Institute of Actuaries of Australia

Risk Margin Monitoring

Prepared by Julie Sims

Presented to the Institute of Actuaries of Australia
17th General Insurance Seminar
7 – 10 November 2010
Gold Coast

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 2 9233 3466 Facsimile: +61 2 9233 3446
Email: actuaries@actuaries.asn.au Website: www.actuaries.asn.au

Abstract

A number of methods are currently in use for determining risk margins for General Insurance. However, there is little quantitative information on whether these methods are performing as intended, for example, whether the value of the insurance liabilities provided under Prudential Standard GPS 310 is sufficient at least 75% of the time. This paper provides a methodology, illustrated with examples, for comparing the "expected" risk margin with the "actual" experience at each valuation. Over a number of valuations, this methodology will provide a check on the adequacy of the risk margin. It can be easily applied for most methods of determining risk margins that are based on stochastic modelling techniques. It can also be applied to short-tail lines whether or not a stochastic technique has been used.

Keywords: risk margins; reserving; hindsight reserving; General Insurance.

Introduction

In General Insurance, insurance liability estimates, inclusive of risk margins, feed directly into the calculation of regulatory capital requirements and regulatory assessment of insurer solvency. They are also an important input to Enterprise Risk Management Frameworks as they provide information about the amount of risk in claim liabilities. Thus, it is important that the calculation is soundly based. However, all of the methods currently in use for assessing risk margins are inevitably based on many untestable assumptions and qualitative assessments of such difficult issues as model error and internal and external systemic risks. It is generally not feasible to objectively assess whether the experience is consistent with each assumption. Thus it would be very useful to have a method of assessing whether the risk margin that has been chosen appears adequate.

Houltram (2005) describes a method of hindsight analysis that is one approach to making this check. However, it requires retrieving a considerable amount of data from previous valuations stretching back several years, together with additional analysis. This paper presents an alternative method of checking risk margin reasonableness that, where a stochastic method has been used to determine the margin, requires only a minor extension to quantify payment variability over a one-year time horizon. The approach can also be applied to short-tail classes, irrespective of what risk margin approach has been applied.

Under PS300, General Insurance liability reports are currently required to include documentation that compares actual experience with that expected based on the previous valuation. Typically, projected experience between valuations does not include explicit quantification of the degree of variability that would be consistent with the valuation basis and the assessed risk margin. Without this, the presentation of actual with expected experience is restricted to a qualitative assessment of whether the differences are consistent with the variability that ought to have been anticipated given the risk margin.

This paper describes how the same method that is used in calculating the risk margin can be used to quantify the amount of variability that is expected in actual claim costs between valuations. Only minor extensions to stochastic methods for calculating risk margins on outstanding claims liabilities are needed to give corresponding "margins" on claim costs for a single period. If the difference between actual and expected claim costs is greater (or less) over a number of valuations than would be reasonable given the expected variability, this would cast doubt on the reasonableness of the risk margin. This provides feedback not only on the goodness of fit of the model, but on whether all sources of risk that are acting on claim costs have been incorporated into the risk margin. Conversely, if qualitative risks (at least those with a reasonable likelihood of occurring) have been over-estimated, this should also become apparent over a number of valuations.

The proposed process is illustrated on actual valuations that include qualitative adjustments of the type recommended by the Risk Margins Taskforce (Marshall et al, 2008). These valuations, made by a number of different actuaries using different stochastic methods, provide case studies of how the actual versus expected comparison can be used to provide feedback on the appropriateness of the risk margin.

Methodology for Stochastic Risk Margins

A stochastic risk margin will be based on a stochastic model capable of predicting the probability distribution of the total outstanding claims, as the 75^{th} percentile must be estimated. That stochastic model will almost certainly have the ability to predict the probability distribution of the claims to be paid in the next transaction period. If we are sufficiently confident in the assumptions underpinning the stochastic model to use it to produce a risk margin, then we are likely to feel confident that its predictions for the next transaction period are reliable. Comparing those predictions with outcomes at each valuation will give regular feedback on the validity of at least some of the assumptions behind the risk margin.

The amount of comfort this gives us about the risk margin will depend on how much of the outstanding claims relates to payments to be made in the next transaction period and how confident we are about the payment pattern. It may be useful to look at a number of transaction periods combined, to give a better test of some of the assumptions that relate to longer term performance. In that case, care must be taken to not use the same "actuals" more than once. For example, if the actuals from two transaction years are compared with the expected claims for those years, the test could only be done every second year.

Some of the more common stochastic models used for assessing risk margins are:

1. Chain ladder bootstrap
2. Other non-parametric bootstraps based on different models
3. Mack's model
4. Generalised linear models
5. Adaptive generalised linear models

The type of modifications needed to these models to produce the information shown in the graphs in this paper is described in the Appendices. Some specific cases are illustrated in the examples below.

In many cases, a subjective adjustment will be applied to a risk margin that has been calculated from a stochastic model, for example, to allow for model error. Some judgment will need to be made on how much of that adjustment should be applied to the next transaction period. For a long-tail line, much of the model uncertainty might relate to uncertainty in future inflation, which will have a negligible effect on the next transaction period.

Testing the Risk Margin

The simplest test that can be performed is: are the actuals greater than the 75^{th} percentile about 25% of the time? A binomial distribution can be used to quantify whether the result is implausible. For example, the probability of getting above the 75^{th} percentile three times out of three valuations is about 0.02 – unlikely to happen by chance, so if it does happen, it suggests that either the risk margin (or the central estimate) is too low. In that case, examining the reasons why those three actuals were high may reveal a source of risk that is not adequately taken into account in the overall risk margin.

The next simplest test requires an assumption about the form of the probability distribution for the claims in the next transaction period. Whatever assumption was made to determine the 75^{th} percentile of the overall risk margin is likely to be appropriate. Then it is straightforward to estimate the 25^{th} percentile, and we can ask: are the actuals between the 25^{th} and 75^{th} percentiles (the **50% range**) about 50% of the time? For example, the probability of getting five out of five actuals inside the 50% range is about 0.03 – again, unlikely to happen by chance, so it would suggest that the risk margin might be too high and should be investigated further. Particularly for long-tail classes, it would be advisable to also examine the actual versus expected by accident period to see if there are patterns in the differences and, if there are, what the cause might be. If there are no obvious patterns, it would be necessary to review the specific risks incorporated in the risk margin in detail to see whether the risk margin could be reduced.

As a further test, if there is a sufficient number of valuations available, we can quantitatively test whether the actual values have the expected spread. For example, if it is reasonable to assume that the probability distribution is lognormal, we can test whether the log of the actual values appears to be normally distributed with the mean and standard deviation associated with the expected value and the risk margin. A chi-squared test can be performed to test whether the actual standard deviation is consistent (or not) with the expected standard deviation.

Checking the Model

If the actual values do not seem to be consistent with the expected values and risk margin, the next step is to check whether the model assumptions are reasonable in the light of the most recent history of the particular business being valued. Plots of model residuals may show patterns that indicate some trends or changes have emerged. Some of the questions that could be asked are:

1. Are the expected values from the stochastic model materially different to the central estimates? Although not ideal, it is not unusual for a simpler model to be used for determining the risk margin than is used for the central estimate. Any simplifications should be examined carefully to see if they are likely to materially affect the risk margin.
2. Does the amount of process variability follow the pattern implied by the stochastic model?
   For example, if the variability is higher than expected according to the model in the later development periods, this may cause the variability to be over-estimated in periods where the claim amounts are higher.
3. Is there systematic variation in the calendar direction that is not taken into account?
   Uncertainty in the future calendar trends can be a major source of risk.
4. Is the development pattern changing over time?
   For example, in a bootstrapping model, are there any trends in the ratios? Not allowing for changes in development pattern may produce considerable bias. Evidence of changes in the historical data may indicate a source of risk.

