# Institute of Actuaries of Australia

# The Mack-Method and Analysis of Variability

## Prepared by Erasmus Gerigk

Presented to the Institute of Actuaries of Australia
Accident Compensation Seminar 28 November to 1 December 2004.

© 2004 Institute of Actuaries of Australia

## Introduction

### The Background

This paper aims to contribute to the current discussion of estimating central estimates and risk margins under APRA's GPS 210. Two statistical reserving methods are presented in detail, namely the "Incremental Loss Ratio Method" and the "Chain-Ladder Method".

Both methods are in fact as old as loss-reserving itself, and the underlying formulas constitute a widely accepted heuristic method of computing IBNR estimates - yet these methods are sometimes not further justified in terms of mathematical statistics.

It was Thomas Mack's achievement in the nineties to interpret both methods within a rigorous statistical framework and to extend these methods to include variability estimates (see Mack [1993, 1994, 1997]). It has now become common practice to refer to this interpretation and extension of the Chain-Ladder-Method as the "Mack-Method" and this article extends this notion to the Incremental Loss Ratio Method, which is not as well-known in Australia, mainly because the only publication so far is written in German (Mack [1997, 2002]).

The key features of both methods in the context of risk margins may be summarised as follows:

- The variability estimates depend only on the insurers own data
- The underlying assumptions can be tested
- The methods are simpler than most other methods, e.g. "Bootstrapping" or sophisticated stochastic modelling

The biggest caveat is probably that in many cases the data provides a bad fit to the proposed models and therefore the results need to be carefully reviewed.

### This paper

This paper intends to give a statistically sound overview of the Incremental Loss Ratio Method and the Chain-Ladder Method without any formal proofs.

The text does not reveal anything genuinely new to the expert and it is not its purpose to add anything to current research, although the Incremental Loss Ratio Method with variability estimates might be unknown to many actuaries. This was the original motivation for presenting this paper in Australia. The method is fairly simple from a statistical point of view at least compared to the Chain-Ladder method.

The whole Chain-Ladder theory has already been developed in Mack [1993] and has been further explored since then at a sophisticated level in many articles. This article gives an introduction to the practitioner into the more difficult Chain-Ladder theory.

**The example spreadsheets**
Instead of adding lengthy numerical examples to the text, this text is accompanied by two spreadsheets, one for either method, which demonstrate how the methods may be implemented. They reader may find all formulas from sections 2-4 there, only the sections on model testing and regression have been left out. The spreadsheets shall serve two purposes: first, to provide a better understanding of the underlying methods. Second, they serve to support the claim that the methods are indeed simple. They were designed for study-purposes only and should not be used if anything depends on the correct result of the calculations.

**The structure of this article**
Sections 1-3 develop the basic theory. The focus lies on motivating the crucial concepts, such as the role of model assumptions and not so much in providing complete technical proofs without becoming too vague or imprecise.

Section 4 briefly discusses a popular method of translating variability estimates into risk margins.

Section 5 sketches how the model assumptions for both methods can be tested.

Section 6 addresses some additional issues that arise in practice.

## SECTION 1: PRELIMINARIES

### 1.1 Review of the theory of point estimation

Although we may assume that the reader is familiar with the basic concepts of statistical inference, it is helpful to recall some concepts and precise terminology from the theory of point estimation.

The problem of point estimation can loosely be described as follows: assume that the characteristics of the elements of a population can be represented by some random vector $X$ with a certain unknown density $f_x$. Further assume that the values $(x_1,...,x_n)$ of the random sample $X_1, ... , X_n$ can be observed. On the basis of the observed $(x_1,...,x_n)$ it is desired to estimate some function, say $t=t(f_x)$, of the underlying distribution. In our case, the primary focus consists of estimating the mean and the variance of the loss distributions for a particular underwriting year or a total portfolio.

In general, point estimation admits two problems: the first, to devise some means of obtaining estimators $T(x_1,...x_n)$ for $t$. The second is to select criteria to define and find a "best" estimator among many possible estimators. The criterion selected for our purposes will be the **mean-squared error (*mse*)**. It is a popular and useful criterion, though it is perhaps crude. In many cases in applied statistics, the *mse* can be calculated by making suitable distribution assumptions about the underlying "true" $X$. It is an intriguing result of the Mack-Method that we will be in a position to estimate the *mse* of our estimators without the need to specify any distributions.

Instead of specifying or parameterising any distributions, our proposed assumptions will only model how the distributions are affected by changes in exposure or volume. This in connection with some independence assumptions will suffice to arrive at estimators and optimality conclusions.

Note that in order to apply the theory of point estimation properly, it is necessary to state explicitly some **model assumptions** about the nature of the random vector $X$. It is probably this point that is sometimes lost sight of when developing actuarial techniques in practice. One of the advantages of making precise model assumptions is the fact that only then we can actually test and possibly reject them, rather than intuitively decide on which model to apply.

### 1.2 Process Error and Estimation Error

Throughout this article, we have to carefully distinguish between several kinds of variability: Given some observed data set $D$ suppose we have some estimator $\hat{C} = \hat{C}(D)$ of the outcome of a "true" but unknown random variable $C$, (in the sequel $C$ can be a loss-distribution and $D$ is the data-triangle that has been observed so far). In a sense that can be made mathematically

precise, the best predictor of $C$, given the data $D$, is the conditional expectation $E(C\,|\,D)$. We do not know $E(C\,|\,D)$ and thus have to estimate it. This gives rise to two sources of variability. First, our estimate $\hat{C}$ could be far away from the true $E(C\,|\,D)$. Second, even if we were lucky enough to exactly determine $E(C\,|\,D)$, the prediction of the individual outcome $C$ would still be subject to the uncertainty $Var(C\,|\,D)$. In order to make this important observation more precise, let us introduce some terminology that will be useful in the subsections to follow:

The **Estimation Error** measures the variability of the $\hat{C}$ around the "true mean value" of the distribution we are trying to estimate. It is defined by $E[\hat{C} - E(C\,|\,D)]^2$. This component is subject to our choice of estimators.

The **Process Error** measures the variability of the random variable $C$. It is defined by $Var(C\,|\,D) = E\{[E(C\,|\,D) - C]^2\,|\,D\}$. As should be intuitively clear, there is nothing we can do to reduce the Process Error, which is inherent to the process we are trying to forecast.

The **Prediction Error** measures the variability of the deviations $C - \hat{C}$. It is defined by $mse(\hat{C}\,|\,D) = E\left((C - \hat{C})^2\,|\,D\right)$.

Looking at the identity $C - \hat{C} = C - E(C\,|\,D) + E(C\,|\,D) - \hat{C}$ we see that the Prediction Error is made up of two components. Indeed, after some calculations it can be shown that

$$mse(\hat{C}\,|\,D) = Var(C\,|\,D) + [E(C\,|\,D) - \hat{C}]^2$$

Thus the important result

**Prediction Error = Process Error + Estimation Error**

also holds when working with conditional variances.

The reader may recall the above discussion from the theory of simple linear regression. When forecasting an individual response, the error of the projection can be split into the Estimation and Process Error. It is an interesting feature of the Mack-Method that it also permits an explicit split of the Prediction Error into its two components. This can give the actuary some insight where the resulting variability of his estimates actually stems from.

### 1.3 Triangulated data and notation

We may assume that the reader is familiar with the basics of run-off triangles. Most of the notation used in the sequel should be intuitively clear. Yet to avoid ambiguity, we give an overview of notations in order to give a formal representation of triangulated data.

