

Introduction to Reserving

Patrik Dahl

Corrected edition, 2003

Introduction to reserving

0. Preface

The following text is just an introduction to reserving methods. In practice reserving will contain important non-mathematical elements. The terminology will often use the word "reserve". In Europe the accounting terminology is nowadays "provisions", maybe even "technical provisions". For reasons of tradition and ease, the former word is most often used in the sequel. The reader that wish to become a "cognoscente" should follow the ongoing discussion within Casualty Actuarial Society, where both practitioners and scholars take active part. Those who believe that all these issues must have been settled long ago will be in for a surprise. There is a list of symbols used, section 5, and a vocabulary, section 14.

1. Introduction

What is a non-life insurance from a financial perspective? Briefly: For a premium an insurance company commits itself to pay a sum if an event has occurred. If we introduce a time axis, we find that first the policyholder signs up for an insurance, then pays a premium and when received by the insurance company, the company starts to earn the premium. During the duration of the policy, as premiums are earned, there might or might not occur a claim. If a claim has occurred, it will eventually be known by the insurer. When the claim is known by the insurer, the insurer reserves the claim and later possibly pays out an amount. Schematically,

Premiums written --> Premiums paid --> Premiums reserved --> Premiums earned --> Claims incurred --> Claims reported --> Claims paid

There are several problems to solve:

- How much premium is earned?
- How much premium is unearned?
- How do we measure the number and size of unknown claims?
- How do we know if the reserves on known claims are sufficient?

The device that solves the two first problems are traditionally called premium reserve. The solution to the two last problems are called "incurred but not reported" reserve, or, IBNR reserve. (sometimes there is a split and we talk about totally unknown claims, "incurred but not yet reported", IBNYR, and "incurred but not enough reported", IBNER, when reported reserves are believed to be insufficient.)

An insurance company has two main reasons for finding out how large the claims on written business are. First, and most important, to feed back this into the pricing. The second reason is to produce financial statistics for analysis and to produce income statements and balance sheets for the company. The value of a correct balance sheet could be found in accounting theory. We will not deal with how to feed back the information into pricing, but concentrate on the estimates produced by a few established methods.

2. Premium reserve

The premium reserve is split into two parts that in accounting terminology are called:

- Provision for unearned premiums
- Provision for unexpired risks

To start with the first of these, it is assumed that written premiums are earned evenly/uniformly over the cover period. If we are inside this period, then the share of the premium that has been earned is the past time's proportion of the total period. This way of apportioning is called "pro rata temporis" (lat.).

If a larger premium has been received, the difference is then the unearned premium. This belongs to the provision for unearned premium.

Example:

Suppose we have an insurance policy starting on 1 September going through 31 August next year. The premium for the total period is 2400. At 31 December we have received two quarterly premiums, or 1200. We have then earned $(4/12) * 2400 = 800$. What is unearned of the received premium is then $1200 - 800 = 400$.

To understand the second part, the unexpired risk reserve, we look upon the whole period covered by the insurance. From a point in time we look forward to all the claims and expenses that could occur after this point. The insurer should reserve funds to cover the expected value of them. If there are future premiums not yet due, these could be deducted. If this amount is larger than the amount given by the pro rata temporis calculation, the difference should be accounted as "unexpired risk reserve". In North America it is, maybe more appropriately, called "premium deficiency reserve". To return to our previous example, assume we at 31 December believe the future costs are 1800. We have two quarterly premiums not yet due or received, each on 600. Thus, we need to have $1800 - 2 * 600 = 600$ in provision for the remaining exposure. As this is higher than the previously calculated 400, we need an unexpired risk reserve of the difference, i.e. 200.

The premiums an insurer receives may contain variable acquisition expenses. According to the law within EU, they should be allocated in time in the same way as the premium. If these acquisition expenses are contingent on an uncertain future, expected values should be used.

3. Claims reserves

In the sequel some methods to calculate the ultimate claims or provision for IBNR will be presented. Before using these methods, not only statistical considerations should be taken. It is important to find out the purpose of the figure. If it should be used in pricing, it should be realistic, only a small bias should be allowed (before loadings). If used in transferring a portfolio of policies, it should also be realistic, but the sign of the bias will depend if it is evaluated for the seller or the buyer. When used to obtain provisions for financial statements, some conservative bias is allowed. Furthermore, it should agree with accounting principles and applicable laws.

One will have to determine if it is gross, i.e. before reinsurance, or net, i.e. after reinsurance, provisions that should be calculated. It is not generally obvious which way to go. We have the equation $(\text{Net}) = (\text{Gross}) - (\text{Ceded})$. But you will probably get different results depending on which two of these three you evaluate. The methods presuppose, or at least favour, homogeneity. This lack of additivity also makes the subdivision of the portfolio of an insurance company a delicate matter. This will be shown in exercise 6.1.

Before using any method one should make appropriate adjustment for inflation, if that is not otherwise considered in the model. It is important that the relevant rate of inflation is used. In fire insurance it could be building cost indices. If there are awards for personal injury, they more often follow the general development of wealth in the society. Furthermore, there is also a trend from the idea that the unfortunate should not be left in poverty towards compensating people for what they might have become if the injury had not happened. That could result in a social or superimposed inflation even higher than both wage cost indices or GNP development in running prices. Self-evidently it is hard to forecast such inflation and immunizing it with the right investment will hardly be possible. If the cost of what is replaced depends on another currency, the exchange rate change should also be considered. It should also be remembered that inflation will have an impact on the number of claims, as deductibles do not usually move in pace with inflation.

The methods assume that all policies in question have the same period of exposure. This is sometimes not true. However, it is usually not a good idea to make the simple adjustment of letting all policies start at the same time, as perils will often vary by season and reporting might also depend on season. The best is often, if possible, to work on accident year.