Methodology for Non-Stochastic Risk Margins on Short-tail Lines

With short-tail lines, it is often the case that a large part of the outstanding claims will be paid in the next transaction period. Then it is likely that the coefficient of variation (CV) of the payments in the next transaction period will be close to the CV of the outstanding claims.

Short-tail Rule of Thumb for Outstanding Claims:
For short-tail lines, use the outstanding claims risk margin (as a %) to test whether the actual amount paid in the next transaction period falls within the 50% range about half the time.

Some assumption will need to be made about the probability distribution of the outstanding claims to estimate the 25th percentile. Normality or lognormality should be reasonable assumptions for most short-tail lines.

For premium liabilities, the valuation report might not contain a prediction of the expected claims in the next transaction period, but it must have a prediction of the ultimate claims. In this case, we may have information in the next report that allows us to update that ultimate prediction based on the additional claim information. For a short-tail line, this updated prediction will be close to the actual ultimate result. This hindsight estimate can be treated as an “actual” to be compared with the expected ultimate claims, with the risk margin for the premium liabilities providing the 75th percentile. By assuming a probability distribution, for example, normal or lognormal, we can also estimate other percentiles such as the 25th.

Short-tail Rule of Thumb for Premium Liabilities:
For short-tail lines, use the premium liabilities risk margin (as a %) to test whether the hindsight estimate after one period falls within the 50% range about half the time.

Methodology for Non-Stochastic Risk Margins on Long-tail Lines

This is more problematic, as the relationship between the overall CV and the CV of the first future transaction period depends on many factors. However, the same could be said about the risk margin, yet "industry" factors have been developed in the papers of Bateup and Reed (2001) and Collings and White (2001). The techniques used in those papers could equally well be used to produce "industry" risk margins on the first future transaction period by repeating the analyses done in 2001 using stochastic models and expert judgment, but with the risk margin on the first future transaction period being estimated in addition to the overall risk margin.

Case Study 1 - Chain Ladder Bootstrap on a Short-tail Line

This example is home buildings insurance, with a small public liability component (about 5% of the total). The valuation has been performed with only minor changes for the last eight years. The risk margin has been estimated using both a chain ladder bootstrap and the Bateup and Reed method at each valuation. The selected risk margin was based on the Bateup and Reed method for the first of those eight valuations (2003). In the remaining years, it was based on the chain ladder bootstrap as described in Appendix 1.

Subjective Adjustments to the Model

There is no explicit allowance for model error. However, the choice of a four period average (from 2005 onwards) for the pseudo-dataset implicitly adds some margin – an eight period average, which is more typical of the models used for the central estimate, gives a margin about 1% lower.

Modifications to the Calculations

The spreadsheet was modified so that only the first future transaction year was projected, and the simulations re-run to give a risk margin, calculated in the same way as for the overall risk margin. Where a judgmental adjustment had been made to the overall lognormal risk margin, a proportional adjustment was made to the first transaction year lognormal risk margin. 25^{th} and 75^{th} percentiles were then calculated from the risk margin, assuming a lognormal distribution.

Testing the Risk Margin

This line is short-tail. 88% of the outstanding claims liability relates to payments projected for the first transaction year after the valuation date. Consequently, the "actual versus expected" comparison over a one-year horizon is close to being a direct test of the adequacy of the model's result and overall risk margin.

Actual and expected values for the amount paid in the each year were taken directly from the report for each valuation. These values are shown in Figure 1, together with the 25^{th} and 75^{th} percentiles from the corresponding first transaction year risk margin. There are two values above the 75^{th} percentile and two below the 25^{th} percentile, which is in excellent agreement with the expected numbers of 1.75 (7 valuations times 1 in 4 probability).

As a further test, we can check whether the log of the actual values appears to be normally distributed with the correct mean and standard deviation. The standardised residuals, assuming the mean and risk margin are correct, have a average of -0.02 and a standard deviation of 0.7. The average is so close to the expected value of 0 that it is not necessary to do a statistical test to tell that it is not significantly different to expected (formally, with 7 valuations, the average times the square root of 6 divided by the standard deviation should be compared with the percentiles of a t-distribution with 6 degrees of freedom). The variance of 0.5 is rather lower than the expected value of 1, but, comparing the variance times 6 with the percentiles of a chi-squared distribution with 6 degrees of freedom, the value is at the 16^{th} percentile, which is not particularly unusual. So again there is good agreement between the estimated risk margins and the actual amounts paid.

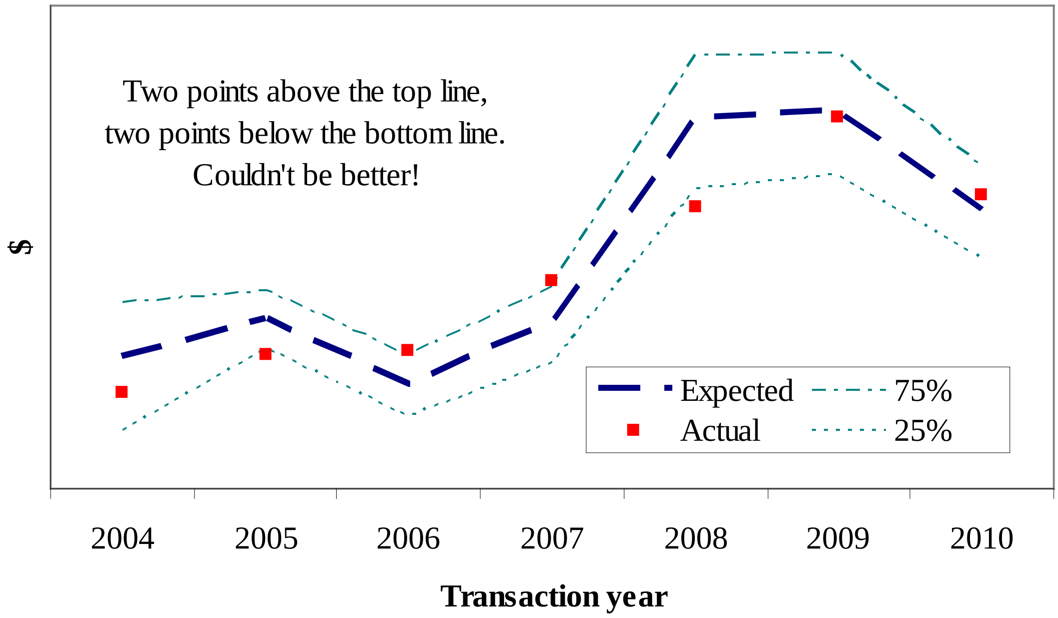

Figure 1: Historical performance of the risk margin for the outstanding claim valuations in 2004 to 2010

Non-stochastic Test of Outstanding Claims Risk Margin

As suggested by the Short-tail Rule of Thumb for Outstanding Claims, the first transaction year risk margin and the overall risk margins were very similar. For the valuations where 10,000 simulations were run, the differences were less than 0.3%, which is close to the accuracy of the simulated estimate. This means that even if we did not have a stochastic risk margin assessment, we could still test the accuracy of the risk margin by assuming the first transaction year risk margin is equal to the overall risk margin.

