Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE

Jaroslav Vosáhlo

Moderní stochastické metody výpočtu škodních rezerv v pojišťovnictví a jejich porovnání

Katedra pravděpodobnosti a matematické statistiky, MFF UK

Vedoucí diplomové práce: RNDr. Michal Pešta, Ph.D.
Studijní program: Finanční a pojistná matematika

Praha, 2013

Charles University in Prague
Faculty of Mathematics and Physics

MASTER THESIS

Jaroslav Vosáhlo

Modern stochastic claims reserving methods in insurance and their comparison

Department of Probability and Mathematical Statistics, MFF UK

Supervisor of the master thesis: RNDr. Michal Pešta, Ph.D.
Study programme: Financial and Insurance Mathematics

Prague, 2013

I would like to thank my supervisor RNDr. Michal Pešta, Ph.D. for precise guidance and useful advice. I also wish to thank Bc. Peter Rusnak for helpful consultations on various topics. My gratitude goes to R-software creators and contributors, Casualty Actuarial Society and Generali Pojišťovna, a.s., which proposed and supported this thesis. Finally, I would like to thank my family and friends for support and companionship during my studies.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Prague date .... Jaroslav Vosáhlo

Contents

Název práce: Moderní stochastické metody tvorby škodních rezerv v pojištění a jejich porovnání
Autor: Jaroslav Vosáhlo
Katedra (ústav): Matematicko-fyzikální fakulta, Univerzita Karlova v Praze
Vedoucí diplomové práce: RNDr. Michal Pešta, Ph.D.
e-mail vedoucího: pesta@karlin.mff.cuni.cz

Abstrakt: Tato práce se zabývá stochastickými rezervovacími metodami používanými v neživotním pojištění. Problém je řešen analytickými metodami a stochastickým modelováním. Nejprve jsou představeny základní rezerovací metody, t.j. Chain-ladder, Bornhuetter-Ferguson, Benktander-Hovinen a Cape-Cod s jejich vlastnostmi a principy. V další části pak hledáme jejich stochastické rozšíření za použití zobecněných lineáních modelů (GLM) a Mackových obecných (nedistribučních) přístupů, zkoumáme druhé momenty odhadnutých škodních rezerv a zavedeme Merz-Wüthrichův způsob měření rizika škodních rezerv. Na závěr je navržen algoritmus na aplikaci bootstrapové simulace a výsledky analytických rezervovacích metod a simulací jsou porovnány.

Klíčová slova: Stochastické modelování rezerv, neživotní pojištění, bootstrapping

Title: Modern stochastic claims reserving methods in insurance and their comparison
Author: Jaroslav Vosáhlo
Department: Faculty of mathematics and physics, Charles University
Supervisor: RNDr. Michal Pešta, Ph.D.
Supervisor's e-mail address: pesta@karlin.mff.cuni.cz

Abstract: This thesis deals with an issue of claims reserving for non-life insurance. The issue is approached in a sense of analytical calculation and stochastic modelling. First, Chain-ladder, Bornhuetter-Ferguson, Benktander-Hovinen and Cape-Cod method are introduced. In following chapters, we try to find related stochastic underlying models including Generalized linear models and Mack's distribution-free approaches, we analyze second moments of claims estimates for each of the methods and examine alternative Merz-Wüthrich approach to reserve risk measurement. At the end, bootstrap algorithm and estimates are suggested and simulation results are compared with analytic ones.

Keywords: Stochastic reserving, non-life insurance, bootstrapping

Chapter 1

Introduction

This thesis presents an overview of common stochastic claims reserving methods in non-life insurance. Examining variability of claims reserves became an important component of reserving process in last decades, for the most part due to more common usage of dynamic risk models, more thorough solvency requirements and rating agencies expectations. Yet, there is a obstacle in expansion of stochastic claims reserving methods usage, since even if we choose a specific analytic reserving method, there is not usually only one generally accepted stochastic approach. This issue is presented in this paper - we show several ways, how common non-stochastic reserving methods can be reproduced/extended to stochastic ones.

We present both distributional and distribution-free models, with each having their own pros and cons. Distributional models can, under certain assumptions, provide us with possible hypothesis testing and outcome distribution (analytic of simulation). On the other hand, there is a considerable loss of generality and it is usually difficult to rationally justify the choice of the specific distribution class. Distribution-free models are essentially more general, but their structure is mostly defined to achieve a specific goal (in our case obtaining an analytic formula for second moments of ultimate claim estimate), which could be difficult to justify as well. Also treatment of higher than second moments is usually out of reach. Great potential is offered with bootstrap simulations, since it provides estimate of the shape of the parameters distribution even in distribution-free models.

Purpose of this paper is to compare more approaches to stochastic reserving and therefore we are not able to show the whole derivations or proofs for every statement. The structure is arranged so that the reader finds for each model its assumptions, including their interpretations, parameter estimates and major model a estimates properties. In case of stochastic methods, we present estimates of claims reserves' second moments with explaining the main steps of the derivation including estimation assumptions, attached with a reference to detailed derivation. The reader should be able to apply all methods based on information from this paper.

Note, that in methods presented in this paper, we do not consider claims development after the last observed development year. There have been suggestions for tail factors for example in Mack Chain-ladder or Mack Bornhuetter-Ferguson method, but since the issue represents more expert or regression problem specific to each data set, the tail factors are not considered.

In Chapter 3 we introduce Chain-ladder, Bornhuetter-Ferguson, Benktander-Hovinen and Cape-Cod methods as purely computational methods to obtain ultimate claims estimate. Principles of individual methods are explained, followed by model assumptions and parameter estimates.

Chapter 4 deals with several stochastic versions of Chain-ladder, Bornhuetter-Ferguson and Benktander-Hovinen. We introduce Generalized linear models framework and connect the Over-dispersed Poisson model and Over-dispersed Negative Binomial model with Chain-ladder technique. Next, we state Mack's versions of Chain-ladder and Bornhuetter-Ferguson method and search for optimal credibity mixture in Benktander-Hovinen method. Conditional and unconditional versions of mean square errors are introduced and estimated for each approach.

In Chapter 5 we subscribe the bootstrap algorithm for structured data and state how it could be applied in claims reserving methods. In Chapter 6 the methods are applied on real data and results are compared.

1.1 Notions and notations

$X_{i,j}$	incremental payments at time j from accident year i
$C_{i,j}$	cumulative payments until time j from accident year i
D_I	available observations at time I (upper triangle)
B_k	set of observations at time I with development year up to k
$R_{i,j}$	outstanding loss liabilities for accident year i at time j
R_i	outstanding loss liabilities for accident year i at time I
R	total outstanding loss liabilities at time I
$\hat{X}$	prediction or estimation of X
X^B	bootstrap pseudo-observation
$\mathbf{b}'$	transposition of row vector b
$\mathsf{E}[\cdot]$	expected value
$\mathsf{Var}(\cdot)$	variance
$\mathsf{msep}_{(\cdot)}(\cdot)$	unconditional mean square of error
$\mathsf{msep}_{(\cdot\|\cdot)}(\cdot)$	conditional mean square of error
BF	Bornhuetter-Ferguson
BH	Benktander-Hovinen
CC	Cape-Cod
CDR	claims development result
CL	Chain-ladder
EDF	Exponential dispersion family
GLM	Generalized linear models
ODP	Over-dispersed Poisson model
ODNB	Over-dispersed Negative Binomial model
MLE	maximum likelihood estimator

Chapter 2

Claims process

Claims reserving issue results from a delay between claim occurrence and settlement. The target is to cover the unsettled claims of past exposures. These are

- A. claims, that have occurred, but have not been reported before the end of insurance period (*IBNR* - incurred but not reported)

- B. claims that have been reported, but we still expect claim costs in the future (*RBNS* - reported but not settled).

For the claims development process description, we use common notation from Merz and Wüthrich [13]. We assume that we are at time I and one period represents one year. Previous accident/origin years are denoted by $i \in \{0, \dots, I\}$ and development years are denoted by $j \in \{0, \dots, J\}$, with $I \geq J$. $X_{i,j}$ denotes incremental payments in period j for claims that occurred at accident year i and $C_{i,j}$ denotes cumulative payments for accident year i until development year j

$$C_{i,j} = \sum_{k=0}^{j} X_{i,k}. \tag{2.1}$$

$C_{i,J}$ is called ultimate claim and we assume, that there is no (significant) claims development after the delay J.

Observations available at time I are given by upper triangle

$$D_I = \{X_{i,j}; i + j \leq I, i \leq I\}. \tag{2.2}$$

In specific cases, we will work with a set of observations with development year lower or equal to k

$$B_k = \{X_{i,j}; i + j \leq I, 0 \leq j \leq k\} \subseteq D_I. \tag{2.3}$$

The outstanding loss liabilities for accident year i at time j are given by

$$R_{i,j} = C_{i,J} - C_{i,j}. \tag{2.4}$$

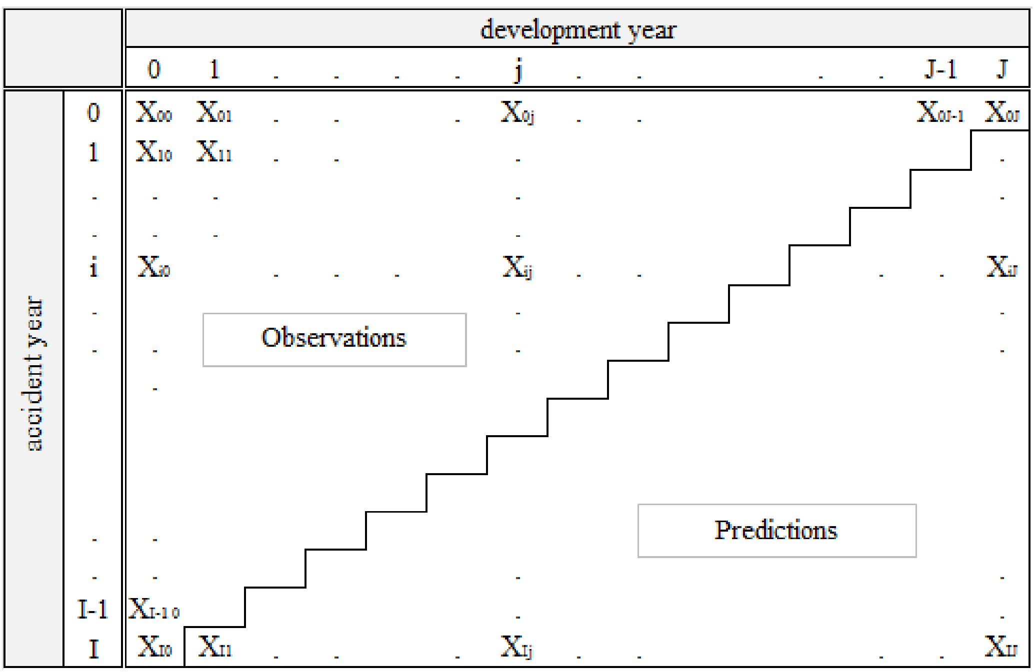
Present outstanding loss liabilities (at time I) for accident year i are indexed only by accident year:

$$R_i = C_{i,J} - C_{i,I-i}. \tag{2.5}$$

Total outstanding loss liabilities at time I are given by

$$R = \sum_{i=0}^{I} R_i. \tag{2.6}$$

Figure 2.1: Development triangle for incremental claims



Chapter 3

Methods overview

In this chapter we introduce four analytic claims reserving methods used in this work. Each method is defined with its assumptions; further, we subscribe model properties and algorithm for parameters a ultimate claim estimation.

3.1 Chain-ladder method

One of the most common methods used for claims development triangle analysis is the Chain-ladder method. Despite its simplicity, this method often shows more accurate results than more complex models and it is commonly used as a basis for comparison with other methods.

Assumptions 1 (Chain-ladder) *We introduce (distribution-free) Chain-ladder model using following assumptions:*

1. *Cumulative claims $C_{i,j}$ of different accident years i are independent.*

2. *There exist development factors $f_0, \ldots, f_{J-1} > 0$ such that for all $0 \leq i \leq I$ and all $0 \leq j \leq J$ we have*

$$\mathsf{E}[C_{i,j} \mid C_{i,0}, \ldots, C_{i,j-1}] = \mathsf{E}[C_{i,j} \mid C_{i,j-1}] = f_{j-1} C_{i,j-1}. \tag{3.1}$$

Lemma 1 (Chain-ladder properties) *Under Assumptions 1 we have:*

- *For $0 \leq i \leq I$*

$$\mathsf{E}[C_{i,J} \mid D_I] = \mathsf{E}[C_{i,j} \mid C_{i,I-i}] = C_{i,I-i} f_{I-i} \ldots f_{J-1}. \tag{3.2}$$

- *Unbiased estimator for Chain-ladder factors $f_j, j = 0, \ldots, J-1$ is given by*

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}. \tag{3.3}$$

- *Unbiased Chain-ladder estimator for $\mathsf{E}[C_{i,J} \mid D_I]$ is given by:*

$$\hat{C}_{i,J}^{CL} = C_{i,I-i}\hat{f}_{I-i} \ldots \hat{f}_{J-1}. \tag{3.4}$$

Proof can be found in Merz and Wüthrich [13].