**Overview of basic notation:**

$i$        accident-year or underwriting year period

$k$        development year

$v_i$       the exposure measure for underwriting year $i$

$S_{i,k}$     incremental payments made in underwriting year $i$ and development year $k$

$C_{i,k}$     cumulative payments made in underwriting year $i$ and development year $k$

$m_{i,k}$     incremental loss ratio in underwriting year $i$ and development year $k$

In most cases the exposure measure $v_i$ will represent the ultimate premium or vehicle count of that particular underwriting year, though these are not the only possible exposure measures.

**<u>An important remark:</u>**
For the sake of simplicity **we will only refer to triangulated data of payments** throughout this article and we will refer to the rows of a run-off triangle as **underwriting years**. This does hopefully not obscure the fact that all methods introduced in the sequel can, in principle, be applied to a much wider range of data, e.g**. incurred claims, claims counts** and possibly other data and that the rows may as well represent **accident years**.

## SECTION 2: THE INCREMENTAL LOSS RATIO METHOD

In this section we formulate the model assumptions for the Incremental Loss Ratio method and discuss the implied estimators without proofs. The key results are the estimators for the model parameters and the recursion formulae for the standard error of the reserves. Sections 2.1 to 2.3 rely heavily on Mack [2002].

### *2.1 Formulation of the model*

Almost all reserving make the assumption that the underwriting years of the run-off triangle are similar to each other. The most simplistic assumption is probably to presume that the claims payments in each underwriting follow exactly the same distribution. Certainly this condition will almost never be fulfilled as the volume of business written in each year will be different and thus the underlying loss distributions in each underwriting year cannot be identical.

Still, an obvious step from here would be to assume that each underwriting year differs *only* with respect to its volume. The influence of the volume on the underlying distributions could be modelled in accordance with a concept that is sometimes called the **individual model**. To be more precise: assume that for each development year *k* there is some increment ratio $m_k$ such that the following relation holds for the expected (incremental) claims payments in each underwriting year $i$:

(AM 1) $\quad E\left(\dfrac{S_{ik}}{v_i}\right) = m_k$

We add another somewhat strong, but still intuitive and statistically testable assumption:

(AM 2) *The individual payments $S_{ik}$ in each underwriting/development year are independent for all $i,k$.*

The individual model also implies how the variance of each $S_{i,k}$ is affected by the exposure-measure. Assume that for each development year $k$ there is some $s_k^2$ such that

(AM 3) $\quad Var\left(\dfrac{S_{ik}}{v_i}\right) = \dfrac{s_k^2}{v_i}$

The condition (AM 3) is intuitively quite appealing if we think of the exposure measure as some form of vehicle- or policy-count. It is just a precise formulation of the well-known fact from statistics that "smaller samples have higher variances".

Once we accept that our portfolio behaves according to the individual model, the assumptions (AM 1), (AM 2) and (AM 3) may not seem to be too bold.

Let us rephrase what (AM 1 ) - (AM 3) say.  We assume:

(AM 1) For a particular development year the expected loss ratio increment is the same for all underwriting years
(AM 2) The payment in each cell is independent of all other payments
(AM 3) The variance of the loss ratio decreases in proportion to the size of volume

It may occur to the reader that (AM 2) is probably the most likely assumption to be violated in reality.

Now any estimator of $m_k$ leads immediately to an estimate of the future payments.  The outstanding payments are naturally projected via $\hat{R}_{ik} = v_i \cdot \hat{m}_k$, and the estimator of the reserve of the year $i$ is given by $\hat{R}_i = \sum_k \hat{R}_{ik}$.  It is clear that these estimators are also unbiased.

By estimating $s_k^2$ we will better understand the variability underlying the above projections.


## 2.2 The estimators of the model parameters

Combining the assumptions (AM 1) - (AM 3) one can deduce the following result:

---

The estimators

$$(2.1) \quad \hat{m}_k = \frac{\sum_{i=1}^{n+1-k} S_{ik}}{\sum_{i=1}^{n+1-k} v_i}$$

$$(2.2) \quad \hat{s}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} v_i \left( \frac{S_{ik}}{v_i} - \hat{m}_k \right)^2$$

are unbiased estimators of $m_k$ and $s_k^2$.

---

The sum in (2.1) is in fact an exposure-weighted mean of the individual increment ratios $m_{ik} = \dfrac{S_{ik}}{v_i}$, as can be seen from rearranging (2.1) to

$$(2.3) \quad \hat{m}_k = \sum_{i=1}^{n+1-k} \left( \frac{v_i}{\sum_{i=1}^{n+1-k} v_i} m_{ik} \right)$$

Note that since each $m_{ik}$ is unbiased one can easily write down plenty of unbiased estimators for $m_k$, e.g. the arithmetic mean of the individual increment ratios $m_{ik}$, but

> *The weighted mean as given in (2.1) has the smallest variance among all linear unbiased estimators of $m_k$.*

We will not prove this result, although the proofs are not particularly difficult. It is the direct consequence of a general principle in statistics, namely that when weighting independent estimators of a random variable, the respective weights should be inversely proportional to the corresponding variances.

For illustration purposes, assume that the "true" increment ratio for a particular portfolio in the 2nd development year was 53%. Then the resulting triangle could look roughly like this:

**Illustration: Incremental Loss Ratio in the 2nd dev. year** $\qquad m_{ik} = \dfrac{S_{ik}}{v_i}$

| | Exposure | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| **1999** | 7,000 | | 50% | | | | |
| **2000** | 5,000 | $v_{2000}$ | 47% $m_{2000,2}$ | | | | |
| **2001** | 6,000 | | 54% | | | | |
| **2002** | 6,400 | | 55% | | | | |
| **2003** | 8,000 | $v_{2003}$ | 55% $m_{2003,2}$ | | | | |
| **2004** | 9,000 | | | | | | |

**Note two things:**
1) We should expect the observed loss ratio increment in UY 2000 to deviate more than the increment in UY 2003 from the true underlying increment of 53%, because of the different exposures.
2) When estimating the loss ratio increment for development year 2, we may put more credibility to the UY 2003 observation than into the UY 2000 observation.

Readers who are already familiar with the *Incremental Loss Ratio Method* will immediately recognise that $\hat{m}_k$ is nothing but the original well-known estimator.

Note that there can be no estimator for the last $s_k^2$, because there is only one data point in the upper right corner of the triangle and thus no estimator for the variability of the individual incremental loss ratios. If we assume that the data are run-off in the $n$-th development year, we can set $\hat{s}_n^2$ to 0, otherwise we may try to interpolate from the preceding $\hat{s}_k^2$, or just take their minimum as an estimator for $s_k^2$. This will usually lie on the safe side as one should expect the $s_k^2$ to decline with $k$.

Up to here we have just reformulated what we may call the *Incremental Loss Ratio Method*. Let us summarise: once we feel confident that our model assumptions apply to our triangulated data, we are almost forced into projecting data in accordance with the Incremental Loss Ratio Method. In fact, we would need good reasons to diverge from that approach. This is already something worth knowing. We may now further investigate the variability of the projected claims payments, which is done in the next subsection.