Figure 2 shows the Short-tail Rule of Thumb for Outstanding Claims applied to this dataset. There is very little difference between this plot and Figure 1, which uses the stochastic calculation of the risk margin. As before, we can conclude that the risk margin is adequate and there is no evidence of bias in the central estimates.

Test of Bateup and Reed Risk Margins

It is common practice to refer to the benchmark risk margins in the IAAust research reported in the paper of Bateup and Reed (2001). If we assume that the risk margin for the expected claims in the first future year can be estimated as the risk margin that would apply to an outstanding claims liability of the same size, adjusted for inflation since 2001, we can test that risk margin against the actual claims at each of the 2004-2010 reports. This is shown in Figure 3. All of the actuals lie within the 50% range. This strongly suggests that the benchmark risk margin is too high in this case.

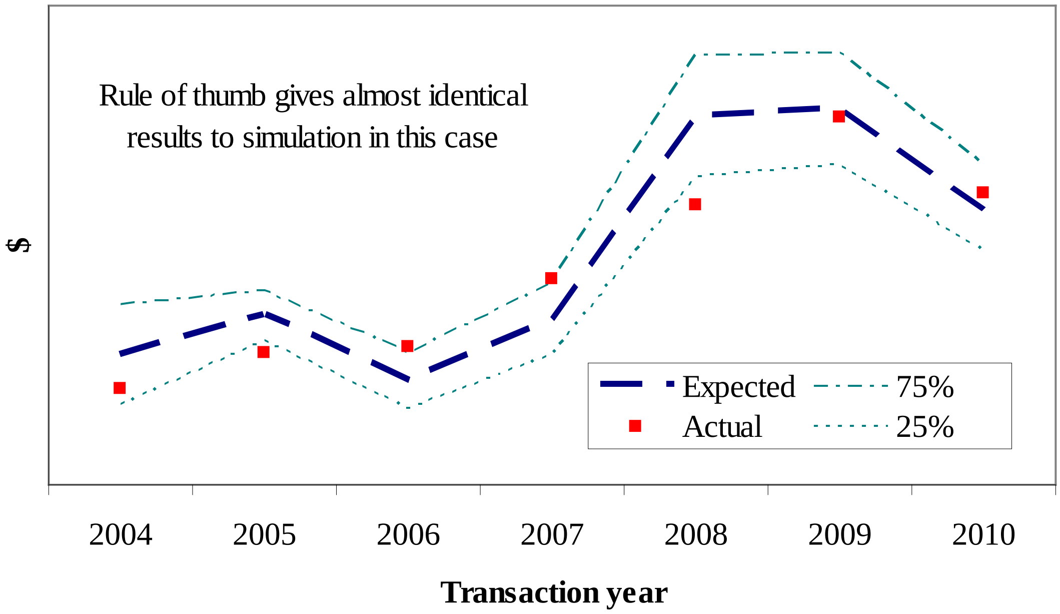

Figure 2: Rule of thumb risk margin version of historical performance of the risk margin for the outstanding claim valuations in 2004 to 2010

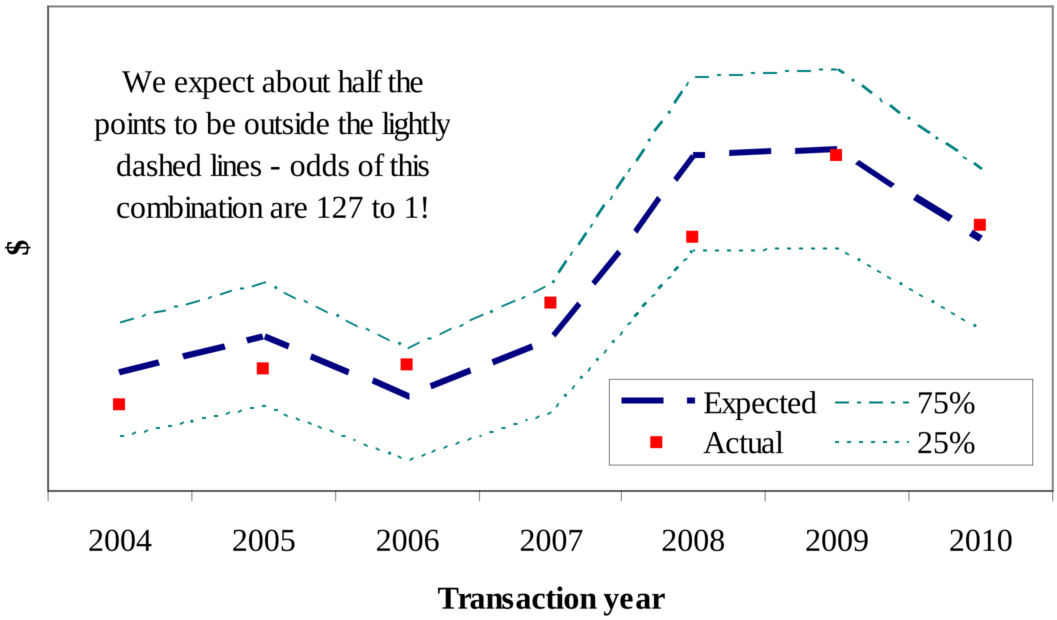

Figure 3: Bateup and Reed risk margin version of historical performance of the risk margin for the outstanding claim valuations in 2004 to 2010

Non-stochastic Test of Premium Liability Risk Margin

Ideally, we would like to perform the same calculations for the premium liabilities. However, for this example, that is not straightforward. This is because the premium liability valuation does not forecast the future payments by transaction year. Also "actual" values have not been calculated for those reports.

However, the short-tail nature of the payments means that after one year, we have very good estimates of the claim frequency and moderately good estimates of claim size. If we assume that the exposure estimate was correct, we can test the estimated premium liabilities (the product of exposure, claim frequency and average claim size) against an updated estimate that should be close to the ultimate value.

Figure 4 shows the premium liabilities as “expected” and the estimates updated one year later as “hindsight after 1 year”, with 25th and 75th percentiles calculated from the risk margin. This risk margin is the outstanding claims risk margin multiplied by 1.5, to allow for occurrence risk.

The 2008, 2009 and 2010 hindsight values are all above the 90th percentile. The chance of one value in two exceeding the 90th percentile is about 1 in 6. The chance of two values in three exceeding the 90th percentile is about 1 in 40. The chance of three values in four exceeding the 90th percentile is about 1 in 300. So warning bells should have begun to ring in 2009 and they are well and truly chiming by 2010.

The hindsight values are high in 2008-2010 because of severe storms in the December or March quarters when the premium liability exposure is still significant. It appears that there may have been a change in the risk applying to the premium liabilities. Unless the events in these years are truly exceptional, either the risk margin or the central estimate of the premium liability needs to be increased in future.