Not only factors outside the insurance company affects the timing and size of claims. There is usually a claims department too. It is of great importance to understand the procedures of the claims department and its staff. How do they set the provisions? When do they set the provision? How long does it take to set the reserve from the reporting of a claim to the recording of a provision? How will the lead time be affected by the size of the claim? Are there any backlogs? Are there any vacancies? When do they change the reserves? When and why do they review the reserves? Is there any, official or unofficial, pressure to keep down not only the payments but also the case reserves? Have there been any changes in procedures? Have there been any changes in staff? Claims people usually talk about open and closed claims. The latter category should actuarially be thought of as “not yet reopened claims”.

The provisions set by the claims people are approximate. Thus, it could be tempting to disregard them and just go on the hard facts, the payments. However, most experience speaks for also, and maybe foremost, using incurred claims, i.e. both the paid and the reserved amount. This does not mean to say that an analysis on paid should not be performed. However, there is no obvious way to reconcile analysis by paid and incurred, respectively, in the general case.

Often the words “long-tailed claims” are mentioned. By this people do not always mean the same thing. They could mean that the claim is reported very late or that it takes a long time to finally settle or pay it after it becomes known.

There could also be changes in the legal environment and in sentiments of courts and society, where events that was not thought to be covered at the time of policy issuance, later is considered to be covered.

It could be reasonable to take out some large claims and estimate them separately, as they might develop differently. If this is done, there should be good criterias for doing it, such as that it will be recovered by an excess-of-loss reinsurance. To just take them out and ignore them by labeling them as “outliers” is unwise.

Expense reserve

An insurance company should also have the funds to handle the claims in the future, in order for the policyholder to receive their rightful amounts even if the company stops to write business. The calculation of future handling expenses is more of an exercise for accountants. But when it is calculated, also the unknown claims should be taken into consideration. In North America they use the concept unallocated loss adjustment expenses, ULAE, for the IBNR claims as opposed to allocated loss adjustment expenses, ALAE, for the known claims.

4. About the methods

The chain ladder method could be said to build purely on past experience. The Bornhuetter-Ferguson method builds on exposure. The Cape Cod method is basically the same as the Bornhuetter-Ferguson method, but uses claims experience to replace the a priori loss ratio. The Benktander/Hovinen Method tries to make a credibility compromise between the chain ladder method and the Bornhuetter-Ferguson method by weighing them together with the assumed proportion known and unknown claims.

The chain ladder method, or versions thereof, has been in use for decades. The presentation given here is based on the article [Mack 1994] by Thomas Mack. The Bornhuetter-Ferguson method is named after two US actuaries and was originally presented in 1975 in [Bornhuetter&Ferguson 1975]. The presentation given here leans on a presentation by [Gluck 1997]. The Cape Cod method was invented independently by Jim Stanard and Hans Bühlmann and is in North America often called the Stanard-Bühlmann method. This presentation is based upon [Patrik 1996] and [Gluck 1997]. The Benktander method, or, Hovinen method is named after Gunnar Benktander [Benktander 1976] and Esa Hovinen [Hovinen 1981], who independently invented it. The separation method was formulated by Greg Taylor [Taylor 1977], on which this presentation builds.

5. Notation

Usually we present the data in the form of a claims triangle. The rows would represent accident years and the columns development periods. If the development periods are years the diagonals from the upper right corner to the lower left corner would represent calendar years. An example of a claims triangle on a cumulative basis, with three accident years and three development periods would thus look like

	1	2	3
1998	C_{11}	C_{12}	C_{13}
1999	C_{21}	C_{22}	
2000	C_{31}		

The following notation will be used in one or more places in the sequel:

$C_{i j}$ Cumulative claims from accident year i , reported through the end of period j .

$D_{i j} = C_{i j} - C_{i j-1}$ Incremental claims from accident year i , reported in period j .

$C_{i m}$ Ultimate claims, where

m the last development period that is known

$R_i = E[C_{i m}] - C_{i j}$ Reserve

f_j One period loss development factor. Also called age-to-age factor or link ratio.

$F_{i j}$ Development factor from accident year i , period j , to ultimate.

L_i Claims relative to an exposure

P_i A measure of exposure

A_k Experience up to development period k

The following notation relates only to section 11 and is more precisely defined there:

c a single claim amount

I_{i+j} A calendar year factors

q_j cumulative claims through period j to total claims for one accident year.

r_j incremental claims in period j to total claims for one accident year. (They sum to unity.)

d_i diagonal sum

N_i Number of claims for accident year i

$n_i = \hat{E}[N_i]$ An estimate of the expected number of claims from accident year i

$B_{i j}$ defined in section 11.

We will use circumflex ($\hat{\ }$) to denote an estimate.

6 The Chain Ladder Method

The chain ladder method builds on that cumulative claims in a period are proportional to the claims in the preceding period. The proportionality factor depends on the number of periods since outset, but is expected to be the same for all accident years.

More formally we assume

$$(CL1) E[C_{i,j+1} | C_{i1}, C_{i2}, \dots, C_{ij}] = C_{ij} * f_j$$

Observe that f_j does not depend on accident year.

(CL2) The vectors $\{C_{i1}, C_{i2}, \dots, C_{im}\}$ and $\{C_{k1}, C_{k2}, \dots, C_{km}\}$ are independent if $i \neq k$

CL1 just brings us one step ahead, whereas we want to get to the end. To get there we are going to utilize the following wellknown result:

LEMMA 6.1

$$\text{If } E[Z] \text{ is finite, then } E[Z] = E[E[Z | X]] \quad (6.1)$$

Using this lemma and CL1, we find

$$\begin{aligned} E[C_{i,j+k} | C_{i1}, \dots, C_{ij}] &= E[E[C_{i,j+k} | C_{i1}, \dots, C_{i,j+k-1}] | C_{i1}, \dots, C_{ij}] = \\ &= E[C_{i,j+k-1} * f_{j+k-1} | C_{i1}, \dots, C_{ij}] = E[C_{i,j+k-1} | C_{i1}, \dots, C_{ij}] * f_{j+k-1} = \\ &= E[E[C_{i,j+k-1} | C_{i1}, \dots, C_{i,j+k-2}] | C_{i1}, \dots, C_{ij}] * f_{j+k-1} = \\ &= E[C_{i,j+k-2} * f_{j+k-2} | C_{i1}, \dots, C_{ij}] * f_{j+k-1} = \\ &= E[C_{i,j+k-2} | C_{i1}, \dots, C_{ij}] * f_{j+k-2} * f_{j+k-1} = \dots = C_{ij} * f_{j+1} * f_{j+2} * \dots * f_{j+k-1} \end{aligned} \quad (6.2)$$

The formula suggests a procedure and we shall indeed show that it could be used.