### *2.3 Analysis of variability*

### The standard error of the Incremental Loss Ratio

Combining (2.3) with (AM 3) immediately tells us that the variance of the estimator $\hat{m}_k$ is given by

$$(2.4) \quad Var(\hat{m}_k) = \frac{s_k^2}{\sum_{i=1}^{n+1-k} v_i}$$

and this expression gives us an estimator for the *s.e.* of the incremental Loss Ratios – note that because of

$$(2.5) \quad \frac{\hat{s}_k^2}{\sum_{i=1}^{n+1-k} v_i} = \frac{1}{n-k} \sum_{i=1}^{n+1-k} \frac{v_i}{\sum_{i=1}^{n+1-k} v_i} \left( \frac{S_{ik}}{v_i} - \hat{m}_k \right)^2$$

the expression in (2.5) offers an intuitive interpretation as a "weighted variance" of the deviations of the individual increment ratios around their weighted mean. At this stage, it is often helpful look at the triangle of the individual incremental loss ratios $m_{i,k}$ in order to better understand where the measured variability comes from.

### The standard error for a single underwriting year

We now show how the standard error of the projected reserves can be calculated. We keep the split between random and Estimation Error.
For the practitioner, the total recursions (2.8) and (2.9) are probably the most important.

The formulas presented below can be derived by direct algebraic manipulations and do not require any higher theorems from statistics. Yet some calculations are quite lengthy, which is why we omit them.

The **Process Error** for a single underwriting year is given by:

$$(2.6) \quad Var(\hat{C}_{i,k+1}) = v_i \cdot \hat{s}_{k+1}^2 + Var(\hat{C}_{i,k})$$

This equation can be made intuitively clear without a formal proof: assumption (AM 3) stated that each individual payment $S_{i,k}$ has variance $v_i \cdot s_k^2$. Thus the variance of each cumulative $\hat{C}_{i,k+1} = \hat{S}_{i,k+1} + \hat{C}_{i,k}$ is estimated by (2.6) since the variances just add up using the independence assumption (AM 2).

The **Estimation Error** for a single underwriting year can be estimated recursively via:

$$(2.7) \quad \left(\hat{C}_{i,k+1} - E(C_{i,k+1})\right)^2 = v_i^2 \cdot s.e.(\hat{m}_{k+1})^2 + \left(\hat{C}_{i,k} - E(C_{i,k})\right)^2$$

We may confirm this equation intuitively with a bit more effort as follows: The term $s.e.(\hat{m}_k)$ measures the variation of the average *estimated* incremental Loss Ratio $\hat{m}_k$ from the "true" incremental Loss Ratio $m_k$, i.e. the Estimation Error. Thus (2.7) says that the Estimation Error we are making is the Estimation Error of the estimate $\hat{S}_{i,k}$ plus the Estimation Error we have made so far. More formally, we may rewrite the Estimation Error as $\left(\hat{C}_{i,k+1} - E(C_{i,k+1})\right)^2 = \left([\hat{S}_{i,k} - E(S_{i,k})] + [\hat{C}_{i,k} - E(C_{i,k})]\right)^2$ and then square out the sum and take expectations on both sides. This leads then to (2.7).

Putting (2.6) and (2.7) together we get immediately the following

> Recursion for the standard error of a single underwriting-year:
>
> $$(2.8) \quad s.e.(\hat{C}_{i,k+1})^2 = v_i \cdot \hat{s}_{k+1}^2 + v_i^2 \cdot s.e.(\hat{m}_{k+1})^2 + s.e.(\hat{C}_{i,k})^2$$

**The total standard error**

The manipulations in the previous apply also for the total standard error – it is basically the same reasoning only with a different volume measure. With the intermediate term $u_k = \sum_{i=n-k+2}^{n} v_i$, which can be thought of as the total

exposure corresponding to development period $k+1$, the following recursion formula for the total standard error of $\hat{C}_k = \sum_{i=1}^{n} C_{i,k}$ holds:

Recursion for the total standard error:

$$(2.9) \quad s.e.\left(\hat{C}_{k+1}\right)^2 = u_k \cdot \hat{s}_{k+1}^2 + u_k^2 \cdot s.e.\left(\hat{m}_{k+1}\right)^2 + s.e.\left(\hat{C}_k\right)^2$$

with the start value $s.e.\left(\hat{C}_2\right)^2 = s.e.\left(\hat{C}_{n,2}\right)^2$. In the same way we may calculate the Process Error and Estimation Error, just by replacing the volumes $v_i$ by $u_k$ in formulas (2.6) and (2.7).

## SECTION 3: THE CHAIN-LADDER METHOD

In this section we formulate the model assumptions for the Chain-Ladder and discuss the implied estimators. The key results are the estimators for the model parameters and the recursion formulae for the standard error of the reserves. This section is basically a very short summary of Mack [1993] and Mack [1997].

### *3.1 Formulation of the model*

We will now turn to what is usually called the Chain-Ladder method. In principle, we proceed in a similar fashion as in the previous section, starting with some (hopefully plausible) model-assumptions and ending up with a set of estimators. But since the model assumptions may not be as easily motivated as in the additive case, we start by taking the CL-method as given and look into the implicit assumptions that we may have made when applying this method. This approach was first introduced by Mack [1993].

Let us start with some general remarks:

In the additive model, the ultimate loss $C_{in}$ was written in the form

(3.1)    $C_{in} = S_{i1} + \ldots + S_{in}$

with the $S_{ik}$ being the "additive increment" in development year $k$, and then we went on looking for estimators of $S_{ik}$. A natural multiplicative approach is to start with the identity

(3.2)    $C_{in} = C_{i1} \, F_{i1} \, F_{i2} \ldots F_{in-1}$

with $F_{ik} = C_{i,k} / C_{i,k-1}$ being the "multiplicative increment" in development year $k$, or the *k-th individual development factor in year i*. In a loose sense, all we need to do is to look for estimators of the $F_{ik}$.

Although the resulting formulas are quite similar to the ones developed in section 2, more care needs to be taken when manipulating terms - some technical difficulties will arise from the simple fact that the Expectations-operator is additively linear, but not compatible with respect to multiplication, since the $C_{ik}$ are not uncorrelated within the underwriting years, and thus in general we would have $E\!\left(C_{i,k+1} / C_{ik}\right) \neq E(C_{i,k+1}) / E(C_{ik})$. To overcome these difficulties we will have to remind ourselves of the notion of *Conditional Expectation* and the role it plays in finding optimal predictors.

Let us see if we can motivate a model-assumption by anticipating the CL-method:

When projecting the ultimate loss $C_{in}$ (or next year's loss $C_{ik}$) in the underwriting period $i$ through some set of development factors $\{\hat{f}_k\}$, we want to apply some formulae like

(3.3)  $\hat{C}_{in} = C_{i,n+1-i} \, \hat{f}_{n-i} \cdot \ldots \cdot \hat{f}_n$

or

(3.4)  $\hat{C}_{i,k} = C_{i,k-1} \, \hat{f}_k$

which do not use any of the loss-information of the older development periods. **Here we are making a strong implicit assumption!** Why, we may ask, don't we use some form of average of the possible projections:

$$
\begin{array}{l}
C_{i1} \, \hat{f}_2 \cdot \ldots \cdot \hat{f}_n \\[4pt]
C_{i2} \, \hat{f}_3 \cdot \ldots \cdot \hat{f}_n \\[4pt]
\ldots \\[4pt]
C_{i,n+1-i} \, \hat{f}_{n+2-i} \cdot \ldots \cdot \hat{f}_n
\end{array}
$$

3.3')

A plausible answer seems to be, that we implicitly assume that we cannot improve our estimator of $C_{in}$ by using the information given in the older $C_{ik}$.

