Section C
PRÉCIS OF PAPERS IN SECTION D

This section provides a précis of each paper included in Section D of Volume 2. The intention is to give a brief summary of the paper, a description of the reserving model on which the paper is based, and a few observations about the applicability of the model. The précis also deal with what data are required and what level of statistical and computational ability is needed, and offer some thoughts on the strengths and weaknesses of the model.

The numerical heading given to each paper refers to the relevant sub-section within Section D of Volume 2, where the full text of that paper is to be found.

[C1]
THE CHAIN LADDER TECHNIQUE — A STOCHASTIC MODEL
Contributed by B Zehnwirth
(9 pages, see [D1])

Summary

The chain ladder technique is one of the oldest actuarial techniques to be applied widely for estimating loss reserves. It appears intuitively natural and was for some time widely regarded as being based on a non-stochastic model: that is, a model which is deterministic and accordingly does not include a random component.

The paper demonstrates the intimate connection between the chain ladder technique and a two-way analysis of variance model applied to the logarithms of the incremental paid losses. Recognition of this connection reveals the merits and defects of the chain ladder technique more clearly.

Description of the model

The basic model is as follows:

$$\text{Log}(P_{ij}) = Y_{ij} = a_i + b_j + e_{ij}$$ (e_{ij} are independent identically distributed normal error terms)

where Pij are the incremental payments for accident year i, development period j. This model implies that each incremental paid loss, P_{ij}, has a lognormal distribution. The model is fitted by least squares regression or by the application of an algorithm (known as "Expectation-Maximisation", or E–M) for the corresponding two-way analysis of variance.

General comments

The basic statistical chain ladder is generally considered to be over-parameterised, and can be criticised for not including any calendar year effects as part of the model. It is, however, a powerful diagnostic tool for exploring payment/calendar year trends.

It can also form a basis for more sophisticated models, which are not so heavily parameterised, and can include calendar year effects and incorporate additional information into the reserving process.

Some of these extensions to the basic statistical chain-ladder are described in the paper by S Christofides in Section D5, which is also summarised in section C5 of this Volume. Further extensions to the basic model are also described in another paper by B Zehnwirth, which is summarised in section E of this Volume. <>

[C2]
EXPONENTIAL RUN-OFF
Contributed by B Ajne
(11 pages, see [D2])

Summary

The paper describes a model of the exponential run-off of the incremental payments after the first few development years, based on observations for personal injuries in motor insurance. A brief example is provided, as well as possible adjustments for the effect of inflation.

Description of the model

The basic model is as follows:

$$C_{ij} = q_i . C_{ij-1}, \; j = \alpha + 1, \; \dots \; , \; A - 1$$
$$C_{ij} = 0, \; j \geq A$$

where C_{ij} are the incremental payments for accident year i, development year j.

The q_i are estimated using an algorithm to maximise a likelihood function. The likelihood function is found assuming that:

$$P(X_i = j) = \beta_i . \, q_i^{j-a}$$

where X_i is the number of years between occurrence and settlement for claims occurred in year i ($X_i > \alpha$).

Each amount of claim payment is assumed to be independent of all others.

The reserves are calculated for each year of origin by multiplying the claims paid to date by a ratio based on q_i.

General comments

The concept of exponential run-off is particularly useful for long-tail lines of business. The method is fairly simple mathematically, and the only data required are incremental payments. Provided an equation "solver" is available, it can be programmed and used very easily in any spreadsheet.

The assumptions made by the model are very strong, and it is doubtful whether assumption (2.2) in the paper can ever be properly met in practice. Since this is central to the exponential run-off assumed by the model, it casts some doubt on the validity of the estimates, although the results of the model may still be useful.

The author suggests that an examination of the residuals would be "useful". In fact, this may more properly be described as "essential".

<>

[C3.a]
THE CURVE FITTING METHOD
Contributed by S Benjamin and L M Eagles
(9 pages, see [D3.a])

Summary

The paper describes the use of curve fitting to the progression of paid and incurred loss ratios. A Craighead curve (otherwise known as a Weibull distribution) is suggested with up to 3 parameters. A least squares method is proposed for the curve fitting, with graphical examples. The use of curve-fitting is compared with other methods.

Description of the model

The progression of loss ratios is considered by dividing the cumulative claims to date by the estimated ultimate premiums for each year of origin. For each year a Craighead curve y(t) is then fitted to the loss-ratios at time t, as follows:

$$y(t) = A(1-e^{-(\frac{t}{b})^c})$$

where A is the estimated ultimate ratio, and b and c are parameters. b and c are fitted to all the years of origin, and A varies for each year. For data consisting of a mixture of short tail and long tail business, a double Craighead curve is proposed.