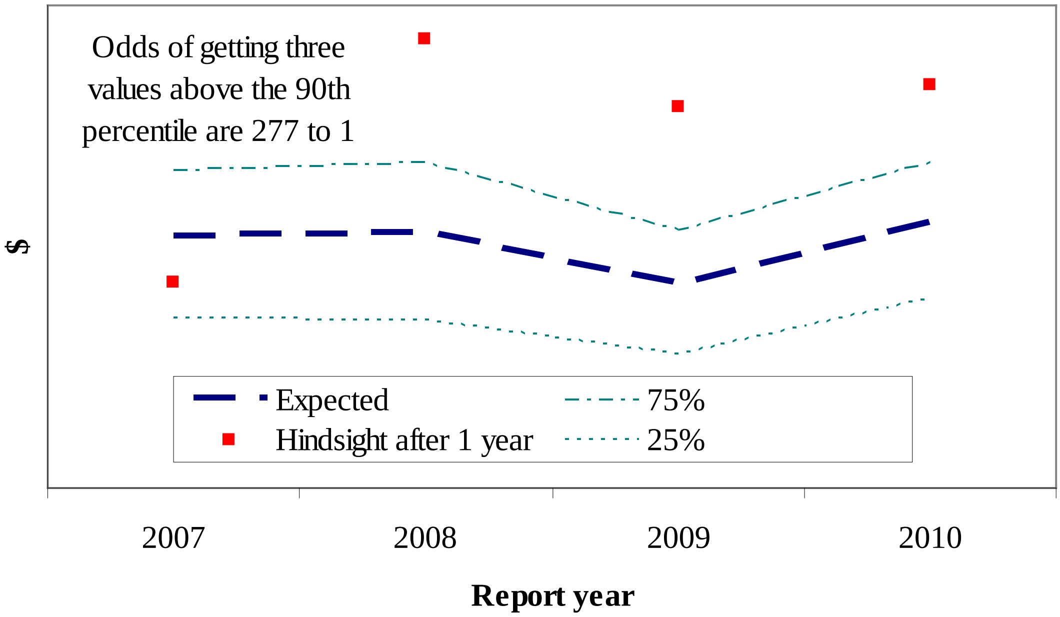

Figure 4: Historical performance of the risk margin on premium liabilities

Case Study 2 - Multiple Bootstrap on a Long-tail Line

This example is Lenders Mortgage Insurance. Central estimates of outstanding claims in this example are obtained from data on when mortgages fall into arrears, their current status and the date and amount of any payments on claims. A claim is considered “reported” when a mortgage falls into arrears. It is “finalised” when either it comes out of arrears or is taken into possession. Some time after being taken into possession, a claim will be paid. Claims are “pending” if they are reported but not finalised.

A policy on a loan currently in arrears that becomes a claim is considered part of the outstanding claims liability. All current policies on loans that go into arrears in the future contribute to the premium liability. This makes it straightforward to determine whether a future payment forms part of the outstanding claims liability or part of the premium liability or neither, simply by looking at the status of the policy at the valuation date.

Ratio models are used to project:

- future active claims from currently active claims;
- future pending claims from active claims;
- future pending claims from current pending claims; and
- future possessions from pending claims.

The ratio models used to forecast central estimates are developed into stochastic bootstrap models as described in Appendix 2.

Bootstrapping had been done on three previous valuations for both outstanding claims and premium liabilities, although the associated risk margin was only used on the last two valuations (the Bateup and Reed calculation was used in the previous valuations). The bootstrapping is mostly set up in Excel, with a Visual Basic routine to copy each simulation to a summary table.

Subjective Adjustments to the Model

The simulated 75th percentile was increased by 30% by the investigating Actuary to allow for model error.

Modifications to the Calculations

The three bootstrap spreadsheets were modified to record an additional data value and re-run. For the outstanding claims, this is the simulated value for the first future transaction year. For the premium liabilities, the expected value in the first future transaction year is very small, as mortgages have to first go into arrears and then eventually become a claim. So the second future transaction year was used instead. This means the feedback on a particular year’s premium liabilities risk margin is not received until two years later.

For the valuations where bootstrapping had not been done, the risk margin for the first or second transaction year was estimated by taking the same proportion of the overall risk margin as the average of the three bootstraps. For the outstanding claims, the risk margin in the first transaction year was 82%, 73% and 78% respectively of the overall risk margin in 2007, 2008 and 2009. 78% was used for 2004-2006. For the premium liabilities, the risk margin in the second transaction year was 134%, 111% and 129% respectively of the overall risk margin in 2007, 2008 and 2009. 124% was used for 2004-2006.

Some additional data extraction had to be done to get the “actuals” needed to do the premium liability risk margin test. There were very significant differences in the historical data between 2008 and 2009 – a number of claims went from $0 to non-zero $ without changing their transaction date. This makes it difficult to know which “actual” value to use!

Testing the Risk Margin

Figure 5 shows the results for the outstanding claims. The risk margin is certainly adequate for all five transaction years where both actual and expected values are available. The last two years are well below the 25th percentile. This may be a result of the forecast of unemployment in Australia due to the Global Financial Crisis being much higher than turned out to be the case.

In this example, a split of the transaction year actual and expected by accident year was not readily available, so it was not possible to get more detail of the pattern of differences across accident years.

Figure 6 shows the results for the premium liability. The premium liability risk margin is also certainly adequate for the three years where both actual and expected values are available. Again, the last two years are well below the 25th percentile.

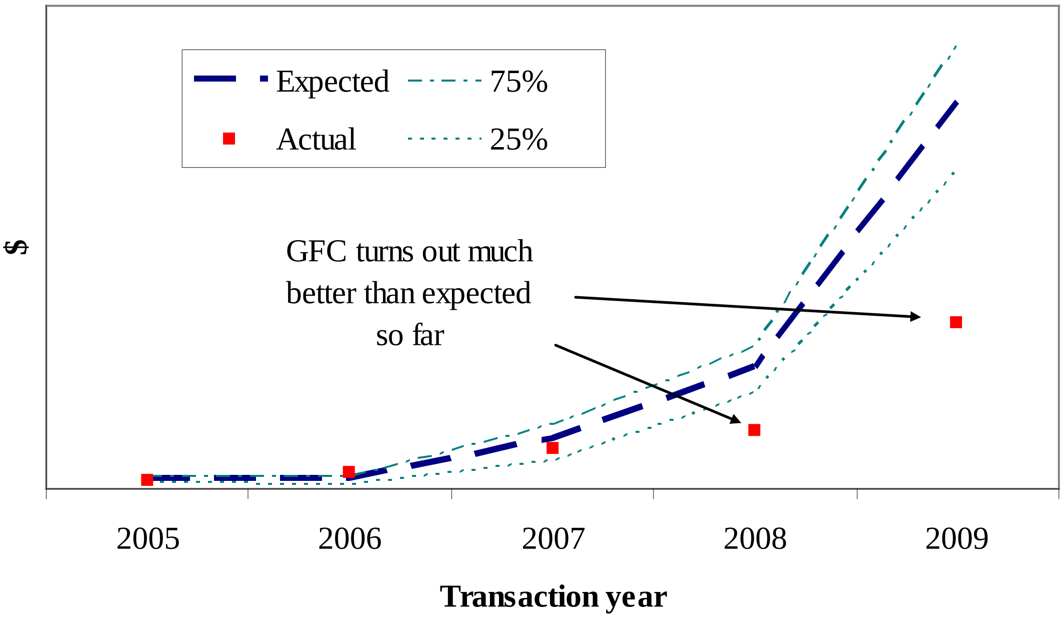

Figure 5: Historical performance of the risk margin for the outstanding claim valuations in 2005 to 2009