3.2 Bornhuetter-Ferguson method

Bornhuetter-Ferguson method considers prior estimates of expected ultimate claim $\mu_i = E[C_{i,J}], 0 \leq i \leq I$. These values are mostly based on expert opinion and should be determined before the end of development year 0.

Assumptions 2 (Bornhuetter-Ferguson) *We introduce Bornhuetter-Ferguson model using following assumptions:*

1. *Cumulative claims $C_{i,j}$ of different accident years i are independent.*

2. *There exist parameters $\mu_0, \ldots, \mu_I > 0$ and a pattern $\beta_0, \ldots, \beta_J > 0$ with $\beta_J = 1$ such that for all $0 \leq i \leq I$ and all $0 \leq j \leq J-1$ and $1 \leq k \leq j-J$ we have*

$$\mathsf{E}[C_{i,0}] = \beta_0 \mu_i, \tag{3.5}$$

$$\mathsf{E}[C_{i,j+k} \mid C_{i,0}, \ldots, C_{i,j}] = C_{i,j} + (\beta_{j+k} - \beta_j)\mu_i. \tag{3.6}$$

Under these assumptions we have:

$$\mathsf{E}[C_{i,j}] = \beta_j \mu_i, \tag{3.7}$$

$$\mathsf{E}[C_{i,J}] = \mu_i. \tag{3.8}$$

To find a relationship between Chain-ladder and Bornhuetter-Ferguson, we have to restrain ourselves only on Assumption 3.7, which is implied by Chain-ladder Assumptions 1. Under Assumption 3.7 the estimators for the pattern $\beta_0, \ldots, \beta_J$ can be obtained using Chain-ladder development factors as follows:

$$\hat{\beta}_j = \prod_{k=j}^{J-1} \frac{1}{\hat{f}_k}. \tag{3.9}$$

Unbiased Bornhuetter-Ferguson estimator for $\mathsf{E}[C_{i,J} \mid D_I]$ is given by

$$\hat{C}_{i,J}^{BF} = C_{i,I-i} + (1 - \hat{\beta}_{I-i})\hat{\mu}_i. \tag{3.10}$$

Proof can be found in Merz and Wüthrich [13].

Note 1 *Another approach to parameters estimation within the Bornhuetter-Ferguson method was suggested by Mack [8] and can be found in Section 4.5.*

3.3 Benktander-Hovinen method

Benktander-Hovinen method combines Chain-ladder and Bornhuetter-Ferguson approach. While Chain-ladder estimate of expected ultimate claim $E[C_{i,J}]$ fully depends on observations, Bornhuetter-Ferguson on the contrary uses only prior estimate μ_i (adjusted by the last-diagonal claim). In this method we use $\hat{C}_{i,J}^{BF}$ as a prior estimate and with increasing knowledge of claim development we iterate it using a credibility mixture of both previous approaches

$$u_i(c) = c\hat{C}_{i,J}^{CL} + (1 - c)\mu_i, \tag{3.11}$$

for $0 \leq c \leq 1$ increasing with obtaining better information on $C_{i,J}$.

Lemma 2 (Benktander-Hovinen properties) *For estimating Benktander-Hovinen ultimate claim $\hat{C}_{i,J}^{BH}$ we choose so called Gunnar Benktander substitution $c = \beta_{I-i}$. Then we have*

$$\hat{C}_{i,J}^{BH} = C_{i,I-i} + (1 - \beta_{I-i})\Big(\beta_{I-i}\hat{C}_{i,J}^{CL} + (1 - \beta_{I-i})\mu_i\Big). \tag{3.12}$$

- *If we assume that $(\beta_j)_{0 \leq j \leq J}$ is known and use identification 3.9, we have*

$$\hat{C}_{i,J}^{BH} = C_{i,I-i} + (1 - \beta_{I-i})\hat{C}_{i,J}^{BF}. \tag{3.13}$$

- *Then, for $m \geq 0$, we can define an iteration process:*

$$\hat{C}^{m+1} = C_{i,I-i} + (1 - \beta_{I-i})\hat{C}^m, \tag{3.14}$$

$$\lim_{m \to \infty} \hat{C}^m = \hat{C}_{i,J}^{CL}, \tag{3.15}$$

where $\hat{C}^0 = \mu_i$.

Proof can be found in Mack [9].

3.4 Cape-Cod method

Cape-Cod method was suggested to deal with a dependency of the estimated ultimate claim on all observations within the particular accident year. For that, we use modified values of diagonal observations, that reflect long-term trend and should eliminate the outliers.

Assumptions 3 (Cape-Cod method) *We introduce Cape-Cod method using following assumptions:*

1. *Cumulative claims $C_{i,j}$ of different accident years i are independent.*
2. *There exist parameters $\pi_0, \ldots, \pi_I > 0, \kappa > 0$ and a claims development pattern $\beta_0, \ldots, \beta_J$ with $\beta_J = 1$ such that for all $0 \leq i \leq I$ and all $0 \leq j \leq J$ we have*

$$\mathsf{E}[C_{i,j}] = \kappa \pi_i \beta_j. \tag{3.16}$$

Lemma 3 (Cape-Cod properties) *Under Assumptions 3 and identification (3.9) we have:*

- *Unbiased Cape-Cod estimator for κ is given by*

$$\hat{\kappa}^{CC} = \frac{\sum_{i=1}^{I} C_{i,I-i}}{\sum_{i=1}^{I} \beta_{I-i}\pi_i}. \tag{3.17}$$

- *For estimating the ultimate claim, we use modified values of diagonal observations*

$$\hat{C}_{i,I-i}^{CC} = \hat{\kappa}^{CC} \pi_i \beta_{I-i}. \tag{3.18}$$

- *Cape-code estimator is then given by*

$$\hat{C}_{i,J}^{CC} = C_{i,I-i} - \hat{C}_{i,I-i}^{CC} + \prod_{j=I-i}^{J-1} f_j \hat{C}_{i,I-i}^{CC}. \tag{3.19}$$

Proof can be found in Merz and Wüthrich [13].

Note 2 (Cape-Cod vs. Bornhuetter-Ferguson method) *We assume that $\hat{\mu}_i = \hat{C}_{i,J} = \hat{\kappa}\pi_i$, where π_i represents premium volume and κ loss ratio. Opposed to BF method, where the loss ratio κ is estimated priori and therefore is independent of claims development, in the CC method the loss ratio (3.17) changes/adjusts with increasing knowledge of the claims development.*

Chapter 4

Stochastic reserving

In the previous chapter several reserving models have been introduced as mechanical methods to obtain an estimate of the ultimate claim $\hat{C}_{i,J}$. Main advantage of these models is their simplicity - they can be easily understood and programmed, but they provide information only about the expected value and do not consider variability of the estimate. When we want to estimate second moments, quantiles or even the full predictive distribution of the ultimate claim, we introduce stochastic reserving models.

In this chapter we will try to expand the analytic methods from Chapter 3 with additional stochastic framework and then examine their stochastic properties. This is done via:

- additional assumptions on variance,
- additional assumptions on distribution,
- defining a new stochastic model that justifies/implies usage of a specific analytic method.

In first two approaches we begin with our non-stochastic model and search for reasonable stochastic extension. Concerning the third approach, the procedure is reversed, since we begin with a stochastic model and try to fit it to our non-stochastic one.

This chapter starts with a definition of a criterion for measuring the variability of prediction. Based on it, we can perform a comparison of models' precision.

4.1 Mean square error of prediction

Common approach to measure quality/uncertainty of prediction is considering second moments, i.e. determining *(conditional) mean square error of prediction* (MSEP). Conditional MSEP for the estimate of ultimate claim $\hat{C}_{i,J}$ is defined as follows

$$\mathsf{msep}_{C_{i,J}|D_I}(\hat{C}_{i,J}) = \mathsf{E}[(C_{i,J} - \hat{C}_{i,J})^2 \mid D_I]. \tag{4.1}$$

For a D_I-measurable estimate $\hat{C}_{i,J}$, the conditional mean square error of prediction can be decomposed

$$\mathsf{msep}_{C_{i,J}|D_I}(\hat{C}_{i,J}) = \underbrace{\mathsf{Var}(C_{i,J} \mid D_I)}_{\text{conditional process variance}} + \underbrace{(\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])^2}_{\text{parameter estimation error}}. \tag{4.2}$$

If we assume that $\hat{R}_i = \hat{C}_{i,J} - C_{i,I-i}$, then all the uncertainty of the estimate $\hat{C}_{i,J}$ is contained in estimated future claims $\hat{X}_{i,j}, j > I-i$ and we can equivalently express the conditional MSEP of $\hat{C}_{i,J}$ as MSEP of estimated outstanding loss liabilities

$$\mathsf{msep}_{C_{i,J}|D_I}(\hat{C}_{i,J}) = \mathsf{msep}_{R_i|D_I}(\hat{R}_i) = \mathsf{E}[(R_i - \hat{R}_i)^2 \mid D_I]. \tag{4.3}$$

The unconditional mean square error of prediction is then defined as expected value of conditional MSEP:

$$\mathsf{msep}_{C_{i,J}}(\hat{C}_{i,J}) = \mathsf{E}[\mathsf{msep}_{C_{i,J}|D_I}(\hat{C}_{i,J})] = \mathsf{E}[\mathsf{Var}(C_{i,J} \mid D_I)] + \mathsf{E}[(\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])^2] \tag{4.4}$$

$$= \mathsf{Var}(C_{i,J}) + \mathsf{E}[(\hat{C}_{i,J} - \mathsf{E}[C_{i,J}])^2] - 2\mathsf{E}\Big[\Big(\hat{C}_{i,J} - \mathsf{E}[C_{i,J}]\Big)\Big(\mathsf{E}[C_{i,J} \mid D_I] - \mathsf{E}[C_{i,J}]\Big)\Big]. \tag{4.5}$$

If $\hat{C}_{i,J}$ is unbiased estimator for $\mathsf{E}[C_{i,J}]$, then we have

$$\mathsf{msep}_{C_{i,J}}(\hat{C}_{i,J}) = \mathsf{Var}(C_{i,J}) + \mathsf{Var}(\hat{C}_{i,J}) - 2\mathsf{Cov}(\hat{C}_{i,J}, \mathsf{E}[C_{i,J} \mid D_I]). \tag{4.6}$$

In all methods introduced in Chapter 3, we have assumed that the claims developments of different accident years are independent. Under this assumption, the total conditional process variance of aggregated ultimate claims $\sum_{i=1}^{I} C_{i,J}$ is a sum of individual process variance errors. This shall not apply in case of parameter estimation errors for different accident years, since there is no general assumption on claims estimates independence and therefore the conditional MSEP for aggregated ultimate claims cannot be simply obtained by summing MSEP of individual accident years. We have following relationship:

$$\mathsf{msep}_{\sum_{i=1}^{I} C_{i,J}|D_I}(\sum_{i=1}^{I} \hat{C}_{i,J}) = \sum_{i=1}^{I} \mathsf{Var}(C_{i,J} \mid D_I) + (\sum_{i=1}^{I} \hat{C}_{i,J} - \sum_{i=1}^{I} \mathsf{E}[C_{i,J} \mid D_I])^2 =$$

$$= \sum_{i=1}^{I} \mathsf{Var}(C_{i,J} \mid D_I) + \sum_{i=1}^{I} (\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])^2 + \\ +2 \sum_{1 \leq i < k \leq J} (\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])(\hat{C}_{k,J} - \mathsf{E}[C_{k,J} \mid D_I]). \tag{4.7}$$

4.2 Distribution-free stochastic Chain-ladder

Due to simplicity and frequent usage, Chain-ladder stochastic models represent the majority of researched stochastic claims reserving models. In this section we state distribution-free stochastic models, that have been suggested for estimation of conditional MSEP. We start with Mack Chain-ladder model and state two approaches based on different additional assumptions; for the whole derivation see Mack [7] and Buchwalder et al. [3].

Assumptions 4 (Mack Chain-ladder)
We introduce Mack Chain-ladder method using following assumptions:

1. *Cumulative claims $C_{i,j}$ of different accident years i are independent.*

2. *$(C_{i,j})_{j \geq 0}$ are Markov processes for all $0 \leq i \leq I$ and there exist development factors $f_0, ..., f_{J-1} > 0$ and variance parameters $\sigma_0^2, \dots, \sigma_{J-1}^2 > 0$ such that for all $0 \leq i \leq I$ and all $0 \leq j \leq J$ we have*

$$\mathsf{E}[C_{i,j} \mid C_{i,j-1}] = f_{j-1} C_{i,j-1}, \tag{4.8}$$

$$\mathsf{Var}(C_{i,j} \mid C_{i,j-1}) = \sigma_{j-1}^2 C_{i,j-1}. \tag{4.9}$$

Lemma 4 (Variance estimate in Mack CL) *Under Assumptions 4 we have unbiased estimate for σ_j^2, $0 \leq j \leq J-1$ in a form of*

$$\hat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2. \tag{4.10}$$

Proof can be found in Merz and Wüthrich [13].