The fitting method is to minimise by iterations $\sum w(t).(y(t)-y_{obs}(t))^2$, where w(t) is the weighting and $y_{obs}(t)$ the observed loss ratio. The use of w(t) allows outliers to be excluded, or the curve to be forced through the most recent data point. Two methods of minimisation by iterations are mentioned, although they are not spelt out in any detail.

General comments

The model was originally intended to be applied to London Market business, but can be used for any type of business, provided that the run-off follows a Craighead curve.

There is no particular reason why the progression of the loss-ratios beyond the data should follow any particular type of curve, so the use of the model to extend the curve beyond the observed data should be treated with some caution.

The data required are paid and incurred claims, together with premiums or some appropriate measure of exposure.

Ideally, the simple visual examination of estimated relative to observed data suggested in the paper should be supplemented by a more formal statistical check of the goodness of fit of the model. The χ^2 test, which can be performed quite easily, is suitable for this purpose.

The paper only requires a few mathematical skills, although implementing the iterative techniques requires a certain level of statistical and computational ability. The model is non-linear with 3 parameters, so it cannot easily be fitted into a formal spreadsheet. However, the existence of equation "solvers" in many spreadsheets may provide a pragmatic solution to the problem of fitting the curve.

[C3.b]
THE REGRESSION METHOD
Contributed by S Benjamin and L M Eagles
(8 pages, see [D3.b])

Summary

The paper describes and illustrates a method of refining the ultimate loss ratios found by some other method (for example the curve fitting method). A suggestion is given as to how, using graphical means, one can assess likely upper and lower bounds for the estimates of ultimate loss-ratios.

Description of the model

Ultimate loss ratios need to be estimated prior to applying this method. For each year of origin and development year, IBNR loss ratios are determined by:

IBNR loss-ratio(development year t) = Ultimate loss ratio – Incurred loss-ratio(development year t)

For a given development year, a regression line is estimated, based on all the years of account, as:

IBNR loss-ratio(development year t) = a × Incurred loss ratio (development year t) + b

for some fixed a and b.

Reserves are then calculated from this formula.

The regression line can actually be reformulated in terms of credibility:

$$\text{Future claims} = Z \times \frac{a}{Z} \times \text{claims to date} + (1-Z) \times \frac{b}{1-Z} \times \text{premiums}$$

Giving no credibility to the premiums, by regressing with b set equal to zero, is equivalent to using a traditional chain-ladder method.

General comments

The method can be used for any type of business, provided that the ultimate loss ratios are already estimated. It is very easy to implement in a spreadsheet. As the method is based on regression, standard errors of the estimates of the parameters can easily be determined by statistical techniques, as well as by the graphical method suggested.

The paper is easily understandable, but the reader has to be familiar with the principle of regression. The user of the method is given many suggestions as to how the method can be presented simply to, for example, an underwriter.

[C4]
REID'S METHOD
Contributed by D H Reid
(20 pages, see [D4])

Summary

This paper describes a class of models, set out in a series of papers written by the author. It considers the case where a relatively complete set of information on individual claims is available, and where past years' claim patterns may be expected to give insight into the more recent years.

This approach was first described by the author in a paper in the Journal of the Institute of Actuaries, "Claims Reserves in General Insurance", Volume 105, Part III (1978). Subsequent papers in this series are set out at the end of this précis.

The method provides the means by which to establish a probability distribution of claim reserves. Emphasis is given to the process of fitting and re-fitting models as necessary, prior to the extrapolation process. The model is very flexible and allows for the tendency of larger claims to take longer to settle, the proportion of nil claims to vary from one origin year to another, the rate of claim settlement to vary both across and within origin years, and for the effect of inflation on claim costs.

Description of the model

The basic model for the claims arising in a particular origin year consists of a number of components:

1. An underlying bivariate distribution of the cost of positive claims by claim settlement amount and development time.

2. A comparable univariate distribution by development time for nil claims.

3. The proportion of all claims represented by nil settlements.

4. A functional transformation of the settlement time axis from fixed calendar period time to real settlement time (i.e. operational time) represented by the underlying distributions (components 1 and 2).

5. A series of claim cost scale parameters intended to represent cost levels for fixed intervals of operational time, relative to the underlying bivariate distribution.

6. A separate treatment of the largest group of claims by size.

The model assumes that the ordering of claim settlements is not affected by the rate of settlement, and that this ordering is represented by the underlying distributions of components 1 and 2. Recent years' data are fitted to these underlying distributions and components 3–6 are estimated. In conjunction with appropriate assumptions, the fitted parameters are used to extrapolate the incomplete portion of recent years' settlements, from which reserves and reserving distributions are derived.