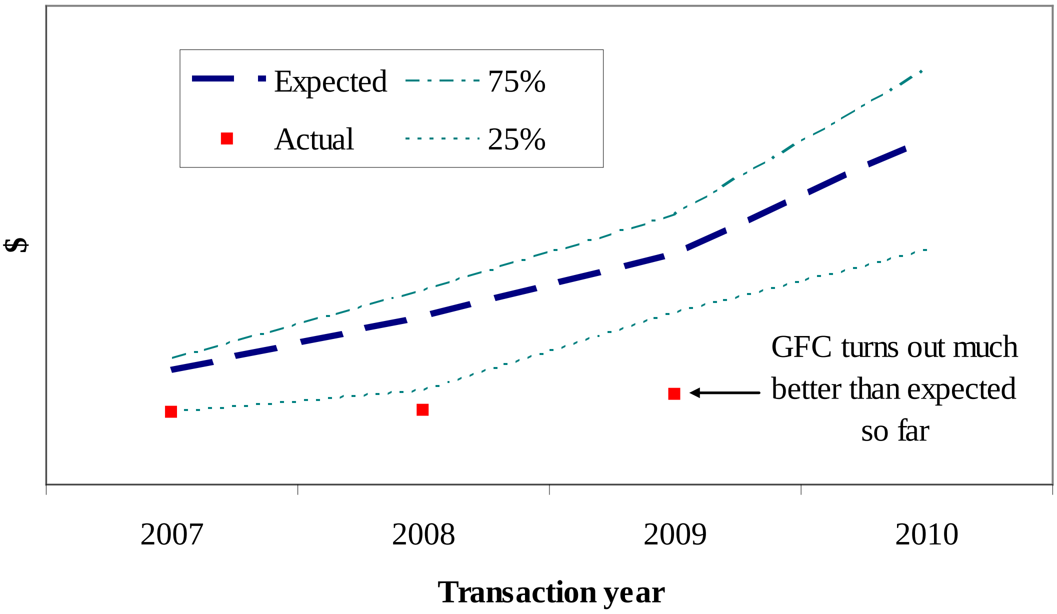

Figure 6: Historical performance of the risk margin on premium liabilities.

Case Study 3 – Adaptive Reserving on a Long-tail Line

This example is Workers' Compensation. The valuation has been performed every six months since December 2003, so there are currently 13 valuations where actual, expected and risk margin data can be compared. However, the risk margin calculation is on an annual basis, so the simplest comparison to do is annual.

Stochastic risk margins were calculated but the Bateup and Reed risk margin was used for the 2004 valuation. The next three valuations used an adaptive Kalman filter, where the development pattern is allowed to change over time, together with a stochastic inflation model. The final two valuations used a more general form of the adaptive filter that uses a gamma probability model for the process error. The stochastic models are largely in SAS, and are the end-product of considerable theoretical and code development.

Subjective Adjustments to the Model

Various subjective adjustments have been made to the risk margins to allow for model error. In 2005-2007, there was a 30% allowance for model error. In 2008, there was a 20% allowance for model error. In 2009, the allowance was increased to 40% to allow for additional model error in the inflation model. Given that there is little inflation model risk in the early periods, and much of the model risk relates to later periods, a judgment needs to be made on what part of this allowance is relevant to the comparisons in this report. The allowance for model risk has been reduced (for the purpose of the following comparisons only) to 10% based on my judgment that all of the inflation model risk and (at least) half of the remaining model risk will relate to later periods.

Modifications to the Calculations

The output of the bootstraps includes values for each cell in the future triangle, so a small SAS program was written to rescale and sum the cells in the required future transaction periods. The 75th percentile was taken directly from the resulting 1000 or 10,000 simulated values. The risk margin was multiplied by 1.1 to allow for part of the model error.

Testing the Risk Margin

Actual and expected values for each year are shown in Figure 7, with 25th and 75th percentiles based on the corresponding first transaction year risk margin. For the five years 2004-2008, the actual values fall just outside the 50% range. This is moderately strong evidence that the risk margin for the first transaction year was too low. However, given that this example is long-tail, one transaction year will only test a small part of the assumptions behind the outstanding claims estimate. Further investigation of the overall risk margin would be indicated, if it had not already been increased. There are two values out of 6 above the 75th percentile, so this is not significantly different to the expected 1.5 values, i.e. the combination of the risk margin and central estimate is sufficient.

If instead, we use the current risk margin for all valuations, all six actual values fall inside the 50% range (Figure 8). This is moderately strong evidence that the risk margin for the first transaction year is now too high. Further investigation of the overall risk margin is indicated, particularly an examination of the breakdown of the actual and expected across accident years.

With valuations being done six-monthly, it is possible that we could get quicker feedback by looking at six-monthly actuals and expecteds. However, this will be an even weaker test of all of the assumptions behind the outstanding claims risk margin than a one transaction year test.

Looking at a 6-monthly period, there is data readily available for 9 valuations. At the June valuation, the data is Oct-March; at the December valuation, the data is Apr-Sep. The risk margins appear appropriate, given that there are only two values outside the 50% range (Figure 9). However, there is some evidence of increased volatility at this scale – possibly due to unmodelled seasonality or variation in the speed at which payments are processed. Variations at this level may be unimportant to the accuracy of the overall valuation, so it may be counter-productive to look at too fine a scale.

Indeed, looking at the data in the groupings Jan-Jun and Jul-Dec, a rather different picture appears (Figure 10). The 75th percentile is certainly adequate. However, it appears that this is because the expected values have a conservative bias. The more recent expected values do not show any evidence of bias.

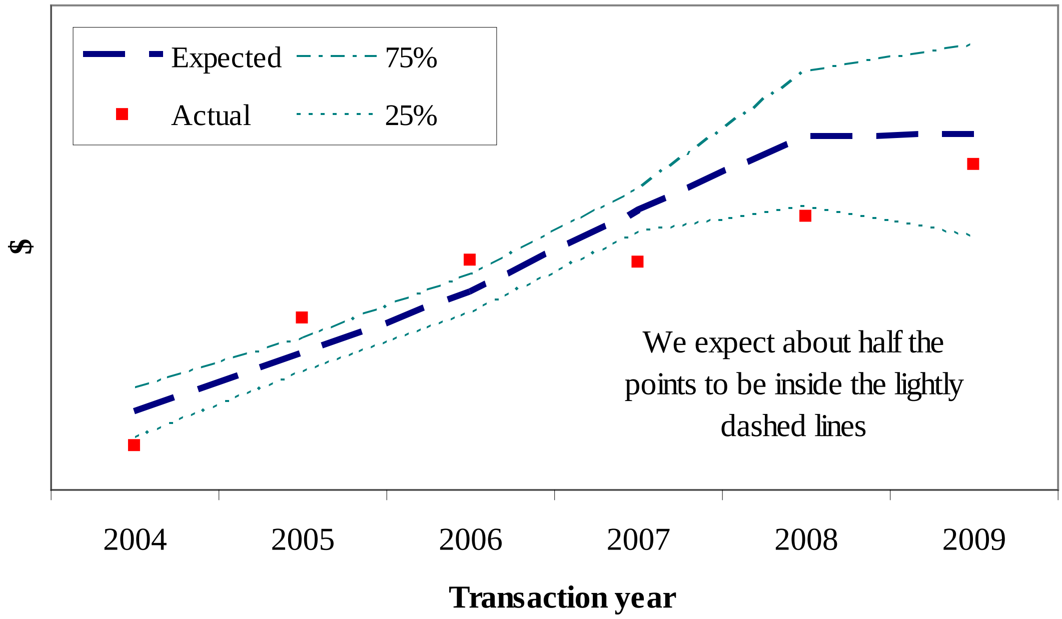

Figure 7: Historical performance of the risk margin for the outstanding claim valuations at 2004 to 2009

