Methods and Models of Loss Reserving Based on Run–Off Triangles: A Unifying Survey

By Klaus D. Schmidt

Lehrstuhl für Versicherungsmathematik
Technische Universität Dresden

Abstract

The present paper provides a unifying survey of some of the most important methods and models of loss reserving which are based on run–off triangles. The starting point is the thesis that the use of run–off triangles in loss reserving can be justified only under the assumption that the development of the losses of every accident year follows a development pattern which is common to all accident years. This assumption can be viewed as a primitive stochastic model of loss reserving.
The notion of a development pattern turns out to be a unifying force in the comparison of methods which to a large extent can be summarized under a general version of the Bornhuetter–Ferguson method. It is shown that the loss–development method and the chain–ladder method as well as the Cape–Cod method and the additive method can be viewed as special cases of the general Bornhuetter–Ferguson method.
Some of these methods can be justified by general principles of statistical inference applied to suitable and more sophisticated stochastic models. It is shown that credibility prediction and Gauss–Markov prediction as well as maximum–likelihood estimation can contribute in a substantial way to the understanding of various methods of loss reserving.

Keywords: Bornhuetter–Ferguson principle, credibility prediction, development pattern, Gauss–Markov prediction, loss reserving, maximum–likelihood estimation.

Contents

1 Introduction

We start with the general modelling of loss development data by a family of random variables representing incremental or cumulative losses and with the run–off triangles representing the observable incremental or cumulative losses (Section 2).

We then introduce the central notion of a development pattern which can be expressed in three different but equivalent ways and turns out to be a powerful and unifying concept for the interpretation and comparison of several methods and models of loss reserving (Section 3).

The subsequent three sections are devoted to methods, least–squares prediction, and maximum–likelihood estimation:

In the section on methods (Section 4), we start with a general version of the Bornhuetter–Ferguson method which provides a general framework into which several other methods, like the loss–development method, the chain–ladder method, the Cape–Cod method and the additive method, can be embedded as special cases. We also consider two variants of the chain–ladder method which have no practical interest but are needed as a link between the chain–ladder method and certain stochastic models.

In the section on least–squares prediction (Section 5), we study credibility prediction and Gauss–Markov prediction. It is shown that, under certain model assumptions, these methods of prediction yield predictors of the Bornhuetter–Ferguson type.

In the section on maximum–likelihood estimation (Section 6), we study maximum–likelihood estimation for a large class of stochastic models for claim counts. It is shown that in many cases, but not always, the maximum–likelihood estimators of the expected ultimate cumulative losses are identical with the chain–ladder predictors of the ultimate cumulative losses.

In the final section (Section 7) we collect some conclusions.

Throughout this paper, let $(\Omega, \mathcal{F}, P)$ be a probability space on which all random variables are defined. We also assume that all random variables are square integrable. Moreover, all equalities and inequalities involving random variables are understood to hold almost surely with respect to the probability measure P.

2 Loss Development Data

We consider a portfolio of risks and we assume that each claim of the portfolio is settled either in the accident year or in the following n development years. The portfolio may be modelled either by incremental losses or by cumulative losses.

2.1 Incremental Losses

To model a portfolio by incremental losses, we consider a family of random variables $\{Z_{i,k}\}_{i,k\in\{0,1,\ldots,n\}}$ and we interpret the random variable $Z_{i,k}$ as the loss of *accident year* i which is settled with a delay of k years and hence in *development year* k and in *calendar year* $i+k$. We refer to $Z_{i,k}$ as the *incremental loss* of accident year i and development year k.

We assume that the incremental losses $Z_{i,k}$ are *observable* for *calendar years* $i+k \leq n$ and that they are *non–observable* for calendar years $i+k \geq n+1$. The observable incremental losses are represented by the following *run–off triangle*:

Accident	**Development Year**									
Year	0	1	$\ldots$	k	$\ldots$	$n-i$	$\ldots$	$n-1$	n
0	$Z_{0,0}$	$Z_{0,1}$	$\ldots$	$Z_{0,k}$	$\ldots$	$Z_{0,n-i}$	$\ldots$	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$	$\ldots$	$Z_{1,k}$	$\ldots$	$Z_{1,n-i}$	$\ldots$	$Z_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$Z_{i,0}$	$Z_{i,1}$	$\ldots$	$Z_{i,k}$	$\ldots$	$Z_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$Z_{n-k,0}$	$Z_{n-k,1}$	$\ldots$	$Z_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$Z_{n-1,0}$	$Z_{n-1,1}$							
n	$Z_{n,0}$								

The problem is to *predict* the non–observable incremental losses.

2.2 Cumulative Losses

To model a portfolio by cumulative losses, we consider a family of random variables $\{S_{i,k}\}_{i,k\in\{0,1,\ldots,n\}}$ and we interpret the random variable $S_{i,k}$ as the loss of *accident year* i which is settled with a delay of *at most* k years and hence *not later than* in *development year* k. We refer to $S_{i,k}$ as the *cumulative loss* of accident year i and development year k, to $S_{i,n-i}$ as a *cumulative loss of the present calendar year* n, and to $S_{i,n}$ as an *ultimate cumulative loss*.

We assume that the cumulative losses $S_{i,k}$ are *observable* for *calendar years* $i+k \leq n$ and that they are *non–observable* for calendar years $i+k \geq n+1$. The observable cumulative losses are represented by the following *run–off triangle*:

Accident Year	Development Year 0	1	…	k	…	$n-i$	…	$n-1$	n
0	$S_{0,0}$	$S_{0,1}$	…	$S_{0,k}$	…	$S_{0,n-i}$	…	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$	…	$S_{1,k}$	…	$S_{1,n-i}$	…	$S_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$S_{i,0}$	$S_{i,1}$	…	$S_{i,k}$	…	$S_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$	…	$S_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
n	$S_{n,0}$								

The problem is to *predict* the non–observable cumulative losses.

2.3 Remarks

Of course, modelling a portfolio by incremental losses is equivalent to modelling a portfolio by cumulative losses:

– The cumulative losses are obtained from the incremental losses by letting

$$S_{i,k} \;:=\; \sum_{l=0}^{k} Z_{i,l}$$

– The incremental losses are obtained from the cumulative losses by letting

$$Z_{i,k} \;:=\; \begin{cases} S_{i,k} & \text{if } k = 0 \\ S_{i,k} - S_{i,k-1} & \text{else} \end{cases}$$

In the sequel we shall switch between incremental and cumulative losses as necessary.

Correspondingly, prediction of non–observable incremental losses is *essentially* equivalent to prediction of non–observable cumulative losses:

– If $\{\widehat{Z}_{i,k}\}_{i,k\in\{0,1,\ldots,n\},i+k\geq n+1}$ is a family of predictors of the non–observable incremental losses, then a family of predictors of the non–observable cumulative losses is obtained by letting

$$\widehat{S}_{i,k} \;:=\; S_{i,n-i} + \sum_{l=n-i+1}^{k} \widehat{Z}_{i,l}$$

– If $\{\widehat{S}_{i,k}\}_{i,k\in\{0,1,\ldots,n\},i+k\geq n+1}$ is a family of predictors of the non–observable cumulative losses, then a family of predictors of the non–observable incremental losses is obtained by letting

$$\widehat{Z}_{i,k} \;:=\; \begin{cases} \widehat{S}_{i,n-i+1} - S_{i,n-i} & \text{if } k = n-i+1 \\ \widehat{S}_{i,k} - \widehat{S}_{i,k-1} & \text{else} \end{cases}$$

For the ease of notation and to avoid the distinction of cases as in the previous definition, we shall also refer to $Z_{i,n-i}$ and $S_{i,n-i}$ as predictors of $Z_{i,n-i}$ and $S_{i,n-i}$, although these random variables are, of course, observable.

Warning: Whenever prediction is subject to an optimality criterion, it cannot be guaranteed in general that the previous formulas lead from optimal predictors of incremental losses to optimal predictors of cumulative losses or vice versa.

The enumeration of accident years and development years starting with 0 instead of 1 is widely but not yet generally accepted; see Taylor [2000] as well as Radtke and Schmidt [2004]. It is useful for several reasons:

- For losses which are settled within the accident year, the delay of settlement is 0. It is therefore natural to start the enumeration of development years with 0.
- Using the enumeration of development years also for accident years implies that the incremental or cumulative loss of accident year i and development year k is observable if and only if $i + k \leq n$. In particular, the cumulative losses $S_{i,n-i}$ are those of the present calendar year n and are crucial in most methods of loss reserving.

After all, the notation used here simplifies mathematical formulas.

3 Development Patterns

The use of run–off triangles in loss reserving can be justified only if it is assumed that the development of the losses of every accident year follows a development pattern which is common to all accident years. This vague idea of a development pattern can be formalized in various ways.

In the present section we consider three types of development patterns which are formally distinct but can easily be converted into each other. These development patterns and their equivalence provide a key to the comparison of several methods of loss reserving.

The assumption of an underlying development pattern can be viewed as a primitive stochastic model of loss reserving.

3.1 Incremental Quotas

The development pattern for incremental quotas compares the expected incremental losses with the expected ultimate cumulative losses:

> **Development Pattern for Incremental Quotas:** *There exist parameters* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ *with* $\sum_{l=0}^{n} \vartheta_l = 1$ *such that the identity*
>
> $$\vartheta_k = \frac{E[Z_{i,k}]}{E[S_{i,n}]}$$
>
> *holds for all* $k \in \{0, 1, \ldots, n\}$ *and for all* $i \in \{0, 1, \ldots, n\}$.

The assumption means that, for every development year $k \in \{0, 1, \ldots, n\}$, the *incremental quotas*

$$\vartheta_{i,k} = \frac{E[Z_{i,k}]}{E[S_{i,n}]}$$

are identical for all accident years.

In the case of a run–off triangle for *paid losses* or *claim counts*, it is usually reasonable to assume in addition that $\vartheta_k > 0$ holds for all $k \in \{0, 1, \ldots, n\}$. In the case of *incurred losses*, however, this additional assumption may be inappropriate since, due to conservative reserving, the (expected) incremental losses of development years $k \in \{1, \ldots, n\}$ may be negative.

3.2 Cumulative Quotas

The development pattern for cumulative quotas compares the expected cumulative losses with the expected ultimate cumulative losses:

Development Pattern for Cumulative Quotas: *There exist parameters* $\gamma_0, \gamma_1, \ldots, \gamma_n$ *with* $\gamma_n = 1$ *such that the identity*

$$\gamma_k = \frac{E[S_{i,k}]}{E[S_{i,n}]}$$

holds for all $k \in \{0, 1, \ldots, n\}$ *and for all* $i \in \{0, 1, \ldots, n\}$.

The assumption means that, for every development year $k \in \{0, 1, \ldots, n\}$, the *cumulative quotas*

$$\gamma_{i,k} = \frac{E[S_{i,k}]}{E[S_{i,n}]}$$

are identical for all accident years.

In the case of a run–off triangle for paid losses or claim counts, it is usually reasonable to assume in addition that $0 < \gamma_0 < \gamma_1 < \cdots < \gamma_n$. In the case of incurred losses, however, this additional assumption may be inappropriate since, due to conservative reserving, the sequence of the (expected) cumulative losses may be decreasing.

The development patterns for incremental and cumulative quotas can be converted into each other:

– If $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ is a development pattern for incremental losses, then a development pattern for cumulative losses is obtained by letting

$$\gamma_k := \sum_{l=0}^{k} \vartheta_l$$

– If $\gamma_0, \gamma_1, \ldots, \gamma_n$ is a development pattern for cumulative losses, then a development pattern for incremental losses is obtained by letting

$$\vartheta_k := \begin{cases} \gamma_0 & \text{if } k = 0 \\ \gamma_k - \gamma_{k-1} & \text{else} \end{cases}$$

Furthermore, the condition $\vartheta_k > 0$ is fulfilled for all $k \in \{0, 1, \ldots, n\}$ if and only if $0 < \gamma_0 < \gamma_1 < \cdots < \gamma_n$.

3.3 Factors

The development pattern for factors compares subsequent expected cumulative losses:

Development Pattern for Factors: *There exist parameters* $\varphi_1, \ldots, \varphi_n$ *such that the identity*

$$\varphi_k = \frac{E[S_{i,k}]}{E[S_{i,k-1}]}$$

holds for all $k \in \{1, \ldots, n\}$ *and for all* $i \in \{0, 1, \ldots, n\}$.

