i - I,\ \gamma_i = \gamma_i + I,\ \alpha_i = \alpha_i + iI, \tag{3.2.17}$$

where I is arbitrary. As a result, the mean level of the inflation over all payment years cannot be determined (it is included in the development factors γ_i); only the deviations from the mean level can be calculated. This situation would change if we knew the accident year exposures so that we could normalize the incremental payments and set all α_i to be equal to each other. Unfortunately, we are unaware of any publicly available information about the exposures or suitable proxies. Accident year premiums, in particular, cannot serve as good proxies for exposures due to the varying premium rates over a business cycle. In the absence of the company's exposure data, we get rid of multi-colinearity by setting

$$\iota_1 = 0; \tag{3.2.18}$$

this means that the rest of the ι parameters measure the *difference* between inflation in their respective years and that in year 1.

Even though Eq. (3.2.15) with restriction (3.2.18) exhibits no theoretical ("perfect") multicolinearity, it still has too many fitting parameters to be of practical use in forecasting. In particular, it is reasonable to require that the development pattern captured by γ_i represents a smooth curve. Yet, as long as all γ_i are estimated independently, this condition cannot be guaranteed. We will both achieve a parsimonious model suitable for forecasting and ensure smooth development patterns if we significantly reduce the number of parameters in (3.2.15) by setting some of them equal to each other or to zero. For instance, we might require that, for a given line of business, $\gamma_1 = \gamma_2$, $\gamma_3 = \gamma_4 = 0$, $\gamma_5 = \gamma_6 = \cdots = \gamma_{s-1}$, and $\iota_2 = \iota_3 = \cdots = \iota_s$. We refer to such specifications as a "model structure". We determine the proper model structure for each line of business based on the industry-wide experience. The concrete values of the parameters remaining in the model and their standard deviations can then be determined for each company; these values can possibly be credibility weighted with the results of the industry cross-sectional analysis.

Because parameters γ_i and ι_i represent trends that will accumulate when we set some of them equal, we need to account for a "model risk" by making the parameters themselves normally distributed random variables. In the example above, in each scenario we would need to simulate two γ parameters, one ι parameter, and only afterwards all the random terms $\varepsilon_{i,j}$. The mean values and the standard deviations for this simulation are the output of the regression (3.2.15). Note that parameters α_i are not trends; these parameters should simply be set to their regression estimates rather than simulated.

Note that parameterization is such that in any forecast (simulation) we set the future values of parameters α_i, γ_i, and ι_i to be the same as in the last year available from the regression.

In order to account for correlations between the insurance lines and the correlations between assets and liabilities, we regress the normalized error terms $\varepsilon_{i,j} / \sigma(P_{i,j})$ against the assets' Principal Components and then perform the Principal Component Analysis on the residuals of that regression. As the end result of this analysis, the random term in (3.2.15) is represented as a linear combination of Principal Components of both assets, PC^{asset}, and liabilities, $PC^{\text{liability}}$, plus an idiosyncratic random term:

$$\varepsilon_{i,j} / \sigma(P_{i,j}) = \sum_{m=1}^{M} C^{\text{asset}}_{i+j,m} \times PC^{\text{asset}}_{i+j,m} + \sum_{n=1}^{N} C^{\text{liability}}_{i+j,n} \times PC^{\text{liability}}_{i+j,n} + \varepsilon_{i,j}, \qquad (3.2.19)$$

where $\varepsilon_{i,j}$ are i.i.d. $N(0,\sigma)$, M is the number of principal factors on the asset side and N is the number of factors on the liability side.

Underwriting Risk (New Business)

The uncertainty of the loss payments associated with the prior accident years (Old Business)—usually referred to as reserve risk—is captured by the model introduced in the previous sections. The same model is used in ERM to describe the loss uncertainty of the new business that will be written between today and the horizon. This future business has one more component that we have not covered yet—the uncertainty of the collected premium. Insurance premium collected per unit of risk (premium rate) is subject to market forces; as a result, the insurance industry has gone through well documented business cycles of "hard" and "soft" markets. Even though the uncertainty of the rate forecast between today and the one-year horizon is usually much less than the rate variations over an entire business cycle, the random element in the future premium cannot be removed.

Historically, the insurance business cycles have not coincided with the economic cycles. Nonetheless, the existence of a relationship between premium rates and the economic environment, in particular, the interest rates, is well known [19]–[22]. Since most collected premiums get invested into fixed income instruments, a hike in the interest rates will result in greater investment income associated with the new policies. In a competitive market environment, this will result in additional pressure towards lowering the premium charged per unit of risk. This negative correlation between interest rates and premium rates is more pronounced in casualty lines where claims are settled long after the premiums are collected than for short-tailed property lines.

Within Seabury ERM, premium rates are modeled as log-normally distributed variables; the correlations between rates in different lines and between rates and the financial risk factors are estimated based on the industry-wide data.

Catastrophe Risk

Losses due to natural catastrophes (CAT losses) are rare events with very high impact. The CAT loss distributions exhibit highly non-normal (fat-tailed) behavior. The company specific distributions depend on the geography of the insured properties and businesses. As a result, such losses are notoriously difficult to model reliably, and the distributions are provided by just a few vendors who specialize in CAT risk. When the distribution is available however, it is relatively easy to include the CAT risk into the overall framework of ERM. Within the simulation model, we need to add catastrophic losses drawn from the CAT distribution to the losses generated in our regular model (3.2.15). This procedure will result in adjusted α_i (and possibly γ_k, if we assume that CAT loss development is different from that of the regular losses) in such catastrophic scenarios. In addition, the CAT scenarios will have significant reinsurance receivables, hence additional credit risk (see below). The real difficulty with catastrophic risk within a Monte Carlo approach stems from the necessity to estimate the risk measures dominated by rare events. In Appendix D:, we show how application of importance sampling and robust estimators can overcome these problems.

Reinsurance Receivables: Credit Risk

Reinsurance receivables, in particular those due the ceded catastrophe risk, constitute the greatest portion of the credit risk faced by an insurance company. Seabury ERM breaks down reinsurance receivables by the reinsurer and then proxies the credit risk of the receivables by the credit risk of a portfolio of risky bonds. The credit rating of the reinsurer gets assigned to the bond, and the amount of each receivable becomes the bond face value.