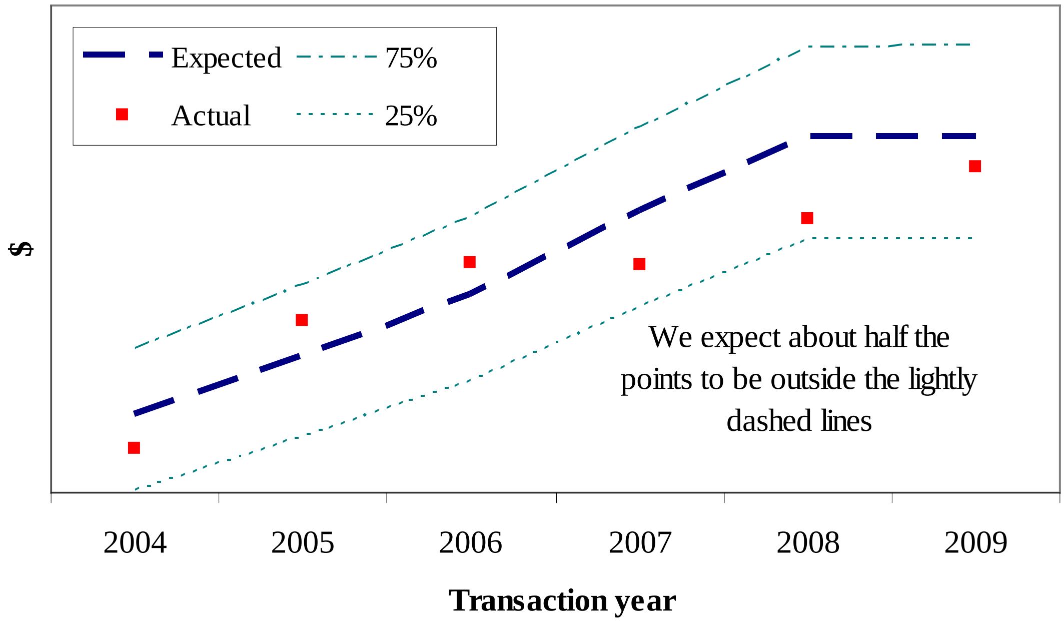

Figure 8: Historical performance of the risk margin for the outstanding claim valuations at 2004 to 2009

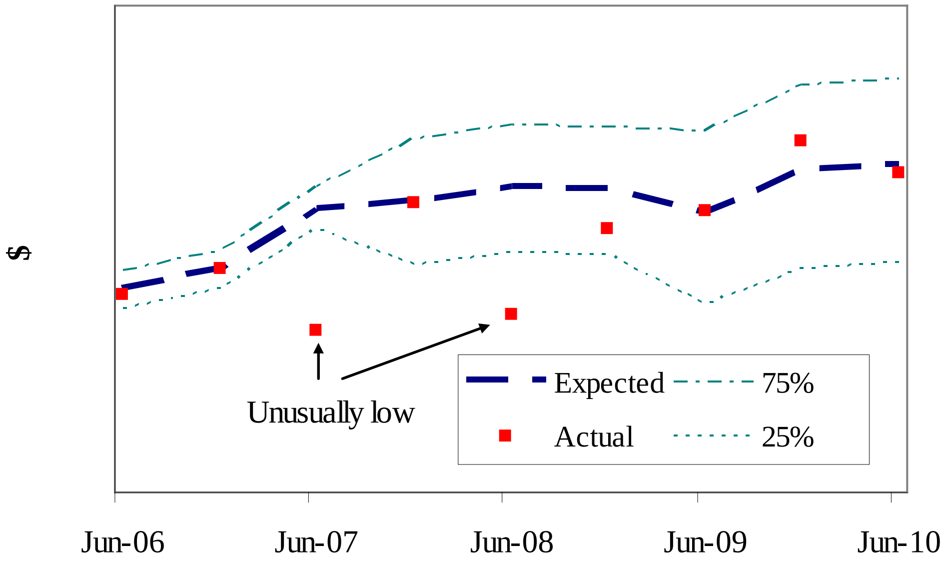

Figure 9: Historical performance of the risk margin for the outstanding claim valuations at June 2006 to June 2010 (data Oct-Mar and Apr-Sep)

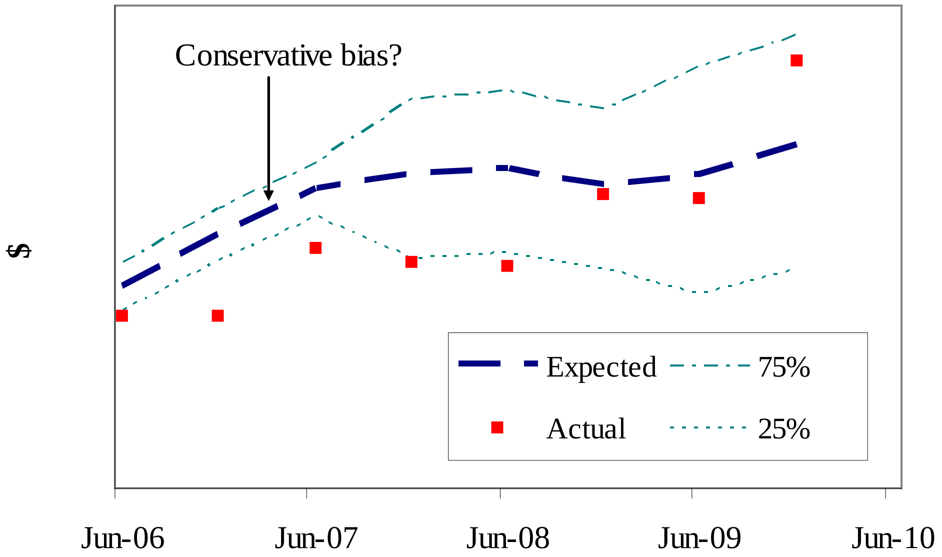

Figure 10: Historical performance of the risk margin for the outstanding claim valuations at June 2006 to June 2010 (data Jan-Jun and Jul-Dec)

As the business is long-tail, it may be useful to look at longer timeframes – two or three years (Figures 11 and 12). Using every second year gives three valuations with non-overlapping data (2004/5, 2006/7 and 2008/9 in Figure 11). Using every third year gives two valuations with non-overlapping data (2004-6 and 2007-9 in Figure 12). With these smaller numbers of valuations, there is not enough information yet to form any view about whether the current risk margin appears too high. There is certainly no indication that it is too low. However, the 2005/6 value is at the 98th percentile, and the 2005-7 value is at the 97th percentile, which suggests that the earlier risk margin (used for 2004 to 2007 valuations) might have been too low.

