ACTUARIAL RESEARCH CLEARING HOUSE
1991 VOL. 1

NON-LIFE INSURANCE CLAIM INCURRAL, ACCRUAL, AND REPORTING ANALYSIS

Progress To Date On Research Funded In Part By

The Actuarial Education and Research Fund

Jim Robinson, FSA, MAAA, Ph.D.
Assistant Professor
Graduate School of Business
University of Wisconsin
1155 Observatory Drive
Madison, WI 53706
(608) 263-2085

ABSTRACT

A stochastic model is presented for non-life claim incurral, accrual and reporting. A review of past work with the model describes distributional properties of the model components and classical parameter estimators. Test statistics are provided for global and local goodness-of-fit, model validation and parameter significance. Forecasts of unreported benefit accruals with estimates of asymptotic error variances are reviewed. A case study is presented to demonstrate the use of the model. Finally, current efforts to generalize the model and to utilize Bayesian inference are discussed.

INTRODUCTION

This review summarizes the key features of a stochastic model underlying the author's dissertation work with non-life insurance claim incurral, accrual and reporting, Robinson (1989). Much of the detail of that paper is omitted from this review, but the interested reader is welcome to contact this author for information on obtaining a copy of the dissertation. The current model was developed as an aid in establishing claim reserve liabilities, but may be useful in other areas such as adjusting premium rates. The author wishes to thank the State Farm Foundation for its financial support during the very early research stages. The claim model continues to evolve thanks to the support of the Actuarial Education and Research Fund.

An excellent review of traditional claim reserve techniques and several of the recently developed stochastic models is found in Taylor (1986).

This presentation provides an overview of the claim model components, classical parameter estimators, test statistics, and forecast point estimates and error variances. A case study demonstrates the application of the model to real data. Finally, we discuss ongoing revisions and extensions to the model.

THE CLAIM MODEL

The claim model is composed of three primary parts, incurral, accrual and reporting. Claim incurral marks the start of an individual claim and generally coincides with the date at which the insurer first becomes contractually obligated to provide benefits associated with an insured event such as an accident or a disability. The number and type of claims incurred within any interval of time is a random quantity. Claim accrual is the benefit accumulation process commencing at incurral with respect to a single claim. Claim accrual is characterized by a claim amount and a development pattern over time, both of which are considered random for our purposes. Claim reporting is the sequence of one or more requests for payments submitted to the insurer arising from past benefit accruals on a claim. For each claim incurred, the number, timing and amounts of such reports are random variables dependent upon the accrual process. The payment of requested claim benefits is not considered by the model.

We let N_{mi} be the number of claims of type m incurred in period i. The model assumes that N_{mi} is independent across m and i, with mean and variance proportional to an observed exposure measure e_{mi}, and is asymptotically normal as total exposure increases. The proportionality constants are functions of an unobserved vector of "generation" parameters, ϕ. Examples satisfying these requirements include Poisson, binomial and negative binomial random variables. Mixtures of these distributions are also possible, that is, the number of claims of one type might be Poisson while the number of claims of a second type might be negative binomial.

Next we let $A_{(mi)kj}$ be the benefit accrued in the j^{th} development period following incurral on the

k^{th} claim of type m incurred in period i. The model assumes that the accruals arising from a claim are independent across m, i and k and are independent of the claim incurral process. The current model further specifies that $A_{(mi)kj}$ has a gamma distribution with shape parameter $\alpha_{(mi)j}$ and inverse scale parameter $\beta_{(mi)j}$ independently across j. Independence across j is possibly the most restrictive assumption of the model.

Finally, we let $R_{(mi)kj}$ be the amount reported to the insurer at the end of development period j on the k^{th} claim of type m incurred in period i. The reporting process is independent across claims. If a report is made at the end of a development period, the model assumes that all unreported accruals are submitted so that no accruals are withheld from the insurer. Lastly, the probability of a report in a development period is roughly proportional to the unreported accruals at the end of that development period. More precisely,

$$\Pr(R_{(mi)kj} > 0 \mid A_{(mi)k1}, \ldots, A_{(mi)kj}, R_{(mi)k1}, \ldots, R_{(mi)k,j-1}) = 1 - \exp\{-\mu_{(mi)j}\, U_{(mi)kj}\},$$

where $U_{(mi)kj} = A_{(mi)k1} + \ldots + A_{(mi)kj} - R_{(mi)k1} - \ldots - R_{(mi)k,j-1}$, the level of unreported accruals at duration j.