Stated differently, we implicitly assume that we cannot improve our estimator of $\hat{f}_k$ by using any of the $C_{i1,\ldots,}\ C_{ik-1}$ . The same reasoning applies to the fact that we do not use any of the information of the preceding underwriting periods.  Translating these thoughts into the language of probability leads directly to the following:  We assume that there are factors $f_k$ such that

(CL 1) $E\left(\dfrac{C_{ik}}{C_{ik-1}} \,\middle|\, C_{i1},\ldots,C_{ik-1}\right) = f_k$    $1 \le i \le n; \quad 2 \le k \le n$

which is the multiplicative Chain-Ladder assumption corresponding to equation (AM 1).

The concept of conditional expectation is crucial to understanding the assumption we are making here.  Bear in mind that a conditional expectation is a *function* and not just a number.  So let us reword (CL 1) once more. In a sense that can be made mathematically precise, the equation

(CL 1) $E\left(C_{i,k} \,\middle|\, C_{i1},\ldots,C_{ik-1}\right) = C_{i,k-1} f_k$    $1 \le i \le n; \quad 2 \le k \le n$

tells us the following: among all functions $g(C_{i1},\ldots,C_{ik-1})$ of the known part of the triangle, the best predictor of $C_{i,k}$ is just the function $g(C_{ik},\ldots,C_{ik-1}) = C_{i,k-1}f_k$ for some unknown parameter $f_k$.

We are not done yet – we will also need a model assumption for the variance that corresponds to (AM 3). Viewing $C_{i,k}$ as an exposure-measure in (CL 1), it may not come as a surprise if we write down in analogy to the additive method:

(CL 3) $\quad Var\left(\dfrac{C_{i,k}}{C_{i,k-1}}\,\middle|\, C_{i1},\ldots,C_{ik-1}\right) = \dfrac{\sigma_k^2}{C_{i,k-1}} \qquad 1 \le i \le n; \quad 2 \le k \le n$

Recall that by definition $Var(C_{i,k} \mid D) = E\{[C_{i,k} - E(C_{i,k} \mid D)]^2 \mid D\}$

It is probably even more difficult to attach an intuitive meaning to (CL 3) than for (CL 1). Let us try, though: we re-write (CL 3) expressing it in the form:

(CL 3) $\quad Var\!\left(C_{i,k}\,\middle|\, C_{i1},\ldots,C_{ik-1}\right) = C_{i,k-1}\sigma_k^2 \qquad 1 \le i \le n; \quad 2 \le k \le n$

We can state (CL 3) in loose words as follows: among all functions $h(C_{i1},\ldots,C_{ik-1})$, the best predictor of the squared distance between the estimate in (CL 1) and $C_{i,k}$ is given by $h(C_{i1},\ldots,C_{ik-1}) = C_{i,k-1}\sigma_k^2$ for some parameter $\sigma_k$.

A justification for (CL 3) may become clearer as soon when we look at the CL-estimators for $f_k$. It is precisely assumption (CL 3) that will make the CL-estimators for $f_k$ optimal.

And finally, we make the assumption

(CL 2) *The underwriting years { $C_{i1}$ , …, $C_{in}$} are globally independent, i.e. the sets { $C_{j1}$ , …, $C_{jn}$} are independent for i≠j .*

Although this assumption is quite strong, it is plausible and – more importantly – statistically testable.


### *3.2 The estimators of the model parameters*

Starting from the model-assumptions (CL 1) – (CL 2), it turns out that similar results like those obtained above for the additive method can be derived.

The estimators

$$(3.5) \quad \hat{f}_k = \frac{\sum\limits_{i=1}^{n+1-k} C_{i,k}}{\sum\limits_{i=1}^{n+1-k} C_{i,k-1}}$$

$$(3.6) \quad \hat{\sigma}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} C_{i,k-1} \left( \frac{C_{i,k}}{C_{i,k-1}} - \hat{f}_k \right)^2$$

are unbiased estimators of $m_k$ and $\sigma_k^2$.

Just like in the additive case, equation (3.5) can be thought of as a weighted mean of the individual development factors $F_{i,k} = \dfrac{C_{i,k}}{C_{i,k-1}}$, this time the weights being the current cumulative loss $C_{i,k}$. Indeed, rearranging the sum gives us

$$(3.7) \quad \hat{f}_k = \sum_{i=1}^{n-k+1} \left( \frac{C_{i,k-1}}{\sum_{i=1}^{n-k} C_{i,k-1}} \cdot F_{i,k} \right)$$

and since each of the individual $F_{i,k}$ is an unbiased estimator of $f_k$, so is their weighted mean $\hat{f}_k$. Again, the particular choice of weights is justified by the following fact, which we state without proof:

*The weighted mean as given in (3.5) has the smallest variance among all linear unbiased estimators of $f_k$.*

This last observation shows why it may be justified to claim that anyone using the Chain-Ladder method with the development factor (3.5) has implicitly relied on the model assumptions (CL 1)-(CL 3).

The interpretation of (3.5) similar as in the additive case, only that the exposure this time is the "cumulative loss in the preceding year":

**Illustration: Cumulative incurred losses up to 2nd development year**

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **1999** | $C_{2000,1}$  716 | 3,501 | | | |
| **2000** | 533 | 2,350 | | | |
| **2001** | 753 | 3,212 | | | |
| **2002** | $C_{2003,1}$  664 | 3,550 | | | |
| **2003** | 1,036 | 4,421 | | | |
| **2004** | 840 | | | | |

**Individual development factors in the 2nd development year**

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **1999** | | 4.89 | | | |
| **2000** | $F_{2000,2}$ | 4.41 | | | |
| **2001** | | 4.26 | | | |
| **2002** | | 5.34 | | | |
| **2003** | $F_{2003,2}$ | 4.27 | | | |
| **2004** | | | | | |

We may have an eye on two things: Firstly, how volatile the individual factors are within each development period and secondly, how much weight is assigned to each individual factor. This check should always be done, as the $\sigma_k^2$ can be quite volatile in the first or last underwriting years and may have erratic high values. These would be carried forward in the calculation of the standard error. In this case, it might make sense to smoothen the extreme $\sigma_k^2$ by means of a regression as will be outlined in the last chapter.

We may also spot trends in the development factors here, although this is more systematically examined by looking at residuals as will be explained in later.

Note that there can be no estimator for the last $\sigma_n^2$, because there is only one data point in the upper right corner of the triangle and thus no estimator for the variability of development factors. We may set $\sigma_n^2$ to 0 if the data are run-off, otherwise we may try to interpolate through a regression from the preceding $\sigma_k^2$.

### 3.3 Analysis of variability

**The standard error of the Development Factor**

The variance of the estimator of the development factors is given by

$$(3.8)\ Var\left(\hat{f}_k\right)= \frac{\sigma_k^2}{\sum_{i=1}^{n+1-k}C_{i,k-1}}$$

This formula can be derived directly and will be justified once more later via regression-analysis in section 5. The estimator in (3.6) for $\sigma_k$ gives us an estimator for $Var\left(\hat{f}_k\right)$. Note that just like in the additive case, $Var\left(\hat{f}_k\right)$ is determined by the fluctuations of the individual $F_{i,k}$ around their weighted mean, which can be seen by inserting (3.6) into (3.8). It is therefore helpful in practice to look at the triangle of individual development factors $F_{i,k}$ in order to understand where the variability comes from.