The assumption means that, for every development year $k \in \{1, \dots, n\}$, the *factors*

$$\varphi_{i,k} = \frac{E[S_{i,k}]}{E[S_{i,k-1}]}$$

are identical for all accident years.

In the case of a run–off triangle for paid losses or claim counts, it is usually reasonable to assume in addition that $\varphi_k > 1$ holds for all $k \in \{1, \dots, n\}$. In the case of incurred losses, however, this additional assumption may be inappropriate since, due to conservative reserving, the sequence of the (expected) cumulative losses may be decreasing.

The development patterns for cumulative quotas and for factors can be converted into each other:

– If $\gamma_0, \gamma_1, \dots, \gamma_n$ is a development pattern for cumulative losses, then a development pattern for factors is obtained by letting

$$\varphi_k := \frac{\gamma_k}{\gamma_{k-1}}$$

– If $\varphi_1, \dots, \varphi_n$ is a development pattern for factors, then a development pattern for cumulative losses is obtained by letting

$$\gamma_k := \prod_{l=k+1}^{n} \frac{1}{\varphi_l}$$

(such that $\gamma_n = 1$).

Furthermore, the condition $\gamma_0 < \gamma_1 < \dots < \gamma_n$ is fulfilled if and only if $\varphi_k > 1$ holds for all $k \in \{1, \dots, n\}$.

Combining this result and that of the previous subsection, it is evident that also the development patterns for incremental quotas and for factors can be converted into each other. We omit the corresponding formulas since they will not be needed in the sequel.

3.4 Estimation

At the first glance, there is little hope to estimate the parameters of the development patterns for incremental or cumulative quotas since the only obvious estimators of ϑ_k and γ_k are the observable quotients $Z_{0,k}/S_{0,n}$ and $S_{0,k}/S_{0,n}$, respectively.

Fortunately, the situation is quite different for the development pattern for factors: For every development year $k \in \{1, \dots, n\}$, each of the *individual development factors*

$$\widehat{\varphi}_{i,k} := \frac{S_{i,k}}{S_{i,k-1}}$$

with $i \in \{0, 1, \ldots, n-k\}$ is a reasonable estimator of φ_k, and this is also true for every weighted mean

$$\widehat{\varphi}_k \;:=\; \sum_{j=0}^{n-k} W_{j,k}\, \widehat{\varphi}_{j,k}$$

with random variables (or constants) satisfying $\sum_{j=0}^{n-k} W_{j,k} = 1$. The most prominent estimator of this large family is the *chain–ladder factor*

$$\widehat{\varphi}_k^{\mathrm{CL}} \;:=\; \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

which can also be written as

$$\widehat{\varphi}_k^{\mathrm{CL}} \;=\; \sum_{j=0}^{n-k} \frac{S_{j,k-1}}{\sum_{h=0}^{n-k} S_{h,k-1}}\, \widehat{\varphi}_{j,k}$$

and is used in the chain–ladder method.

Due to the correspondence between the three development patterns, it is then clear that in the same way estimators of factors can be converted into estimators of cumulative quotas and hence into estimators of incremental quotas.

3.5 Remarks

In the case of a run–off triangle for paid losses or claim counts, the intuitive interpretation of the development patterns of incremental or cumulative quotas would be their interpretation as incremental or cumulative probabilities. This interpretation is helpful, but it is not quite correct since the parameters of the development pattern are defined as *quotients of expectations* instead of *expectations of quotients* and since these quantities are in general distinct.

One may thus argue that the definitions of development patterns are inconvenient since they do not exactly correspond to intuition. In the following two sections, however, it will be shown that the definitions given here are nevertheless reasonable since they provide a powerful and unifying concept for the interpretation and comparison of several methods and models of loss reserving.

4 Methods

The present section provides a unifying presentation of the most important methods of loss reserving. The starting point is a general version of the Bornhuetter–Ferguson method which is closely related to the notion of a development pattern for cumulative quotas and turns out to be a unifying principle under which various other methods of loss reserving can be subsumed.

4.1 Bornhuetter–Ferguson Method

The Bornhuetter–Ferguson method is based on the assumption that there exist parameters $\alpha_0, \alpha_1, \ldots, \alpha_n$ and $\gamma_0, \gamma_1, \ldots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] \;=\; \gamma_k \, \alpha_i$$

holds for all $i, k \in \{0, 1, \ldots, n\}$. Then we have

$$E[S_{i,n}] \;=\; \alpha_i$$

and hence

$$E[S_{i,k}] \;=\; \gamma_k \, E[S_{i,n}]$$

such that the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The Bornhuetter–Ferguson method is also based on the additional assumption that *prior estimators*

$$\widehat{\alpha}_0, \widehat{\alpha}_1, \ldots, \widehat{\alpha}_n$$

of the expected ultimate cumulative losses $E[S_{i,n}]$ and *prior estimators*

$$\widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_n$$

of the development pattern are given and that $\widehat{\gamma}_n = 1$.

Comment: Prior estimators may be obtained from information provided by various sources:

- *Internal information*: This is any information which is *contained* in the run–off triangle of the portfolio under consideration. Internal information could be used, e. g., by estimating the development pattern from the given run–off triangle.
- *External information*: This is any information which is *not contained* in the run–off triangle of the portfolio under consideration. External information could be obtained, e. g., from market statistics, from other portfolios which are judged to be similar to the given one, or from premiums or other volume measures of the portfolio under consideration; see Section 4.6.

Of course, prior estimators may also be obtained by combining internal and external information. In any case, the choice of prior estimators is an important decision to be made by the actuary.

The *Bornhuetter–Ferguson predictors* of the cumulative losses $S_{i,k}$ with $i+k \geq n$ are defined as

$$\widehat{S}_{i,k}^{\mathrm{BF}} \quad := \quad S_{i,n-i} + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\widehat{\alpha}_i$$

The definition of the Bornhuetter–Ferguson predictors reminds of the identity

$$E[S_{i,k}] \quad = \quad E[S_{i,n-i}] + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\widehat{\alpha}_i$$

which is a consequence of the model assumption.

The definition of the Bornhuetter–Ferguson predictors shows that the prior estimators are dominant for young accident years whereas they are less important for old development years. Also, in the extreme case where the prior estimators are completely determined by external information, the major part of the run–off triangle is ignored and only the cumulative losses of the present calendar year are used. This is reasonable when the quality of the data from older calendar years is poor.

Example A. We consider the following reduced run–off triangle for cumulative losses which contains the cumulative losses of the present calendar year and is complemented by the prior estimators of the expected ultimate cumulative losses and of the development pattern:

Accident		**Development Year** k					
Year i	$\widehat{\alpha}_i$	0	1	2	3	4	5
0	3517						3483
1	3981					3844	
2	4598				3977		
3	5658			3880			
4	6214		3261				
5	6325	1889					
$\widehat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

Computing now the Bornhuetter–Ferguson predictors, the run–off triangle is completed as follows:

Accident		**Development Year** k					
Year i	$\widehat{\alpha}_i$	0	1	2	3	4	5
0	3517						3483
1	3981					3844	*4043*
2	4598				3977	*4391*	*4621*
3	5658			3880	*4785*	*4389*	*5577*
4	6214		3261	*4442*	*5436*	*5995*	*6306*
5	6325	1889	*3344*	*4546*	*5558*	*6127*	*6443*
$\widehat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

When the cumulative losses of the present calendar year are judged to be reliable, it may be desirable to modify the Bornhuetter–Ferguson predictors in order to strengthen the weight of the cumulative losses of the present calendar year and to reduce that of the prior estimators of the expected ultimate cumulative losses. This goal can be achieved by iteration.

For example, if on the right hand side of the previous formula the prior estimators $\widehat{\alpha}_i$ are replaced by the Bornhuetter–Ferguson predictors $\widehat{S}^{\mathrm{BF}}_{i,n}$, then the resulting predictors are the *Benktander–Hovinen predictors*

$$\widehat{S}^{\mathrm{BH}}_{i,k} \;:=\; S_{i,n-i} + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\,\widehat{S}^{\mathrm{BF}}_{i,n}$$

which in the case $\widehat{\gamma}_{n-i} < \widehat{\gamma}_k$ increase the weight of the cumulative losses of the present calendar year and reduce that of the prior estimators of the expected ultimate cumulative losses.

More generally, the *iterated Bornhuetter–Ferguson predictors* of order $m \in \mathbb{N}_0$ are defined by letting

$$\widehat{S}^{(m)}_{i,k} \;:=\; \begin{cases} S_{i,n-i} + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\,\widehat{\alpha}_i & \text{if} \quad m = 0 \\ S_{i,n-i} + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\,\widehat{S}^{(m-1)}_{i,n} & \text{if} \quad m \geq 1 \end{cases}$$

Then we have $\widehat{S}^{(0)}_{i,k} = \widehat{S}^{\mathrm{BF}}_{i,k}$ and $\widehat{S}^{(1)}_{i,k} = \widehat{S}^{\mathrm{BH}}_{i,k}$, and induction yields

$$\begin{aligned} \widehat{S}^{(m)}_{i,k} &= \Big(1 - \big(1 - \widehat{\gamma}_{n-i}\big)^m\Big)\,\widehat{\gamma}_k\,\frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}} + \big(1 - \widehat{\gamma}_{n-i}\big)^m\,\widehat{S}^{\mathrm{BF}}_{i,k} \\ &= \widehat{\gamma}_k\,\frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}} + \big(1 - \widehat{\gamma}_{n-i}\big)^m \left(\widehat{S}^{\mathrm{BF}}_{i,k} - \widehat{\gamma}_k\,\frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}}\right) \\ &= \widehat{\gamma}_k\,\frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}} + \big(1 - \widehat{\gamma}_{n-i}\big)^m \big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\big)\left(\widehat{\alpha}_i - \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}}\right) \end{aligned}$$

for all $m \in \mathbb{N}_0$. In the particular case where $\widehat{\alpha}_i = S_{i,n-i}/\widehat{\gamma}_{n-i}$ or $\widehat{\gamma}_{n-i} = 1$, the iteration is without interest since in that case the identity

$$\widehat{S}^{(m)}_{i,k} \;=\; \widehat{\gamma}_k\,\frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}}$$

holds for all $m \in \mathbb{N}_0$. By contrast, the iteration is of considerable interest in the case where $0 < \widehat{\gamma}_{n-i} < 1$ since in that case we obtain

$$\lim_{m\to\infty} \widehat{S}^{(m)}_{i,k} \;=\; \widehat{\gamma}_k\,\frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}}$$

and convergence of the sequence of the iterated Bornhuetter–Ferguson predictors is monotone but may be increasing or decreasing.

Example B. The following table contains the prior estimators of the expected ultimate cumulative losses, the iterated Bornhuetter–Ferguson predictors

$$\widehat{S}_{i,n}^{(m)} = \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}} + \bigl(1-\widehat{\gamma}_{n-i}\bigr)^{m+1}\left(\widehat{\alpha}_i - \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}}\right)$$

and their limits:

Accident Year i	**Prior** $\widehat{\alpha}_i$	**Iterated Bornhuetter–Ferguson Predictors** $\widehat{S}_{i,5}^{(0)}$	$\widehat{S}_{i,5}^{(1)}$	$\widehat{S}_{i,5}^{(2)}$	$\widehat{S}_{i,5}^{(3)}$	$\widehat{S}_{i,5}^{(4)}$	$\widehat{S}_{i,5}^{(5)}$	...	$\widehat{S}_{i,5}^{(10)}$	...	**Limit**
0	3517	*3483*	*3483*	*3483*	*3483*	*3483*	*3483*	...	*3483*	...	*3483*
1	3981	*4043*	*4046*	*4046*	*4046*	*4046*	*4046*	...	*4046*	...	*4046*
2	4598	*4621*	*4623*	*4624*	*4624*	*4624*	*4624*	...	*4624*	...	*4624*
3	5658	*5577*	*5553*	*5546*	*5544*	*5543*	*5543*	...	*5543*	...	*5543*
4	6214	*6306*	*6351*	*6373*	*6384*	*6389*	*6392*	...	*6394*	...	*6394*
5	6325	*6443*	*6528*	*6589*	*6633*	*6664*	*6687*	...	*6730*	...	*6746*

The iteration steps 0 and 1 correspond to the Bornhuetter–Ferguson method and to the Benktander–Hovinen method, respectively. The table illustrates that convergence is monotone but may be increasing or decreasing, and that convergence is usually fast for old accident years and slow for young accident years.