The collection of accrual shape and inverse scale parameters, $\alpha_{(mi)j}$ and $\beta_{(mi)j}$, along with the reporting intensity parameters, $\mu_{(mi)j}$, are functions of an unknown vector of "development" parameters, $\boldsymbol{\theta}$. Thus, the incurral, accrual and reporting model is indexed by the combination of generation and development parameters, $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$. Sufficient conditions are placed upon the functions acting on these parameters to provide parameter identifiability from observed data.

Figure 1 below provides a visual image of the development of a hypothetical claim. The change in the accrued line represents $A_{(mi)kj}$ for $j=1,2,\ldots,10$. The only nonzero reports occur at durations 2, 5 and 9. Notice that the amount of the report brings cumulative reports up to cumulative accruals.

Although the current model is restrictive in many senses, it is a reasonable starting point in the evolution of more realistic models. This model attempts to balance flexibility against computational complexity, with more weight given to the latter in early stages. Some characteristics of the current model follow.

1. The model parameters are interpretable. The generation parameters control claim incurral rates and types over time. The development parameters dictate the amount and pattern of claim accruals and the frequency of claim reporting. We hope that such interpretability will simplify the incorporation of subjective information into the claim forecasting process, will provide useful ancillary data on process trends that might be used in rerating and management of the business, and will provide the parameter base for future Bayesian inference.
2. The simplicity of the current model provides for easily computed conditional and unconditional moment formulas for observed reports in terms of the generation and development parameters. This allows for the practical calculation of model

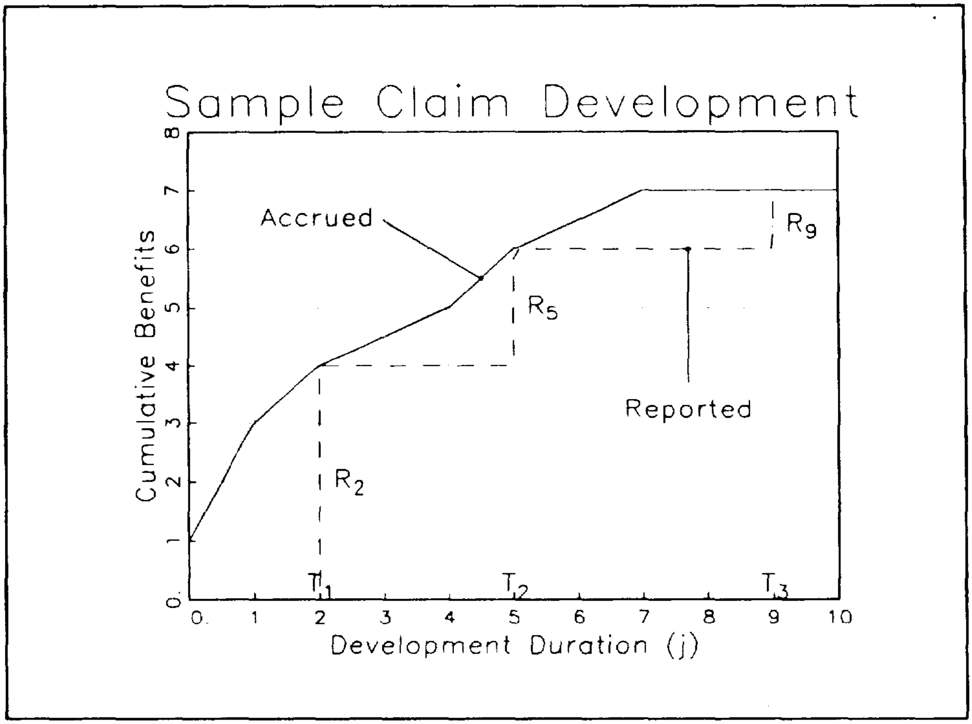

Figure 1 - Sample Claim Development

parameter estimates, test statistics and forecasts on a personal computer. Relaxing some of the independence assumptions, generalizing from the gamma distribution for accruals or changing the characteristics of the reporting process will likely require more intense computations. Although future model enhancements will inevitably be necessary, we hope that the current model will provide a sufficiently flexible structure to capture the first and second moment behavior of some claim processes and demonstrate the viability of this technique.

3. The model allows for multiple claim reports from a single claim. Few stochastic claim models provide for more than one report or settlement per claim.