**The standard error for a single underwriting year**

The whole exercise from the additive case can be carried out in the multiplicative case in a similar fashion. There are recursion formulae for the Process and Estimation Error, respectively, and for the Prediction Error of the projected loss. We try to give an intuitive understanding of the formulas without even attempting any formal proof. This shall not conceal the fact that the results quoted in this section are far from obvious and that the rigorous proofs form the hardest bit of the whole theory. The interested reader may refer to Mack [1994] and Mack (1999).

The **Process Error** for a single underwriting year is:

$$(3.9)\ Var(C_{i,k+1}\mid D) = E(C_{i,k}\mid D)\sigma_{k+1}^2 + Var(C_{i,k}\mid D)f_{k+1}^2$$

which allows a similar interpretation as in the additive case: the first term contains $s.e.(F_{i,k})^2 = \sigma_k^2/C_{i,k}$ which measures the variation of the individual factors $F_{i,k}$ around their "true" mean $f_k$, i.e. the Process Error. And we add the Process Error of the preceding period increased by the development factor.

The **Estimation Error** for a single underwriting year is given by:

$$(3.10)\ \left(\hat{C}_{i,k+1} - E(C_{i,k+1}\mid D)\right)^2 = \hat{C}_{i,k}^2 s.e.(\hat{f}_{k+1})^2 + \left(\hat{C}_{i,k} - E(C_{i,k}\mid D)\right)^2 \hat{f}_{k+1}^2$$

Again, the term $s.e.(\hat{f}_k)$ measures the variation of the average *estimated* development factor $\hat{f}_k$ from the "true" factor $f_k$, i.e. the Estimation Error.

For the **Prediction Error** we thus write:

(3.11)

$$s.e.\left(\hat{C}_{i,k+1}\right)^2 = \hat{C}_{i,k}^2\left(\frac{\hat{\sigma}_{k+1}^2}{\hat{C}_{i,k}} + s.e.(\hat{f}_{k+1})^2\right) + \hat{f}_{k+1}^2\left(Var(C_{i,k} \mid D) + \left(\hat{C}_{i,k} - E(C_{i,k} \mid D)\right)^2\right)$$

We may reformulate the above formula once more in a more simplified notation:

> Recursion for the standard error of a single underwriting-year:
>
> (3.12) $s.e.\left(\hat{C}_{i,k+1}\right)^2 = \hat{C}_{i,k}\hat{\sigma}_{k+1}^2 + \hat{C}_{i,k}^2 s.e.(\hat{f}_{k+1})^2 + \hat{f}_{k+1}^2 \cdot s.e.\left(\hat{C}_{i,k}\right)^2$

**The total standard error**

With the exposure term $h_k = \sum_{i=n-k+1}^{n} C_{i,k}$ corresponding to the development year $k+1$ the following recursion formula for the total standard error of $\hat{C}_k = \sum_{i=1}^{n} C_{i,k}$ holds for $k = 2,\dots, n\text{-}1$:

> Recursion for the total standard error:
>
> (3.14) $s.e.\left(\hat{C}_{k+1}\right)^2 = h_k\hat{\sigma}_{k+1}^2 + h_k^2 \cdot s.e.(\hat{f}_{k+1})^2 + \hat{f}_{k+1}^2 \cdot s.e.\left(\hat{C}_k\right)^2$

with the start value $s.e.\left(\hat{C}_2\right)^2 = s.e.\left(\hat{C}_{n,2}\right)^2$.

The reader may note the formal similarity between the recursion formulae of the Chain-Ladder and the Incremental Loss Ratio estimates. It is indeed possible to design a complex set-up in a spreadsheet that does the job in both cases, though this does not enhance transparency.

## SECTION 4: CALCULATION OF A RISK MARGIN

Once the mean and the standard error of the underlying loss-distribution have been estimated, it is possible to further use this information in order to calculate a so-called risk-margin corresponding to a given level of adequacy. It cannot be recommended to stick to a distribution-free approach at this point, e.g. with a one-sided version of Chebyshev's inequality such as $P(X - E(X) \geq \alpha \cdot s.e.(X)) \leq 1/(1 + \alpha^2)$. Such a general inequality cannot lead to sharp results. In practice often an additional assumption with respect to the shape of the underlying loss-distribution is made.

A popular and simple approach consists of assuming that the total outstanding claims provisions follow a Log-Normal-Distribution, i.e. $OCP \sim \log N(\mu, \sigma^2)$, where the parameters $\mu$ and $\sigma$ are uniquely determined by the mean and the coefficient of variation of the underlying distribution. For the sake of completeness we recite the basic formulas for the required parameters. With *E(OCP)* and *CV* denoting the mean and coefficient of variation, respectively, we have:

$$\sigma^2 = \log(1 + CV^2)$$
$$\mu = \log(E(OCP)) - \tfrac{1}{2}\sigma^2$$

If we now require the level of adequacy to be *x*%, we first calculate the *x*%-percentile *p* of the standard normal distribution. The *x*% percentile *P* for the original loss distribution is then given by the expression

$$P = \exp(\mu + p \cdot \sigma^2)$$

The resulting risk margin is then often expressed as a percentage of the central estimate.

Two comments may seem appropriate here:

1. The log-normal model is an assumption that may not be justified in each single case and one should be aware of some peculiar features of this distribution. In particular, the Log-Normal-Distribution implies a maximum possible risk margin of approximately 25.5%, which is reached by a coefficient of variation of 75% and is then declining as a function of the *CV*. The pure formulaic approach as outlined above may therefore not be suitable for highly volatile classes.

2. It is important to bear in mind that the thus derived risk margin applies to the respective triangle as a single entity. If the underlying class of business forms a subclass of a insurer's portfolio then, on an overall basis, the insurer will benefit from correlation effects. As a result, the sum of the individual risk margins of all subclasses can be substantially higher than the actually required risk margin on the aggregated total

level. In order to take those effects into account in accordance with APRA's GPS 210, one may discount the individual segment margin by some estimated "diversification benefit". This critical issue is not further addressed in this paper.

## SECTION 5: TESTING THE MODEL ASSUMPTIONS

In the sequel, we will introduce some means with which to test whether the model assumptions for either the additive or the multiplicative method are fulfilled. This should form an important step in any proper statistical analysis. It turns out that the model-assumptions of the additive and the multiplicative method can be interpreted as a linear regression, an area in which statistical tests are well understood and are included in most statistical software packages. We quote the most basic tests. A nice overview of the deeper connections between the presented Chain-Ladder-type methods and their interpretation as weighted regressions can be found in Murphy [1994]. This section relies heavily on Mack [1994]. The interested reader will find much more on testing in this article.

### *5.1 Testing the Additive Model*

Let us recall the model assumptions for the Incremental Loss Ratio method:

(AM 1) $E(S_{ik}) = v_i m_k$

(AM 2) *The individual payments $S_{ik}$ in each underwriting/development year are independent for all $i,k$.*

(AM 3) $Var(S_{ik}) = v_i s_k^2$

For a fixed $k$, the model-assumptions can be viewed as a linear regression of the form

(5.1) $Y_i = a + x_i b + \varepsilon_i$

with $a = 0$ and $m_k = b$. Looking at the error term $\varepsilon_i$ we see that because of

(5.2) $Var(\varepsilon_i) = v_i s_k^2$

the variances for each observation are not constant. Such regression models are called *heteroskedastic* (as opposed to the equal variance models that are called homoskedastic). Since the parameter $s_k^2$ is unknown, but the weights $v_i$ are known, textbook theory has it that the *best linear unbiased estimator* for $m_k$ is determined via *weighted least squares*. It is the unique value that minimises the expression

(5.3) $\displaystyle\sum_{i=1}^{n+1-k} \frac{(S_{i,k} - v_i m_k)^2}{v_i}$

The formula means that we "downweight" the squared residuals for those observations with large variances. It turns out that **the solution to (5.3) is**

**indeed the estimate** $\hat{m}_k$ from equation (2.1), and the variance of this estimator coincides with the variance given by (2.4)! Thus the whole machinery known from statistical textbooks to test the linearity (AM 1) and the variance assumption (AM 3) can be applied. This is usually done by plotting the data and residuals.