We could rewrite CL1 on the following form

$$(CL1') E[C_{i,j+1} / C_{ij} | C_{i1}, \dots, C_{ij}] = f_j$$

Thus, we could use observed ratios $C_{i,j+1}/C_{ij}$ as unbiased estimators of f_j . Before combining estimates of the same f_j we make a further assumption,

$$(CL3) \text{Var}[C_{i,j+1} | C_{i1}, C_{i2}, \dots, C_{ij}] = C_{ij} * \mathbf{s}_j^2$$

Observe that the last factor in the variance is not depending on accident year. We will also use the following lemma,

LEMMA 6.2

Suppose X_i are n uncorrelated random variables with the same mean, but with variances \mathbf{s}_i^2 . Then the best linear unbiased estimator of the mean is given by

$$\sum_{i=1}^n w_i * X_i \quad (6.3)$$

where $w_i \propto \mathbf{s}_i^{-2}$ and $\sum_{i=1}^n w_i = 1$

Proof: (sketch) Form the Lagrangian

$$L(w_1, \dots, w_n, I) = \sum_{i=1}^n w_i^2 \mathbf{s}_i^2 - I(1 - \sum_{i=1}^n w_i) \quad (6.4)$$

Solve the system

$$\begin{cases} \frac{\partial}{\partial w_k} L(w_1, \dots, w_n, \mathbf{I}) = 0 & k = 1, \dots, n \\ \frac{\partial}{\partial \mathbf{I}} L(w_1, \dots, w_n, \mathbf{I}) = 0 \end{cases} \quad (6.5)$$

$$\text{which gives } w_i = \mathbf{s}_i^{-2} / \sum_{k=1}^n \mathbf{s}_k^{-2} \quad (6.6) \blacksquare$$

Rewriting CL3 gives (why?)

$$(\text{CL3}') \text{Var}[C_{i,j+1}/C_{i,j} | C_{i1}, \dots, C_{ij}] = \mathbf{s}_j^2 / C_{ij}$$

and the weights are thus by 6.6

$$w_i = (\mathbf{s}_j^2 / C_{ij})^{-1} / \sum_{k=1}^{m-j} (\mathbf{s}_j^2 / C_{kj})^{-1} = C_{ij} / \sum_{k=1}^{m-j} C_{kj} \quad (6.7)$$

and

$$\hat{f}_j = \sum_{i=1}^{m-j} w_i \hat{f}_{ij} = \sum_{i=1}^{m-j} \left\{ \frac{C_{ij}}{\sum_{k=1}^{m-j} C_{kj}} * \frac{C_{i,j+1}}{C_{ij}} \right\} = \frac{\sum_{i=1}^{m-j} C_{i,j+1}}{\sum_{i=1}^{m-j} C_{i,j}} \quad (6.8)$$

To be able to use the algorithm suggested by formula (6.2) with the estimators from (6.8) we need to prove that the estimates are uncorrelated. Define the set of experience up to development period k by

$$A_k = \{C_{ij} | j \leq k, i \leq m\}$$

Then we have

$$E[\hat{f}_j * \hat{f}_k] = E[E[\hat{f}_j * \hat{f}_k | A_k]] = \quad (6.9a)$$

$$= E[\hat{f}_j * E[\hat{f}_k | A_k]] = \quad (6.9b)$$

$$= E\left[\hat{f}_j * E\left[\left(\frac{\sum_{n=1}^{m-k} C_{n,k+1}}{\sum_{n=1}^{m-k} C_{n,k}}\right) | A_k\right]\right] = \quad (6.9c)$$

$$= E\left[\hat{f}_j * \left(\frac{\sum_{n=1}^{m-k} E[C_{n,k+1} | A_k]}{\sum_{n=1}^{m-k} C_{n,k}}\right)\right] = \quad (6.9d)$$

$$= E\left[\hat{f}_j * \left(\frac{\sum_{n=1}^{m-k} E[C_{n,k+1} | C_{n,k}]}{\sum_{n=1}^{m-k} C_{n,k}}\right)\right] = \quad (6.9d)$$

$$= E\left[\hat{f}_j * \left(f_k * \left(\frac{\sum_{n=1}^{m-k} C_{n,k}}{\sum_{n=1}^{m-k} C_{n,k}}\right)\right)\right] = \quad (6.9e)$$

$$= E[\hat{f}_j] * f_k = \quad (6.9f)$$

$$= E[\hat{f}_j] * E[\hat{f}_k] \quad (6.9g)$$

Where we repeatedly used the lemma 6.1, the CL1 assumption, the estimator (6.8), and that we could take out what is known.

This just proves that the two different estimates are uncorrelated. But the reader will observe that we never did anything with f_j . In fact, f_j could be replaced by a product of several f 's and we could repeat the procedure that we used for f_k for them one by one, mutatis mutandum.

If we combine this with (6.2) it shows that the following ultimate estimator is unbiased,

$$\hat{E}[C_{i_m} | C_{i_j}] = C_{i_j} * \hat{f}_j * \dots * \hat{f}_{m-1} \quad (6.10)$$

Example

	1	2	3
1998	30	50	65
1999	40	90	
2000	55		

$$f_1 = \frac{50 + 90}{30 + 40} = 2,00$$

$$f_2 = \frac{65}{50} = 1,30$$

Ultimate year 1999: $90 * 1,30 = 117$

Reserve year 1999: $117 - 90 = 27$

Ultimate year 2000: $55 * 2,00 * 1,30 = 143$

Reserve year 2000: $143 - 55 = 88$

Reserve both years: 115

Observandum

When we use the product of the factors, it is made up from the last factors. If we invert this cumulative factor we will get the percentage that is reported. To see this, you could think of the ultimate as 100%, how do we get there? By multiplying with the last factor, hence by dividing 100% by the last factor we get the percentage reported up to this period. This is a convenient notion when communicating with non-actuaries. In the example you will find that we thus have $1/(2,00 * 1,30) = 38,5\%$ reported after one year and $1/1,30 = 76,9\%$ after two years.