Note 3 *We can see, that previous lemma does not provide an instruction of how to estimate the last variance parameter* $\hat{\sigma}^2_{J-1}$*. This is caused by the fact that there is not enough data to estimate variability of claims development with such a delay. Mack [7] suggested putting* $\hat{\sigma}^2_{J-1} = \min(\hat{\sigma}^4_{J-2}/\hat{\sigma}^2_{J-3}, \min(\hat{\sigma}^2_{J-2}, \hat{\sigma}^2_{J-3}))$*. Since the lack of data for bigger* j *can generally bring less representative variance estimates, we can go even further and extrapolate more than just the last element using for example fitted regression model.*

Theorem 1 (Conditional process variance for Mack CL)
Under Assumptions 4, using the unbiasedness of estimates $\hat{f}_j$ *(3.3) and* $\hat{\sigma}^2_j$ *(4.10), the conditional process variance for ultimate claim* $C_{i,J}$ *is given by*

$$\mathsf{Var}(C_{i,J} \mid D_I) = (\mathsf{E}[C_{i,J} \mid C_{i,I-i}])^2 \sum_{j=I-i}^{J-1} \frac{\sigma^2_j / f^2_j}{\mathsf{E}[C_{i,j} \mid C_{i,I-i}]} \tag{4.11}$$

and estimated by

$$\widehat{\mathsf{Var}}(C_{i,J} \mid D_I) = (\hat{C}^{CL}_{i,J})^2 \sum_{j=I-i}^{J-1} \frac{\hat{\sigma}^2_j / \hat{f}^2_j}{\hat{C}^{CL}_{i,j}}. \tag{4.12}$$

Proof can be found in Mack [7].

Parameter estimation error is given by:

$$\begin{aligned}(\hat{C}^{CL}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])^2 &= C^2_{i,I-i}(\hat{f}_{I-i} \cdots \hat{f}_{J-1} - f_{I-i} \cdots f_{J-1})^2 = \\ &= C^2_{i,I-i}\Big(\prod_{j=I-i}^{J-1} \hat{f}^2_j + \prod_{j=I-i}^{J-1} f^2_j - 2 \prod_{j=I-i}^{J-1} \hat{f}_j f_j \Big).\end{aligned} \tag{4.13}$$

Since $\mathsf{E}[C_{i,J} \mid D_I]$ as well as true values of development factors f_j are unknown, the explicit analytical determination of parameter estimation error is not possible and approximation through additional assumptions is needed. We start with Mack's derivation following Mack [7].

Assumptions 5 (Mack CL estimate of parameter estimation error)
We define S_k *as*

$$S_k = \hat{f}_{I+1-i} \cdots \hat{f}_{k-i}(f_k - \hat{f}_k) f_{k+1} \cdots f_{I-1}. \tag{4.14}$$

Then, we assume that:

1. *S_k^2 can be estimated by $\mathsf{E}[S_k^2 \mid B_k]$,*

2. *$S_j S_k, j < k$ can be estimated by $\mathsf{E}[S_j S_k \mid B_k]$.*

Theorem 2 (Parameter estimation error for Mack CL)
Under Assumptions 5 the parameter estimation error for Mack Chain-ladder estimate $C_{i,J}$ is given by

$$(\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])^2 = (\hat{C}_{i,J}^{CL})^2 \sum_{j=I-i}^{J-1} \frac{\hat{\sigma}_j^2/\hat{f}_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}}. \tag{4.15}$$

Proof can be found in Mack [7].

In the following theorem we bring previous estimators together to obtain conditional MSEP and we complete it with an estimator for aggregated accident years following (4.7).

Theorem 3 (Mack CL MSEP - single and aggregated accident years)
Under Assumptions 5 we have following estimators for conditional MSEP of outstanding claims for single and aggregated accident years:

$$\widehat{\mathsf{msep}}_{C_{i,J}|D_I}^{Mack}(\hat{C}_{i,J}^{CL}) = (\hat{C}_{i,J}^{CL})^2 \sum_{j=I-i}^{J-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2}\Big(\frac{1}{\hat{C}_{i,j}^{CL}} + \frac{1}{\sum_{i=0}^{I-j-1} C_{i,j}}\Big), \tag{4.16}$$

$$\begin{aligned} \widehat{\mathsf{msep}}_{\sum_{i=1}^{I} C_{i,J}|D_I}^{Mack}\Big(\sum_{i=1}^{I} \hat{C}_{i,J}^{CL}\Big) &= \sum_{i=1}^{I} \widehat{\mathsf{msep}}_{C_{i,J}|D_I}^{Mack}(\hat{C}_{i,J}^{CL}) + \\ &+ 2 \sum_{1 \le i < k \le I} \hat{C}_{i,J}^{CL} \hat{C}_{k,J}^{CL} \sum_{j=I-i}^{J-1} \frac{\hat{\sigma}_j^2/\hat{f}_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}}. \end{aligned} \tag{4.17}$$

Proof can be found in Mack [7].

Buchwalder et al. [3] deal with another approach of estimating the parameter estimation error. It is based on introduction of a stronger model which implies Mack Chain-ladder model:

Assumptions 6 (Autoregressive Mack Chain-ladder)

1. *Cumulative claims $C_{i,j}$ of different accident years i are independent.*

2. *There exist development factors* $f_0, \dots, f_{J-1} > 0$*, variance parameters* $\sigma_0^2, \dots, \sigma_{J-1}^2 > 0$ *and random variables* $\varepsilon_{i,j}$*, such that for all* $0 \leq i \leq I$ *and all* $0 \leq j \leq J$ *we have*

$$C_{i,j+1} = f_j C_{i,j} + \sigma_j \sqrt{C_{i,j}} \varepsilon_{i,j+1}, \tag{4.18}$$

where $\varepsilon_{i,j}$ *are independent a identically distributed with* $\mathsf{E}[\varepsilon_{i,j}] = 0$ *and* $\mathsf{Var}(\varepsilon_{i,j}) = 1$.

Theorem 4 (Parameter estimation error for Autoregressive CL)
Under Assumptions 6 the parameter estimation error for $C_{i,J}$ *is given by:*

$$(\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])^2 = C_{i,I-i}^2 \Big(\prod_{j=I-i}^{J-1} \Big(\hat{f}_j^2 + \frac{\hat{\sigma}_j^2}{\sum_{j=0}^{I-j-1} C_{i,j}} \Big) - \prod_{j=I-i}^{J-1} \hat{f}_j^2 \Big). \tag{4.19}$$

Proof can be found in Buchwalder et al. [3].

If slightly modified, we could see that Mack's estimation (4.15) is always lower than the autoregressive model estimation (4.19) and Buchwalder et al. [3] showed, that Mack estimate is a linear approximation of autoregressive estimate. Next theorem summarizes MSEP for autoregressive approach.

Theorem 5 (AR CL MSEP - single and aggregated accident years)
Under Assumptions 6 we have following estimators for conditional MSEP of outstanding claims for single and aggregated accident years:

$$\widehat{\mathsf{msep}}^{AR}_{C_{i,J}|D_I}(\hat{C}^{CL}_{i,J}) = (\hat{C}^{CL}_{i,J})^2 \Big(\sum_{j=I-i}^{J-1} \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{\hat{C}^{CL}_{i,j}} + \prod_{j=I-i}^{J-1} \Big(\frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + 1 \Big) - 1 \Big), \tag{4.20}$$

$$\begin{aligned} \widehat{\mathsf{msep}}^{AR}_{\sum_{i=1}^{I} C_{i,J}|D_I} \Big(\sum_{i=1}^{I} \hat{C}^{CL}_{i,J} \Big) &= \sum_{i=1}^{I} \widehat{\mathsf{msep}}^{AR}_{C_{i,J}|D_I}(\hat{C}^{CL}_{i,J}) + \\ &+ 2 \sum_{1 \leq i < k \leq I} C_{i,I-i} \hat{C}^{CL}_{k,I-i} \Big(\prod_{i=I-i}^{J-1} \Big(\hat{f}_j^2 + \frac{\hat{\sigma}_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} \Big) - \prod_{i=I-i}^{J-1} \hat{f}_j^2 \Big). \end{aligned} \tag{4.21}$$

4.3 Merz-Wüthrich method

Merz and Wüthrich [12] presented another approach to measure uncertainty within Mack Chain-ladder model (Assumptions 4). Their target is not the variability of ultimate claims estimate but variability of one-year change of the estimate - *claims development result* (*CDR*).

If we assume that $\mathsf{E}[C_{i,J} \mid D_I]$ is a predictor for $C_{i,J}$ as well as $\mathsf{E}[C_{i,J} \mid D_{I+1}]$ and if $\hat{C}^I_{i,J}$ and $\hat{C}^{I+1}_{i,J}$ are appropriate estimators, then we define claims development result after one year for accident year i as follows

$$CDR_i(I+1) = \mathsf{E}[C_{i,J} \mid D_I] - \mathsf{E}[C_{i,J} \mid D_{I+1}], \tag{4.22}$$

with estimate

$$\widehat{CDR}_i(I+1) = \hat{C}^I_{i,J} - \hat{C}^{I+1}_{i,J}. \tag{4.23}$$

It is apparent (using martingale property), that the expected value of CDR equals to zero. One of Solvency II targets is to estimate uncertainty of CDR to determine the amount of required risk capital. Natural way is considering mean square error of CDR. In contrast to estimating MSEP of ultimate claims, this basically means transition from long-term risk valuation to short-term.

Merz and Wüthrich [12] quantified two criteria of variability within the CDR. These are

$$\mathsf{msep}_{\widehat{CDR}_i(I+1)|D_I}(0) = \mathsf{E}\Big[\Big(\widehat{CDR}_i(I+1) - 0\Big)^2 \mid D_I\Big], \tag{4.24}$$

$$\mathsf{msep}_{CDR_i(I+1)|D_I}(\widehat{CDR}_i(I+1)) = \mathsf{E}\Big[\Big(\widehat{CDR}_i(I+1) - CDR_i(I+1)\Big)^2 \mid D_I\Big]. \tag{4.25}$$

In the first approach we quantify the variability of estimated CDR from its expected zero value. In the second approach we retrospectively analyze the difference between estimated and observable CDR. We call the expression (4.24) the prospective conditional MSEP of CDR and (4.25) the retrospective conditional MSEP of CDR for accident year i. For the estimation of these two expressions, we need to state some additional assumptions. We start with definitions, that will simplify the form of MSEP.

$$\hat{\Delta}_{i,J} = \frac{\hat{\sigma}^2_{I-i}/\hat{f}^2_{I-i}}{\sum_{k=0}^{i-1} C_{k,I-i}} + \sum_{j=I-i+1}^{J-1} \Big(\frac{C_{I-j,j}}{\sum_{k=0}^{I-j} C_{k,j}}\Big)^2 \frac{\hat{\sigma}^2_j/\hat{f}^2_j}{\sum_{k=0}^{I-j-1} C_{k,j}}, \tag{4.26}$$

$$\hat{\Phi}_{i,J} = \sum_{j=I-i+1}^{J-1} \left(\frac{C_{I-j,j}}{\sum_{k=0}^{I-j} C_{k,j}} \right)^2 \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{C_{I-j,j}}, \tag{4.27}$$

$$\hat{\Psi}_{i,J} = \frac{\hat{\sigma}_{I-i}^2 / \hat{f}_{I-i}^2}{C_{i,I-i}}, \tag{4.28}$$

$$\hat{\Lambda}_{i,J} = \frac{C_{i,I-i}}{\sum_{k=0}^{i} C_{k,I-i}} \frac{\hat{\sigma}_{I-i}^2 / \hat{f}_{I-i}^2}{\sum_{k=0}^{i-1} C_{k,I-i}} + \sum_{j=I-i+1}^{J-1} \left(\frac{C_{I-j,j}}{\sum_{k=0}^{I-j} C_{k,j}} \right)^2 \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}}. \tag{4.29}$$

Assumptions 7 *For $1 >> a_j > 0$ we can use following approximation*

$$\prod_{j=0}^{J} (1 + a_j) - 1 \approx \sum_{j=0}^{J} a_j. \tag{4.30}$$

This approximation is applied to

$$a_j^{(1)} = \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{C_{I-j,j}} \frac{C_{I-j,j}}{\sum_{k=0}^{I-j} C_{k,j}}, a_j^{(2)} = \frac{\sigma_j^2 / f_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}},$$

$$a_j^{(3)} = \frac{\sum_{i=0}^{I-j-1} C_{i,j} (\sigma_j^2 / f_j^2)}{(\sum_{k=0}^{I-j} C_{k,j})^2}, a_j^{(4)} = \frac{\sigma_j^2 / f_j^2}{\sum_{k=0}^{I-j} C_{k,j}}.$$

Now, we can state Merz-Wüthrich estimators for prospective and retrospective conditional MSEP of claims development result.