As a start we may plot the $v_i$ against the $S_{i,k}$ in order to confirm that there is indeed a reasonable linear relationship.

As a next step we plot the residuals, i.e. the deviations of the known outcomes to the fitted curve. In general, if the chosen model is a good one, then the residuals should reflect only random fluctuation of the data around the fitted line. To be precise, we plot the *weighted* residuals, since we are dealing with a weighted regression where the deviations from the fitted curve are expected to "fan out".

Thus, we plot the residuals

$$(5.4) \quad \frac{S_{i,k} - v_i \hat{m}_k}{\sqrt{v_i}}$$

against the exposure $v_i$. Ideally, this graph will show no pattern but only "noise".

In practice, it is hard to draw conclusions from this check for small triangles or development years with thin data, i.e. the later development years.


### 5.2 Testing the Multiplicative Model

The Chain-Ladder model can be tested in the same way. For the sake of completeness we repeat the basic findings. First, let us recall the three basic Chain-Ladder assumptions:

(CL 1) $E\big(C_{i,k}\,\big|\,C_{i1},\ldots,C_{ik-1}\big) = C_{i,k-1} f_k$
(CL 2) *Independence of the underwriting years*
(CL 3) $Var\big(C_{i,k}\,\big|\,C_{i1},\ldots,C_{ik-1}\big) = C_{i,k-1}\sigma_k^2$

Again, for a fixed $k$, the model-assumptions can be viewed as a *heteroskedastic* linear regression.

The *best linear unbiased estimator* for $f_k$ is determined via *weighted least squares* by minimising the expression

$$(5.5) \quad \sum_{i=1}^{n+1-k} \frac{(C_{i,k} - C_{i,k-1} f_k)^2}{C_{i,k-1}}$$

and its unique solution is the Chain-Ladder estimate $\hat{f}_k$ from equation (3.5), and its variance coincides with the variance given by (3.8)!

We may plot the $C_{i,k-1}$ against the $C_{i,k}$ in order to confirm that there is indeed a reasonable linear relationship. In practice, it might be better to plot $S_{i,k} = C_{i,k} - C_{i,k-1}$ against $C_{i,k}$ instead, the required slope of the regression line is then $f_k - 1$.

The corresponding residuals become

$$(5.6) \quad \frac{C_{i,k} - C_{i,k-1}\hat{f}_k}{\sqrt{C_{i,k-1}}} = \frac{S_{i,k} - C_{i,k-1}(\hat{f}_k - 1)}{\sqrt{C_{i,k-1}}}$$

which are plotted against the known $C_{i,k-1}$. This graph should show no pattern but only "noise".

It is worth noting that the residual analysis can be further refined to test if other variance assumptions than (CL 3) are more in line with the given data. A thorough discussion is beyond of the scope of this article. The interested reader is referred to Mack [1994].

### 5.3 Checking for calendar-year effects

Before we could to apply the regression-analysis in the preceding subsection, we had to assume that (CL 2) is already fulfilled (or that at least the underwriting years are uncorrelated). As mentioned before, this is in fact a quite strong statement because in reality there is a range of factors which may violate such an assumption. It is out of the scope of this article to introduce a thorough test for independence. The reader who is seeking more depth is referred to Mack [1994]. Here, we illustrate a simple but useful test to check for so-called calendar-year effects. And indeed, one of the major reasons why the independence of the underwriting years (CL 2) is violated can be attributed to calendar-year effects. Calendar-year effects may be caused by changes of legislation or claims handling or in irregular inflation, to name but a few.

Calendar-year effects are acting on the diagonals of the run-off triangle, i.e. they affect several underwriting years in different development periods.

Statistically speaking, a calendar-year effect will produce increment ratios or cumulative losses that are *systematically* higher (or lower) than what would have been predicted by the increment ratios or development factors by the model.

Thus, if we plot the standardised residuals

(5.8) $\dfrac{S_{i,k} - v_i \hat{m}_k}{\hat{s}_k \sqrt{v_i}}$    $1 \le i + k \le n$

or

(5.7) $\dfrac{C_{i,k} - C_{i,k-1}\hat{f}_k}{\hat{\sigma}_k \sqrt{C_{i,k-1}}}$    $1 \le i + k \le n$ ,

against the corresponding calendar-year $i+k$-1 we expect only "white noise" if the model-assumptions are fulfilled. If, on the other hand, these residuals all have the same sign for a fixed calendar-year, then this could be an indication for a calendar-year effect and must be further investigated since a mechanical application of the Incremental Loss Ratio or Chain-Ladder method would now use increment ratios that are systematically too high in every development period! It is probably safest to disregard the whole diagonal by means of setting "weights" as will be explained in the next section.

Apart from spotting calendar-year effects, plotting the residuals (5.7) and (5.8) can also help detecting large deviations of single individual development factors from their weighted mean. The graph then becomes a useful tool in connection with the setting individual weights and can be quickly analysed. A more formal statistical test for calendar-year effects is developed in Mack [1994].

## SECTION 6: PRACTICAL CONSIDERATIONS

### 6.1 Dealing with outliers

In almost cases in practice, it will not make sense to apply the Mack too mechanically, because the data may be distorted by outliers and calendar-year effects – or one may feel that it is more appropriate to use only the last few diagonals because the underlying business has changed, to name but a few reasons.   Consequently, one may decide to manually adjust the respective development factors and in principle there is nothing wrong with that approach.  However, this procedure leaves the actuary with the problem of also adjusting the corresponding sigmas or standard errors, which is less intuitive.  Yet there is a more precise way of how to tackle this issue.  We only need to "rewrite" all formulas, this time omitting those individual increment ratios or development factors which we regard as outliers. The formal procedure and underlying theory are in fact quite simple, however it seems that it has not been fully appreciated by the actuarial literature.  An example of how to realise the proposed procedure in practice can be found in the accompanying Excel spreadsheets.

### Formalism of setting weights

Formally, the approach looks as follows:  We start by looking again at the formula for the incremental loss ratio, expressed as a weighted mean.   In principle there is nothing stopping us from using a different set of weights, say $w_{i,k} \cdot S_{i,k}$  in order to reflect the credibility we put into $S_{i,k}$.  In this case, the corresponding formulas for the increment ratios translate naturally to

$$(6.1) \quad \hat{m}_k = \sum_{i=1}^{n+1-k} \left( \frac{w_{i,k} \cdot v_i}{\sum_{i=1}^{n+1-k}(w_{i,k} v_i)} \hat{m}_{ik} \right) = \frac{\sum_{i=1}^{n+1-k}(w_{i,k} \cdot S_{ik})}{\sum_{i=1}^{n+1-k}(w_{i,k} \cdot v_i)}$$

and for the parameters relating to the s.e. we have

$$(6.2) \quad \hat{s}_k^2 = \frac{1}{|I|-1} \sum_{i=1}^{n+1-k} w_{i,k} \cdot v_i \left( \frac{S_{ik}}{v_i} - \hat{m}_k \right)^2$$

where $|I|$ denotes the number of incremental loss ratios that are taken into account in year $k$ and

$$(6.3) \quad Var(\hat{m}_k) = \frac{s_k^2}{\sum_{i=1}^{n+1-k}(w_{i,k} \cdot v_i)}$$

From here we can proceed in the usual way, say via the recursion formula etc.