Exercises

6.1 Generalize the assumption (CL3), so that the conditional variance is equal to $C_{i_j}^c \mathbf{s}_j^2$ for a $c > 0$.

Determine the best linear estimation \hat{f}_j . Examine especially the cases $c = 0, 1, 2$.

6.2 Determine the variance of the chain ladder estimation. Express it as an expected value (which we cannot calculate). Hint: Condition on the same A_k as on page 7.

7. The Naive Loss Ratio Method

This method assumes that we a priori know the ultimate losses share of the premium, P_i . This is usually referred to as the ultimate loss ratio, L_i . (One could, of course, use some other, preferably better, measure of exposure, but this is the standard one). How this share is known is outside the

method, it could come from the pricing calculations or from “guesstimates” by e.g. account executives or fire engineers according to their experience (the infamous “underwriting judgment”),

$$\hat{E}[C_{i m}] = L_i * P_i \quad (7.1)$$

Thus, the necessary IBNR reserve will be the difference between the ultimate losses and the reported claims,

$$R_{i j} = L_i * P_i - C_{i j} \quad (7.2)$$

It is obvious that this method does not presuppose anything about the claims location in time, nor does it differentiate between actual claims or expected claims, it simply sees them as communicating vessels. It is true that this method is simplistic and have its most proponents among the “practical men”. It has limited value outside the case in the early life of an accident year when just a few and small claims are known.

8. The Bornhuetter-Ferguson Method

The Bornhuetter-Ferguson method is more sophisticated than the Naïve Loss Ratio method. It looks on where in time claims will be reported or paid. It is very similar to an ordinary budgeting model used by businesses. You could say that you budget for future claims by period. The sum of these future budgeted claims is the IBNR reserve. As time goes on, estimated claims for past periods are replaced by outcomes without affecting the estimate of future claims. This could also be cast in a statistical framework.

More formally, the following principles apply:

(BF1) Expected claims are considered known, i.e. we have a predictor of final claims identical to the mean $\hat{C}_{i m}^N = E[C_{i m}]$.

(In practice $\hat{C}_{i m}^N$ could be computed by the Naïve Loss Ratio method.)

(BF2) Unemerged claims are independent of emerged claims, or,

$$C_{i j} \text{ is independent of } C_{i m} - C_{i j}$$

Let $F_{i j}$ be the factor that would develop losses from development period j to the end for accident year i .

(BF3) The $F_{i j}$ are known, in the meaning that we know $F_{i j} = \frac{E[C_{i m}]}{E[C_{i j}]}$

How the $F_{i j}$ are determined is outside the method, but in practice the $F_{i j}$ could have been determined by the Chain Ladder technique.

Now it is readily seen that the following is an unbiased predictor of final claims.

$$\hat{C}_{i m}^{BF} = C_{i j} + \left(1 - \frac{1}{F_{i j}}\right) * \hat{C}_{i m} \quad (8.1)$$

This predictor, known as the Bornhuetter-Ferguson method has the merit over the Naïve Loss Ratio method that it takes emerged claims into account as it swaps past expected with real emergence.

We could also rewrite (8.1) on the following form,

$$\begin{aligned}\hat{C}_{im}^{BF} &= \frac{1}{F_{ij}} * C_{ij} * F_{ij} + \left(1 - \frac{1}{F_{ij}}\right) * \hat{C}_{im}^N = \\ &= W_{ij} * C_{ij} * F_{ij} + (1 - W_{ij}) * \hat{C}_{im}^N\end{aligned}\quad (8.2)$$

With $W_{ij} = 1/F_{ij}$.

We find that this is a weighting of of a chain ladder-type estimate and the 'known' expected claims.

We make a further assumption,

$$(BF4) \text{Var}[C_{ij} * F_{ij}] = F_{ij} * \text{Var}[C_{im}]$$

Theorem

The weights implicitly defined in (8.2) produces the best combination of the two predictors

$$C_{im}^{CL} = C_{ij} * F_{ij} \text{ and } \hat{C}_{im}^N \text{ (in the meaning of minimizing the quadratic loss)}$$

Proof

We shall find the weights W_{ij} that solves the problem

$$\min_{W_{ij}} E \left[\left(C_{im} - \left(W_{ij} * (C_{im} - C_{ij} * F_{ij}) + (1 - W_{ij}) * (C_{im} - \hat{C}_{im}^N) \right) \right)^2 \right] \quad (8.3)$$

Due to the unbiasedness this is equivalent to minimizing the following variance,

$$\text{Var} \left[\left(W_{ij} * C_{ij} * F_{ij} + (1 - W_{ij}) * \hat{C}_{im}^N \right)^2 \right] \quad (8.4)$$

To show the uncorrelatedness of the components we need the auxillary result,

$$\begin{aligned}\text{Cov}[-C_{ij} * F_{ij}, C_{im}] &= -F_{ij} * \text{Cov}[C_{ij}, C_{im}] = \\ &= -F_{ij} * \text{Cov}[C_{ij}, (C_{im} - C_{ij}) + C_{ij}] = \\ &= -F_{ij} * \text{Cov}[C_{ij}, (C_{im} - C_{ij})] - F_{ij} * \text{Cov}[C_{ij}, C_{ij}] = \\ &= -F_{ij} * 0 - F_{ij} * \text{Var}[C_{ij}] = \\ &= -F_{ij} * \text{Var}[C_{im}] / F_{ij} = -\text{Var}[C_{im}] := -v^2\end{aligned}\quad (8.5)$$