4.2 Loss–Development Method

The loss–development method is based on the assumption that there exist parameters $\gamma_0, \gamma_1, \dots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] = \gamma_k \, E[S_{i,n}]$$

holds for all $i, k \in \{0, 1, \dots, n\}$. Then the parameters $\gamma_0, \gamma_1, \dots, \gamma_n$ form a development pattern for cumulative quotas.

The loss–development method is also based on the additional assumption that *prior estimators*

$$\widehat{\gamma}_0, \widehat{\gamma}_1, \dots, \widehat{\gamma}_n$$

of the development pattern are given and that $\widehat{\gamma}_n = 1$.

The *loss–development predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\widehat{S}_{i,k}^{\mathrm{LD}} := \widehat{\gamma}_k \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}}$$

The definition of the loss–development predictors reminds of the identity

$$E[S_{i,k}] = \gamma_k \frac{E[S_{i,n-i}]}{\gamma_{n-i}}$$

which is a consequence of the model assumption.

When compared with the Bornhuetter–Ferguson predictors, the importance of the cumulative losses of the present calendar year and of the prior estimators of the development pattern is increased in the loss–development predictors since the latter do not involve any prior estimators of the expected ultimate cumulative losses.

Example C. We consider the following reduced run–off triangle for cumulative losses which contains the cumulative losses of the present calendar year and is complemented by the prior estimators of the development pattern:

Accident Year i	**Development Year** k 0	1	2	3	4	5
0						3483
1					3844	
2				3977		
3			3880			
4		3261				
5	1889					
$\widehat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

Computing now the loss–development predictors, the run–off triangle is completed as follows:

Accident Year i	**Development Year** k 0	1	2	3	4	5
0						3483
1					3844	*4046*
2				3977	*4393*	*4624*
3			3880	*4767*	*5266*	*5543*
4		3261	*4476*	*5499*	*6074*	*6394*
5	1889	*3440*	*4722*	*5802*	*6409*	*6746*
$\widehat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

The loss–development predictors can be written as

$$\widehat{S}_{i,k}^{\mathrm{LD}} = S_{i,n-i} + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\widehat{S}_{i,n}^{\mathrm{LD}}$$

This shows that the loss–development predictors are nothing else than the Bornhuetter–Ferguson predictors with respect to the prior estimators

$$\widehat{\alpha}_i^{\mathrm{LD}} := \widehat{S}_{i,n}^{\mathrm{LD}}$$

of the expected ultimate cumulative losses. In other words, the loss–development method is a particular case of the Bornhuetter–Ferguson method with prior estimators of the expected ultimate cumulative losses which are based on internal and external information.

Moreover, in the case where $0 < \widehat{\gamma}_{n-i} < 1$, the loss–development predictors are precisely the limits of the sequences of the iterated Bornhuetter–Ferguson predictors with respect to arbitrary prior estimators of the expected ultimate cumulative losses, as has been shown in Section 4.1.

4.3 Chain–Ladder Method

The chain–ladder method is based on the assumption that there exist parameters $\varphi_1, \ldots, \varphi_n$ such that the identity

$$E[S_{i,k}] \;=\; \varphi_k \, E[S_{i,k-1}]$$

holds for all $i \in \{0, 1, \ldots, n\}$ and $k \in \{1, \ldots, n\}$. Then the parameters $\varphi_1, \ldots, \varphi_n$ form a development pattern for factors.

The *chain–ladder predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\widehat{S}_{i,k}^{\mathrm{CL}} \;:=\; S_{i,n-i} \prod_{l=n-i+1}^{k} \widehat{\varphi}_l^{\mathrm{CL}}$$

where

$$\widehat{\varphi}_k^{\mathrm{CL}} \;:=\; \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

is the *chain–ladder factor* introduced in Section 3. The definition of the chain–ladder predictors reminds of the identity

$$E[S_{i,k}] \;=\; E[S_{i,n-i}] \prod_{l=n-i+1}^{k} \varphi_l$$

which is a consequence of the model assumption.

When compared with the loss–development predictors, it is remarkable that the chain–ladder predictors are not determined by the cumulative losses of the present calendar year but involve, via the chain–ladder factors, *all* cumulative losses of the run–off triangle.

Example D. We consider the following run–off triangle for cumulative losses:

Accident	**Development Year** k					
Year i	0	1	2	3	4	5
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	
2	1265	2433	3233	3977		
3	1490	2873	3880			
4	1725	3261				
5	1889					

Computing first the chain–ladder factors and then the chain–ladder predictors, the run–off triangle is completed as follows:

Accident	**Development Year** k					
Year	0	1	2	3	4	5
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	*4013*
2	1265	2433	3233	3977	*4454*	*4650*
3	1490	2873	3880	*4780*	*5354*	*5590*
4	1725	3261	*4334*	*5339*	*5980*	*6243*
5	1889	*3587*	*4767*	*5873*	*6578*	*6867*
$\widehat{\varphi}_k^{\text{CL}}$		*1.899*	*1.329*	*1.232*	*1.120*	*1.044*

It has been pointed out in Section 3 that the different development patterns and their estimators can be converted into each other. In particular, letting

$$\gamma_k \;:=\; \prod_{l=k+1}^{n} \frac{1}{\varphi_l}$$

converts a development pattern for factors into a development pattern for cumulative quotas and letting

$$\widehat{\gamma}_k \;:=\; \prod_{l=k+1}^{n} \frac{1}{\widehat{\varphi}_l}$$

converts the estimators of a development pattern for factors into estimators of a development pattern for cumulative quotas. Thus, letting

$$\widehat{\gamma}_k^{\text{CL}} \;:=\; \prod_{l=k+1}^{n} \frac{1}{\widehat{\varphi}_l^{\text{CL}}}$$

the chain–ladder predictors can be written as

$$\widehat{S}_{i,k}^{\text{CL}} \;=\; \widehat{\gamma}_k^{\text{CL}} \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}^{\text{CL}}}$$

This shows that the chain–ladder predictors are nothing else than the loss–development predictors with respect to the *chain–ladder cumulative quotas* $\widehat{\gamma}_k^{\text{CL}}$ as prior estimators of the cumulative quotas. Furthermore, we have

$$\widehat{S}_{i,k}^{\text{CL}} \;=\; S_{i,n-i} + \left(\widehat{\gamma}_k^{\text{CL}} - \widehat{\gamma}_{n-i}^{\text{CL}}\right) \widehat{S}_{i,n}^{\text{CL}}$$

This shows that the chain–ladder predictors are precisely the Bornhuetter–Ferguson predictors with respect to the prior estimators $\widehat{\gamma}_k^{\text{CL}}$ of the cumulative quotas and the prior estimators

$$\widehat{\alpha}_i^{\text{CL}} \;:=\; \widehat{S}_{i,n}^{\text{CL}}$$

of the expected ultimate cumulative losses. In other words, the chain–ladder method is a particular case of the loss–development method and hence of the Bornhuetter–Ferguson method with prior estimators of the development pattern and the expected ultimate cumulative losses which are completely based on internal information.

The chain–ladder method can be modified by replacing the chain–ladder factors $\widehat{\varphi}_k^{\mathrm{CL}}$ by any other estimators of the form

$$\widehat{\varphi}_k \;=\; \sum_{j=0}^{n-k} W_{j,k}\,\widehat{\varphi}_{j,k}$$

with random variables (or constants) satisfying $\sum_{j=0}^{n-k} W_{j,k} = 1$.

4.4 Grossing–Up Method

The grossing–up method is based on the assumption that there exist parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] \;=\; \gamma_k\, E[S_{i,n}]$$

holds for all $i, k \in \{0, 1, \ldots, n\}$. Then the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The *grossing–up predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\widehat{S}_{i,k}^{\mathrm{GU}} \;:=\; \widehat{\gamma}_k^{\mathrm{GU}}\, \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}^{\mathrm{GU}}}$$

where

$$\widehat{\gamma}_k^{\mathrm{GU}} \;:=\; \begin{cases} 1 & \text{if } k = n \\ \dfrac{\sum_{j=0}^{n-k-1} S_{j,k}}{\sum_{j=0}^{n-k-1} \widehat{S}_{j,n}^{\mathrm{GU}}} & \text{else} \end{cases}$$

is the *grossing–up cumulative quota* of development year k. The definition of the grossing–up predictors reminds of the identity

$$E[S_{i,k}] \;=\; \gamma_k\, \frac{E[S_{i,n-i}]}{\gamma_{n-i}}$$

which is a consequence of the model assumption.

The computation of the grossing–up cumulative quotas and of the grossing–up predictors for the ultimate cumulative losses proceeds by recursion along the accident years, which yields

$$
\begin{array}{rcl@{\qquad}c@{\qquad}rcl}
\widehat{\gamma}_n^{\mathrm{GU}} & = & 1 & \text{and} & \widehat{S}_{0,n}^{\mathrm{GU}} & = & S_{0,n} \\
\widehat{\gamma}_{n-1}^{\mathrm{GU}} & = & \dfrac{S_{0,n-1}}{\widehat{S}_{0,n}^{\mathrm{GU}}} & \text{and} & \widehat{S}_{1,n}^{\mathrm{GU}} & = & \dfrac{S_{1,n-1}}{\widehat{\gamma}_{n-1}^{\mathrm{GU}}} \\
\widehat{\gamma}_{n-2}^{\mathrm{GU}} & = & \dfrac{S_{0,n-2}+S_{1,n-2}}{\widehat{S}_{0,n}^{\mathrm{GU}}+\widehat{S}_{1,n}^{\mathrm{GU}}} & \text{and} & \widehat{S}_{2,n}^{\mathrm{GU}} & = & \dfrac{S_{2,n-2}}{\widehat{\gamma}_{n-2}^{\mathrm{GU}}} \\
 & \vdots & & & & \vdots &
\end{array}
$$

As can be seen from the definition, the grossing–up predictors are nothing else than the loss–development predictors with respect to the grossing–up cumulative quotas $\widehat{\gamma}_k^{\mathrm{GU}}$ as prior estimators of the cumulative quotas. Furthermore, we have

$$\widehat{S}_{i,k}^{\mathrm{GU}} \;=\; S_{i,n-i} + \Bigl(\widehat{\gamma}_k^{\mathrm{GU}} - \widehat{\gamma}_{n-i}^{\mathrm{GU}}\Bigr)\widehat{S}_{i,n}^{\mathrm{GU}}$$

which shows that the grossing–up predictors are precisely the Bornhuetter–Ferguson predictors with respect to the prior estimators $\widehat{\gamma}_k^{\mathrm{GU}}$ of the cumulative quotas and the prior estimators

$$\widehat{\alpha}_i^{\mathrm{GU}} \;:=\; \widehat{S}_{i,n}^{\mathrm{GU}}$$

of the expected ultimate cumulative losses. In other words, the grossing–up method is a particular case of the loss–development method and hence of the Bornhuetter–Ferguson method with prior estimators of the development pattern and the expected ultimate cumulative losses which are completely based on internal information.

Since the previous remark applies as well to the chain–ladder predictors, the question arises whether there is any difference between the grossing–up predictors and the chain–ladder predictors. The answer to this question is that there is no difference at all since it can be shown that the grossing–up cumulative quotas and the chain–ladder cumulative quotas are identical for all development years; see e. g. Lorenz and Schmidt [1999].

The grossing–up method thus provides a computational alternative to the chain–ladder method, but this alternative seems to be of little practical interest if any. The reformulation of the chain–ladder method provided by the grossing–up method is, however, of considerable interest with regard to the comparison of methods:

First, among all methods for cumulative losses considered here, the chain–ladder method appears to be somewhat singular since it uses estimators of a development pattern for factors instead of cumulative quotas, but its equivalence with the grossing–up method shows that this singularity is only due to the most intelligent formulation of an algorithm which avoids recursion and is hence more easily understood.

Second, the grossing–up method provides an substantial link between the chain–ladder method and the marginal–sum method; see Subsection 4.5.