General comments

The methodology is intended for situations where a detailed analysis of claims behaviour can be obtained. It is likely to be of most relevance for Direct business, where data on amounts and numbers of claims are available by claim size. The method is quite complex, and requires a considerable amount of effort to implement in its fullest form.

The method is very flexible, and can be adapted to embrace more (or less) elaborate models of claim development. It can also be used to help develop sub-models relating claim cost movements to extraneous variables, such as inflation.

A considerable amount of statistical knowledge is required. Some steps in the process require the user to be able to use numerical techniques, for example finding parameters that maximise a likelihood function, without setting out explicitly how this may be achieved.

The original 1978 paper introduced the idea of Operational Time to the context of claim reserving. Although the detailed modelling of the underlying bivariate distribution has now been much simplified in the light of experience, the remainder of the original approach remains valid. The papers in the series, all by D H Reid, are set out below:

1. Claim reserves in general insurance, *Journal of the Institute of Actuaries*, 105, pp 211–296, 1978.

2. Reserves for outstanding claims in non-life insurance, *Transactions of the International Congress of Actuaries*, Zurich and Lausanne, 2, pp 229–241, 1980.

3. A method of estimating outstanding claims in motor insurance with applications to experience rating, *Cahiers du CERO*, Bruxelles, 23, pp 275–289, 1981.

4. Discussion of methods of claim reserving in non-life insurance, *Insurance: Mathematics and Economics* 5, pp 45–56, North Holland, Amsterdam, 1986.

5. Operational time and a fundamental problem of insurance in a data-rich environment, *Applied Stochastic Models and Data Analysis*, 1995, pp 257–269.

[C5]
REGRESSION MODELS BASED ON LOG-INCREMENTAL PAYMENTS
Contributed by S Christofides
(54 pages, see [D5])

Summary

The paper describes a statistical reserving model, based on the logs of the incremental payments. It shows, by a simple example, how such models can be fitted and results derived using a spreadsheet. A more realistic example is then considered, and refinements of the model are described. Because it is a statistical model, standard errors (a measure of the variability of the estimate) for the future incremental payments can be calculated and statistical techniques used to test the fit of the model.

Description of the model

The basic model is as follows:

$$\text{Log}(P_{ij}) = Y_{ij} = a_i + b_j + e_{ij} \quad (e_{ij} \text{ are independent identically distributed normal error terms})$$

where P_{ij} are the incremental payments for accident year i, development period j.

The a_i and b_j are fitted by regression, which can be done automatically in most spreadsheets. The future payments and standard errors are then calculated using matrix manipulation.

Refinements to the basic model are illustrated, including fitting a curve for the development parameters, and adjusting for claims volume and inflation. Models based on curves for the development factors can be useful for estimating tails, as they can be used to project beyond the existing data set.

General comments

The method is of general use and is not restricted to any particular class of business. The only data required are incremental payments. The basic method can be easily programmed in any spreadsheet, although the matrix manipulation necessary to calculate the standard errors may be somewhat time-consuming. Once the basic model has been set up in a spreadsheet, however, the model can be fitted and future payments predicted with very little time or effort for any data set of the same size.

The method does not work for negative incremental payments. There is also a limit to the number of future payments (n) that can be predicted in a spreadsheet, to the largest nxn matrix that a given spreadsheet package can manipulate.

The paper requires a basic level of statistical knowledge. Familiarity with matrix manipulation and regression in a spreadsheet would be helpful, although the worked example sets out all the steps clearly enough for this not to be a necessity.

◇

[C6]
MEASURING THE VARIABILITY OF CHAIN LADDER RESERVE ESTIMATES
Contributed by T Mack
(65 pages, see [D6])

Summary

The author has written a series of papers on the subject of the variability of chain-ladder estimates, most notably the CAS prize-winning paper "Measuring The Variability Of Chain Ladder Reserve Estimates". The paper in Section D6 is a reproduction of this paper with some modifications and additions.

The paper derives a formula for the standard error of chain-ladder reserve estimates without assuming any specific claim amount distribution function. For ease of reference, these techniques are described as the "Distribution-free approach".