Using the calculation rules for covariances and remembering that covariances between a r.v. and a constant vanishes, we find

$$\begin{aligned}\text{Cov}[C_{im} - C_{ij} * F_{ij}, C_{im} - \hat{C}_{im}^N] &= \\ &= \text{Cov}[C_{im}, C_{im}] + \text{Cov}[-C_{ij} * F_{ij}, C_{im}] = v^2 - v^2 = 0\end{aligned}\quad (8.6)$$

Which shows the uncorrelatedness. Calculating the variance of the components gives,

$$\text{Var}[C_{ij} - \hat{C}_{im}^N] = v^2 \quad (8.7)$$

and

$$\begin{aligned}
& \text{Var}[C_{im} - C_{ij} * F_{ij}] = \\
& = \text{Var}[C_{im}] + F_{ij}^2 * \text{Var}[C_{ij}] + 2 * \text{Cov}[C_{im}, -C_{ij} * F_{ij}] = \\
& = v^2 + F_{ij}^2 * v^2 / F_{ij} - 2 * v^2 = v^2 * (F_{ij} - 1) \quad (8.8)
\end{aligned}$$

By using Lemma 6.2 we find that the optimal weights are

$$W_{ij} = \frac{\frac{1}{v^2 * (F_{ij} - 1)}}{\frac{1}{v^2 * (F_{ij} - 1)} + \frac{1}{v^2}} = \frac{1}{F_{ij}} \quad (8.9)$$

which is what was asserted. ■

It should be noted from (8.6) that we should have $F_{ij} \geq 1$.

As in the Naïve Loss Ratio Method, section 7, it is often convenient to express $\hat{E}[C_{im}]$ as an a priori loss ratio times a premium, cf formula 7.1.

Example

We have the same claims triangle as in the chain ladder example, but also supplemented with risk premium.

	Premiums	1	2	3
1998	70	30	50	65
1999	115	40	90	
2000	140	55		

Assume that we think that expected claims to risk premium should be 100% for both 1999 and 2000. Let us also accept that 38,5% is expected to be reported through period one and 76,9% through period two (cf the observandum at the end of section 6) for both years.

Then we have

$$\text{Reserve 1999: } (1 - 0,769) * 1,00 * 115 = 26,6$$

$$\text{Ultimate 1999: } 90 + 26,6 = 116,6$$

$$\text{Reserve 2000: } (1 - 0,385) * 1,00 * 140 = 86,1$$

$$\text{Ultimate 2000: } 55 + 86,1 = 141,1$$

Reserve both years: 112,7

Exercises

8.1 Prove that if X and Y are random variables with existing variances and if X and (Y-X) are uncorrelated, then $\text{Cov}[X, Y] = \text{Var}[X]$

8.2 By introducing the notation, $D_{ij} = C_{ij} - C_{i,j-1}$, prove that, given i, $\text{Var}[D_{ij}] \propto E[D_{ij}]$ and

$$\text{that } \text{Var}[C_{im}] = \sum_{j=1}^m \text{Var}[D_{ij}].$$

8.3 Suppose we have estimates of F_{ij} , for all j, and we think we could estimate the C_{im} (e.g. by the naïve loss ratio method). Determine the formulas for the annual costs $D_{ij}, j=1, \dots, m$.

9. The Cape Cod Method

This method is similar to the B-F method. Instead of requiring an a priori loss ratio, it estimates one with the help of a measure of exposure and claims to date.

Let us return to formula (8.1). By using (7.1) we could rewrite it as

$$C_{i\ m-i+1} + \left(1 - \frac{1}{F_{i\ m-i+1}}\right) * \hat{E}[C_{i\ m}] = C_{i\ m-i+1} + \left(1 - \frac{1}{F_{i\ m-i+1}}\right) * L_i * P_i \quad (9.1)$$

We now assume that the $F_{i\ j}$ do not depend on accident year, i.e. $F_{i\ j} = F_j$

Thus, an IBNR for all years is given by

$$\sum_{i=1}^m IBNR_i = \sum_{i=1}^m \left(1 - \frac{1}{F_{m-i+1}}\right) * L_i * P_i \quad (9.2)$$

Now, assume that $L_i = L$ for all i , i.e.

$$\sum_{i=1}^m IBNR_i = L * \sum_{i=1}^m \left(1 - \frac{1}{F_{m-i+1}}\right) * P_i \quad (9.3)$$

At the same time we have

$$\sum_{i=1}^m C_{i\ m-i+1} + \sum_{i=1}^m IBNR_i = L * \sum_{i=1}^m P_i \quad (9.4)$$

From (9.3) and (9.4) we could conclude

$$\hat{L} = \frac{\sum_{i=1}^m C_{i\ m-i+1}}{\sum_{i=1}^m (P_i / F_{m-i+1})} \quad (9.5)$$

(This straightforward exercise is left to the reader)

A closer look will reveal that we instead of blowing up the claims as in chain ladder, we instead apportion the premiums.

We could rewrite formula (9.5) as

$$\hat{L} = \frac{\sum_{i=1}^m \{(C_{i\ m-i+1} * F_{m-i+1} / P_i) * (P_i / F_{m-i+1})\}}{\sum_{i=1}^m (P_i / F_{m-i+1})} \quad (9.5')$$

We could interpret this as an indication of a variance structure, c.f. Lemma 6.2. Thus, formula (9.5') says the variance is proportional to how far we are from the ultimate, measured on a development scale, and inversely proportional to the exposure. This will further be commented upon in chapter 12.