4.5 Marginal–Sum Method

The marginal–sum method is based on the assumption that there exist parameters $\alpha_0, \alpha_1, \ldots, \alpha_n$ and $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ with $\sum_{l=0}^n \vartheta_l = 1$ such that the identity

$$E[Z_{i,k}] \;=\; \vartheta_k \, \alpha_i$$

holds for all $i, k \in \{0, 1, \ldots, n\}$. Summation yields

$$E[S_{i,n}] \;=\; \alpha_i$$

and hence

$$E[Z_{i,k}] \;=\; \vartheta_k \, E[S_{i,n}]$$

such that the parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ form a development pattern for incremental quotas.

Observable random variables $\widehat{\alpha}_0^{\rm MS}, \widehat{\alpha}_1^{\rm MS}, \ldots, \widehat{\alpha}_n^{\rm MS}$ and $\widehat{\vartheta}_0^{\rm MS}, \widehat{\vartheta}_1^{\rm MS}, \ldots, \widehat{\vartheta}_n^{\rm MS}$ are said to be *marginal–sum estimators* if they are solutions to the *marginal–sum equations*

$$\sum_{l=0}^{n-i} \widehat{\alpha}_i \, \widehat{\vartheta}_l \;=\; \sum_{l=0}^{n-i} Z_{i,l}$$

for $i \in \{0, 1, \ldots, n\}$ and

$$\sum_{j=0}^{n-k} \widehat{\alpha}_j \, \widehat{\vartheta}_k \;=\; \sum_{j=0}^{n-k} Z_{j,k}$$

for $k \in \{0, 1, \ldots, n\}$ as well as

$$\sum_{l=0}^{n} \widehat{\vartheta}_l \;=\; 1$$

The marginal–sum equations remind of the identities

$$\sum_{l=0}^{n-i} \alpha_i \, \vartheta_l \;=\; \sum_{l=0}^{n-i} E[Z_{i,l}]$$

and

$$\sum_{j=0}^{n-k} \alpha_j \, \vartheta_k \;=\; \sum_{j=0}^{n-k} E[Z_{j,k}]$$

as well as

$$\sum_{k=0}^{n} \vartheta_k \;=\; 1$$

which are immediate from the model assumptions.

The question arises whether marginal–sum estimators exist and are unique. The answer to this question is affirmative: Marginal–sum estimators exist and are unique, and they satisfy

$$\widehat{\alpha}_i^{\mathrm{MS}} \;=\; \widehat{S}_{i,n}^{\mathrm{GU}}$$

and

$$\widehat{\vartheta}_k^{\mathrm{MS}} \;=\; \begin{cases} \widehat{\gamma}_0^{\mathrm{GU}} & \text{if } k=0 \\ \widehat{\gamma}_k^{\mathrm{GU}} - \widehat{\gamma}_{k-1}^{\mathrm{GU}} & \text{if } k \geq 1 \end{cases}$$

In view of the discussion of the grossing–up method, the previous identities imply that the marginal–sum estimators satisfy

$$\widehat{\alpha}_i^{\mathrm{MS}} \;=\; \widehat{S}_{i,n}^{\mathrm{CL}}$$

and

$$\widehat{\vartheta}_k^{\mathrm{MS}} \;=\; \begin{cases} \widehat{\gamma}_0^{\mathrm{CL}} & \text{if } k=0 \\ \widehat{\gamma}_k^{\mathrm{CL}} - \widehat{\gamma}_{k-1}^{\mathrm{CL}} & \text{if } k \geq 1 \end{cases}$$

Thus, letting

$$\widehat{\gamma}_k^{\mathrm{MS}} \;:=\; \sum_{l=0}^{k} \widehat{\vartheta}_l^{\mathrm{MS}}$$

we obtain

$$\widehat{\gamma}_k^{\mathrm{MS}} \;=\; \widehat{\gamma}_k^{\mathrm{CL}}$$

for all $k \in \{0, 1, \ldots, n\}$.

The *marginal–sum predictors* of the cumulative losses $S_{i,k}$ with $i+k \geq n$ are defined as

$$\widehat{S}_{i,k}^{\mathrm{MS}} \;:=\; \widehat{\gamma}_k^{\mathrm{MS}} \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}^{\mathrm{MS}}}$$

Then we have

$$\widehat{S}_{i,k}^{\mathrm{MS}} \;=\; \widehat{S}_{i,k}^{\mathrm{CL}}$$

This shows that the marginal–sum method is equivalent to the chain–ladder method.

4.6 Cape–Cod Method

The Cape–Cod method is based on the assumption that there exist parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] \;=\; \gamma_k \, E[S_{i,n}]$$

holds for all $i, k \in \{0, 1, \ldots, n\}$. Then the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The Cape–Cod method is also based on the additional assumption that *premiums* or other *volume measures* $\pi_0, \pi_1, \ldots, \pi_n \in (0, \infty)$ of the accident years are known, that the *expected ultimate cumulative loss ratios*

$$\kappa_i \;:=\; E\left[\frac{S_{i,n}}{\pi_i}\right]$$

are identical for all accident years, and that *prior estimators* $\widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_n$ of the development pattern are given and satisfy $\widehat{\gamma}_n = 1$.

The *Cape–Cod predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\widehat{S}_{i,k}^{\mathrm{CC}} \;:=\; S_{i,n-i} + \Bigl(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Bigr)\pi_i \, \widehat{\kappa}^{\mathrm{CC}}$$

where

$$\widehat{\kappa}^{\mathrm{CC}} \;:=\; \frac{\sum_{j=0}^{n} S_{j,n-j}}{\sum_{j=0}^{n} \widehat{\gamma}_{n-j}\, \pi_j}$$

is the *Cape–Cod loss ratio*, which is an estimator of the expected ultimate cumulative loss ratio (common to all accident years).

The Cape–Cod predictors are nothing else than the Bornhuetter–Ferguson predictors with respect to the prior estimators

$$\widehat{\alpha}_i^{\mathrm{CC}} \;:=\; \pi_i \, \widehat{\kappa}^{\mathrm{CC}}$$

of the expected ultimate cumulative losses. In other words, the Cape–Cod method is a particular case of the Bornhuetter–Ferguson method with prior estimators of the expected ultimate cumulative losses which are based on both internal and external information.

Example E. We consider the following reduced run–off triangle for cumulative losses which contains the cumulative losses of the present calendar year and is complemented by the premiums and the prior estimators of the development pattern:

Accident Year i	π_i	**Development Year** k 0	1	2	3	4	5
0	4025						3483
1	4456					3844	
2	5315				3977		
3	5986			3880			
4	6939		4261				
5	8158	1889					
$\widehat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

The previous triangle differs from those considered before since the value of $S_{4,1}$ is 4261 instead of 3261, which indicates that there might be an outlier in accident year 4. Using the table

i	$S_{i,5-i}$	$\widehat{\gamma}_{5-i}$	π_i	$\widehat{\gamma}_{5-i}\,\pi_i$
0	3483	1.000	4025	*4025*
1	3844	0.950	4456	*4233*
2	3977	0.860	5315	*4571*
3	3880	0.700	5986	*4190*
4	4261	0.510	6939	*3539*
5	1889	0.280	8158	*2284*
$\sum$	21334			*22842*

we obtain $\widehat{\kappa}^{\text{CC}} = 0.934$. Computing now the prior estimators of the expected ultimate cumulative losses and the Cape–Cod predictors, the run–off triangle is completed as follows:

Accident Year i	$\widehat{\alpha}_i$	**Development Year** k 0	1	2	3	4	5
0	*3758*						3483
1	*4162*					3844	*4052*
2	*4964*				3977	*4424*	*4672*
3	*5591*			3880	*4775*	*5278*	*5557*
4	*6481*		4261	*5492*	*6529*	*7113*	*7437*
5	*7619*	1889	*3641*	*5089*	*6308*	*6994*	*7375*
$\widehat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

The previous table should be compared with the following one which is the same run–off triangle completed with the loss–development predictors:

Accident Year i	**Development Year** k 0	1	2	3	4	5
0						3483
1					3844	*4046*
2				3977	*4393*	*4624*
3			3880	*4767*	*5266*	*5543*
4		4261	*5848*	*7185*	*7937*	*8355*
5	1889	*3440*	*4722*	*5802*	*6409*	*6746*
$\widehat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

The example indicates that the development of the Cape–Cod predictors over the accident years is much smoother than the development of the loss–development predictors which means that the Cape–Cod method reduces outlier effects. The smoothing effect is of course due to and depends on the premiums or other volume measures which are used instead.

The following considerations may help to understand the smoothing effect of the Cape–Cod method:

Assume that, for every accident year i, the expected ultimate cumulative loss ratio is estimated by

$$\widehat{\kappa}_i \;:=\; \frac{\widehat{S}_{i,n}^{\text{LD}}}{\pi_i} \;=\; \frac{S_{i,n-i}}{\widehat{\gamma}_{n-i}\,\pi_i}$$

Then the Cape–Cod loss ratio can be written as a weighted mean

$$\widehat{\kappa}^{\text{CC}} \;=\; \frac{\sum_{j=0}^n S_{j,n-j}}{\sum_{j=0}^n \widehat{\gamma}_{n-j}\,\pi_j} \;=\; \sum_{j=0}^n \frac{\widehat{\gamma}_{n-j}\,\pi_j}{\sum_{h=0}^n \widehat{\gamma}_{n-h}\,\pi_h}\,\widehat{\kappa}_j$$

and the identity

$$S_{i,n-i} \;=\; \widehat{\gamma}_{n-i}\,\pi_i\,\widehat{\kappa}_i$$

suggests to decompose the cumulative loss $S_{i,n-i}$ of the present calendar year into its *regular part*

$$T_{i,n-i} \;:=\; \widehat{\gamma}_{n-i}\,\pi_i\,\widehat{\kappa}^{\text{CC}}$$

and its *outlier effect*

$$X_{i,n-i} \;:=\; S_{i,n-i} - T_{i,n-i}$$

and then to apply the loss–development method to the regular part while keeping the outlier effect fixed over all subsequent development years. Since

$$\begin{aligned}
\widehat{T}_{i,k}^{\text{LD}} + X_{i,n-i} &= \widehat{\gamma}_k\,\frac{T_{i,n-i}}{\widehat{\gamma}_{n-i}} + \Big(S_{i,n-i} - T_{i,n-i}\Big) \\
&= S_{i,n-i} + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\,\frac{T_{i,n-i}}{\widehat{\gamma}_{n-i}} \\
&= S_{i,n-i} + \Big(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\Big)\,\pi_i\,\widehat{\kappa}^{\text{CC}} \\
&= \widehat{S}_{i,k}^{\text{CC}}
\end{aligned}$$

we see that the resulting predictors are precisely the Cape–Cod predictors.

The Cape–Cod method can be modified by replacing the Cape–Cod loss ratio $\widehat{\kappa}^{\text{CC}}$ by any other estimator of the form

$$\widehat{\kappa} \;=\; \sum_{j=0}^n W_j\,\widehat{\kappa}_j$$

with random variables (or constants) satisfying $\sum_{j=0}^n W_j = 1$.

4.7 Additive Method

The additive method is based on the assumption that there exist known parameters $\pi_0, \pi_1, \ldots, \pi_n \in (0, \infty)$ and unknown parameters $\zeta_0, \zeta_1, \ldots, \zeta_n$ such that the identity

$$E[Z_{i,k}] \;=\; \zeta_k \, \pi_i$$

holds for all $i, k \in \{0, 1, \ldots, n\}$.

If the parameters $\pi_0, \pi_1, \ldots, \pi_n$ are interpreted as *premiums* or other *volume measures* of the accident years, then the assumption means that, for every development year k, the *expected incremental loss ratios*

$$\zeta_{i,k} \;:=\; E\left[\frac{Z_{i,k}}{\pi_i}\right]$$

are identical for all accident years. Letting

$$\alpha_i \;:=\; \pi_i \sum_{k=0}^{n} \zeta_k$$

and

$$\gamma_k \;:=\; \frac{\sum_{l=0}^{k} \zeta_l}{\sum_{l=0}^{n} \zeta_l}$$

we obtain

$$E[S_{i,k}] \;=\; \gamma_k \, \alpha_i$$

for all $i, k \in \{0, 1, \ldots, n\}$ such that $\alpha_i = E[S_{i,n}]$ and the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The *additive predictors* of the incremental losses $Z_{i,k}$ with $i + k \geq n$ are defined as

$$\widehat{Z}_{i,k}^{\text{AD}} \;:=\; \widehat{\zeta}_k^{\text{AD}} \, \pi_i$$

and the *additive predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\widehat{S}_{i,k}^{\text{AD}} \;:=\; S_{i,n-i} + \sum_{l=n-i+1}^{k} \widehat{Z}_{i,l}^{\text{AD}}$$

where

$$\widehat{\zeta}_k^{\text{AD}} \;:=\; \frac{\sum_{j=0}^{n-k} Z_{j,k}}{\sum_{j=0}^{n-k} \pi_j}$$

is the *additive incremental loss ratio* of development year k.