Theorem 6 (MSEP of CDR for single accident year) *Under Assumptions 4 and approximation Assumptions 7 we have following estimators for prospective and retrospective conditional MSEP of claims development result*

$$\mathsf{msep}_{\widehat{CDR}_i(I+1)|D_I}(0) = (\hat{C}_{i,J}^{CL})^2 (\hat{\Delta}_{i,J} + \hat{\Phi}_{i,J} + \hat{\Psi}_{i,J}), \tag{4.31}$$

$$\mathsf{msep}_{CDR_i(I+1)|D_I}(\widehat{CDR}_i(I+1)) = (\hat{C}_{i,J}^{CL})^2 (\hat{\Delta}_{i,J} + \hat{\Phi}_{i,J}). \tag{4.32}$$

Proof can be found in Merz and Wüthrich [12].

Theorem 7 (MSEP of CDR for aggregated accident years)
Under Assumptions 4 and approximation Assumptions 7 we have following estimators for prospective and retrospective conditional MSEP of aggregated claims development results

$$\begin{aligned}\mathsf{msep}_{\sum_{i=1}^{I}\widehat{CDR}_i(I+1)|D_I}(0) = &\sum_{i=1}^{I}\mathsf{msep}_{\widehat{CDR}_i(I+1)|D_I}(0)+\\ &+2\sum_{0<i<k\leq J}\hat{C}_{i,J}^{CL}\hat{C}_{k,J}^{CL}\Big(\hat{\Phi}_{i,J}+\hat{\Lambda}_{i,J}+\frac{\hat{\sigma}_{I-i}^2/\hat{f}_{I-i}^2}{\sum_{k=0}^{i}C_{k,I-i}}\Big),\end{aligned} \tag{4.33}$$

$$\begin{aligned}\mathsf{msep}_{\sum_{i=1}^{I}CDR_i(I+1)|D_I}(\sum_{i=1}^{I}\widehat{CDR}_i(I+1)) = &\sum_{i=1}^{I}\mathsf{msep}_{CDR_i(I+1)|D_I}(\widehat{CDR}_i(I+1))+\\ &+2\sum_{0<i<k\leq J}\hat{C}_{i,J}^{CL}\hat{C}_{k,J}^{CL}(\hat{\Phi}_{i,J}+\hat{\Lambda}_{i,J}).\end{aligned} \tag{4.34}$$

Proof can be found in Merz and Wüthrich [12].

There is a question of whether one should concentrate on the variability of ultimate claims estimate or CDR estimate. The MSEP of ultimate claim currently represents the main target in majority of stochastic claims reserving methods, but as in other of risk management branches, one should pursue both. Short term risk should be considered in one-year based processes like product pricing/premium volume determination or business plans processing.

There is a difficulty with measuring CDR variability that since it is quite new approach, there is only one well-known analytic formula for MSEP of CDR and under different model than Mack Chain-ladder, we cannot manage with analytic approach. But we can expect that with solvency requirements and increasing usage of simulation methods in claims reserving the proportional representation of short-term risk measurement in claims reserving will gradually grow.

4.4 Chain-ladder Generalized linear models

An intuitive way to measure the variability of ultimate claim is to expand non-stochastic reserving method with a distributional underlying framework, i.e. to find a distributional model which returns same estimates for the expected value as our original analytic method.

Majority of commonly used distributions for claims modelling (with the exception of log-normal) belongs to so called *exponential dispersion family* (EDF). This provides us the use of *generalized linear models* (GLM). Generalized linear model is a generalization of ordinary linear regression and it unifies various statistical models. Unlike the linear regression, it provides more complex (nonlinear) relationship between the response and explanatory variables and it expands the distribution boundaries of these variables.

In this section we introduce general framework of GLM for incremental claims with its basic properties and then follow England & Verrall [5] and focus on the Over-dispersed Poisson and Over-dispersed Negative Binomial model and their connection to Chain-ladder method.

Assumptions 8 (Generalized linear models) *GLM in based on the framework consisting of following assumptions:*

1. *Distribution of incremental claims $X_{i,j}$ belongs to exponential dispersion family (EDF) with following density, mean and variance:*

$$f(x, \theta, \phi, w_{i,j}) = a(x, \frac{\phi}{w_{i,j}}) exp\Big\{\frac{x\theta - b(\theta)}{\phi / w_{i,j}}\Big\}, \tag{4.35}$$

$$\mathsf{E}[X_{i,j}] = x_{i,j} = b'(\theta), \tag{4.36}$$

$$\mathsf{Var}(X_{i,j}) = \frac{\phi}{w_{i,j}} b''(\theta), \tag{4.37}$$

where $\phi > 0$ is a dispersion parameter, $w_{i,j} > 0$ are known weights, $b(\cdot)$ is twice continuously differentiable function with invertible second derivation and $a(\cdot, \cdot)$ is a real-valued function ensuring the integral over all x is equal to one.

2. *For each $x_{i,j}$, $0 \leq i \leq I$,$0 \leq j \leq J$, we have a linear predictor $\eta_{i,j}$*

$$\eta_{i,j} = \mathbf{\Gamma}_{i,j}\mathbf{b}, \tag{4.38}$$

linked by monotonic a differentiable response function $h(\cdot)$ and link function $g(\cdot)$:

$$x_{i,j} = h(\eta_{i,j}), \tag{4.39}$$

$$g(x_{i,j}) = \eta_{i,j} = \mathbf{\Gamma}_{i,j}\boldsymbol{b}. \tag{4.40}$$

We can see that the generalization of linear regression therefore brings new challenges in a sense of required choice of:

1. distribution of $X_{i,j}$ defined by function $b(\cdot)$ and parameters ϕ and θ,
2. systematic component structure $\mathbf{\Gamma}_{i,j}\mathbf{b}$ and number of regressors,
3. response $h(\cdot)$ or link function $g(\cdot)$.

Table 4.1: Exponential dispersion family overview

EDF distribution	$\mathsf{Var}(X_{i,j})$
Poisson	$x_{i,j}$
Over-dispersed Poisson	$\phi x_{i,j}$
Gamma	$\phi x_{i,j}^2$
Tweedie's compound Poisson	$\phi x_{i,j}^p$, $p \in (0,1)$
Over-dispersed negative binomial	$\phi \lambda_{i,j} x_{i,j}$
Normal (and all Gaussian)	ϕ

We can find a summary overview of common EDF distributions used for claims reserving in Table 4.1. Further discussion of appropriate parametrization and model adjustment can be found in McCullagh and Nelder [11] and Merz and Wüthrich [13]. In the following section we concentrate on two GLM models, that are related to the Chain-ladder model.

Assumptions 9 (Over-dispersed Poisson model)
We introduce Over-dispersed Poisson model using following assumptions:

1. *$X_{i,j}$ are independent and have Over-dispersed Poisson distribution with following expected value and variance:*

$$\mathsf{E}[X_{i,j}] = x_{i,j}, \quad \mathsf{Var}(X_{i,j}) = \phi x_{i,j}. \tag{4.41}$$

2. *Expected value $x_{i,j}$ is given by linear component structure linked through a logarithmic link*

$$log(x_{i,j}) = c + m_i + g_j \tag{4.42}$$

with restriction $m_0 = 0$, $g_0 = 0$.

The over-dispersion is introduced due to deficiency of Poisson distribution when applied to data with higher variability, since it has the same variance and expected value. Over-dispersion is in GLM adopted through quasi-likelihood approach, where the first two moments are used to fit data to GLM and quasi-distribution parameters remain proportionally unchanged. The dispersion parameter ϕ is estimated via weighted Pearson residuals (see Section 5.2). Quasi-likelihood approach also facilitates working with non-integer data, since the discrete random variables can be treated as continuous, while distribution parameters remain unchanged. For detailed estimation of parameters $c, m_i, g_j, 0 \leq i \leq I$, $0 \leq j \leq J$ via quasi-maximum likelihood see McCullagh and Nelder [11] and England and Verrall [5].

Although the requirement for observations to be integer is solved in quasi-likelihood approach, the requirement for observation to be positive on the other hand is quite restrictive and it cannot be easily circumvented (for suggestions, see Verrall [15]).

The Over-dispersed Poisson distribution is not a distribution in a strict analytic sense. It cannot be assigned with analytical form of density of cumulative distribution function. Another way to increase the variance of Poisson distribution without changing the expected value is using the Negative Binomial distribution as Gamma-Poisson mixture. Assume that we have a random variable with Poisson distribution $X \sim \mathsf{Po}(\lambda)$ and parameter λ is appropriately parametrized Gamma distributed random variable, then the X has Negative Binomial distribution. The connection between the Poisson and Negative Binomial distribution encourages the usage of Over-dispersed Negative Binomial distribution in claims reserving. The Over-dispersed Negative Binomial model is defined more directly, since there is a requirement for Chain-ladder behavior on the expected value of incremental claims.

Assumptions 10 (Over-dispersed Negative Binomial model)
We introduce Over-dispersed Negative Binomial model using following assumptions:

1. *Cumulative claims $C_{i,j}$ of different accident years i are independent.*
2. *Incremental claims are conditionally independent of the cumulative claims at the previous time period.*
3. *$X_{i,j}$ are over-dispersed negative binomial distributed with following expected*

value and variance:

$$\mathsf{E}[X_{i,j} \mid C_{i,j-1}] = (f_{j-1}-1)C_{i,j-1}, \mathsf{Var}(X_{i,j} \mid C_{i,j-1}) = \phi f_{j-1}(f_{j-1}-1)C_{i,j-1}, \tag{4.43}$$

If we want to estimate the development factors f_j via GLM maximum likelihood, we need to specify systematic component structure and link function for $x_{i,j}$. We follow England and Verrall [4], where logarithmic link is used:

$$log(x_{i,j}) = log(f_{j-1} - 1) + log(C_{i,j-1}) = c + g_{j-1} + log(C_{i,j-1}), \tag{4.44}$$

where

$$c + g_j = log(f_j - 1), \text{with} \quad g_0 = 0. \tag{4.45}$$

Now we can state the theorem about Over-dispersed Poisson and Negative Binomial Chain-ladder properties.

Theorem 8 (Over-dispersed Poisson & Negative Binomial properties) *Under Assumptions 9 or 10 the Over-dispersed Poisson and Over-dispersed Negative Binomial model lead to the same estimate of ultimate claim $C_{i,J}$ as Chain-ladder method (Assumptions 1).*

Proof can be found in Renshaw and Verrall [14] and Verrall [15].

Note that in Chapter 3, the Chain-ladder was formulated for cumulative claims, while the two GLM models are defined for incremental claims, but since the Chain-ladder method can be applied to cumulative and incremental claims with the same result, this difference has no impact.

Following theorem defines the estimator for unconditional MSEP of Over-dispersed Poisson model from England and Verrall [4]

Theorem 9 (ODP CL MSEP - single & aggregated accident years) *Under Assumptions 9 the unconditional MSEP of estimated outstanding claims for accident year i is approximated by:*

$$\mathsf{msep}^{ODP}_{C_{i,J}}(\hat{C}^{CL}_{i,J}) \approx \sum_{j=I-i+1}^{J} \mathsf{Var}(X_{i,j}) + \sum_{j=I-i+1}^{J} \mathsf{Var}(\hat{X}_{i,j}) + 2 \sum_{I-i<k<l\leq J} \mathsf{Cov}(\hat{X}_{i,k}, \hat{X}_{i,l}) \tag{4.46}$$

and estimated by

$$\widehat{\mathsf{msep}}^{ODP}_{C_{i,J}}(\hat{C}^{CL}_{i,J}) = \sum_{j=I-i+1}^{J} \Big(\hat{\phi}\hat{x}_{i,j} + \hat{x}^2_{i,j}\mathsf{Var}(\hat{\eta}_{i,j})\Big) + 2\sum_{I-i<j<l\leq J} \hat{x}_{i,j}\hat{x}_{i,l}\mathsf{Cov}(\hat{\eta}_{i,j},\hat{\eta}_{i,l}). \quad (4.47)$$

The estimate of unconditional MSEP for aggregated accident years is given by

$$\widehat{\mathsf{msep}}^{ODP}_{\sum_{i=1}^{I} C_{i,J}}(\sum_{i=1}^{I}\hat{C}^{CL}_{i,J}) = \sum_{i=1}^{I}\sum_{j=I-i+1}^{J}\hat{\phi}\hat{x}_{i,j} + \sum_{i=1}^{I}\sum_{j=I-i+1}^{J}\hat{x}^2_{i,j}\mathsf{Var}(\hat{\eta}_{i,j}) + \\ +2\sum_{i=1,k=1}^{I}\sum_{\substack{I-i<j\leq J\\ I-k<l\leq J\\ j\leq l\\ (i,j)\neq(k,l)}}\hat{x}_{i,j}\hat{x}_{k,l}\mathsf{Cov}(\hat{\eta}_{i,j},\hat{\eta}_{k,l}), \quad (4.48)$$

where $\eta_{i,j}$ is GLM linear predictor, for which $\eta_{i,j} = c + m_i + g_j = \log(x_{i,j})$.