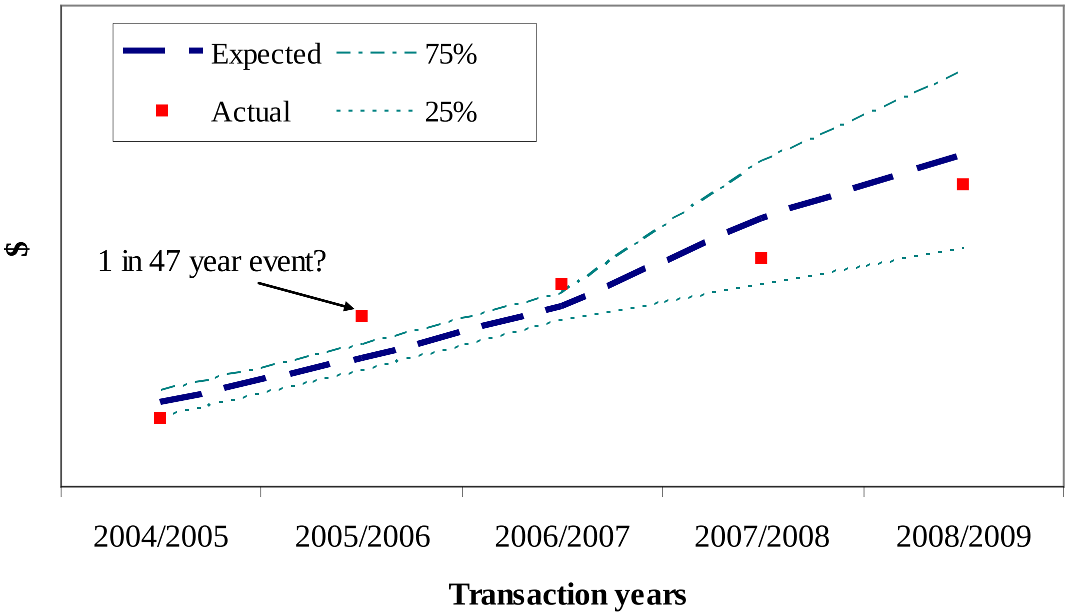

Figure 11: Historical performance of the risk margin for the outstanding claim valuations at 2005 to 2009 based on two years of data

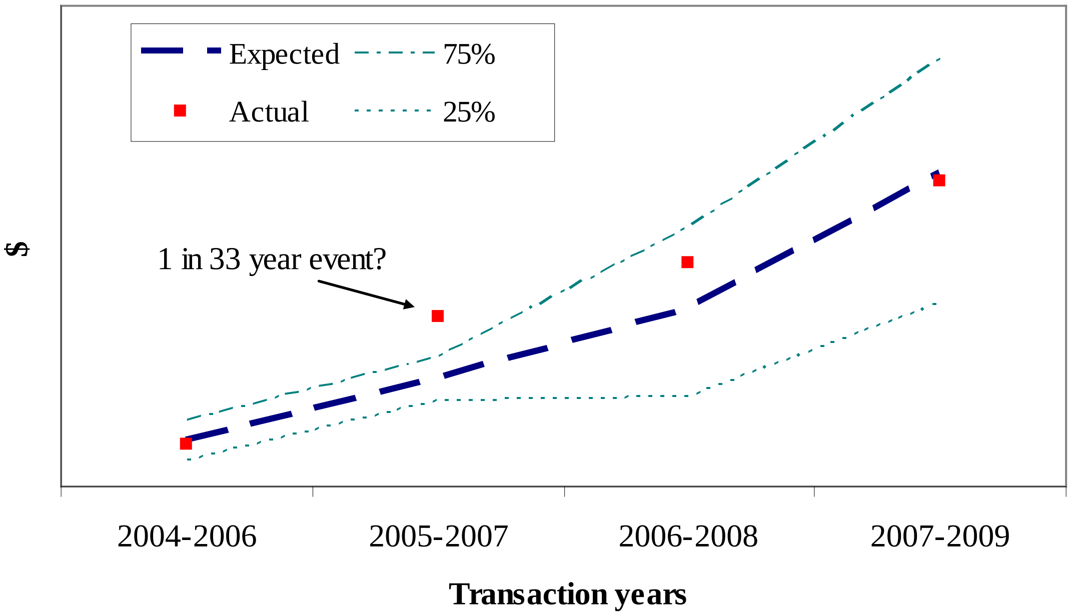

Figure 12: Historical performance of the risk margin for the outstanding claim valuations at 2006 to 2009 based on three years of data

Overall, there is some evidence that the 75th percentile was too small in the past but it appears to be (possibly more than) adequate now.

Conclusion

Using the risk margin to draw confidence limits around expected experience in the presentation of actual versus expected results in insurance liability valuation reports can provide useful information about whether the extent of deviation is within the bounds anticipated by the previous valuation basis.

Where a stochastic method has been used to determine the risk margin, this is likely to be straightforward to do, although judgement is required on the extent to which any subjective adjustments to the overall risk margin should be attributed to the transaction period following the valuation date.

Where the risk margin is based on mainly on judgement or industry standards, the conclusions are less clear-cut, particularly for long-tail lines. However, it is hoped that some rules of thumb can be developed to be used in such cases.

Acknowledgements

This research was support in part by an Australian Actuarial Research Grant from the Institute of Actuaries of Australia. Thanks are also due to Tim Brown and Susana Pui, who helped collate the data used in the examples.

References

Bateup R and Reed I, 2001, *Research and Data Analysis Relevant to the Development of Standards and Guidelines on Liability Valuation for General Insurance*, 13th General Insurance Seminar, The Institute of Actuaries of Australia.

Collings S and White G, 2001, *APRA Risk Margin Analysis*¸ 13th General Insurance Seminar, The Institute of Actuaries of Australia.

Houltram A, 2005, *An Empirical Approach to Insurance Liability Prediction Error Assessment*, 15th General Insurance Seminar, The Institute of Actuaries of Australia.

Mack T, 1993, *Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates*, Astin Bulletin, Vol 23, No 2, 213-225.

Mack T, 1999, *The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor*, Astin Bulletin, Vol 29, No 2, 361-366.

Marshall K, Collings S, Hodson M and O'Dowd C, 2008, *A Framework for Assessing Risk Margins*, 16th General Insurance Seminar, The Institute of Actuaries of Australia.

Appendix 1: Chain Ladder Bootstrap

There are many variations on the chain ladder bootstrap. One possible version is described below, together with the modifications required to perform the suggested test of the risk margin.

The risk margin can be estimated as follows:

1. A chain ladder model is fitted to the total paid loss triangle in current dollars. Any selected ratios can be used, but typically the "all" period weighted average is used, with no adjustments. Fitted cumulative values are calculated by applying the selected ratios backwards from the diagonal. Incremental fitted values are calculated from the cumulative values.
2. Scaled residuals are calculated for each cell as the actual incremental value minus the fitted incremental value, divided by the square root of the fitted incremental value. The variance of the scaled residuals is used to calculate the process variance, assuming that the process variance is proportional to the mean.
3. A random selection from the scaled residuals is rescaled and added to the fitted values to form a new "pseudo" dataset.
4. A chain ladder model is fitted to the new dataset. A choice must be made automatically of the number of periods and type of ratio to be used, typically a fixed number of periods. The projection of outstanding claims from this model gives one simulated IBNR value. This process is repeated many times, typically 1000 or 10,000 times.