Example F. We consider the following run–off triangle for cumulative losses which is complemented by the premiums:

Accident Year i	π_i	Development Year k 0	1	2	3	4	5
0	4025	1001	1855	2423	2988	3335	3483
1	4456	1113	2103	2774	3422	3844	
2	5315	1265	2433	3233	3977		
3	5986	1490	2873	3880			
4	6939	1725	3261				
5	8158	1889					

We thus obtain the following run–off triangle for incremental losses which is complemented by the additive incremental loss ratios:

Accident Year i	π_i	Development Year k 0	1	2	3	4	5
0	4025	1001	854	568	565	347	148
1	4456	1113	990	671	648	422	
2	5315	1265	1168	800	744		
3	5986	1490	1383	1007			
4	6939	1725	1536				
5	8158	1889					
$\widehat{\zeta}_k$		*0.243*	*0.222*	*0.154*	*0.142*	*0.091*	*0.037*

Computing now the additive predictors of the non–observable incremental losses, the run–off triangle of incremental losses is completed as follows:

Accident Year i	π_i	Development Year k 0	1	2	3	4	5
0	4025	1001	854	568	565	347	148
1	4456	1113	990	671	648	422	*165*
2	5315	1265	1168	800	744	*484*	*197*
3	5986	1490	1383	1007	*850*	*545*	*221*
4	6939	1725	1536	*1069*	*985*	*631*	*257*
5	8158	1889	*1811*	*1256*	*1158*	*742*	*302*
$\widehat{\zeta}_k$		*0.243*	*0.222*	*0.154*	*0.142*	*0.091*	*0.037*

Accordingly, the run–off triangle of cumulative losses is completed as follows:

Accident Year i	π_i	Development Year k 0	1	2	3	4	5
0	4025	1001	1855	2423	2988	3335	3483
1	4456	1113	2103	2774	3422	3844	*4009*
2	5315	1265	2433	3233	3977	*4461*	*4658*
3	5986	1490	2873	3880	*4730*	*5275*	*5496*
4	6939	1725	3261	*4330*	*5315*	*5946*	*6203*
5	8158	1889	*3700*	*4956*	*6114*	*6856*	*7158*

Letting

$$\widehat{\gamma}_k^{\mathrm{AD}} \;:=\; \frac{\sum_{l=0}^{k} \widehat{\zeta}_l^{\mathrm{AD}}}{\sum_{l=0}^{n} \widehat{\zeta}_l^{\mathrm{AD}}}$$

and

$$\widehat{\alpha}_i^{\mathrm{AD}} \;:=\; \pi_i \sum_{l=0}^{n} \widehat{\zeta}_l^{\mathrm{AD}}$$

the additive predictors of the non–observable cumulative losses may be written as

$$\widehat{S}_{i,k}^{\mathrm{AD}} \;:=\; S_{i,n-i} + \left(\widehat{\gamma}_k^{\mathrm{AD}} - \widehat{\gamma}_{n-i}^{\mathrm{AD}}\right)\widehat{\alpha}_i^{\mathrm{AD}}$$

This shows that the additive predictors of the cumulative losses are nothing else than the Bornhuetter–Ferguson predictors with respect to the *additive cumulative quotas* $\widehat{\gamma}_k^{\mathrm{AD}}$ and the prior estimators $\widehat{\alpha}_i^{\mathrm{AD}}$ of the expected ultimate cumulative losses. In other words, the additive method is a particular case of the Bornhuetter–Ferguson method with prior estimators of the cumulative quotas and of the expected ultimate cumulative losses which are based on both internal and external information.

The *expected cumulative loss ratios*

$$\kappa_i \;:=\; E\left[\frac{S_{i,n}}{\pi_i}\right]$$

satisfy

$$\kappa_i \;=\; \sum_{l=0}^{n} \zeta_{i,l}$$

Since the expected incremental loss ratios are identical for all accident years, it follows that also the expected cumulative loss ratios are identical for all accident years. Therefore, the *additive loss ratio*

$$\widehat{\kappa}^{\mathrm{AD}} \;:=\; \sum_{l=0}^{n} \widehat{\zeta}_l^{\mathrm{AD}}$$

can be interpreted as an estimator of the expected ultimate cumulative loss ratio

$$\kappa \;=\; \sum_{l=0}^{n} \zeta_l$$

common to all accident years. Moreover, the prior estimators $\widehat{\alpha}_i^{\mathrm{AD}}$ can be written as

$$\widehat{\alpha}_i^{\mathrm{AD}} \;=\; \pi_i\, \widehat{\kappa}^{\mathrm{AD}}$$

and it can be shown that

$$\widehat{\kappa}^{\mathrm{AD}} \;=\; \frac{\sum_{j=0}^{n} S_{j,n-j}}{\sum_{j=0}^{n} \widehat{\gamma}_{n-j}^{\mathrm{AD}}\, \pi_j}$$

This shows that the additive predictors of the non–observable cumulative losses are nothing else than the Cape–Cod predictors with respect to the additive cumulative quotas $\widehat{\gamma}_k^{\mathrm{AD}}$. In other words, the additive method is a particular case of the Cape–Cod method with prior estimators of the cumulative quotas which are based on both internal and external information.

The observation that the additive method is a special case of the Cape–Cod method is due to Zocher [2005].

4.8 Remarks

The following table compares the different methods of loss reserving considered in this section with regard to the choices of the prior estimators of the expected ultimate cumulative losses α_i and of the cumulative quotas γ_k:

Expected Ultimate	**Cumulative Quotas**		
Cumulative Losses	Arbitrary	$\widehat{\gamma}_k^{\mathrm{CL}}$	$\widehat{\gamma}_k^{\mathrm{AD}}$
Arbitrary	Bornhuetter–Ferguson Method		
$\widehat{S}_{i,n}^{\mathrm{LD}}$	Loss–Development Method	Chain–Ladder Method	
$\pi_i \, \widehat{\kappa}^{\mathrm{CC}}$	Cape–Cod Method		Additive Method

Note that the prior estimators $S_{i,n}^{\mathrm{LD}}$ and $\pi_i \, \widehat{\kappa}^{\mathrm{CC}}$ depend on the choice of the prior estimators $\widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_n$.

Of course, the four other combinations which apparently have not been given a name in the literature could be used as well, and even other choices of the prior estimators of the expected ultimate cumulative losses and of the cumulative quotas could be considered.

The discussion of the present section and, in particular, the above table shows that the Bornhuetter–Ferguson method provides a general principle under which several methods of loss reserving can be subsumed. The focus

– on prior estimators of the expected ultimate cumulative losses and
– on prior estimators of the cumulative quotas

provides a large variability of loss reserving methods. The above table contains important special cases but could certainly be enlarged.

Moreover,

– any convex combination of prior estimators of the expected ultimate cumulative losses yields new prior estimators of the expected ultimate cumulative losses, and

– any convex combination of prior estimators of the development pattern for cumulative quotas yields new prior estimators of the development pattern.

This point is made precise in the following example:

Example G. Let $\widehat{\alpha}_0, \widehat{\alpha}_1, \ldots, \widehat{\alpha}_n$ be prior estimators of $\alpha_0, \alpha_1, \ldots, \alpha_n$ and let $\widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_n$ be prior estimators of $\gamma_0, \gamma_1, \ldots, \gamma_n$ such that each of these prior estimators is completely based on external information. Then the prior estimators

$$\widetilde{\alpha}_i \quad := \quad a_1 \, \widehat{\alpha}_i + a_2 \, \widehat{S}_{i,n}^{\mathrm{LD}} + a_3 \, (\pi_i \, \widehat{\kappa}^{\mathrm{CC}})$$

with $a_1 + a_2 + a_3 = 1$ and

$$\widetilde{\gamma}_k \quad := \quad b_1 \, \widehat{\gamma}_k + b_2 \, \widehat{\gamma}_k^{\mathrm{CL}} + b_3 \, \widehat{\gamma}_k^{\mathrm{AD}}$$

with $b_1 + b_2 + b_3 = 1$ are prior estimators of $\alpha_0, \alpha_1, \ldots, \alpha_n$ and $\gamma_0, \gamma_1, \ldots, \gamma_n$, respectively, which through the weights a_1, a_2, a_3 and b_1, b_2, b_3 express the reliability attributed to the prior estimators $\widehat{\alpha}_i, \widehat{S}_{i,n}^{\mathrm{LD}}, \pi_i \, \widehat{\kappa}^{\mathrm{CC}}$ and $\widehat{\gamma}_k, \widehat{\gamma}_k^{\mathrm{CL}}, \widehat{\gamma}_k^{\mathrm{AD}}$, respectively.

5 Least–Squares Prediction

Least–squares prediction is one of the general principles of statistical inference. It is similar to least–squares estimation but differs from the latter since the target quantity is a non–observable random variable instead of a model parameter.

The main aspects of least–squares prediction are credibility prediction and Gauss–Markov prediction; in either case, the problem is to determine optimal predictors with respect to the expected squared prediction error.

An extension of Gauss–Markov prediction is conditional Gauss–Markov prediction in which unconditional first and second order moments are replaced by conditional moments.

5.1 Credibility Prediction

In the context of loss reserving, credibility prediction aims at predicting any linear combination T of (observable or non–observable) incremental losses by a predictor of the form

$$\widehat{T} \;=\; a + \sum_{j=0}^{n} \sum_{l=0}^{n-j} a_{j,l} Z_{j,l}$$

These predictors are said to be *admissible*. Note that
– the class of all admissible predictors does not depend on the sum to be predicted,
– the admissible predictors are not necessarily linear in the observable incremental losses since the coefficient a may be distinct from 0, and
– the admissible predictors are not assumed to be unbiased.

The general form of the prediction problem is reasonable since it includes, e. g., prediction of the ultimate cumulative losses $S_{i,n}$ which are sums of the observable incremental losses $Z_{i,0}, Z_{i,1}, \dots, Z_{i,n-i}$ and the non–observable incremental losses $Z_{i,n-i+1}, \dots, Z_{i,n}$.

For a sum T of incremental losses, an admissible predictor is said to be a *credibility predictor* of T if it minimizes the *expected squared prediction error*

$$E[(\widehat{T} - T)^2]$$

over all admissible predictors $\widehat{T}$.

The following results are well–known:
(1) For every sum T of incremental losses, there exists a credibility predictor $\widehat{T}^{\mathrm{CR}}$ and the credibility predictor is unique.

(2) If T_1 and T_2 are sums of incremental losses and if c_1 and c_2 are real numbers, then the credibility predictor of

$$T := c_1 T_1 + c_2 T_2$$

satisfies

$$\widehat{T}^{\mathrm{CR}} = c_1 \widehat{T}_1^{\mathrm{CR}} + c_2 \widehat{T}_2^{\mathrm{CR}}$$

which means that credibility prediction is linear.

(3) If T is a sum of incremental losses, then an admissible predictor $\widehat{T}^*$ is the credibility predictor of T if and only if it satisfies the *normal equations*

$$E[\widehat{T}^*] = E[T]$$

and

$$E[\widehat{T}^* Z_{j,l}] = E[T Z_{j,l}]$$

for all $j, l \in \{0, 1, \ldots, n\}$ such that $j + l \leq n$.

(4) The credibility predictor of any sum of incremental losses is unbiased.