Proof can be found in England and Verrall [4].

The variance of the linear predictor $\eta_{i,j}$ is a standard output of GLM software packages and can be obtained via Fisher information matrix, which, if inversed, represents the asymptotic variance of MLE estimates.

Detailed expression of MSEP using Fisher information matrix for Over-dispersed Poisson model can be found in Chapter 6 in Merz and Wüthrich [13]. We deal with further derivation of GLM MSEP in Section 4.6, where the estimation of linear predictor variance is not sufficient and additional expressions need to be estimated.

Now, we state the procedure for estimation of Negative Binomial model MSEP using the recursive property (4.43).

Theorem 10 (ODNB CL MSEP - single & aggregated accident years) *Under Assumptions 10 the unconditional MSEP of estimated outstanding claims for accident year i is approximated by:*

$$\mathsf{msep}^{ODNB}_{C_{i,J}}(\hat{C}^{CL}_{i,J}) \approx \mathsf{Var}(C_{i,J}) + \mathsf{Var}(\hat{C}_{i,J}), \quad (4.49)$$

and estimated by

$$\widehat{\mathsf{msep}}^{ODNB}_{C_{i,J}}(\hat{C}^{CL}_{i,J}) = \hat{\phi}C_{i,I-i}\prod_{j=I-i}^{J-1}\hat{f}_j\Big(\prod_{j=I-i}^{J-1}\hat{f}_j - 1\Big) + C^2_{i,I-i}\mathsf{Var}\Big(\prod_{j=I-i}^{J-1}\hat{f}_j\Big). \quad (4.50)$$

The estimate of unconditional MSEP for aggregated accident years is given by

$$\widehat{\mathsf{msep}}^{ODNB}_{\sum_{i=0}^{I} C_{i,J}}\Big(\sum_{i=0}^{I} \hat{C}^{CL}_{i,J}\Big) = \sum_{i=0}^{I} \widehat{\mathsf{msep}}^{ODNB}_{C_{i,J}}(\hat{C}^{CL}_{i,J}) + 2 \sum_{1 \le i < k \le I} \mathsf{Cov}(\hat{C}_{i,J}, \hat{C}_{k,J}), \tag{4.51}$$

where $\mathsf{Var}\Big(\prod_{j=I-i+1}^{J} \hat{f}_j\Big)$ *and covariances* $\mathsf{Cov}(\hat{C}_{i,J}, \hat{C}_{k,J})$ *are estimated via recursive algorithm following Assumptions 11.*

Proof can be found in England and Verrall [4].

Assumptions 11 (ODNB recursive algorithm)
Under Assumptions 10 the expression $\mathsf{Var}(\hat{f}_{J-2}\hat{f}_{J-1})$ *is given by*

$$\mathsf{Var}(\hat{f}_{J-2}\hat{f}_{J-1}) = \Big(\mathsf{E}[\hat{f}_{J-2}]\Big)^2 \mathsf{Var}(\hat{f}_{J-1}) + \Big(\mathsf{E}[\hat{f}_{J-1}]\Big)^2 \mathsf{Var}(\hat{f}_{J-2}) + \mathsf{Var}(\hat{f}_{J-1})\mathsf{Var}(\hat{f}_{J-2}) \tag{4.52}$$

and estimated by

$$\widehat{\mathsf{Var}}(\hat{f}_{J-2}\hat{f}_{J-1}) = \hat{f}^2_{J-2}\mathsf{Var}(\hat{f}_{J-1}) + \hat{f}^2_{J-1}\mathsf{Var}(\hat{f}_{J-2}) + \mathsf{Var}(\hat{f}_{J-1})\mathsf{Var}(\hat{f}_{J-2}), \tag{4.53}$$

$$\widehat{\mathsf{Var}}(\hat{f}_j) = \exp(\hat{c} + \hat{g}_j)^2 \mathsf{Var}(\hat{c} + \hat{g}_j). \tag{4.54}$$

For estimation of $\mathsf{Var}(\hat{f}_{J-k} \cdots \hat{f}_{J-1})$, $k > 2$, *recursive formula is used:*

$$\begin{aligned} \mathsf{Var}(\hat{f}_{J-k} \cdots \hat{f}_{J-1}) &= \Big(\mathsf{E}[\hat{f}_{J-k}]\Big)^2 \mathsf{Var}(\hat{f}_{J-k+1} \cdots \hat{f}_{J-1}) + \\ &+ \Big(\mathsf{E}[\hat{f}_{J-k+1} \cdots \hat{f}_{J-1}]\Big)^2 \mathsf{Var}(\hat{f}_{J-k}) + \mathsf{Var}(\hat{f}_{J-k+1} \cdots \hat{f}_{J-1})\mathsf{Var}(\hat{f}_{J-k}). \end{aligned} \tag{4.55}$$

Assuming that the development factors are independent, the covariance $\mathsf{Cov}(\hat{C}_{i,J}, \hat{C}_{k,J})$, $i < k$ *is given by*

$$\mathsf{Cov}(\hat{C}_{i,J}, \hat{C}_{k,J}) = \mathsf{Cov}(C_{i,I-i}\hat{f}_{I-i} \cdots \hat{f}_{J-1}, C_{k,I-k}\hat{f}_{I-k} \cdots \hat{f}_{J-1}) = \tag{4.56}$$

$$= C_{i,I-i}C_{k,I-k} \prod_{j=I-k}^{I-i-1} \hat{f}_j \mathsf{Var}\Big(\prod_{j=I-i}^{J-1} \hat{f}_j\Big). \tag{4.57}$$

Note 4 (Mack Chain-ladder as GLM) *Even though Mack Chain-ladder was introduced as distribution-free model, England and Verrall [4] presented it in a GLM framework with assumption that cumulative claims $C_{i,j}$ are normally distributed with expected value and variance:*

$$\mathsf{E}[C_{i,j} \mid C_{i,j-1}] = f_{j-1}C_{i,j-1}, \quad \mathsf{Var}(C_{i,j} \mid C_{i,j-1}) = \sigma_j^2 C_{i,j-1} \tag{4.58}$$

and showed, that these assumptions returns the same estimates of $\hat{f}_j$ and $\hat{\sigma}_j^2$, $j = 1 \dots J-1$ as original estimates (3.3), (4.10).

And since the GLM with normally distributed cumulative claims can be used as approximation to Over-dispersed Negative Binomial model, we can expect both models to return similar MSEP estimates.

4.5 Distribution-free Bornhuetter-Ferguson

In this section we follow Mack [8], who introduced stochastic framework for Bornheutter-Ferguson method, that corresponds with Mack Chain-ladder Assumptions 4. Mack also suggested estimation algorithm for Bornheutter-Ferguson parameters, that are independent of Chain-ladder development factors to present BF method as a standalone technique.

This model is designed for incremental claims, therefore we are transitioning from cumulative pattern $\beta_0, \dots, \beta_J$ to incremental pattern $\gamma_0, \dots, \gamma_J$, where $\gamma_0 = \beta_0$ and $\gamma_j = \beta_j - \beta_{j-1}$, for $1 \le j \le J$.

For simplicity, the variance parameters σ_j^2 are denoted as in Mack Chain-ladder, even though they represent different variables.

Assumptions 12 (Mack Bornhuetter-Ferguson) *We introduce Mack stochastic Bornhuetter-Ferguson method using following assumptions:*

1. *All incremental claims $X_{i,j}$ are independent.*
2. *There exist parameters $\mu_0, \dots, \mu_I > 0$ and a pattern $\gamma_0, \dots, \gamma_J > 0$ with $\sum_{j=0}^{J} \gamma_j = 1$ such that for all $0 \le i \le I$ and all $0 \le j \le J-1$ we have*

$$\mathsf{E}[X_{i,j}] = \mu_i \gamma_j. \tag{4.59}$$

3. *There are unknown variance parameters $\sigma_0^2, \dots, \sigma_{J-1}^2 > 0$ such that for all $0 \le i \le I$ and all $0 \le j \le J$ we have*

$$\mathsf{Var}(X_{i,j}) = \mu_i \sigma_j^2. \tag{4.60}$$

Now, we focus at the parameter estimates following Mack [9] and [10]. As in Section 3.2, we assume that we have a prior (expert) estimate of ultimate claim $\hat{\mu}_i$. BF parameters are then estimated by following procedure:

1. Obtain first-phase estimates of γ_j, $0 \le j \le J$, from

$$\widetilde{\gamma}_j = \sum_{i=0}^{I-i} X_{i,j} / \sum_{i=0}^{I-i} \hat{\mu}_i. \tag{4.61}$$

2. Apply selected smoothing regression on $\widetilde{\gamma}_j \to \widetilde{\gamma}_j^\star$ and normalize them.

3. Estimate variance parameters σ_j^2, $0 \le j \le J-1$, from

$$\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=0}^{I-j-1} \frac{(X_{i,j} - \hat{\mu}_i \widetilde{\gamma}_j^\star)^2}{\hat{\mu}_i}. \tag{4.62}$$

4. If needed, apply smoothing regression on $\hat{\sigma}_j^2 \to \hat{\sigma}_j^{2\star}$.

5. Obtain second-phase estimates of γ_j, $0 \le j \le J$ by minimizing

$$Q = \sum_{i+j \le I} \frac{(X_{i,j} - \hat{\mu}_i \hat{\gamma}_j)^2}{\hat{\mu}_i \hat{\sigma}_j^2}. \tag{4.63}$$

After normalization, we obtained estimates $\hat{\gamma}_j, \hat{\sigma}_j^2, 0 \le j \le J$ (tail factor $\hat{\sigma}_J^2$ can be obtained by extrapolation consistent with applied smoothing model). Next to the prior estimation of ultimate claim $\hat{\mu}_i$, the actuary should be also able to assign its uncertainty. Mack [8] suggested an estimate of $\widehat{\mathsf{Var}}(\hat{\mu}_i)$ using premium volume π_i connected through prior estimate of loss ratio $\hat{\kappa}_i$

$$\widehat{\mathsf{Var}}(\hat{\mu}_i) = \frac{\pi_i}{I-1} \sum_{i=0}^{I} \pi_i \Big(\frac{\hat{\mu}_i}{\pi_i} - \frac{\bar{\hat{\mu}}}{\bar{\pi}} \Big)^2, \tag{4.64}$$

where $\bar{\hat{\mu}}, \bar{\pi}$ are appropriate average values over all $0 \le i \le I$.

Using (4.3), we can replace MSEP of ultimate claim estimate with MSEP of reserve estimate $\hat{R}_i^{(BF)}$. This adjustment simplifies MSEP expression, since under BF assumptions the future incremental claims are independent of past development. The estimated reserve is given by

$$\hat{R}_i^{(BF)} = \hat{\mu}_i (1 - \hat{\beta}_{I-i}). \tag{4.65}$$

Assuming the independence of incremental claims, we can express its conditional MSEP as follows

$$\mathsf{msep}_{R_i|D_I}(\hat{R}_i^{(BF)}) = \mathsf{E}[(R_i - \hat{R}_i^{(BF)})^2 \mid D_I] = \mathsf{Var}(R_i) + \mathsf{Var}(\hat{R}_i^{(BF)}). \quad (4.66)$$

Again, we will estimate conditional process variance and parameter estimation error separately.

Theorem 11 (Conditional process variance for distribution-free BF) *Under Assumptions 12 the conditional process variance for Bornheutter-Ferguson ultimate claim estimate is given by*

$$\mathsf{Var}(R_i \mid D_I) = \mu_i(\sigma^2_{I-i+1} \dots \sigma^2_J) \quad (4.67)$$

and estimated by

$$\widehat{\mathsf{Var}}(R_i \mid D_I) = \hat{\mu}_i(\hat{\sigma}^2_{I-i+1} \dots \hat{\sigma}^2_J). \quad (4.68)$$

Proof - direct consequence of Assumptions 12.

Theorem 12 (Parameter estimation error for distribution-free BF) *Under Assumptions 12 the conditional process variance for Bornheutter-Ferguson ultimate claim estimate is given by*

$$\mathsf{Var}(\hat{R}_i \mid D_I) = \Big(\mu_i^2 + \mathsf{Var}(\hat{\mu}_i)\Big)\mathsf{Var}(\hat{\beta}_{I-i}) + \mathsf{Var}(\hat{\mu}_i)(1-\beta_{I-i})^2 \quad (4.69)$$

and estimated by

$$\widehat{\mathsf{Var}}(\hat{R}_i \mid D_I) = \Big(\hat{\mu}_i^2 + \widehat{\mathsf{Var}}(\hat{\mu}_i)\Big)\widehat{\mathsf{Var}}(\hat{\beta}_{I-i}) + \widehat{\mathsf{Var}}(\hat{\mu}_i)(1-\hat{\beta}_{I-i})^2, \quad (4.70)$$

where

$$\widehat{\mathsf{Var}}(\hat{\beta}_j) = \min(\widehat{\mathsf{Var}}(\hat{\gamma}_0) + \dots + \widehat{\mathsf{Var}}(\hat{\gamma}_{j-1}), \widehat{\mathsf{Var}}(\hat{\gamma}_j) + \dots + \widehat{\mathsf{Var}}(\hat{\gamma}_J)), \quad (4.71)$$

$$\widehat{\mathsf{Var}}(\hat{\gamma}_j) = \frac{\hat{\sigma}_j^2}{\sum_{i=0}^{I-j} \hat{\mu}_i}. \quad (4.72)$$

Proof can be found in Mack [8].