The modification through weights corresponds to slightly modified versions of the original model assumptions. A more rigorous discussion for the Chain-Ladder method can be found in Mack [1999]. As long as we are only "deselecting" outliers we are only interested in the case where $w_{i,k} \in \{0,1\}$ and the modifications stay simple.

In a similar fashion the formulas in the Chain-Ladder method can be amended to account for outliers. The formulas for the development factors and the sigmas become then, respectively:

$$(6.4) \quad \hat{f}_k = \sum_{i=1}^{n+1-k} \left( \frac{w_{i,k} \cdot C_{i,k-1}}{\sum_{i=1}^{n+1-k} (w_{i,k} C_{i,k-1})} \cdot \hat{F}_{i,k} \right) = \frac{\sum_{i=1}^{n+1-k} (w_{i,k} \cdot C_{i,k})}{\sum_{i=1}^{n+1-k} (w_{i,k} \cdot C_{i,k-1})}$$

$$(6.5) \quad \hat{\sigma}_k^2 = \frac{1}{|I|-1} \sum_{i=1}^{n+1-k} w_{i,k} C_{i,k-1} \left( \frac{C_{i,k}}{C_{i,k-1}} - \hat{f}_k \right)^2$$

where $|I|$ denotes the number of individual development factors that are taken into account and

$$(6.6) \quad Var(\hat{f}_k) = \frac{\sigma_k^2}{\sum_{i=1}^{n+1-k} (w_{i,k} \cdot C_{i,k-1})}$$

The weighting procedure as introduced herein may look harmless and it is indeed mathematically trivial. Yet it is probably one of the key mechanisms to be employed when working in practice and can also be incorporated when working with spreadsheets as the accompanying workbook shall demonstrate.


### 6.2 Deciding between the multiplicative and additive model

**General remarks**

So far we have introduced two different methods of projecting claims, and have given some basic statistical tests to test the underlying model assumptions. Apart from the statistical tests, it is also important to have a less formal understanding of how these projection-methods work and where the caveats lie. The reality will often be that one has to decide which model is the "lesser evil".

In an imprecise way we may first note the following: if the outstanding reserve is mainly determined by the run-off of already reported claims, this is an indication that the multiplicative method may lead to more stable results. If on

the other hand the portfolio is more affected by genuinely late-reported IBNR-claims, the additive method might be favoured, as the additive method is indifferent to the loss development that has been observed so far.

Furthermore, for the most recent underwriting years (the data in the lower left corner of the triangle) the multiplicative method may not applicable because no losses have been reported, and thus the multiplicative projection will forecast 0 claims. In this case one might be forced either to apply some additive projection for the most recent underwriting years, or estimate the lost in the first development year by other means.

Another important remark must be made with respect to the Incremental Loss Ratio method: **if the underwriting-year premium is chosen as an exposure measure, then the Incremental Loss-ratio Method makes only sense if the premium-volume does in fact reflect the size of the underlying business**. In other words, premium cycles or rate hardenings need to be corrected before the additive method can be applied.

Overall, the choice of model and the degree of trust that the actuary may put into their results will always be of judgement and experience.


**Graphical interpretation**

Another very simple yet quite useful tool in analysing the appropriateness of the respective model-assumptions consists of plotting appropriate graphs of the run-off data.

The rational behind this procedure is nothing but the fundamental assumption that the future payments will be somehow similar to what has happened in the past. This is so obvious that it seems hardly worth mentioning. Yet this remark leads to a very simplistic but useful rule of thumb when the actuary has to decide which of the two proposed methods could be given preference.

The key assumption in the additive case is that *the loss ratios increase by a certain amount* in a particular development year, the amount being the same for all underwriting years, except for random noise. In formulas:

$$(6.7) \quad \frac{C_{i,k}}{v_i} = \frac{C_{i,k-1}}{v_i} + m_k$$

Therefore, a plot of the development periods against loss ratios for all underwriting years should show that the resulting graphs are parallel to each other.
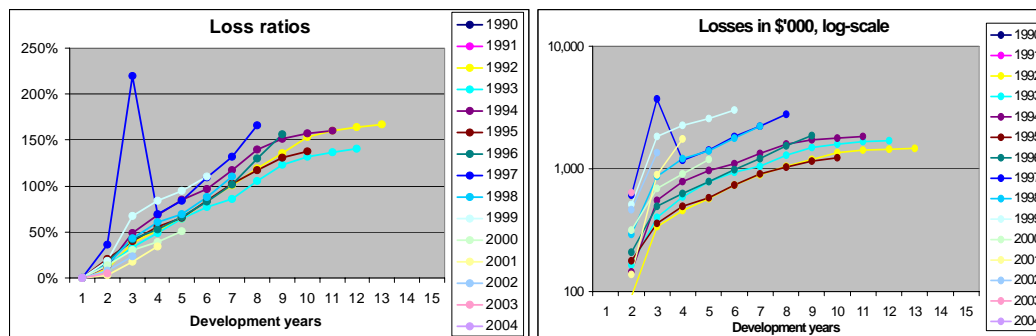
In the multiplicative case the key assumption was that within a specified development year, the cumulative loss amounts of each underwriting year are multiplied by the same factor except for random noise. This in turn means

that *the logarithms of the cumulative losses increase by the same amount.* Speaking in formulas, we have for the Chain-Ladder model:

$$(6.8) \quad \log C_{i,k} = \log C_{i,k-1} + \log f_k$$

As a result, the graphs of the underwriting-years in a dev-year/logarithmic cumulative-loss-space should be parallel.

To sum it up, a rough check as to wether the data has behaved according to either model consists of plotting two graphs looking roughly like the following:



If we intend to apply the Incremental Loss Ratio method, then the graph on the left hand side should show a parallel pattern. And if we intend to apply the Chain-Ladder method, then the graph on the right hand side is the appropriate interpretation of the data.

In fact, the additive or multiplicative projection methods can be interpreted as "parallelly projecting" the loss experience of the older underwriting years in a loss ratio graph or logarithmic dollar-graph, respectively. The incremental loss ratio or development factor in year $k$ is nothing but an average of the observed slopes in that development year.

The graphical representations are often helpful in many respects, not only when deciding between the two models. It may be possible to spot trends, outliers or other peculiarities. For example, in the above data-example, we can spot immediately an erratic spike in years 2 and 3 of the underwriting year 1997. As a consequence, these two individual development factors must be taken out by means of the weighting procedure described above.

Furthermore, it is sometimes possible to immediately identify an appropriate range of underwriting periods and calendar years in which the data show the most homogenous pattern and just "get a feel" for the run-off behaviour.

Overall, the usefulness of these two graphs when interpreting either method can hardly be overestimated.

## *6.3 Reduction of parameters*

As a final refinement we discuss how the sigmas and development factors might be smoothened or the run-off extended via regression techniques. This section is based on Mack [1997, 2002]. The reader who is not familiar with weighted regressions may find this section difficult, although the basic idea of fitting a curve through the parameters is quite straightforward.