Example

The data is still the same:

	Premiums	1	2	3
1998	70	30	50	65
1999	115	40	90	
2000	140	55		

We will also use 38,5% as the figure expected to be reported through period one and 76,9% through period two, of the observandum at the end of chapter 6.

$$\hat{L} = \frac{65 + 90 + 55}{70 * 1,00 + 115 * 0,769 + 140 * 0,385} = 0,989$$

Reserve 1999: $(1 - 0,769) * 0,989 * 115 = 26,3$

Ultimate 1999: $90 + 26,3 = 116,3$

Reserve 2000: $(1 - 0,385) * 0,989 * 140 = 85,2$

Ultimate 2000: $55 + 85,2 = 140,2$

Reserve both years: 111,5

Exercise

9.1 Verify formula (9.5)

10. The Benktander Hovinen Method

An advantage with the Bornhuetter-Ferguson method compared to the chain ladder method is that it does not let early claims drive the reserve. In fact, the reserve does not take them into account at all. If a year has developed quite differently from the a priori expected, it does not seem wise to ignore that this might have a bearing on the reserves. The general experience also tells us that things get more stable over time and accident years tend to develop more alike, which is the primary assumptions in the chain ladder method (CL 1). Thus, a weighting, credibility theory style, of the estimates would be tempting:

$$R_w = w * R_{CL} + (1 - w) * R_{BF} \quad (10.1)$$

Moreover, we wish the weight, w , to increase by time. Instead of selecting time on a calendar scale, we select it as the expected proportion of known claims. This gives the formula:

$$R_{p_k} = q_k * R_{CL} + (1 - q_k) * R_{BF} \quad (10.2)$$

This weighting could not be justified by Lemma 6.2, as there is no common variance structure. The value of this pragmatic method will further be discussed in section 12.

Example

We use the figures from the chain ladder and Bornhuetter-Ferguson example.

Reserve 1999: $0,769 * 27,0 + (1 - 0,769) * 26,6 = 26,9$

Ultimate 1999: $90 + 26,9 = 116,9$

Reserve 2000: $0,385 * 88,0 + (1 - 0,385) * 86,1 = 86,8$

Ultimate 2000: $55 + 86,8 = 141,8$

Reserve both years: 113,7

11. The Separation Method

The separation method assumes that incremental claims are products of factors that depend on the accident year, the development year and the calendar year.

Let

c the mean single claim amount

i the accident year

j the development year

r_j the expected proportion of claims development year j , if no calendar year effect exists. Thus, the r_j 's sums to unity.

I_{i+j} the calendar year effect factor (e.g. inflation)

n_i is the number of claims for the accident year i . (The estimation of n_i is outside this method, but one method could be found in exercise 11.1)

then

$$E[D_{ij}] = c * n_i * r_j * I_{i+j} \quad (11.1)$$

$$\frac{E[D_{ij}]}{n_i} = c * r_j * I_{i+j} \quad (11.1')$$

This last expression could be seen as entry (i,j) in a triangle. For this triangle we define the diagonal sums

$$d_0 = c * r_0 * I_0$$

$$d_1 = c * r_0 * I_1 + c * r_1 * I_1 = c * (r_0 + r_1) * I_1$$

$$d_2 = c * r_0 * I_2 + c * r_1 * I_2 + c * r_2 * I_2 = c * (r_0 + r_1 + r_2) * I_2$$

...

$$d_{k-1} = c * (r_0 + r_1 + \dots + r_{k-1}) * I_{k-1} = c * (1 - r_k) * I_{k-1}$$

$$d_k = c * (r_0 + r_1 + \dots + r_k) * I_k = c * I_k$$

We have an observed triangle with entries D_{ij} , we divide each row with the predicted \hat{n}_i . Define

$$B_{ij} = D_{ij} / \hat{n}_i.$$

We can now form observed diagonal sums d_k . Starting from the last equation we could recursively

calculate $c * \hat{I}_{i+j}$ and \hat{r}_j

$$c * \hat{I}_k = \hat{d}_k$$

$$\hat{r}_k = B_{1k} / (c * \hat{I}_k)$$

$$c * \hat{I}_{k-1} = \hat{d}_{k-1} / (1 - \hat{r}_k)$$

$$\hat{r}_{k-1} = (B_{1k-1} + B_{2k-1}) / (c * \hat{I}_k + c * \hat{I}_{k-1})$$

$$c * \hat{I}_{k-2} = \hat{d}_{k-2} / (1 - \hat{r}_k - \hat{r}_{k-1})$$

$$\hat{r}_{k-2} = (B_{1k-2} + B_{2k-2} + B_{3k-2}) / (c * \hat{I}_k + c * \hat{I}_{k-1} + c * \hat{I}_{k-2})$$

etc.

We have now obtained almost all parameters for predicting the lower half of the triangle. The remaining factor would be the I 's that represent future inflation, or any other similar calendar year effect. This could build upon the ratios $(c * \hat{I}_{i+j}) / (c * \hat{I}_{i+j-1})$, and, possibly on macroeconomic considerations.

Example

We have the same data as before, but this time in incremental amounts. We also have the estimated ultimate claim numbers:

	0	1	2	est. no.
1998	30	20	15	8,00
1999	40	50		14,67
2000	55			17,00

Dividing by the estimated number gives:

	0	1	2
1998	3,75	2,50	1,88
1999	2,73	3,41	
2000	3,24		
Col. sums	9,71	5,91	1,88

Observed diagonal sums:

$$d_0 = 3,75 = 3,75$$

$$d_1 = 2,73 + 2,50 = 5,23$$

$$d_2 = 3,24 + 3,41 + 1,88 = 8,52$$

recursively we calculate

$$c\hat{I}_2 = 8,52$$

$$\hat{r}_2 = 1,88 / 8,52 = 0,2201$$

$$c\hat{I}_1 = 5,23 / (1 - 0,2201) = 6,70$$

$$\hat{r}_1 = 5,91 / (8,52 + 6,70) = 0,3882$$

$$c\hat{I}_0 = 3,75 / (1 - 0,2201 - 0,3882) = 9,57$$

$$\hat{r}_0 = 9,57 / (8,52 + 6,70 + 9,57) = 0,3917$$

To make predictions we need future cI 's. We choose to take them from the last development,

$$c\hat{I}_3 = c\hat{I}_2 * (c\hat{I}_2 / c\hat{I}_1) = 8,52 * (8,52 / 6,70) = 10,83$$

$$c\hat{I}_4 = c\hat{I}_2 * (c\hat{I}_2 / c\hat{I}_1)^2 = 8,52 * (8,52 / 6,70)^2 = 13,77$$

Now, we are able to calculate the incremental claims

$$\hat{D}_{12} = \hat{n}_1 * \hat{r}_2 * c\hat{I}_3 = 14,67 * 0,2201 * 10,83 = 34,97$$

$$\hat{D}_{21} = \hat{n}_2 * \hat{r}_1 * c\hat{I}_3 = 17,00 * 0,3882 * 10,83 = 71,47$$

$$\hat{D}_{22} = \hat{n}_2 * \hat{r}_2 * c\hat{I}_4 = 17,00 * 0,2201 * 13,77 = 51,52$$

From which we find a reserve of 157,96.