We can see that the estimation is straightforward. The only issue is the estimate $\widehat{\mathsf{Var}}(\hat{\beta}_j)$, which cannot be simply chosen as $\widehat{\mathsf{Var}}(\hat{\beta}_j) = \widehat{\mathsf{Var}}(\hat{\gamma}_1) + \dots + \widehat{\mathsf{Var}}(\hat{\gamma}_j)$, since the requirement $\sum_{j=0}^{J} \hat{\gamma}_j = 1$ causes dependency between estimates $\hat{\gamma}_j$. Therefore estimates (4.71) were suggested, so that for $\widehat{\mathsf{Var}}(\hat{\beta}_k)$ with small k, we use variance of row parameters $\hat{\gamma}_j$, $j = 0 \dots k$, and for large k, we use variance of their supplement $\hat{\gamma}_j$, $j = k+1 \dots J$, which should reduce the dependency.

Theorem 13 (DF BF MSEP - single and aggregated accident years) *Under Assumptions 10, we have following estimator for the conditional MSEP of estimated outstanding loss reserve for accident year i*

$$\widehat{\mathsf{msep}}_{R_i|D_I}(\hat{R}_i) = \hat{\mu}_i(\hat{\sigma}^2_{I-i+1}\dots\hat{\sigma}^2_{J-1}) + \Big(\hat{\mu}_i^2 + \widehat{\mathsf{Var}}(\hat{\mu}_i)\Big)\widehat{\mathsf{Var}}(\hat{\beta}_{I-i}) + \widehat{\mathsf{Var}}(\hat{\mu}_i)(1-\hat{\beta}_{I-i})^2. \tag{4.73}$$

The estimate of conditional MSEP for aggregated accident years is given by

$$\widehat{\mathsf{msep}}_{\sum_{i=1}^I R_i|D_I}(\sum_{i=1}^I \hat{R}_i) + \sum_{i=1}^I \widehat{\mathsf{msep}}_{R_i|D_I}(\hat{R}_i) + \sum_{i<k} \widehat{\mathsf{Cov}}(\hat{R}_i, \hat{R}_k), \tag{4.74}$$

where $\widehat{\mathsf{Cov}}(\hat{R}_i, \hat{R}_k)$ *is obtained as*

$$\begin{aligned}\widehat{\mathsf{Cov}}(\hat{R}_i, \hat{R}_k) = \hat{\rho}^{\mu}_{i,k}\sqrt{\widehat{\mathsf{Var}}(\hat{\mu}_i)}\sqrt{\widehat{\mathsf{Var}}(\hat{\mu}_k)}(1-\hat{\beta}_{I-i})(1-\hat{\beta}_{I-k})+ \\ +\hat{\rho}^{\beta}_{i,k}\sqrt{\widehat{\mathsf{Var}}(\hat{\beta}_{I-i})}\sqrt{\widehat{\mathsf{Var}}(\hat{\beta}_{I-k})}\hat{\mu}_i\hat{\mu}_k\end{aligned} \tag{4.75}$$

and $\hat{\rho}^{\mu}_{i,j}$, $\hat{\rho}^{\beta}_{i,j}$, $i < k$ *are estimated correlation coefficients (assuming constant correlation coefficient of ultimate claims estimates* $\hat{\mu}_i$*)*

$$\hat{\rho}^{\mu}_{i,k} = \sqrt{(J+1)}/(J+1), \tag{4.76}$$

$$\hat{\rho}^{\beta}_{i,k} = \sqrt{\frac{\hat{\beta}_{I-k}(1-\hat{\beta}_{I-i})}{\hat{\beta}_{I-i}(1-\hat{\beta}_{I-k})}}. \tag{4.77}$$

Proof can be found in Mack [8].

4.6 Bornheutter-Ferguson GLM

In Section 4.4 we have met Over-dispersed Poisson model, remarking, that it has the same first moment estimates as Chain-ladder method. In Section 3.2 we have expressed Bornheutter-Ferguson ultimate claim estimate by Chain-ladder development factors. Alai et al. [1] actually omitted the Chain-ladder link and connected Over-dispersed Poisson model directly to Bornheutter-Ferguson method.

We slightly modify the Over-dispersed Poisson model from Section 4.4 by following transformations/substitutions:

$$c = 0, \quad \mu_i = \exp(m_i), \quad \gamma_j = \exp(g_j). \tag{4.78}$$

After this adjustment, we obtain following model:

Assumptions 13 (BF Over-dispersed Poisson model)
We introduce Over-dispersed Poisson model using following assumptions:

1. *$X_{i,j}$ are independent and have Over-dispersed Poisson distribution with following expected value and variance:*
$$\mathsf{E}[X_{i,j}] = x_{i,j}, \quad \mathsf{Var}(X_{i,j}) = \phi x_{i,j}. \tag{4.79}$$

2. *Expected value $x_{i,j}$ is given by multiplicative structure*
$$x_{i,j} = \mu_i \gamma_j, \tag{4.80}$$
which is in GLM point of view understood as logarithmic linked component structure
$$\log(x_{i,j}) = \log(\mu_i) + \log(\gamma_j) \quad (= m_i + g_j), \tag{4.81}$$
with restriction $\mu_0 = 1$.

Let us review the Bornheutter-Ferguson assumption from Section 3.2
$$\mathsf{E}[C_{i,j+k} \mid C_{i,0}, \ldots, C_{i,j}] = C_{i,j} + (\beta_{j+k} - \beta_j)\mu_i.$$
We can see that under identification $\beta_j = \sum_{k=0}^{j} \gamma_k$ is the assumption satisfied and usage of Over-dispersed Poisson model as stochastic basis is therefore in place.

It is clear that the use of estimates μ_i under Assumptions 13 would be inconsistent with the fundamental principle of the Bornheutter-Ferguson method - the prior estimate $\hat{\mu}_i$ being independent of the claims development. This assumption is preserved by modifying model parameters. It means, that μ_i, γ_j are estimated under Assumptions 13 and then adjusted, so that the row factors are equal to prior estimates $\hat{\mu}_i$.

Estimation of parameters μ_i, $0 \le i \le I$ and γ_j, $0 \le j \le J$ via maximum likelihood is divided in two points of view: in GLM approach, we use above mentioned restriction $\mu_0 = 1$, in general MLE approach, there is more common assumption that $\sum_{j=0}^{J} \gamma_j = 1$, which is consistent with classic BF estimates. Using one approach or another makes no difference, since the parameters estimates are proportional, but we need both parameter estimates and therefore we use $\hat{\mu}_i^{(GLM)}, \hat{\gamma}_j^{(GLM)}$ for GLM approach and $\hat{\mu}_i^{(MLE)}, \hat{\gamma}_j^{(MLE)}$ for MLE approach.

As expected, using the ODP-CL connection (Theorem 8), we have the same estimate for ultimate claim as we defined in Section 3.3:
$$\hat{C}_{i,J}^{BF} = C_{i,I-i} + \Big(1 - \sum_{j=0}^{I-i} \hat{\gamma}_j^{(MLE)}\Big)\hat{\mu}_i = C_{i,I-i} + (1 - \hat{\beta}_{I-i})\hat{\mu}_i. \tag{4.82}$$

Now, we will take a closer look at conditional MSEP of this estimate.

Theorem 14 (Conditional process variance for BF ODP model) *Under Assumptions 13 the conditional process variance for Bornheutter-Ferguson ultimate claim estimate is given by*

$$\mathsf{Var}(C_{i,J} \mid D_I) = \sum_{j=I-i+1}^{J} \mathsf{Var}(X_{i,j}) \tag{4.83}$$

and estimated by

$$\widehat{\mathsf{Var}}(C_{i,J} \mid D_I) = \hat{\phi}\hat{\mu}_i(1 - \hat{\beta}_{I-i}^{(CL)}). \tag{4.84}$$

Proof - direct consequence of Assumptions 13 and Theorem 8.

Parameter estimation error of Bornheutter-Ferguson ODP model deals with uncertainty of estimates $\hat{\gamma}_j^{(MLE)}$, $j = 0 \dots J$ and $\hat{\mu}_i$, $i = 0 \dots I$.

For ultimate claim estimate $\hat{\mu}_i$, alternatively to Mack's formula (4.64), we can use expert estimate via coefficient of variation

$$\widehat{\mathsf{Var}}(\hat{\mu}_i) = \hat{\mu}_i^2 \widehat{\mathsf{Vco}}(\hat{\mu}_i)^2, \tag{4.85}$$

where Alai et al. [1] describe, that reasonable rate for $\widehat{\mathsf{Vco}}(\hat{\mu}_i)$ is considered between 5% and 10%.

Concerning the variability of row pattern, we need to estimate

$$\Big((1 - \beta_{I-i}) - (1 - \hat{\beta}_{I-i})\Big)^2 = \Big(\sum_{j=I-i+1}^{J} \gamma_j - \sum_{j=I-i+1}^{J} \hat{\gamma}_j^{(MLE)}\Big)^2. \tag{4.86}$$

This expression is estimated by $\mathsf{Var}(\sum_{j=I-i+1}^{J} \hat{\gamma}_j^{(MLE)})$ but unlike the Mack's estimate (4.71), Alai et al. [1] used GLM framework to put bigger emphasis to dependency between parameters $\hat{\gamma}_j^{(MLE)}$. Estimation of correlations between $\hat{\gamma}_j^{(MLE)}$ is done by converting from $\hat{\gamma}_j^{(GLM)}$ using Taylor approximation.
Covariance of GLM parameters $\hat{\mathbf{b}} = (\hat{m}_1^{(GLM)}, \dots, \hat{m}_I^{(GLM)}, \hat{g}_0^{(GLM)}, \dots, \hat{g}_J^{(GLM)})$ is given by inverse Fisher information matrix

$$\mathsf{Cov}(\hat{\mathbf{b}}, \hat{\mathbf{b}}) = \mathbf{H}^{-1}(\hat{\mathbf{b}}). \tag{4.87}$$

In exponential dispersion family the Fisher information matrix $\mathbf{H}^{-1}$ represents the covariance matrix of asymptotically multivariate normal MLE estimates. Fisher information matrix is a standard output of GLM software package. Detailed derivation of Over-dispersed Poisson BF MSEP can be found in Alai et al. [1], Fisher matrix properties can be found in Chapter 6 in Merz and Wüthrich [13].

Theorem 15 (Parameter estimation error for BF ODP model) *Under Assumptions 13 the parameter estimation error for for BF ultimate claim estimate is given by*

$$(\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I])^2 = (1 - \hat{\beta}^{(CL)}_{I-i})\mathsf{Var}(\hat{\mu}_i) + \mu_i^2 \Big(\sum_{j=I-i+1}^{J} \gamma_j - \sum_{j=I-i+1}^{J} \hat{\gamma}_j^{(MLE)} \Big)^2 \tag{4.88}$$

and estimated by

$$(1 - \hat{\beta}^{(CL)}_{I-i})^2 \hat{\mu}_i^2 \widehat{\mathsf{Vco}}(\hat{\mu}_i)^2 + \hat{\mu}_i^2 \sum_{j,k=I-i+1}^{J} \Omega_{j,k}, \tag{4.89}$$

where

$$\Omega_{j,k} = c_{j,k} \Big(\hat{\boldsymbol{\gamma}}' \mathbf{H}(\hat{\mathbf{b}})^{-1} \hat{\boldsymbol{\gamma}} - \hat{\mu}_0^{(MLE)} \hat{\boldsymbol{\gamma}}' \mathbf{H}(\hat{\mathbf{b}})^{-1} (\widetilde{\Gamma}_j + \widetilde{\Gamma}_k) + (\hat{\mu}_0^{(MLE)})^2 \mathbf{H}(\hat{\mathbf{b}})^{-1}_{j,k} \Big), \tag{4.90}$$

$$c_{j,k} = \hat{\gamma}_j^{(MLE)} \hat{\gamma}_k^{(MLE)} (\hat{\mu}_0^{(MLE)})^{-4}, \tag{4.91}$$

$$\hat{\mathbf{b}} = (\hat{m}_1^{(GLM)}, \ldots, \hat{m}_I^{(GLM)}, \hat{g}_0^{(GLM)}, \ldots, \hat{g}_J^{(GLM)})', \tag{4.92}$$

$$\hat{\boldsymbol{\gamma}} = (\underbrace{0, \ldots, 0}_{I}, \hat{\gamma}_0, \ldots, \hat{\gamma}_J)', \tag{4.93}$$

$$\widetilde{\Gamma}_j = (\underbrace{0, \ldots, 0}_{I+j}, 1, \underbrace{0, \ldots, 0}_{J-j}). \tag{4.94}$$

Proof can be found in Alai et al. [1].