4. The current model focuses on individual claim data rather than aggregate claim data. Traditional claim reserve techniques consider only aggregate or average claim development experience. With such "macro" models, it may be very difficult to determine whether a slowdown in claim activity is due to reduced claim incurral and accrual rates or is due to shifts in accrual or reporting patterns.

The implications of these component trends on claim reserve estimates is obvious. Since the current "micro" model is designed to estimate each of these component processes, it will automatically distinguish among these possibilities. We should caution, however, that the ability of any model, micro or macro, to uniquely identify claim process component trends is highly dependent upon parameter identifiability which, in turn, is obtained by restrictions on the form of the model. If these restrictions are deemed unrealistic, then the credibility of the model results is suspect. Nonetheless, we believe that the additional information contained within individual claim data makes micro models more likely to accurately capture component trends than aggregate-claim-based macro models.

MODEL ESTIMATION, TESTING AND FORECASTING

The model estimation is accomplished in two stages. First, the development parameters are estimated from observed claim reports. Second, the generation parameters are estimated from the observed number and type of claims using estimates the probability of an incurred claim being reported prior to valuation from the first step. One step estimation techniques are discussed in Robinson (1989), but are not pursued. The two step structure is simpler and allows the generation and development parameters to be separately fitted. Although the two-step procedure is not as efficient as a one-step process, the two-step[technique is more robust in the following sense. If the claim incurral process is incorrectly specified, then the one-step estimation will produce biased results of the development parameters. The two-step development parameter estimates will not be contaminated by errors in the generation model. Errors in specifying the accrual or reporting components of the model will bias the generation parameter estimates in both the one-step and the two-step models.

The development parameter estimation procedure minimizes a quadratic form involving differences among the parametric and empirical probabilities of a report occurring and the first two moments of the amount reported at each observed development duration prior to valuation. The estimates are shown to be consistent, asymptotically normal and efficient with this class of estimators. We might describe the technique as a generalized method of moments (GMM) approach based upon the zeroeth, first and second moments of observed reports. Point estimates of the parameters and error variance estimates are available.

The development parameter estimates determine the probability that an incurred claim remains unreported (IBNR) at the valuation date. Taken along with the number and type of observed claims, we estimate the generation parameters by a GMM technique.

Test statistics and their asymptotic distributions are available for global and local model significance, for model validation using claim data obtained after estimation, and for parameter significance. These test statistics are helpful in model refinement and detection of outliers.

For reserve estimation, we are interested in the distribution of accruals, past and future, which are unreported at valuation. Once estimated and tested, the model provides point estimates and the variance-covariance structure of unreported accruals by incurral period, claim type and development period from incurral. These may be summed or discounted as necessary, with a corresponding error variance computation. Separate estimates for observed and IBNR claim accruals are available. Finally, the error variance is easily decomposed according to source of uncertainty, i.e. development parameter error, generation parameter error and process variance. We can use these forecasts in conjunction with other information to establish claim reserve liabilities with some assessment of the reliability of the estimates.

CASE STUDY - EXCESS MAJOR MEDICAL

We demonstrate the form of the model results by the following example taken from Robinson (1989). We use individual claim payment histories for several thousand excess major medical claims incurred from 1980 to 1983. Since report histories were not available, payment histories were substituted. This coverage form is characterized by a large deductible and a 10 year benefit period. Lags from incurral to first payment were usually in excess of one year. The claim experience is very heterogeneous and requires substantial partitioning into claim types before estimation is possible. Claims are grouped by incurral quarter, small vs. large first payments, single vs. multiple payment, and by lag from incurral to first payment. This severe partitioning results in several claim type cells with only a few claims. As a result, the validity of the asymptotic error variance estimates is suspect. Table 1 below summarizes the estimates of cumulative unreported accruals by development year lag from incurral.

Column (2) displays the point estimates of unreported accruals by years from incurral at the end of 1983. Recall that the oldest claims were incurred in 1980 and are less than 4 years old at the end of 1983. So, all of the numbers beyond lag 4 are based upon an extrapolation of the functional forms applied to the generation and development parameters. Any uncertainty in these functional forms may not be adequately represented in the error variance estimates. Columns (3) through (6) display the standard deviation of the error associated with the estimates in column (2). Column (3) is the standard deviation in the absence of any uncertainty in the parameter estimates. Columns (4) and (5) show the additional uncertainty arising from the estimation of the generation and development parameters. Finally, column (6) is the standard deviation arising from all sources except uncertainty in the model form. Note that the majority of uncertainty arises from parameter estimation error. Although this policy form is unusually volatile, the level of uncertainty in unreported accruals is probably larger than most actuaries would estimate using traditional techniques.