Description of the model

The foundation of the Distribution-free approach is the observation of three main assumptions which are shown to underlie traditional chain-ladder techniques. These are:

(i) $E(C_{i,k+1}|C_{i1}, ..., C_{ik}) = C_{ik}f_k$, $1 \leq i \leq I$, $1 \leq k \leq I-1$,

(ii) $\{C_{i1}, ..., C_{iI}\}$, $\{C_{j1}, ..., C_{jI}\}$, $i \neq j$, are independent,

(iii) $Var(C_{i,k+1}|C_{i1}, ..., C_{ik}) = C_{ik}\sigma_k^2$, $1 \leq i \leq I$, $1 \leq k \leq I-1$.

Where Cik denotes the accumulated total claims amount of accident year i up to development year k, f_k is the development factor from k to k+1, and σ_k are parameters.

The first two assumptions seem intuitively sensible, although these can be demonstrated to be the implicit assumptions of the formal chain-ladder model. The third assumption is deduced from the fact that the estimator of f_k, is the C_{ik}-weighted mean of the individual development factors.

An important corollary of assumption (i) is that the development factors are not correlated. That is, if we have a particularly high development factor in one period, there is no tendency for the subsequent factor to be particularly low (or high).

The main results of the paper are as follows. The estimate of the standard error of the reserve estimate for accident year i, $\hat{R}_i$, is:

$$\hat{SE}(\hat{R}_i)^2 = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2}\left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}}\right)$$

The estimate of the standard error of the reserve estimate for all accident years combined, $\hat{R}$, is:

$$\hat{SE}(\hat{R})^2 = \sum_{i=2}^{I}\left(\hat{SE}(\hat{R}_i)^2 + \hat{C}_{iI}\left(\sum_{j=i+1}^{I}\hat{C}_{jI}\right)\sum_{k=I+1-i}^{I-1}\frac{2\hat{\sigma}_k^2}{\left(\hat{f}_k^2\sum_{n=1}^{I-k}C_{nk}\right)}\right)$$

A hat indicates an estimator of the particular figure. The derivations of the estimators of C_{ik}, f_k and σ_k are straightforward, and are set out in the paper.

General comments

Although the above formulae look quite daunting, they consist of nothing more than basic arithmetic — addition, multiplication and so on — and are in fact quite easy to implement in a spreadsheet. Once the formulae have been set up, a new set of data can be brought into a spreadsheet. Activating a "calc" to the spreadsheet will then yield the estimates of the standard errors of the reserves for each accident year, and the reserve as a whole, for the new set of data. This is probably one of the easiest ways of obtaining estimates of reserve variability.

There are many potential drawbacks to simple chain-ladder reserve estimates, which are discussed in Volume 1. The approach in this paper does, however, have the significant benefit of making clear the assumptions one is making. Also, because it is a statistical model, it provides a series of diagnostic tools to test whether these assumptions are valid, as well as giving estimates of reserve variability. The use of these diagnostic tools in discussed further in Section F of Volume 2.

To understand fully the proofs in the paper requires a considerable amount of statistical knowledge. However, the general reasoning involved and the final formulae for the standard errors of the reserve estimates are quite simple, and within the reach of most people with a basic grasp of statistics.

[C7]
PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY FROM INDIVIDUAL PAYMENTS DATA

Contributed by T S Wright

(20 pages, see [D7]

Summary

The paper describes an approach to estimating future claims using data on individual claim payments, rather than the more usual aggregate data. The approach provides an estimate of the whole probability distribution of the outstanding liability, rather than just the first two moments. This additional information may be used to assess safety loadings of reserve estimates, allowing for the skewness of the distribution of the outstanding liability.

Description of the approach

The approach may be summarised as follows:

(i) Estimate the distribution functions, $F_i(x)$, for the size of payments made in development period i.

(ii) Use a weighted combination of the $F_i(x)$ to estimate the distribution of future payments, F(x).

(iii) Fit a curve to F(x) and discretise the fitted curve so it can be used in a compounding algorithm in step (v).

(iv) Construct a probability distribution for the number of future payments.

(v) Calculate the compound distribution of the amount of future payments based on the estimated probability distribution functions in (iii) and (iv). This is done using Panjer's recursive method.

General comments

The approach relies on the availability of individual claim size information, and is capable of implementation in a spreadsheet. To do so, one needs to be able to fit curves to distributions. The curve-fitting and calculating of the compound distribution would probably be quite time-consuming to implement. The approach is probably of most use for situations where one is not considering a very large number of claims.

The paper does make a few sweeping assumptions, which are not fully spelt out. It is intended, however, to illustrate a pragmatic approach to the use of individual claim size information. The paper illustrates the calculation of a safety loading using the

Proportional Hazards criterion , suggested by Wang, which may not be widely known. The use of Panjer's recursive method may also be new to many readers.

The paper requires a moderate level of statistical and computational ability.

◇