Exercises

11.1 Assume that the incremental number of claims notified in period j of accident year i is $Po(a_i * b_j)$ and that each cell (i,j) is independent of each other. Derive the maximum likelihood equations for the parameters. Confirm that you without loss of generality could assume $\sum_j b_j = 1$.

Find a recursive procedure for the parameters.

(Hint: reuse the mentioned identity and think about how the triangle goes into the ML equations).

Apply this to the following triangle of incremental claim numbers:

		0	1	2
1998		3	3	2
1999		5	6	
2000		6		

12. Model selection and tests

All prediction rests on the assumption that the unknown has something in common with the known.

If the future outcome is not the same as predicted by the provision it could emanate from several sources. No estimation procedure in the world could force a random variable to stop being random. Thus, one should always try to see if the deviation is from the process or from an estimation error. Moreover, there might also be a model error. It should also be borne in mind that figures produced by an actuary will not always be taken for granted. It is hard to argue for the result of a model if the assumptions are not reasonably met, or, if it ignores relevant factors. A formula might impress some people, but not all, and an argument that starts from data description usually is more convincing.

There are a number of models for IBNR published in actuarial papers. They might contain advanced mathematics and many parameters. Some authors may have very good academic credentials. However, many have not been tested in practice. It is important to evaluate both their explicit and implicit assumptions, as well as their robustness both with regard to data and to parameters. If there is a bias, both sign and size should be checked. Of course, all this also goes for any model you build yourselves.

There is very often a demand for “early warning” models that at the same time ignore sudden random fluctuations. It is known from control theory that such demands cannot both be met in an uncompromising way.

The important thing is to make a good prediction, by which one usually mean something that is MVUE. Including more parameters give a better fit with the past, but worsens the precision of the prediction. (To decide on the size of models, the pragmatic Akaike Information Criteria, AIC, came up some decades ago. Since then several authors have tried to improve and justify similar models. The interested reader is referred to the modern statistical theory literature.)

At a first look it would be easy to discard the naive loss ratio method, but it has its advantages when the experience is limited in time and numbers and too few claims have occurred for any conclusion. However, managers in an insurance company might also be tempted to reduce the IBNR reserve with an increase in incurred, as “that is what it is for”. But, then it is important for an actuary to defend the difference between an average and outcome.

The development factors could be seen as a regression link between to development periods. It could be possible that claims develop after the following models, where x is the past value and y the future value:

$$y = b * x + e \quad (12.1)$$

$$y = a + b * x + e \quad (12.2)$$

$$y = b * (x + \mathbf{e}) \quad (12.3)$$

$$y = b * x + \sqrt{x} * \mathbf{e} \quad (12.4)$$

For each of these models one assumes that the error term has expected value zero, that they are uncorrelated, with the same variance across accident years but maybe not between periods. The reader (immediately?) recognizes that (12.4) gives the ordinary chain ladder factor.

Some further thoughts on these models could be found in [Murphy 1994]. [Gogol 1995] have some objections to [Murphy 1994], that would hit anybody that ignore Jensen's inequality in performing estimations and calculating expectations. [Venter 1998] has tried to summarize which tests should be passed in order to use a model.

If one believe that one of these models are valid, minimum variance estimation is the most useful. As not all are standard models, the burden of proof rests with the user. To check, regression diagnostics could be used. It is also important to avoid a mindless use of tests to test whether factors differ significantly from one. Only using those ones would lead to a bias, as distributions most likely are skew. (That the Gauss-Markov theorem is applicable should not lead to the conclusion that distributions are symmetrical.) If several loss development factors are on their own considered not to be different from one, that does not necessarily go for their product.

Using models based on several unbiased estimators do not necessarily lead to an unbiased outcome. One should always bear in mind that the function of an expected value is not equal to the expected value of the function, unless the function is linear.

The Benktander-Hovinen method, was as said in section 10, not justifiable by weighing together with inverse variances, as the components did not have a common variance structure. This does not mean that the results of the model are bad. In fact it could be shown that it under a variance structure like the Bornhuetter-Ferguson is quite close to a credibility estimate. The article [Mack 2000] shows this in detail and it even claims that the method beats the components in most cases.

It is important to check the emergence of claims to what was predicted. If these result from process error, estimation error or model error should be considered. If errors of the later kinds are not determinable it could be worthwhile to fence off overinterpretation by laymen, by simple tests such as using the binomial distribution for the numbers of "ups and downs".

What is presented here are methods for determining reserves or ultimate claims. Implicitly they restrict the number of models that would fit with these methods. To illustrate this, the separation method and methods that build on that the logarithm of an incremental claim relies on a distribution with a row factor and a column factor, have sometimes been called overdispersed Poisson models. In some cases they would have the same estimators as the chain ladder, but in other not. There has also been quite a controversy about which model that underlies the chain ladder method.

Whether it is better to model the emergence of claims from bottom up or to use a model will have to be decided based on evidence.

We cannot know which method/model is the true one. We should however see to that the methods/models used are testable and hence falsifiable.