Theorem 16 (ODP BF MSEP - single & aggregated accident years) *Under Assumptions 13 we have following estimators for conditional MSEP of estimated outstanding claims for accident year i*

$$\widehat{\mathsf{msep}}^{ODP}_{C_{i,J}|D_I}(\hat{C}^{BF}_{i,J}) = \hat{\phi}(1 - \hat{\beta}_{I-i})\hat{\mu}_i + (1 - \hat{\beta}_{I-i})^2 \hat{\mu}_i^2 \widehat{\mathsf{Vco}}(\hat{\mu}_i)^2 + \hat{\mu}_i^2 \sum_{j,k=I-i+1}^{J} \Omega_{j,k}, \tag{4.95}$$

The estimate of conditional MSEP for aggregated accident years is given by

$$\begin{aligned} \widehat{\mathsf{msep}}^{ODP}_{\sum_{i=1}^I C_{i,J}|D_I}\Big(\sum_{i=1}^{I} \hat{C}^{BF}_{i,J}\Big) = \sum_{i=1}^{I} \widehat{\mathsf{msep}}^{ODP}_{C_{i,J}|D_I}(\hat{C}^{BF}_{i,J}) + \\ + 2 \sum_{1 \le i < k \le I} \hat{\mu}_i \hat{\mu}_k \Big(\sum_{j=I-i+1}^{J} \sum_{l=I-k+1}^{J} \Omega_{j,l} \Big). \end{aligned} \tag{4.96}$$

4.7 Stochastic Benktander-Hovinen

We continue with Benktander-Hovinen method, that was defined in Section 3.3 as credibility mixture of Chain-ladder and Bornhuetter-Ferguson. Estimate of ultimate claim is given by

$$\hat{C}_{i,J}^{BH} = c\hat{C}_{i,J}^{CL} + (1-c)\hat{C}_{i,J}^{BF}. \tag{4.97}$$

It is clear, that the same relationship also holds for reserves R_i

$$\hat{R}_i^{BH} = c\hat{R}_i^{CL} + (1-c)\hat{R}_i^{BF}. \tag{4.98}$$

Credibility factor c was set as $c = \beta_{I-i}$. Following Mack [9], we examine, what is the optimal credibility factor c^*. We define a function $\hat{R}_i(c)$ depending on c

$$\hat{R}_i(c) = c\hat{R}_i^{CL} + (1-c)\hat{R}_i^{BF}, \tag{4.99}$$

for which, we want to minimize its unconditional mean square error of prediction

$$\mathsf{msep}_{R_i}(\hat{R}_i(c)) = \mathsf{E}[(R_i(c) - R_i)^2]. \tag{4.100}$$

Theorem 17 (Benktander-Hovinen optimal credibility factor)
If we assume, that the payout pattern is given by

$$\mathsf{E}[C_{i,j}] = \beta_j \mathsf{E}[C_{i,J}] = \beta_j \mathsf{E}[\mu_i], \tag{4.101}$$

where μ_i is a random variable independent of $C_{i,I-i}$ and $C_{i,J}$, then the optimal credibility factor minimizing (4.100) is given by

$$c^* = \frac{\beta_{I-i}}{1-\beta_{I-i}} \frac{\mathsf{Cov}(C_{I,I-i}, R_i) + \beta_{I-i}(1-\beta_{I-i})\mathsf{Var}(\mu_i)}{\mathsf{Var}(C_{I,I-i}) + \beta_{I-i}^2 \mathsf{Var}(\mu_i)}. \tag{4.102}$$

Proof can be found in Mack [9].

We can see that additional assumptions are necessary to estimate c^*, since $\mathsf{Cov}(C_{I,I-i}, R_i)$ and $\mathsf{Var}(C_{I,I-i})$ are unknown. Mack [9] suggested following model:

Assumptions 14 (Mack Benktander-Hovinen model)
We introduce Mack [9] Benktander-Hovinen model using following assumptions:

1. *Cumulative claims $C_{i,j}$ of different accident years i are independent.*

2. The payout pattern is given by

$$\mathsf{E}[C_{i,j}/C_{i,J} \mid C_{i,J}] = \beta_j, \tag{4.103}$$

$$\mathsf{Var}[C_{i,j}/C_{i,J} \mid C_{i,J}] = \beta_j(1-\beta_j)\vartheta^2(C_{i,J}). \tag{4.104}$$

Theorem 18 (Mack Benktander-Hovinen optimal credibility factor)
Under Assumptions 14 and assumptions of Theorem 17, we have following optimal credibility factor

$$c^* = \frac{\beta_{I-i}}{\beta_{I-i} - t_i}, \text{ with } t_i = \frac{\mathsf{E}[\vartheta^2(C_{i,J})]}{\mathsf{Var}(\mu_i) + \mathsf{Var}(C_{i,J}) - \mathsf{E}[\vartheta^2(C_{i,J})]}. \tag{4.105}$$

Proof can be found in Mack [9].

Theorem 19 (Mack Benktander-Hovinen model MSEP)
Under Assumptions 14 and assumptions of Theorem 17, the unconditional MSEP of outstanding loss liabilities estimate is given by

$$\mathsf{msep}_{R_i}(\hat{R}_i(c)) = \mathsf{E}[\vartheta^2(C_{i,J})]\Big(\frac{c^2}{1-\beta_{I-i}} + \frac{1}{\beta_{I-i}} + \frac{(1-c)^2}{t_i}\Big)(1-\beta_{I-i})^2. \tag{4.106}$$

Proof can be found in Mack [9].

Table 4.7 summarizes the form of unconditional MSEP depending on the choice of credibility factor c.

Table 4.2: Mack Benktander-Hovinen model MSEP

Model	c	$\mathsf{msep}_{R_i}(\hat{R}_i(c))$
Optimal credibility mixture	c^*	$\mathsf{E}[\vartheta^2(C_{i,J})](1+t_i)(1-\beta_{I-i})/(t_i+\beta_{I-i})$
Chain-ladder	$c=1$	$\mathsf{E}[\vartheta^2(C_{i,J})](1-\beta_{I-i})/(\beta_{I-i})$
Bornhuetter-Ferguson	$c=0$	$\mathsf{E}[\vartheta^2(C_{i,J})](1-\beta_{I-i})(1+(1-\beta_{I-i})/t_i)$

Note 5 (Mack Benktander-Hovinen parameter assessment)
In order to apply the model, one needs to define the function $\vartheta^2(C_{i,J})$ and estimate uncertainties in $\widehat{\mathsf{Var}}(\mu_i)$ and $\widehat{\mathsf{Var}}(C_{i,j})$. Mack suggested form of

$$\vartheta^2(C_{i,J}) = \frac{1}{1+\vartheta} C_{i,j}^2, \tag{4.107}$$

which is for example implied if $(C_{i,j}/C_{i,J} \mid C_{i,J})$ *is conditionally Beta distributed.*

For $\widehat{\mathsf{Var}}(\hat{\mu}_i)$*, the coefficient of variation* $\widehat{\mathsf{Vco}}(\hat{\mu}_i)^2$ *(4.85) can be used.*

Concerning the estimate of $\widehat{\mathsf{Var}}(C_{i,j})$*, Merz and Wüthrich [13] use again coefficient of variation* $\widehat{\mathsf{Vco}}(\hat{\mu}_i)$ *adjusted by process error.*

$$\widehat{\mathsf{Vco}}(C_{i,j}) = \sqrt{\widehat{\mathsf{Vco}}(\mu_i)^2 + \%procces\ error^2}. \tag{4.108}$$

Note, that in Mack Benktander-Hovinen MSEP estimation (4.106) we assumed, that the parameters β_j *are known, therefore the variability of this parameter is not included and the MSEP value is not comparable with other models.*

Note 6 (Benktander-Hovinen Bayesian models)

Bayesian models belong to the class of distributional models and are used when dealing with a combination of prior and posterior distribution of examined variable. The structure of the model again is set so that we start with a prior distribution and with increasing knowledge of the system (number of observations), we adjust the posterior distribution step by step.

Overview of commonly used Bayesian models in claims reserving can be found in Merz and Wüthrich [13]. Mack [9] stated log-normal Bayesian model, that was originally suggested by Gogol [6] and showed on a numerical example, that such a strong assumption on distribution does not seem to be fully compensated by model improvement. On the other hand, there is a clear benefit in possibility of obtaining ultimate reserve predictive distribution.

This is after all key issue for actuary when choosing a stochastic model, whether the distributional assumption is justified and distributional benefits are worth the loss of generality.

Chapter 5

Bootstrapping

Bootstrap is a very powerful simulation tool to obtain stochastic information about a sample of data without additional distributional assumption. Assume that we have a sample $X_1, \ldots, X_k$; following diagram shows, how the bootstrap works:

$$\begin{array}{cccccc} X_1, \ldots, X_k & \stackrel{\text{resampling}}{\longrightarrow} & (X_1^B, \ldots, X_k^B)_1 & \stackrel{\text{estimating}}{\longrightarrow} & f(X_1^B, \ldots, X_k^B)_1 = & \hat{\theta}_1^B \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_1, \ldots, X_k & \stackrel{\text{resampling}}{\longrightarrow} & (X_1^B, \ldots, X_k^B)_n & \stackrel{\text{estimating}}{\longrightarrow} & f(X_1^B, \ldots, X_k^B)_n = & \hat{\theta}_n^B \end{array}$$

Basic idea of bootstrapping is resampling data from itself, i.e. generating new values from empirical distribution of the original sample $X_1, \ldots, X_k$. By repeating this process, we obtain an arbitrary amount of new pseudo-samples $(X_1^B, \ldots, X_k^B)_m, m \leq n$ and for each of them, we can estimate a parameter of our interest $\hat{\theta}_m^B = f(X_1^B, \ldots, X_k^B)_m$. Putting these estimates together, we get bootstrap distribution of our parameter

$$\hat{\theta}_m^B \stackrel{(\mathrm{d})}{\sim} \hat{F}_n^{\theta}.$$

Since we need $X_1, \ldots, X_k$ to be independent and identically distributed for the bootstrap to make sense, the basic approach is not designed for structured data as it is in our case. In such cases, the bootstrap objects are the residuals, which we can assume to be i.i.d., if appropriately prepared. Following algorithm is used to obtain bootstrap distribution of ultimate claim $C_{i,J}$ for development triangle of cumulative claims:

1. Select the model and estimate model parameters and fitted values.
2. Calculate the residuals $r_{i,j} = r_{model}(C_{i,j}, \mathsf{E}[C_{i,J}])$ from upper triangle D_I.
3. Resample the residuals and get new pseudo-observations by reversing the residual obtaining process $C^B_{i,j} = r^{-1}_{model}(r^B, \mathsf{E}[C_{i,J})]$ using randomly selected residual.
4. Estimate $C_{i,J}$ from pseudo-observations $C^B_{i,j}$, $i + j \leq I, i \leq I$.
5. Repeat steps 3 - 5.

We can see that the bootstrap applied to ultimate claim estimation could be considered more "powerful" than the analytical estimated MSEP, since it provides predictive distribution including the opportunity of obtaining quantiles or the expected shortfall. So even though in this paper we use bootstrap results mainly to be compared with the analytical MSEP, this does not mean that we used the maximum potential that bootstrap possesses. Using the analytical estimated MSEP (if available) and bootstrap technique and combining the benefits of both approaches seems the most promising.

Bootstrap can also enlarge our stochastic understanding of specific claims reserving method, in case that we have no stochastic underlying model like in our case with Cape-Cod method. The main challenge is to ensure the bootstrap assumptions are satisfied, in particular residuals being i.d.d., since we have no assumptions on claims variance. The bootstrap is in these cases more intuitive than mathematical, but this does not mean, that the result is not interpretable.

Note 7 *Another simulation method used in claims reserving is the Monte Carlo simulation. It is used exclusively for distributional models, since it is based on generating pseudo-samples from selected distribution with estimated parameters. Typical application of Monte Carlo methods are the Bayesian models (see England and Verrall [4]).*

5.1 Bootstrapping CL,BF,BH,CC

Bootstrap defined in the previous section can be applied to all analytic methods from Chapter 3. Finding i.i.d. residuals poses the biggest challenge. Merz and Wüthrich [13] looked for the most suitable residual function for Mack Chain-ladder and came with following suggestion:

$$\hat{r}_{i,j} = \left(1 - \frac{C_{i,j-1}}{\sum_{i=0}^{I-j} C_{i,j-1}}\right)^{-1/2} \frac{C_{i,j}/C_{i,j-1} - \hat{f}_{j-1}}{\hat{\sigma}_{j-1} C_{i,j-1}^{-1/2}}. \tag{5.1}$$

Residuals in this form have conditional variance equal to one, therefore they could be resampled ($\hat{r}_{i,j} \to \hat{r}_{CL}$) and used in Autoregressive Mack CL equation as residuals $\varepsilon_{i,j}$:

$$C_{i,j+1}^B = \hat{f}_j C_{i,j}^B + \hat{r}_{CL} \hat{\sigma}_j \sqrt{C_{i,j}^B}, \tag{5.2}$$

with $C_{i,0}^B = C_{i,0}$.