TABLE 1					
Cumulative Unreported Accruals By Years From Incurral (000's)					
(1)	(2)	(3)	(4)	(5)	(6)
Lag Years	Cumulative Accruals	St Dev w/o Parameter Error	St Dev Generation Parameter Error	St Dev Development Parameter Error	Total St Dev
1	1630	283	155	196	377
2	9749	681	969	1339	1788
3	21274	1067	2349	4082	4829
4	30127	1386	3759	7343	8365
5	34982	1604	4790	9271	10557
6	37907	1743	5380	10213	11674
7	40351	1863	5796	10977	12552
8	42664	1977	6177	11771	13440
9	44920	2088	6557	12623	14377
10	47143	2197	6940	13530	15364

MODEL EXTENSIONS AND MODIFICATIONS

The current model assumptions are restrictive in the following ways.

1. $A_{(mi)kj}$ is independent across j. That is, the behavior of past accruals do not modify the distribution of future accruals arising from an individual claim. If claims are severely partitioned into homogeneous types with little variation in development pattern within claim type, then this independence assumption may be approximately correct. This requires that the data be spread over a large number of claim types, however, reducing the applicability of asymptotic results. If only a few observable claim types are available, with substantial variation of development patterns within type, then the independence assumption is not likely to be acceptable.

2. The probability of a report taking place is only a function of the current level of

unreported accruals and does not depend upon the number or frequency of past reports on the claim or on the accrual pattern. We might expect the probability of a report to be greater after the first report has occurred, all else being equal.

3. Characteristics 1 and 2 combine to imply that the conditional expected moments of the amount reported at a specified development duration given the past history of reports on that claim only depends on the number of development periods since the last report. This also implies that the unconditional covariance between reports (zero or nonzero) at adjacent development durations is negative. That is, if the amount reported at duration j is above the unconditional expected value, then the amount reported at duration j+1 is expected to be below its unconditional mean. Empirical correlations of many claim report histories show positive correlation between reports at adjacent development durations. We can sometimes circumvent this problem to some extent by aggressively partitioning the data into numerous claim types. Reports may be unconditionally positively correlated, while being negatively correlated conditional upon claim type. But again, this claim typing introduces its own additional problems.

4. Claim accruals have a gamma distribution. If we are only interested in first and second moments of claim accruals, this assumption is not a serious constraint. The two-parameter gamma distribution allows great freedom in the level of the variance vs. the mean. If we are interested in the tail behavior of claim accruals, however, the gamma assumption is not likely to be acceptable. The gamma distribution provides for simple expressions for the moments of claim reports, but may not possess sufficiently heavy tails.

5. The model estimation and forecasting techniques do not allow the valuation actuary to consistently incorporate subjective information into the establishment of unreported accruals. The actuary, for example, may have information about component claim incurral, accrual or reporting trends and the actuary can adjust the generation and development parameters in an ad hoc fashion to modify the point estimate of unreported accruals. However, the same information may already be reflected in the data or may be refuted to a great extent by the data. Consequently, it is difficult to determine the weight to give to additional subjective information when making these ad hoc adjustments. Furthermore, there is no guidance in determining the modified level of uncertainty in the point estimate.

Current research efforts are as follows.

1. Incorporate some dependency among accruals. There are several interesting approaches to this extension of the current model.

   a. We may arrive at the current assumption of independent gamma distributions through a series of steps. First, assume A^* is total deflated accruals over the life of the claim with a gamma distribution with parameters γ and β. Next, assume that the proportions of A^* arising from the various development periods of the claim are generated by a Dirichlet distribution with parameters $\alpha_1,\ldots,\alpha_J$. Third, assume that $\gamma=\alpha_1+\ldots+\alpha_J$. Then, if A_j^* is the deflated accruals from period j, we have A_j^* with independent gamma distributions across j. Finally, we define A_j as the inflated accrual during period j and let $A_j = A_j^* \beta/\beta_j$ so that A_j are independent gammas with varying shape and inverse scale parameters.