5. The standard deviation of the IBNR values is used to calculate the parameter variance, allowing for the number of parameters estimated, typically assumed to be the number of accident periods plus the number of development periods less one. The total variance is the sum of the process and parameter variance.
6. The average of the IBNR values and the total variance are fitted to a normal and/or lognormal distribution. The 75th percentile of the distribution is calculated. The selected risk margin is based on the estimate(s) of the 75th percentile (or half the standard deviation if larger).

This method assumes:

1. Process variance is proportion to the mean. If diagnostic plots show this is inappropriate, alternative assumptions can be used, for example, variance proportional to mean squared.
2. Constant trend (possibly zero) in the calendar direction, in both the past and the future.
3. No changes in the development pattern over time.

The process above is modified at step 4, to produce one additional value: the projection of the next (or more than one) transaction period. This produces an additional series of values. Steps 5 and 6 are also applied to this series, giving a 75th percentile (and any other desired percentiles) of the probability distribution for the next transaction period's claims.

Appendix 2: Other Non-Parametric Bootstraps

Many standard actuarial models can be converted to stochastic models in a simple fashion, by replacing best estimate selections with a selection of a mean and standard deviation. To make the description more concrete, I will refer to the selections as ratios, but the same method can be applied to payments per claim or other model parameters.

1. For each development period, select a mean ratio and a standard deviation of ratios, based on historical data or any other method.
2. The means and standard deviations are used to scale the observed ratios in that development period, giving a triangle of scaled ratios.
3. A random sample (with replacement) is taken from the scaled ratios. This sample is used to generate a pseudo-triangle of ratios, which is used to create a pseudo-triangle. This pseudo-triangle is used to estimate new parameters for each development period.
4. Another random sample is taken from the scaled ratios. It is rescaled using the new parameters to give a future triangle of ratios. These ratios are used for projection to give one simulated IBNR value. This process is repeated many times, typically 1000 or 10,000 times.
5. The average of the IBNR values and the total variance are fitted to a normal and/or lognormal distribution. The 75th percentile of the distribution is calculated. The selected risk margin is based on the estimate(s) of the 75th percentile. Alternatively, the 75th percentile of the simulated values may be used directly.

This process can be applied to a number of related or independent models, in the same way that the central estimate of the outstanding claims is calculated.

This method typically assumes:

1. Constant trend (possibly zero) in the calendar direction, in both the past and the future.
2. No changes in the development pattern over time.

The process above is modified at step 4, to produce one additional value: the projection of the next (or more than one) transaction period. This produces an additional series of values. Step 5 is also applied to this series, giving a 75th percentile (and any other desired percentiles) of the probability distribution for the next transaction period's claims.

Appendix 3: Mack's Model

The usual Mack model is described in Mack (1993). Theorem 2 in that paper shows that the estimators of the ratios are uncorrelated, so the variance of the sum of the cells in the first future transaction period is the sum of the variances of the those cells. The calculation of those variances is mostly simply done with the recursive formulae in Mack (1999), using the same notation as in that paper ($\alpha = 1$ and $w_{ik} = 1$ for all i and k for the usual chain ladder):

$$s.e.(\hat{C}_{i,k+1})^2 = \hat{C}_{ik}^2\left(s.e.(F_{ik})^2 + s.e.(\hat{f}_k)^2\right) + s.e.(\hat{C}_{i,k})^2\,\hat{f}_k^2$$

with starting value $s.e.(\hat{C}_{i,n+1-i}) = 0$, $s.e.(F_{ik})^2 = \dfrac{\hat{\sigma}_k^2}{w_{ik}C_{ik}^{\alpha}}$ and $s.e.(\hat{f}_k)^2 = \dfrac{\hat{\sigma}_k^2}{\sum_{j=1}^{n-k} w_{jk}C_{jk}^{\alpha}}$ (1)

So, using (1) with $k=n+1-i$, the variance of the first future transaction period is:

$$\sum_{i=2}^{n} s.e.\left(\hat{C}_{i,n+2-i)}\right)^2 = \sum_{i=2}^{n} C_{i,n+1-i}^2\left(s.e.(F_{i,n+1-i})^2 + s.e.(\hat{f}_{n+1-i})^2\right) \qquad (2)$$

It only remains to make an assumption about the type of probability distribution to estimate the 75th percentile of the first future transaction period. A typical assumption would be lognormality.

For more than one future transaction period, there is correlation between the accident periods so there are additional terms to add to the sum of the variances of the accident periods. Following the same method as Mack (1993) uses for Theorem 3 and its Corollary, the variance of the sum of the first d future transaction periods can be shown to be:

$$\sum_{i=2}^{n} s.e.\left(\hat{C}_{i,\min(n,n+d+1-i)}\right)^2 + \sum_{i=2}^{n-1}\hat{C}_{i,\min(n,n+d+1-i)}\sum_{j=i+1}^{n}\hat{C}_{j,\min(n,n+d+1-j)}\sum_{k=n+1-i}^{\min(n-1,n+d-j)}\frac{s.e.(\hat{f}_k)^2}{\hat{f}_k^2} \qquad (3)$$

The first summation can be calculated iteratively using (1) d times for each i. Note that the final summation is empty when j is greater than $(d-1+i)$. As j is always greater than i, the final term is always zero when d=1, i.e. for a single future transaction period. So only the first sum is needed when d=1, and in that case (3) reduces to (2).

Appendix 4: Generalised Linear Models

For a lognormal generalised linear model, it is possible to produce an analytical formula for the mean and standard deviation of the sum of any combination of future cells. This means that it is only necessary to make an assumption about the type of probability distribution to estimate the 75th percentile of any number of future transaction years. Alternatively, simulation can be used as described below.

For other types of generalised linear models, one method of producing the probability distribution of the outstanding claims is to do a parametric bootstrap:

1. The output of the fitted model should contain a covariance matrix for the parameters. Assuming the joint distribution of the parameters is multivariate normal, generate a random sample of the parameters. Use that sample to forecast means of the future triangle.
2. The output should also contain an estimate of the process variance. Use the process variance, the cell mean and the assumed cell probability distribution to generate a random value in each cell of the future triangle. Sum the values in the future triangle to get the total IBNR. This process is repeated many times, typically 1000 or 10,000 times.
3. The average of the IBNR values and the total variance are fitted to a normal and/or lognormal distribution. The 75th percentiles of the distribution is calculated. The selected risk margin is based on the estimate(s) of the 75th percentile. Alternatively, the 75th percentile of the simulated values may be used directly.

This method typically assumes:

1. No changes in the development pattern over time.

The process above is modified at step 2, to produce one additional value: the sum of the random values in the next (or more than one) transaction period. This produces an additional series of values. Step 3 is also applied to this series, giving a 75th percentile (and any other desired percentiles) of the probability distribution for the next transaction period's claims.

Appendix 5: Adaptive Generalised Linear Models

Adaptive generalised linear models allow the development pattern to change over time, i.e. the shape of the development may be different for different accident periods. A functional form is specified for the development pattern, with as many parameters as are needed to give a good fit to the data. The fitted parameters will vary gradually with accident period.

Purpose-built software is required for this type of model. The probability distribution of outstanding claims is produced by simulation, which typically will produce a value for each future cell. Rescaling may be done to match the overall average of the values in each cell to the central estimate from the valuation, if different. The total of all cells gives the IBNR. The selected risk margin is based on the estimate(s) of the 75th percentile of the simulated values.

The process above is modified by summing only the cells in the required future transaction periods. The selected risk margin is based on the estimate(s) of the 75th percentile of the simulated values.