Because of (2) it is sufficient to determine the credibility predictors of the incremental losses $Z_{i,k}$. In the case where $i + k \leq n$, we have

$$\widehat{Z}_{i,k}^{\mathrm{CR}} = Z_{i,k}$$

In the case where $i + k \geq n + 1$, we write

$$\widehat{Z}_{i,k}^{\mathrm{CR}} = a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} Z_{h,m}$$

and determine the coefficients from the *normal equations*

$$E\left[a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} Z_{h,m}\right] = E[Z_{i,k}]$$

and

$$E\left[\left(a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} Z_{h,m}\right) Z_{j,l}\right] = E[Z_{i,k} Z_{j,l}]$$

which may equivalently be written as

$$a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m}\, E[Z_{h,m}] = E[Z_{i,k}]$$

and

$$\sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} \operatorname{cov}[Z_{h,m}, Z_{j,l}] \;=\; \operatorname{cov}[Z_{i,k}, Z_{j,l}]$$

for all $j, l \in \{0, 1, \ldots, n\}$ such that $j + l \leq n$.

We thus see that the credibility predictor of a non–observable incremental loss is completely determined by the first and second order moments of the incremental losses. Solving the normal equations proceeds in two steps:

– The normal equations involving covariances form a system of linear equations for the coefficients $a_{i,k,h,m}$. The fact that a credibility predictor of $Z_{i,k}$ exists implies that this system of linear equations has at least one solution.
– Inserting any such solution into the normal equation involving expectations yields the coefficient $a_{i,k}$.

It should be noted that the system of linear equations may have several solutions (which is the case if and only if the covariance matrix of the observable cumulative losses is singular). This means that the credibility predictor of $Z_{i,k}$, which is known to be unique, can be *represented* in several ways.

In most credibility models for loss reserving which have been considered in the literature, it is assumed that any two incremental losses from different accident years are uncorrelated. In this case, the credibility predictor of a non–observable incremental loss $Z_{i,k}$ can be written as

$$\widehat{Z}_{i,k}^{\text{CR}} \;=\; a_{i,k} + \sum_{m=0}^{n-i} a_{i,k,i,m} Z_{i,m}$$

and its coefficients can be determined from the reduced normal equations

$$a_{i,k} + \sum_{m=0}^{n-i} a_{i,k,i,m} \, E[Z_{i,m}] \;=\; E[Z_{i,k}]$$

and

$$\sum_{m=0}^{n-i} a_{i,k,i,m} \operatorname{cov}[Z_{i,m}, Z_{i,l}] \;=\; \operatorname{cov}[Z_{i,k}, Z_{i,l}]$$

for all $l \in \{0, 1, \ldots, n-i\}$.

As an example, let us now consider credibility prediction in the credibility model of Witting, which is a model for claim counts:

Credibility Model of Witting:

(i) *Any two incremental losses of different accident years are uncorrelated.*

(ii) *There exist parameters* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n \in (0,1)$ *with* $\sum_{l=0}^n \vartheta_l = 1$ *such that, for every accident year* $i \in \{0,1,\ldots,n\}$*, the conditional joint distribution of the family* $\{Z_{i,k}\}_{k\in\{0,1,\ldots,n\}}$ *with respect to the ultimate cumulative loss* $S_{i,n}$ *is the multinomial distribution with parameters* $S_{i,n}$ *and* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$*.*

For the remainder of this subsection we assume that the assumptions of the credibility model of Witting are fulfilled. Then we have

$$\begin{aligned} E(Z_{i,k}|S_{i,n}) &= S_{i,n}\,\vartheta_k \\ \mathrm{cov}(Z_{i,k}, Z_{i,l}|S_{i,n}) &= \begin{cases} -\,S_{i,n}\,\vartheta_k^2 + S_{i,n}\,\vartheta_k & \text{if } k = l \\ -\,S_{i,n}\,\vartheta_k\vartheta_l & \text{else} \end{cases} \end{aligned}$$

Letting

$$\begin{aligned} \alpha_i &:= E[S_{i,n}] \\ \sigma_i &:= \mathrm{var}[S_{i,n}] \end{aligned}$$

we obtain

$$\begin{aligned} E[Z_{i,k}] &= \alpha_i\,\vartheta_k \\ \mathrm{cov}[Z_{i,k}, Z_{i,l}] &= \begin{cases} (\sigma_i - \alpha_i)\,\vartheta_k^2 + \alpha_i\,\vartheta_k & \text{if } k = l \\ (\sigma_i - \alpha_i)\,\vartheta_k\vartheta_l & \text{else} \end{cases} \end{aligned}$$

The first of the previous identities shows that the parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ form a development pattern for incremental quotas. Inserting the previous identities into the normal equations, we obtain, for all $i, k \in \{0,1,\ldots,n\}$ such that $i + k \geq n+1$,

$$\widehat{Z}_{i,k}^{\mathrm{CR}} = \vartheta_k \left(\frac{1}{1+\gamma_{n-i}\,\tau_i}\,\alpha_i + \frac{\gamma_{n-i}\,\tau_i}{1+\gamma_{n-i}\,\tau_i}\,\frac{S_{i,n-i}}{\gamma_{n-i}} \right)$$

and hence

$$\begin{aligned} \widehat{S}_{i,k}^{\mathrm{CR}} &= S_{i,n-i} + \sum_{l=n-i+1}^{k} \widehat{Z}_{i,l}^{\mathrm{CR}} \\ &= S_{i,n-i} + \Big(\gamma_k - \gamma_{n-i}\Big) \left(\frac{1}{1+\gamma_{n-i}\,\tau_i}\,\alpha_i + \frac{\gamma_{n-i}\,\tau_i}{1+\gamma_{n-i}\,\tau_i}\,\frac{S_{i,n-i}}{\gamma_{n-i}} \right) \end{aligned}$$

where $\gamma_k := \sum_{l=0}^k \vartheta_l$ and $\tau_i := (\sigma_i - \alpha_i)/\alpha_i$. This shows that the credibility predictor of the non–observable cumulative loss $S_{i,k}$ is the Bornhuetter–Ferguson predictor with respect to the prior estimators

$$\widehat{\gamma}_k := \gamma_k$$

of the development pattern for cumulative quotas and the prior estimators

$$\widehat{\alpha}_i^{\mathrm{CR}} \;:=\; \frac{1}{1+\gamma_{n-i}\,\tau_i}\,\alpha_i + \frac{\gamma_{n-i}\,\tau_i}{1+\gamma_{n-i}\,\tau_i}\,\frac{S_{i,n-i}}{\gamma_{n-i}}$$

of the expected ultimate cumulative losses, which are weighted means of external information provided by the unknown parameter α_i and internal information provided by the loss–development predictor $\widehat{S}_{i,n}^{\mathrm{LD}} = S_{i,n-i}/\gamma_{n-i}$.

Example H. If, in addition to the assumptions of the model of Witting, it is assumed that every ultimate cumulative loss $S_{i,n}$ has the Poisson distribution with expectation α_i, then we have $\tau_i = 0$ and the credibility predictors of every non–observable cumulative loss $S_{i,k}$ satisfy

$$\widehat{S}_{i,k}^{\mathrm{CR}} \;=\; S_{i,n-i} + \Big(\gamma_k - \gamma_{n-i}\Big)\alpha_i$$

and are thus identical with the Bornhuetter–Ferguson estimators with respect to the prior estimators $\widehat{\gamma}_k := \gamma_k$ and $\widehat{\alpha}_i := \alpha_i$. In this case, the assumptions of the Poisson model are fulfilled and maximum–likelihood estimation could be used as an alternative to credibility prediction; see subsection 6.1 below.

Similar results obtain in the credibility model of Mack [1990] and in a special case of the credibility model of Hesselager and Witting [1998]; see Radtke and Schmidt [2004].

5.2 Gauss–Markov Prediction

A predictor $\widehat{T}$ of a linear combination T of (observable or non–observable) incremental losses is said to be

– a *linear predictor* if there exists a family $\{a_{j,l}\}_{j,l\in\{0,1,\ldots,n\},l+j\le n}$ of coefficients such that

$$\widehat{T} \;=\; \sum_{j=0}^{n}\sum_{l=0}^{n-j} a_{j,l}Z_{j,l}$$

– an *unbiased predictor* of T if

$$E[\widehat{T}] \;=\; E[T]$$

– a *Gauss–Markov predictor* of T if it is an unbiased linear predictor of T which minimizes the *expected squared prediction error*

$$E[(\widehat{T}-T)^2]$$

over all unbiased linear predictors $\widehat{T}$ of T.

The existence of a Gauss–Markov predictor of T cannot be guaranteed in general. (For example, if $E[Z_{i,k}] = 0$ holds for every observable every incremental loss and if T is such that $E[T] \neq 0$, then there exists no unbiased linear estimator of T.) Therefore, we consider Gauss–Markov prediction only under the assumptions of the linear model.

Let $\mathbf{Z}_1$ denote a random vector consisting of the observable incremental losses and let $\mathbf{Z}_2$ denote a random vector consisting of the non–observable incremental losses (arranged in any order).

Linear Model:

(i) *There exist matrices* $\mathbf{A}_1$ *and* $\mathbf{A}_2$ *and a vector* $\boldsymbol{\beta}$ *such that*

$$\begin{aligned} E[\mathbf{Z}_1] &= \mathbf{A}_1\boldsymbol{\beta} \\ E[\mathbf{Z}_2] &= \mathbf{A}_2\boldsymbol{\beta} \end{aligned}$$

(ii) *The matrix* $\mathbf{A}_1$ *has full column rank.*

(iii) *The matrix*

$$\boldsymbol{\Sigma}_{11} := \mathrm{var}[\mathbf{Z}_1]$$

is invertible.

For the remainder of this subsection, we assume that the assumptions of the linear model are fulfilled.

Under the assumptions of the linear model. the following results are well–known:

(1) For every sum T of incremental losses, there exists a Gauss–Markov predictor $\widehat{T}^{\mathrm{GM}}$ and the Gauss–Markov predictor is unique.

(2) If T_1 and T_2 are sums of incremental losses and if c_1 and c_2 are real numbers, then the Gauss–Markov predictor of

$$T := c_1T_1 + c_2T_2$$

satisfies

$$\widehat{T}^{\mathrm{GM}} = c_1\widehat{T}_1^{\mathrm{GM}} + c_2\widehat{T}_2^{\mathrm{GM}}$$

which means that Gauss–Markov prediction is linear.

Because of (2) it is sufficient to determine the Gauss–Markov predictors of the incremental losses $Z_{i,k}$. In the case where $i + k \leq n$, we have

$$\widehat{Z}_{i,k}^{\mathrm{GM}} = Z_{i,k}$$

In the case where $i + k \geq n + 1$, we obtain

$$\widehat{Z}_{i,k}^{\mathrm{GM}} = \mathbf{a}'_{i,k}\widehat{\boldsymbol{\beta}}^{\mathrm{GM}} + \mathrm{cov}[Z_{i,k}, \mathbf{Z}_1]\,\boldsymbol{\Sigma}_{11}^{-1}\Big(\mathbf{Z}_1 - \mathbf{A}_1\widehat{\boldsymbol{\beta}}^{\mathrm{GM}}\Big)$$

where $\mathbf{a}'_{i,k}$ is the row vector of the matrix $\mathbf{A}_2$ satisfying $E[Z_{i,k}] = \mathbf{a}'_{i,k}\boldsymbol{\beta}$,

$$\widehat{\boldsymbol{\beta}}^{\mathrm{GM}} \;:=\; (\mathbf{A}'_1\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}'_1\boldsymbol{\Sigma}_{11}^{-1}\mathbf{Z}_1$$

is the *Gauss–Markov estimator* of $\boldsymbol{\beta}$ (based on the observable incremental losses) and $\mathrm{cov}[Z_{i,k}, \mathbf{Z}_1]$ is the row vector with entries $\mathrm{cov}[Z_{i,k}, Z_{j,l}]$ with $j, l \in \{0, 1, \ldots, n\}$ and $j + l \leq n$; see Goldberger [1962] and Rao and Toutenburg [1995] as well as Halliwell [1996, 1999], Hamer [1999] and Schmidt [1998, 1999a, 2004].

As an example, let us now consider Gauss–Markov prediction in the linear model of Mack:

> **Linear Model of Mack:** *There exist parameters* $\pi_0, \pi_1, \ldots, \pi_n \in (0, \infty)$ *and* $\zeta_0, \zeta_1, \ldots, \zeta_n$ *as well as* $\sigma_0, \sigma_1, \ldots, \sigma_n \in (0, \infty)$ *such that*
>
> $$E[Z_{i,k}] \;=\; \pi_i\,\zeta_k$$
>
> *and*
>
> $$\mathrm{cov}[Z_{i,k}, Z_{j,l}] \;=\; \begin{cases} \pi_i\,\sigma_k & \text{if } i = j \text{ and } k = l \\ 0 & \text{else} \end{cases}$$
>
> *holds for all* $i, j, k, l \in \{0, 1, \ldots, n\}$.