### The Sigma-Regression

We have already indicated that the "sigmas" of both methods can be quite volatile without indicating how to possibly deal with this difficulty.

The considerations in this subsection can in principle be applied to both the additive and the multiplicative method. However, it seems that in the additive case it is much harder to fit a curve through the $\hat{s}_k^2$ than through the $\hat{\sigma}_k^2$, which is why we restrict our discussion to the Chain-Ladder method. Yet it must be pointed out that for both methods there is no logical connection leading from the model assumptions to the smoothening-approach suggested here.

As a starting point, one might first want to plot the $\hat{\sigma}_k^2$ on a log-linear scale and investigate any extreme outliers by analysing again the triangle of individual development factors. If the $\hat{\sigma}_k^2$ seem to decline in a log-linear fashion from a certain development period $k$ onwards, a formal curve fitting approach becomes:

(6.9) $Var\big(C_{i,k}\big|C_{i1},\ldots,C_{ik-1}\big) = C_{i,k-1}\sigma^2 e^{-ck}$

This approach does not affect the optimal estimators for $f_k$ since the $\hat{\sigma}_k^2$ differ from each other only through a constant factor. In theory, the "most suitable" regression for $c$ and $\sigma^2$ depends on the respective variances of the $\hat{\sigma}_k^2$ - which are unknown. Yet some intuitive reasoning can help: it is plausible to expect the respective unknown variances of the $\hat{\sigma}_k^2$ to decrease, the more data points are included in their estimation, with the decrease proportional to the number of data points taken into account (e.g. the variance of the estimator of a sample mean shows exactly this behaviour). Therefore it seems at least plausible to use the reciprocal of ($n$-$k$-1) as a weight for $\hat{\sigma}_k^2$ in a regression. The corresponding expression to be minimised in order to estimate $\sigma$ and $c$ can then be written as

(6.10) $\sum\limits_{k=1}^{n-2}(n-k-1)\Big(\log(\hat{\sigma}_k^2) - \log(\sigma^2 e^{-ck})\Big)^2$

There is no neat expression for the solutions for $\sigma$ and $c$, and some statistical computer package or a solver function might be the method of choice.

The resulting smoothed sigmas may then be used to recalculate the standard errors of the factors and estimators, in particular for the first and last development years.


**Regression of development factors**

In a similar way, we can fit a curve thorough the development factors in order to smoothen or extend the run-off. Again, experience shows that in the multiplicative case the factors often decline in a log-linear fashion, at least from a certain development year on (e.g. from the 2nd or 3rd development year). So we may again start with a relation like

(6.11) $f_k = 1 + ae^{-bk}$

and then determine $a$ and $b$. As with the sigmas, we should weight the regression with the variances of the $f_k$, which we have already calculated at this stage.

Therefore, the expression to be minimised becomes

(6.12) $\sum_{k=2}^{n-1} \dfrac{(\hat{f}_k - 1 - ae^{-bk})^2}{Var(\hat{f}_k)}$

In order to keep the expressions as simple as possible, we may use the $\hat{\sigma}_k^2$ from the sigma-regression, in which case minimising (6.12) is equivalent to minimising

(6.13) $\sum_{k=2}^{n-1} \dfrac{(\hat{f}_k - 1 - ae^{-bk})^2}{\sum_{i=1}^{n+1-k} C_{i,k}}$

Again, there is no neat explicit expression to solve (6.13) but some solver-function will do the trick.

The biggest difficulty that arises from here consists of the derivation of a theoretically correct standard error that corresponds to the factors from the regression. This is far from trivial and may depend on additional assumptions. We will not address this difficult issue in this paper. One may use formulas like (3.8) together with the sigma-regression, and then apply the recursion-formulae, although as a word of warning it must be stressed that this approach is not easily justified theoretically.

## Appendix A: Formulas for the Incremental Loss Ratio method

**The increment ratio**

$$\hat{m}_k = \frac{\sum_{i=1}^{n+1-k} S_{ik}}{\sum_{i=1}^{n+1-k} v_i}$$

**Projected loss in underwriting year *i* and development year *k+1*:**

$$\hat{C}_{i,k+1} = \hat{C}_{i,k} + v_i \cdot \hat{m}_k$$

**Sigma:**

$$\hat{s}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} v_i \left( \frac{S_{ik}}{v_i} - \hat{m}_k \right)^2$$

**Standard error of the incremental Loss Ratio:**

$$s.e.(\hat{m}_k)^2 = \frac{\hat{s}_k^2}{\sum_{i=1}^{n+1-k} v_i}$$

**Standard error for the projected loss underwriting year *i* and development year *k*:**

$$s.e.(\hat{C}_{i,k+1})^2 = v_i \cdot \hat{s}_k^2 + v_i^2 \cdot s.e.(\hat{m}_k)^2 + s.e.(\hat{C}_{i,k})$$

**The standard error of the total reserve:**

$$u_k = \sum_{i=n-k+2}^{n} v_i$$

$$s.e.(\hat{C}_{k+1})^2 = u_k \cdot \hat{s}_k^2 + u_k^2 \cdot s.e.(\hat{m}_k)^2 + s.e.(\hat{C}_k)^2$$

## Appendix B: Formulas for the Chain-Ladder method

**The development factor:**

$$\hat{f}_k = \frac{\sum_{i=1}^{n+1-k} C_{i,k}}{\sum_{i=1}^{n+1-k} C_{i,k-1}}$$

**Projected loss:**

$$\hat{C}_{i,k} = \hat{f}_k \cdot \hat{C}_{i,k-1}$$

**Sigma:**

$$\hat{\sigma}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} C_{i,k-1} \left( \frac{C_{i,k}}{C_{i,k-1}} - \hat{f}_k \right)^2$$

**Standard error of the development factor:**

$$s.e.\left( \hat{f}_k \right)^2 = \frac{\hat{\sigma}_k^2}{\sum_{i=1}^{n+1-k} C_{i,k-1}}$$

**Standard error for the projected loss underwriting year $i$ and development year $k$:**

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

**The standard error of the total reserve:**

$$h_k = \sum_{i=n-k+1}^{n} C_{i,k}$$

$$s.e.\left( \hat{C}_{k+1} \right)^2 = h_k \hat{\sigma}_{k+1}^2 + h_k^2 \cdot s.e.(\hat{f}_{k+1})^2 + \hat{f}_{k+1}^2 \cdot s.e.\left( \hat{C}_k \right)^2$$

## Bibliography

**Mack, T.**
[1993] Distribution–free calculation of the standard error of chain–ladder reserve estimates. ASTIN Bull. 23, 213–225.
[1994] Which stochastic model is underlying the chain–ladder method ? Insurance Math. Econom. 15, 133–138.
[1994] Measuring the variability of chain–ladder reserve estimates. In: CAS Forum Spring 1994, pp. 101–182.
[1997] Measuring the variability of chain–ladder reserve estimates. In: Claims Reserving Manual, vol. 2. London: Institute of Actuaries.
[1997, 2002] Schadenversicherungsmathematik. Karlsruhe: Verlag Versicherungswirtschaft.
[1999] The standard error of chain–ladder reserve estimates: Recursive calculation and inclusion of a tail factor. ASTIN Bull. 29, 361–365.

**Murphy, D. M.**
[1994] Unbiased loss development factors. Proc. CAS 81, 154–222.