In practice it is not uncommon to ignore observations that are many calendar years back. The reason for doing this is that one might no longer be able to assume time homogeneity. Sometimes one also finds that 'extreme' factors are discarded, e.g. "mid three of last five". Apart from the always questionable practice of doing away with true data, one might also introduce a bias, which should be corrected. The interested reader could check what correction would apply in this case if the link ratios are assumed to come from an exponential distribution.

Many practitioners also 'polish' factors to obtain smoothness and robustness. Of course this violates some independence assumptions. If this is done, one should do it considering that it is a geometric average that should apply and that their effect will be weighted. The more towards the end the share in

the product is pushed, the more accident years will be affected, and vice versa. Doing this requires craftsmanship with numbers, experience and knowledge. Of course, it should always be justifiable.

Exercise

12.1 Find minimum variance estimators for the parameters of the models (12.1)-(12.4).

13. Regulation

Insurance companies are required to be able to meet all future obligations arising from all insurance policies written by them, this is actually the product sold. To show that this is the case, the provisions for that should also be shown in the balance sheet of the company.

The Swedish Insurance Companies' Act, Försäkringsrörelselagen (FRL), starts the chapter on the actual business, chapter 7, §1 by:

”Ett försäkringsbolags försäkringstekniska avsättningar skall motsvara belopp som erfordras för att bolaget vid varje tidpunkt skall kunna uppfylla alla åtaganden som skäligen kan förväntas uppkomma med anledning av ingångna försäkringsavtal. ...”

The way this is shown is governed by the accounts directive from EU (91/674/EEC), to which the Swedish law adheres:

”Article 60

Provisions for claims outstanding

1. Non-life insurance

- (a) A provision shall in principle be computed separately for each case on the basis of the cost still expected to arise. Statistical methods may be used if they result in an adequate provision having regard to the nature of the risks; Member States may, however, make the application of such methods subject to prior approval.
 - (b) This provision shall also allow for claims incurred but not reported by the balance sheet date; its amount shall be determined having regard to past experience as to the number and magnitude of claims reported after the balance sheet day.
 - (c) Claims settlement costs shall be included in the provision irrespective of their origin.
- ...”

More or less the same principles apply in the International Accounting Standard (IAS) and the US GAAP (Generally Agreed Accounting Principles), more specifically FAS 60. When it comes to statutory accounting in the US, this is on state level, not federal level, although most states have regulations that conform closely with the North American Insurance Commissioners model law.

In accounting there are standards and principles that should be followed. These would be found in academic textbooks on accounting. They do not always follow from mathematical reasoning, but from centuries of commercial experience and practice. These principles are important to understand. If you violate them, you will get into trouble with accountants and auditors.

An example is that one should never set a lower reserve for balance sheet purposes than you now believe will be necessary some time in the future for the now existing business.

14. Glossary

Accident year = Skadeår

ALAE = Allocated loss adjustment expenses = Allokerade skaderegleringsomkostnader

Acquisition expenses = anskaffningskostnader

ATA = Age-to-age = Period till period

ATU = Age-to-ultimate = Period till ultimo

Balance sheet = Balansräkning

Case reserve = Skadereglerarnas uppskattning av framtida skadekostnader för kända skadefall

Earn = intjäna

Excess-of-loss = Återförsäkring som innebär att man ersätter bara den del av en skada som överstiger ett visst belopp

Expense reserve = Omkostnadsreserv (för framtida administrativa kostnader)
Income statement = Resultaträkning
Incurred = Inträffade
Infamous = Ökänd(a)
IBNER = Inträffade, ej fullt rapporterade, skador
IBNR = Inträffade, ej rapporterade, skador
Loss ratio = Skadekvot
Perils = faror
Policyholder = Försäkringstagare
Premium deficiency reserve = Se *Unexpired risk provision*
Provision = Avsättning
Reconcile = Förena, Samstämma
ULAE = Unallocated loss adjustment expenses = Oallokerade skaderegleringsomkostnader
Underwriting year = Teckningsår
Unexpired risk provision = Avsättning för kvardröjande risker
Written premium = Tecknad premie

15. References

- Benktander, Gunnar (1976): An Approach to Credibility in Calculating IBNR for Casualty Excess Insurance. The Actuarial Review, April 1976, p. 7.
- Bornhuetter, R.L. and Ferguson, R.E. (1972): The Actuary and IBNR. Proceedings of the Casualty Actuarial Society, Vol. LIX, pp 181-195.
- European Union (1991): Accounts Directive 91/674/EEC
- Försäkringsrörelselagen, SFS 1982:713.
- Gogol, Daniel E.: Discussion of "Unbiased Loss Development Factors". PCAS LXXXII 1995, p72-77.
- Gluck, Spencer M. (1997): Balancing Development and Trend in Loss Reserve Analysis. Proceedings of the Casualty Actuarial Society, Vol. LXXXIV, pp 482-532.
- Hovinen, Esa (1981): Additive and Continuous IBNR. Proceedings of the XV ASTIN Coll., Loen, Norway 1981.
- Mack, Thomas (1994): Measuring the Variability of Chain Ladder Reserve Estimates. Casualty Actuarial Forum, Spring 1994, Vol. 1
- Mack, Thomas (2000): Credible Claims Reserve: The Benktander Method, ASTIN Bull. Vol. 30, No.2.
- Mack, Thomas and Venter, Gary (1999): A Comparison of Stochastic Models that Reproduce ChainLadder Reserve Estimates. Proceedings of the XXXth ASTIN Coll, Tokyo 1999.
- Murphy, Daniel M.(1994): Unbiased Loss Development Factors. PCAS LXXXI 1994, p154-207.
- Patrik, Gary S. (1996): Reinsurance, in "Foundations of Casualty Actuarial Science", 3rd edition.
- Taylor, G.C. (1977): Separation of Inflation and Other Effects from the Distribution of Non-Life Insurance Claim Delays. ASTIN Bull. Vol. 9, Issue C.
- Venter, Gary (1998): Testing the assumptions of age to age factors. Proceedings of the Casualty Actuarial Society, Vol. LXXXV, pp 807-847.