The independence among the accruals results from the balancing of variance of aggregate deflated accruals with the variance of the Dirichlet development pattern. If we increase the variance of accrual amount relative to the variance of development patterns, then we produce positive correlation among the accruals. That is, a large accrual in one development period is likely to be a sign that the claim amount is larger than expected and, therefore, we should expect other accruals to be larger than otherwise expected. Similarly, if we reduce the variance of total accrual amount relative to the variance of development patterns, then we produce negative correlation among the accruals. This is especially clear when the variance of A^* is set to zero. Thus, by removing the constraint that $\gamma=\alpha_1+\ldots+\alpha_J$, we introduce an additional model parameter controlling the level of correlation among development period accruals.

b. A somewhat different approach is to let $A_j = S\, A_j^* \beta/\beta_j$, where S is an independent random variable with a fixed mean of unity. It is easy to show that the resulting unconditional correlation between accruals is positive and proportional to the variance of S. Thus, we cannot model negative correlation with this extension, but we may be able to select simple distributions for S which yield tractable results, e.g. discrete distributions.

c. A third approach is to assume the existence of unobservable claim types dictating the parameters of conditionally independent gamma accruals. That is, assume that random variable T indexes the hidden claim type and that, conditional upon T, the A_j's are independent gammas with parameters α_{jT} and β_{jT}. We might take T to be a discrete or continuous random variable. The previous extension is a special case of this construction. This extension allows for both positive and negative correlation among accruals.

d. A simpler modification to obtain correlation among the accruals is to hypothesize that the conditional distribution of A_j given prior accruals is gamma with shape and/or scale parameters which depend upon the past accruals. For example, assume that A_j is gamma with shape parameter α_j and inverse scale parameter $\beta_j/(A_1+\ldots+A_{j-1})$. This is equivalent to the assumption that the growth rates of cumulative accruals are independent gammas with parameters α_j and β_j. This is similar to the structure implicit by traditional chain and ladder techniques. Of course, many other functional forms are possible producing a variety of correlation structures.

Early attempts to work with each of these extensions indicate that d is most likely to produce workable formulas for moments of observed reports. The identifiability of development parameters is also a concern with any of these modifications. That is, with increased model flexibility, it may be possible that two or more sets of development parameters will result in the same moment structure for observed claim patterns.

2. Allow the reporting intensity to depend on more of the past accrual and report history than merely the unreported claims. For example, we might include a constant term, μ_0, in the intensity as follows,

$$\Pr(R_{(mi)kj}>0 \mid A_{(mi)k1},\ldots,A_{(mi)kj},R_{(mi)k1},\ldots,R_{(mi)k,j-1}) = 1-\exp\{-\mu_0-\mu_{(mi)j}\,U_{(mi)kj}\}.$$

This allows for a minimum level of reporting activity regardless of the amount of unreported accruals.

Another possibility is to allow the reporting intensity to vary with the number of prior reports. In particular, $\mu_{(mi)j}$ might be extended to $\mu_{(mi)jz}$, where $z=0$ if the claim is unreported at duration j-1 and $z=1$ otherwise.

As with the discussion of accrual modification. these extensions of the reporting model may cause identifiability problems for development parameters.

3. A major goal of the ongoing research with the current model is the eventual application of Bayesian inference techniques. Model parameters have been made interpretable with this possibility in mind. If the previous modifications are successful, severe claim type partitioning will be reduced and a more parsimonious set od generation and development parameters may be possible. The application of a Bayesian prior on this reduced parameter space may then be computationally feasible. This will allow for a consistent melding of subjective information with observed claim data to create a Bayesian posterior distribution for the parameters and the unreported accruals. Bayesian analysis has not been attempted as yet due to the difficulties involved in working with the likelihood function of the observed claim data and the large number of parameters in the current model.

4. Other areas of investigation for future consideration include generalizing from the

gamma distribution for claim incurrals, correlation among the number of claims incurred by type and period of incurral, and small sample properties of the estimates and forecasts.

The range of possible model extensions is tremendous. With increasing computer power many new avenues will open up. We hope that this first simple model will provide the basis for useful refinements and help to identify the key problems in building micro models of individual claim behavior.

BIBLIOGRAPHY

Robinson, James M. (1989), "Modeling, Inference and Forecasting Techniques for the Analysis of Non-life Insurance Claim Reserves," unpublished Ph.D. dissertation, University of Wisconsin - Madison.

Taylor, G. C. (1986), **CLAIMS RESERVING FOR NON-LIFE INSURANCE**, Amsterdam, The Netherlands: Elsevier Science Publishing B.V.