For the remainder of this subsection we assume that the assumptions of the linear model of Mack are fulfilled. Define

$$\boldsymbol{\beta} \;:=\; \begin{pmatrix} \zeta_0 \\ \zeta_1 \\ \vdots \\ \zeta_n \end{pmatrix}$$

and, for all $i, k \in \{0, 1, \ldots, n\}$,

$$\mathbf{a}'_{i,k} \;:=\; \begin{pmatrix} 0 & \ldots & 0 & \pi_i & 0 & \ldots & 0 \end{pmatrix}$$

where π_i occurs in position $1 + k$. This shows that the linear model of Mack satisfies indeed the assumptions of the linear model. For the Gauss–Markov estimator of $\boldsymbol{\beta}$ we obtain

$$\widehat{\boldsymbol{\beta}}^{\mathrm{GM}} \;=\; \begin{pmatrix} \dfrac{\sum_{j=0}^{n} Z_{j,0}}{\sum_{j=0}^{n} \pi_j} \\[2ex] \dfrac{\sum_{j=0}^{n-1} Z_{j,1}}{\sum_{j=0}^{n-1} \pi_j} \\ \vdots \\ \dfrac{Z_{0,n}}{\pi_0} \end{pmatrix}$$

Since $\text{cov}[Z_{i,k}, Z_{j,l}] = 0$ holds for all $i, j, k, l \in \{0, 1, \dots, n\}$ such that $i + k \geq n + 1$ and $j + l \leq n$, it follows that the Gauss–Markov predictor of the non–observable incremental loss $Z_{i,k}$ satisfies

$$\widehat{Z}_{i,k}^{\text{GM}} = \pi_i \frac{\sum_{j=0}^{n-k} Z_{j,k}}{\sum_{j=0}^{n-k} \pi_j}$$

and hence

$$\widehat{Z}_{i,k}^{\text{GM}} = \widehat{Z}_{i,k}^{\text{AD}}$$

and linearity of Gauss–Markov prediction yields

$$\widehat{S}_{i,k}^{\text{GM}} = \widehat{S}_{i,k}^{\text{AD}}$$

This shows that the additive method is justified by Gauss–Markov prediction in the linear model of Mack.

5.3 Conditional Gauss–Markov Prediction

In the present subsection we consider a sequential model for the chain–ladder method. This model is a sequential model since it involves successive conditioning with respect to the σ–algebras $\mathcal{G}_0, \mathcal{G}_1, \dots, \mathcal{G}_{n-1}$ where, for each $k \in \{1, \dots, n\}$, the σ–algebra

$$\mathcal{G}_{k-1}$$

represents the information provided by the cumulative losses $S_{j,l}$ of accident years $j \in \{0, 1, \dots, n-k+1\}$ and development years $l \in \{0, 1, \dots, k-1\}$, which is at the same time the information provided by the incremental losses $Z_{j,l}$ of accident years $j \in \{0, 1, \dots, n-k+1\}$ and development years $l \in \{0, 1, \dots, k-1\}$.

Sequential Chain–Ladder Model: *For each $k \in \{1, \dots, n\}$, there exists a random variable φ_k and a strictly positive random variable σ_k such that*

$$E^{\mathcal{G}_{k-1}}(S_{i,k}) = S_{i,k-1}\, \varphi_k$$

and

$$\text{cov}^{\mathcal{G}_{k-1}}(S_{i,k}, S_{j,k}) = \begin{cases} S_{i,k-1}\, \sigma_k & \textit{if } i = j \\ 0 & \textit{else} \end{cases}$$

holds for all $i, j \in \{0, 1, \dots, n-k+1\}$.

In the case where the random variables $\varphi_1, \dots, \varphi_n$ are all constant, integration yields $E[S_{i,k}] = \varphi_k\, E[S_{i,k-1}]$ such that the parameters $\varphi_1, \dots, \varphi_n$ form a development pattern for factors. In the general case, the random parameters $\varphi_1, \dots, \varphi_n$ may be interpreted as a *random development pattern* for factors.

The sequential chain–ladder model may be considered as a sequence of n conditional linear models corresponding to the development years $k \in \{1, \ldots, n\}$. Each of these conditional linear models consists of an observable part

$$\begin{pmatrix} E^{\mathcal{G}_{k-1}}(S_{0,k}) \\ E^{\mathcal{G}_{k-1}}(S_{1,k}) \\ \vdots \\ E^{\mathcal{G}_{k-1}}(S_{n-k,k}) \end{pmatrix} = \begin{pmatrix} S_{0,k-1} \\ S_{1,k-1} \\ \vdots \\ S_{n-k,k-1} \end{pmatrix} \varphi_k$$

and a non–observable part

$$E^{\mathcal{G}_{k-1}}(S_{n-k+1,k}) = S_{n-k+1,k-1}\,\varphi_k$$

Then the $\mathcal{G}_{k-1}$–conditional Gauss–Markov estimator $\widehat{\varphi}_k^{\mathrm{GM}}$ of the random parameter φ_k satisfies

$$\widehat{\varphi}_k^{\mathrm{GM}} = \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

and hence coincides with the chain–ladder factor $\widehat{\varphi}_k^{\mathrm{CL}}$.

Furthermore, for every accident year $i \geq n-k+1$, the $\mathcal{G}_{k-1}$–conditional Gauss–Markov predictor $\widehat{S}_{i,k}^{\mathrm{GM}}$ of the non–observable cumulative loss $S_{i,k}$ satisfies

$$\begin{aligned} \widehat{S}_{i,k}^{\mathrm{GM}} &= S_{i,k-1}\,\widehat{\varphi}_k^{\mathrm{GM}} \\ &= S_{i,k-1}\,\widehat{\varphi}_k^{\mathrm{CL}} \end{aligned}$$

The previous formula, however, is only useful when $S_{i,k-1}$ is observable, which is the case if and only if $i+k-1 \leq n$ and hence $i = n-k+1$.

Turning the point of view from development years to accident years, we see that the $\mathcal{G}_{n-i}$–conditional Gauss–Markov predictors of the first non–observable cumulative losses $S_{i,n-i+1}$ satisfy

$$\widehat{S}_{i,n-i+1}^{\mathrm{GM}} = S_{i,n-i}\,\widehat{\varphi}_{n-i+1}^{\mathrm{CL}}$$

and hence coincide with the chain–ladder predictors.

In the case $i+k = n+1$, the chain–ladder predictors are thus justified by conditional Gauss–Markov estimation, but another justification is needed in the case $i+k \geq n+2$. This can be achieved by minimizing the $\mathcal{G}_{k-1}$–conditional expected prediction error

$$E^{\mathcal{G}_{k-1}}\Big(\Big(\widehat{S}_{i,k} - S_{i,k}\Big)^2\Big)$$

over the collection of all predictors $\widehat{S}_{i,k}$ of $S_{i,k}$ satisfying

$$\widehat{S}_{i,k} = \widehat{S}_{i,k-1}^{\mathrm{CL}}\widehat{\varphi}_k$$

for some $\mathcal{G}_{k-1}$–conditionally unbiased linear estimator $\widehat{\varphi}_k$ of φ_k, and it turns out that the minimum over this restricted class of predictors is attained for the chain–ladder predictor $\widehat{S}_{i,k}^{\mathrm{CL}}$. The sequential optimality criterion adopted here reflects very well the sequential character of the chain–ladder method and of the chain–ladder model. The criterion is also reasonable since prediction for the first non–observable calendar year is much more important than prediction for subsequent calendar years: Predictors for the first non–observable calendar year cannot be corrected later whereas predictors for subsequent calender years will be corrected anyway since already one year later additional loss experience and hence a new run–off triangle will be available.

The sequential chain–ladder model is due to Schnaus and was proposed by Schmidt and Schnaus [1996] where it is studied in detail; see also Schmidt [1997, 1999b, 2006]. The sequential chain–ladder model is a slight but convenient extension of the chain–ladder model of Mack [1993]. A systematic comparison of several models for the chain–ladder method is given in Hess and Schmidt [2002].

5.4 Remarks

Although least–squares prediction is a central topic in econometrics, it appears that this method has been ignored in loss reserving until recently. It is the merit of Halliwell [1996] that least–squares prediction is by now considered as a most useful tool in loss reserving; see also Schmidt [1999a], Hamer [1999], Halliwell [1999], Radtke and Schmidt [2004], and Schmidt [2006].

6 Maximum–Likelihood Estimation

Another general principle of statistical inference is maximum–likelihood estimation. The maximum–likelihood principle is applicable only if the joint distribution of all observable random variables is known with the exception of certain parameters.

The models considered here are models for claim counts. The basic model is the Poisson model which is a special case of the general multinomial model.

6.1 Poisson Model

The Poisson model is a model for claim counts and consists of the following assumptions:

Poisson model:

(i) *The family* $\{Z_{i,k}\}_{i,k\in\{0,1,\ldots,n\}}$ *of all incremental losses is independent.*
(ii) *There exist parameters* $\alpha_0,\alpha_1,\ldots,\alpha_n \in (0,\infty)$ *and* $\vartheta_0,\vartheta_1,\ldots,\vartheta_n \in (0,1)$ *with* $\sum_{l=0}^{n}\vartheta_l = 1$ *such that, for all* $i,k \in \{0,1,\ldots,n\}$*, the incremental loss* $Z_{i,k}$ *has the Poisson distribution with expectation* $\alpha_i\,\vartheta_k$*.*

We assume in this subsection that the assumptions of the Poisson model are fulfilled.

Because of (ii) we have

$$E[Z_{i,k}] \;=\; \alpha_i\,\vartheta_k$$

Summation yields

$$E[S_{i,n}] \;=\; \alpha_i$$

and hence

$$E[Z_{i,k}] \;=\; \vartheta_k\,E[S_{i,n}]$$

such that the parameters $\vartheta_0,\vartheta_1,\ldots,\vartheta_n$ form a development pattern for incremental quotas.

In the Poisson model the joint distribution of all incremental losses is known except for the parameters. In fact, we have

$$P\left[\bigcap_{i=0}^{n}\bigcap_{k=0}^{n}\{Z_{i,k}=z_{i,k}\}\right] \;=\; \prod_{i=0}^{n}\prod_{k=0}^{n}\left(e^{-\alpha_i\vartheta_k}\,\frac{(\alpha_i\vartheta_k)^{z_{i,k}}}{z_{i,k}!}\right)$$

To estimate the parameters we can thus use the maximum–likelihood method. The maximum–likelihood method is based in the joint distribution of *all observable incremental losses* which is given by

$$P\left[\bigcap_{i=0}^{n}\bigcap_{k=0}^{n-i}\{Z_{i,k}=z_{i,k}\}\right] \;=\; \prod_{i=0}^{n}\prod_{k=0}^{n-i}\left(e^{-\alpha_i\vartheta_k}\,\frac{(\alpha_i\vartheta_k)^{z_{i,k}}}{z_{i,k}!}\right)$$

It follows that the likelihood function L is given by

$$L\Big(\alpha_0, \alpha_1, \ldots, \alpha_n, \vartheta_0, \vartheta_1, \ldots, \vartheta_n \;\Big|\; \mathcal{Z}\Big) \;:=\; \prod_{i=0}^{n} \prod_{k=0}^{n-i} \left(e^{-\alpha_i \vartheta_k} \, \frac{(\alpha_i \vartheta_k)^{Z_{i,k}}}{Z_{i,k}!} \right)$$

where $\mathcal{Z} := \{Z_{i,k}\}_{i,k \in \{0,1,\ldots,n\}, i+k \leq n}$. Interpreting the maximum–likelihood principle in a wide sense (which ignores the second order conditions for a maximum), observable random variables

$$\widehat{\alpha}_0^{\mathrm{ML}}, \widehat{\alpha}_1^{\mathrm{ML}}, \ldots, \widehat{\alpha}_n^{\mathrm{ML}}$$

and

$$\widehat{\vartheta}_0^{\mathrm{ML}}, \widehat{\vartheta}_1^{\mathrm{ML}}, \ldots, \widehat{\vartheta}_n^{\mathrm{ML}}$$

are said to be *maximum–likelihood estimators* if they annihilate all first order partial derivatives of the likelihood function (or, equivalently, of the log–likelihood function) and satisfy the constraint

$$\sum_{l=0}^{n} \widehat{\vartheta}_l^{\mathrm{ML}} \;=\; 1$$

Straightforward computation shows that the maximum–likelihood estimators satisfy the marginal–sum equations

$$\sum_{l=0}^{n-i} \widehat{\alpha}_i \, \widehat{\vartheta}_l \;=\; \sum_{l=0}^{n-i} Z_{i,l}$$

with $i \in \{0, 1, \ldots, n\}$ and

$$\sum_{j=0}^{n-k} \widehat{\alpha}_j \, \widehat{\vartheta}_k \;=\; \sum_{j=0}^{n-k} Z_{j,k}$$

with $k \in \{0, 1, \ldots, n\}$ and, of course, the constraint

$$\sum_{l=0}^{n} \widehat{\vartheta}_l \;=\; 1$$

Therefore, the maximum–likelihood estimators coincide with the marginal–sum estimators. It now follows from the properties of the marginal–sum estimators that in the Poisson model the maximum–likelihood estimators of the expected ultimate cumulative losses are identical with the chain–ladder predictors of the ultimate cumulative losses. This was first observed by Hachemeister and Stanard [1975].

However, if, in addition to the assumptions of the Poisson model, it is assumed that the expected ultimate cumulative losses are all identical such that

$$\alpha_i \;=\; \alpha$$

holds for all $i \in \{0, 1, \ldots, n\}$, then maximum–likelihood estimation is still possible but the maximum–likelihood estimators turn out to satisfy

$$\widehat{\alpha} \;=\; \sum_{l=0}^{n} \frac{1}{n-l+1} \sum_{j=0}^{n-k} Z_{j,l}$$

and

$$\widehat{\vartheta}_k \;=\; \frac{\displaystyle \frac{1}{n-k+1} \sum_{j=0}^{n-k} Z_{j,k}}{\displaystyle \sum_{l=0}^{n} \frac{1}{n-l+1} \sum_{j=0}^{n-l} Z_{j,l}}$$

In particular, the maximum–likelihood estimators of the expected ultimate cumulative losses are *not* identical with the chain–ladder estimators of the ultimate cumulative losses; see Schmidt and Zocher [2005].

6.2 Multinomial Model

The multinomial model is a model for claim counts and consists of the following assumptions:

Multinomial Model:

(i) *The accident years are independent.*
(ii) *There exist parameters* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n \in (0, 1)$ *with* $\sum_{l=0}^{n} \vartheta_l = 1$ *such that, for every accident year* $i \in \{0, 1, \ldots, n\}$*, the conditional joint distribution of the family* $\{Z_{i,k}\}_{k \in \{0,1,\ldots,n\}}$ *with respect to the ultimate cumulative loss* $S_{i,n}$ *is the multinomial distribution with parameters* $S_{i,n}$ *and* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$.

We assume in this subsection that the assumptions of the multinomial model are fulfilled.

Because of (ii) we have

$$E(Z_{i,k} | S_{i,n}) \;=\; \vartheta_k \, S_{i,n}$$

and hence

$$E[Z_{i,k}] \;=\; \vartheta_k \, E[S_{i,n}]$$

such that the parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ form a development pattern for incremental quotas.

The multinomial model is appealing since it suggests that every claim of any accident year is reported or settled with probability ϑ_k in development year k. It thus reminds of the urn model in which $S_{i,n}$ balls are drawn with replacement from an urn consisting of balls with $1+n$ different colours corresponding to the development years.

Letting

$$\alpha_i \ := \ E[S_{i,n}]$$

it is easy to see that the multinomial model contains the Poisson model as the special case in which every ultimate cumulative loss $S_{i,n}$ has the Poisson distribution with expectation α_i. Moreover, under the assumptions of the multinomial model, it can be shown that the incremental losses of any accident year are independent if and only if the family of all incremental losses is independent and every incremental loss has the Poisson distribution with expectation $\alpha_i \vartheta_k$. Therefore, the main advantage of the multinomial model over the Poisson model is the fact that it allows for dependence between the incremental losses of a given accident year.

If, in addition to the assumptions of the multinomial model, the distributions of the ultimate cumulative losses are assumed to belong to a parametric family of distributions, then the joint distribution of all incremental losses is known except for the parameters and maximum–likelihood estimation can be used to estimate the expected ultimate cumulative losses.

In the case where each of the ultimate cumulative losses has a Poisson distribution, we are back to the Poisson model and the maximum–likelihood estimators of the expected ultimate cumulative losses are identical with the chain–ladder predictors of the ultimate cumulative losses.

The same result obtains in the case where each of the ultimate cumulative losses has a negativebinomial distribution; see Schmidt and Wünsche [1998]. Negativebinomial distributions are of interest since they are mixed Poisson distributions (with respect to a mixing gamma distribution), and mixed Poisson distributions in turn are of interest since their variances exceed their expectations, which is the case for most empirical claim count distributions.

In fact, a much more general result is true: If, in addition to the assumptions of the multinomial model, each of the ultimate cumulative losses has a Hofmann distribution, then the maximum–likelihood estimators of the expected ultimate cumulative losses are identical with the chain–ladder predictors of the ultimate cumulative losses; see Schmidt and Zocher [2005]. The definition and the discussion of Hofmann distributions are beyond the scope of this paper, but we remark that Hofmann distributions were introduced by Hofmann [1955] and that every Hofmann distribution is at the same time a mixed Poisson distribution and a compound Poisson distribution and can be computed by recursion; see e. g. Hess, Liewald and Schmidt [2002].

Since the class of all Hofmann distributions is a wide class of mixed Poisson distributions, the multinomial model with ultimate cumulative loss counts having a Hofmann distribution is a very general model for claim counts in which the maximum–likelihood estimators of the expected ultimate cumulative losses are identical with the chain–ladder predictors of the ultimate cumulative losses.

6.3 Remarks

Alternatively, the Poisson model can be extended to a general stochastic model in which the family $\{Z_{i,k}\}_{i,k\in\{0,1,\dots,n\}}$ is independent and the distribution of every incremental loss belongs to an exponential family. In such models, the theory of generalized linear models can be applied.

7 Conclusions

The notion of a development pattern, which can be expressed in three different but equivalent ways, provides a powerful tool for the comparison of different methods and of different models of loss reserving.

The general Bornhuetter–Ferguson method provides a general framework into which several methods of loss reserving can be embedded via
- a particular choice of the prior estimators of the development pattern for cumulative quotas and/or
- a particular choice of the prior estimators of the expected ultimate cumulative losses.

Moreover, there are many stochastic models in which
- the credibility predictors or
- the Gauss–Markov predictors or
- the maximum–likelihood estimators of the expected ultimate cumulative losses

can be interpreted as Bornhuetter–Ferguson predictors.

The choice of a stochastic model or a method of prediction is a choice which has to be made by the actuary and may have a considerable impact on the result. In the Poisson model, e. g., credibility prediction and maximum–likelihood estimation are possible but lead to different results; here the choice of the statistical method could be based on the judgement that either external information or internal information is more reliable. Still in the Poisson model, the form of the maximum–likelihood estimators of the expected ultimate cumulative losses depends on the assumption that the expected ultimate cumulative losses may be different or are all identical.

We also remark that the chain–ladder method and the additive method can be extended to the multivariate case which corresponds to a portfolio consisting of several subportfolios representing dependent lines of business. Moreover, the multivariate chain–ladder method and the multivariate additive method can be justified by multivariate models extending the univariate models considered in the present paper. A detailed discussion of these multivariate methods and models may be found in Schmidt [2006].

References

Benktander, G. [1976]: *An approach to credibility in calculating IBNR for casualty excess reinsurance.* The Actuarial Review, April **7**.

Goldberger, A. S. [1962]: *Best linear unbiased prediction in the generalized linear regression model.* J. Amer. Statist. Assoc. **57**, 369–375.

Hachemeister, C. A., and Stanard, J. N. [1975]: *IBNR claims count estimation with static lag functions.* Unpublished.

Halliwell, L. J. [1996]: *Loss prediction by generalized least squares.* Proc. CAS **83**, 436–489.

Halliwell, L. J. [1999]: *Loss prediction by generalized least squares – Author's response.* Proc. CAS **86**, 764–769.

Hamer, M. D. [1999]: *Loss prediction by generalized least squares – Discussion.* Proc. CAS **86**, 748–763.

Hess, K. T., Liewald, A., and Schmidt, K. D. [2002]: *An extension of Panjer's recursion.* ASTIN Bull. **32**, 283–297.

Hess, K. T., and Schmidt, K. D. [2002]: *A comparison of models for the chain–ladder method.* Insurance Mathematics and Economics **31**, 351–364.

Hesselager, O., and Witting, T. [1988]: *A credibility model with random fluctuations in delay probabilities for the prediction of IBNR claims.* ASTIN Bull. **18**, 79–90.

Hofmann, M. [1955]: *Über zusammengesetzte Poisson–Prozesse und ihre Anwendungen in der Unfallversicherung.* Mitt. VSVM **55**, 499–575.

Hovinen, E. [1981]: *Additive and continuous IBNR.* Paper presented at the International ASTIN Colloquium in Loen.

Lorenz, H., and Schmidt K. D. [1999]: *Grossing–up, chain–ladder and marginal–sum estimation.* Blätter DGVM **24**, 195–200.

Mack, T. [1990]: *Improved estimation of IBNR claims by credibility theory.* Insurance Mathematics and Economics **9**, 51–57.

Mack, T. [1993]: *Distribution–free calculation of the standard error of chain–ladder reserve estimates.* ASTIN Bull. **23**, 213–225.

Mack, T. [2000]: *Credible claims reserves – The Benktander method.* ASTIN Bull. **30**, 333–347.

Neuhaus, W. [1992]: *Another pragmatic loss reserving method or Bornhuetter–Ferguson revisited.* Scand. Actuar. J., 151–162.

Radtke, M., and Schmidt, K. D. (eds.) [2004]: *Handbuch zur Schadenreservierung.* Karlsruhe: Verlag Versicherungswirtschaft.

Rao, C. R., and Toutenburg, H. [1995]: *Linear Models – Least Squares and Alternatives.* Berlin – Heidelberg – New York: Springer.

Schmidt, K. D. [1997]: *Non–optimal prediction by the chain–ladder method.* Insurance Mathematics and Economics **21**, 17–24.

Schmidt, K. D. [1998]: *Prediction in the linear model – A direct approach.* Metrika **48**, 141–147.

Schmidt, K. D. [1999a]: *Loss prediction by generalized least squares – Discussion.* Proc. CAS **86**, 736–747.

Schmidt, K. D. [1999b]: *Chain–ladder prediction and asset–liability management.* Blätter DGVM **24**, 1–9.

Schmidt, K. D. [2004]: *Prediction.* In: *Encyclopedia of Actuarial Science*, vol. 3, pp. 1317–1321. Chichester: Wiley.

Schmidt, K. D. [2006]: *Optimal and additive loss reserving for dependent lines of business.* CAS 2006 Call Paper Program.

Schmidt, K. D., and Schnaus, A. [1996]: *An extension of Mack's model for the chain–ladder method.* ASTIN Bull. **26**, 247–262.

Schmidt, K. D., and Wünsche, A. [1998]: *Chain–ladder, marginal–sum and maximum–likelihood estimation.* Blätter DGVM **23**, 289–307.

Schmidt, K. D., and Zocher, M. [2005]: *Loss reserving and Hofmann distributions.* Bull. SAV, 127–162.

Taylor, G. C. [2000]: *Loss Reserving – An Actuarial Perspective.* Boston – Dordrecht – London: Kluwer.

Vaughn, T. R., and Tinney, P. [2004]: *The modified Bornhuetter–Ferguson approach to IBNR allocation.* Proc. CAS **91**, 1–13.

Witting, T. [1987]: *Kredibilitätsschätzungen für die Anzahl IBNR–Schäden.* Blätter DGVM **18**, 45–58.

Zocher, M. [2005]: Personal communication.

Klaus D. Schmidt
Lehrstuhl für Versicherungsmathematik
Technische Universität Dresden
D–01062 Dresden

e–mail: `klaus.d.schmidt@tu-dresden.de`

July 15, 2006