From pseudo-observations $C_{i,j}^B \in D_I$ we can estimate pseudo-development factors $\hat{f}_j^B$ and obtain ultimate claim estimate. By repeating this process, we obtain bootstrap distribution of ultimate claim.

Obtained pseudo-development factors f_j^B immediately offer the possibility of application to Bornhuetter-Ferguson, Benktander-Hovinen and Cape-Cod method with help of identification

$$\hat{\beta}_j = \prod_{k=j}^{J-1} \frac{1}{\hat{f}_k}. \tag{5.3}$$

But, there is a large dependency on Chain-ladder technique, which could be considered as a big disadvantage, not to mention that using autoregressive model assumption is not justified by Bornhuetter-Ferguson, Benktander-Hovinen or Cape-Cod method.

An alternative approach is based on Distribution-free BF Assumptions 12. Residuals have following form

$$r_{i,j} = r_{BF}(X_{i,j}, \hat{x}_{i,j}, \hat{\phi}, w_{i,j}) = \frac{X_{i,j} - \hat{\mu}_i \hat{\gamma}_j}{\sqrt{\mu_i \sigma_j^2}}. \tag{5.4}$$

Pseudo-observations are again obtained by inverting the process and resampling the residuals:

$$X_{i,j}^B = \hat{\mu}_i \hat{\gamma}_j + r_{BF} \sqrt{\mu_i \sigma_j^2}. \tag{5.5}$$

Based on these pseudo-observations, we can estimate ultimate claims $\hat{C}_{i,J}^{BF}$, $\hat{C}_{i,J}^{BH}$ and $\hat{C}_{i,J}^{CC}$. Therefore, we use one approach for distribution-free CL method and two approaches (one based on CL (5.2) and one based on BF (5.5) method) to distribution-free Bornhuetter-Ferguson, Benktander-Hovinen, Cape-Cod.

5.2 Bootstrapping GLM

As in previous sections, the choice of bootstrap residuals is the main challenge. We begin with definition of unscaled Pearson residuals, which are used to estimate the dispersion factor

$$r_{i,j} = r_P(X_{i,j}, \hat{x}_{i,j}, w_{i,j}) = \frac{X_{i,j} - \hat{x}_{i,j}}{\sqrt{\frac{V(\hat{x}_{i,j})}{w_{i,j}}}}. \tag{5.6}$$

The scale parameter is then estimated by

$$\hat{\phi}^{(ODP)} = \frac{\sum_{D_I} r_P(X_{i,j}, \hat{x}_{i,j}, w_{i,j})^2}{\#observations - \#parameters}. \tag{5.7}$$

For bootstrap simulation we use scaled Pearson residuals:

$$r^s_{i,j} = r_{PS}(X_{i,j}, \hat{x}_{i,j}, \hat{\phi}^{(ODP)}, w_{i,j}) = \frac{X_{i,j} - \hat{x}_{i,j}}{\sqrt{\frac{V(\hat{x}_{i,j})\hat{\phi}^{(ODP)}}{w_{i,j}}}}. \tag{5.8}$$

Pseudo-observations are then obtained by inverting the process:

$$X^B_{i,j} = r_{PS}\sqrt{\frac{V(\hat{x}_{i,j})\hat{\phi}^{(ODP)}}{w_{i,j}}} + \hat{x}_{i,j}, \tag{5.9}$$

which guarantees, that residuals are i.i.d. in case of Over-dispersed Poisson model.

For the Over-dispersed Negative Binomial model, England & Verrall [5] suggested to bootstrap the development factors and then apply the recursive property of the model. Unscaled residuals are given by

$$r_{i,j} = r_{NB}(C_{i,j+1}, C_{i,j}, \hat{f}_j, w_{i,j}) = \frac{\sqrt{C_{i,j}}(C_{i,j+1}/C_{i,j} - \hat{f}_j)}{\sqrt{\hat{f}_j(\hat{f}_j - 1)}}. \tag{5.10}$$

Dispersion parameter $\hat{\phi}^{(NB)}$ is obtained as (5.7), but in this case, the number of observation means number of known values $C_{i,j+1}/C_{i,j}$. Bootstrap residual function has following form

$$r^s_{i,j} = r_{NBS}(C_{i,j+1}, C_{i,j}, \hat{f}_j, \hat{\phi}^{(NB)}) = \frac{\sqrt{C_{i,j}}(C_{i,j+1}/C_{i,j} - \hat{f}_j)}{\sqrt{\hat{\phi}^{(NB)}\hat{f}_j(\hat{f}_j - 1)}}. \tag{5.11}$$

Pseudo-development factors are then obtained by inverting the process:

$$f_j^B = r_{NBS}\frac{\sqrt{\hat{\phi}^{(NB)}\hat{f}_j(\hat{f}_j - 1)}}{\sqrt{C_{i,j}}} + \hat{f}_j, \tag{5.12}$$

Pseudo-observations for $j > 0$ are naturally given by

$$C_{i,j+1}^B = f_j^B C_{i,j}^B, \tag{5.13}$$

with $C_{i,0}^B = C_{i,0}$ and $C_{i,j}$ fulfills the role of weights $w_{i,j}$.

5.3 Bootstrapping Merz-Wüthrich method

In Merz-Wüthrich method, we need to find stochastic properties of claims development result. We use an intuitive way of getting CDR bootstrap distribution, which is to bootstrap estimates of claims $\hat{C}_{i,I+1-i}$, $0 \leq i \leq I$, and then compare the common estimate of ultimate claim from available observations D_I with the one using in addition also pseudo-observations from $I + 1$st diagonal. For that cause, we use adjusted residual function (5.1).

$$C_{i,I-i+1}^B = \hat{f}_j^B C_{i,I-i}, \tag{5.14}$$

$$\widehat{CDR}_i^B(I+1) = \widehat{\mathsf{E}}[C_{i,J} \mid D_I] - \widehat{\mathsf{E}}[C_{i,J} \mid D_I \cup C_{i,I-i+1}^B]. \tag{5.15}$$

Finally, we will compare the standard deviation of bootstrap CDR pseudo-observations with Merz-Wüthrich estimator for CDR variance

$$\widehat{\mathsf{Var}}\Big(CDR_i(I+1) \mid D_I\Big) = (\hat{C}_{i,J}^{CL})^2\hat{\Psi}_{i,J}. \tag{5.16}$$

Chapter 6

Results

Claims reserving methods are tested on product liability data presented in Merz and Wüthrich [13]. These data show relative stability within accident years and possess non-negative increment claims, which facilitates fitting of Over-dispersed Poisson a Negative Binomial model. Triangle of incremental claims divided by thousand is shown in Table 6.

For the practical part of the thesis, R-software was used, including Chain-Ladder package, which offers basic claims triangle operations, CL parameters and MSEP estimation and several GLM reserving techniques (not Negative Binomial). CD containing R-source codes is attached.

Table 6.1: Incremental claims (divided by 1000)

acc \ dev	0	1	2	3	4	5	6	7	8	9
0	5 947.0	3 721.2	895.7	207.8	206.7	62.1	65.8	14.9	11.1	15.8
1	6 346.8	3 246.4	723.2	151.8	67.8	36.6	52.8	11.2	11.6	
2	6 269.1	2 976.2	847.1	262.8	152.7	65.4	53.5	8.9		
3	5 863.0	2 683.2	722.5	190.7	133.0	88.3	43.3			
4	5 778.9	2 745.2	653.9	273.4	230.3	105.2				
5	6 184.8	2 828.3	572.8	244.9	105.0					
6	5 600.2	2 893.2	563.1	225.5						
7	5 288.1	2 440.1	528.0							
8	5 290.8	2 357.9								
9	5 675.6									

Table 6.2 shows the reserve estimates $\hat{R}_i$ of analytic methods defined in Chapter 3. For prior BF estimates $\hat{\mu}_i$, we take 75% of the premium Π_i, which is quite

Table 6.2: Reserves estimates

acc	Π_i	$\hat{\beta}_{I-i}$	$\hat{R}_i^{CL}$	$\hat{R}_i^{BF}$	$\hat{R}_i^{BH}$	$\hat{R}_i^{CC}$
0	15 473 558	1.0000	0	0	0	0
1	14 882 436	0.9986	15 126	15 833	15 127	14 204
2	14 456 039	0.9975	26 257	26 701	26 259	23 954
3	14 054 917	0.9965	34 538	37 308	34 548	33 470
4	14 525 373	0.9914	85 302	94 131	85 378	84 446
5	15 025 923	0.9845	156 494	174 748	156 777	156 770
6	14 832 965	0.9701	286 121	332 668	287 513	298 442
7	14 550 359	0.9484	449 167	563 061	455 043	505 131
8	14 461 781	0.8800	1 043 242	1 301 817	1 074 278	1 167 882
9	15 210 363	0.5896	3 950 815	4 681 925	4 250 874	4 200 234
$\sum$	147 473 714	-	6 047 064	7 228 192	6 385 797	6 484 533

pessimistic in comparison with other methods. Chain-ladder reserve is the lowest as a result of diagonal observation being below average. Therefore there is a possibility of underestimation of $\hat{R}_i^{CL}$. A suggestion for adjustment can be found in Chain-ladder boostrapping results, where zero residuals are used and upper triangle of expected values is created (see Table 6.5).

Table 6.3 summarizes MSEP values 4.2 under Mack Assumptions 4 and Autoregressive CL Assumptions 6. We can see, that parameter estimation errors are almost the same for both approaches.

Table 6.3: Distribution-free Chain-ladder MSEP

		$\widehat{\mathsf{Var}}(C_{i,J} \mid D_I)^{\frac{1}{2}}$	$\lvert\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I]\rvert$		$(\widehat{\mathsf{msep}}_{C_{i,J}\mid D_I})^{1/2}$	
acc	$\hat{R}_i^{CL}$	Mack & AR	Mack	AR	Mack	AR
0	0	0	0	0	0	0
1	15 126	191	187	187	268	268
2	26 257	742	535	535	915	915
3	34 538	2 669	1 493	1 493	3 059	3 059
4	85 302	6 832	3 392	3 392	7 628	7 628
5	156 494	30 478	13 517	13 517	33 341	33 341
6	286 121	68 212	27 286	27 286	73 467	73 467
7	449 167	80 076	29 675	29 675	85 398	85 398
8	1 043 242	126 960	43 903	43 903	134 336	134 337
9	3 950 815	389 783	129 769	129 770	410 817	410 818
agg	6 047 064	424 380	185 024	185 026	462 960	462 961

Table 6.4: GLM Chain-ladder MSEP

	$\widehat{\mathsf{Var}}(C_{i,J} \mid D_I)^{\frac{1}{2}}$		$\lvert\hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I]\rvert$		$(\widehat{\mathsf{msep}}_{C_{i,J}\mid D_I})^{1/2}$	
acc	ODP	ODNB	ODP	ODNB	ODP	ODNB
0	0	0	0	0	0	0
1	14 919	15 105	14 612	1 003	20 883	15 138
2	19 656	19 912	17 160	1 121	26 093	19 943
3	22 543	22 849	17 159	1 744	28 331	22 915
4	35 428	36 000	22 040	3 513	41 724	36 171
5	47 986	48 931	27 108	13 549	55 114	50 773
6	64 885	66 652	32 927	27 300	72 761	72 026
7	81 296	84 460	38 935	29 686	90 139	89 526
8	123 897	133 630	66 176	43 910	140 462	140 659
9	241 107	317 699	227 661	129 774	331 606	343 182
agg	298 290	367 699	309 564	185 203	429 892	411 707

Numerical values of estimated process estimation error a parameter estimation error for Over-dispersed Poisson and Over-dispersed Negative Binomial model are reported in Table 6.4. Estimated dispersion factors are following: $\hat{\phi}^{(ODP)} = 14714$, $\hat{\phi}^{(NB)} = 15062$. We can see that MSEP values for ODP and ODNB are quite similar, but there is a difference in decomposition between process and parameter risk, where Over-dispersed Negative Binomial decomposition more closely resembles the decomposition in distribution-free CL (see Note 4). Over-dispersed Poisson has due to higher number of model parameters lower parameter estimation error but higher process error.

Table 6.5: CL, BF, BH, CC bootstrapping under AR Chain-ladder assumption

method	analytic $\hat{R}_i$	zero-res. mean$(\hat{R}_i^B)$	mean$(\hat{R}_i^B)$	sd$(\hat{R}_i^B)$
CL	6 047 064	6 105 882	6 096 152	208 084
BF	7 228 192	7 228 986	7 224 331	160 858
BH	6 385 797	6 525 530	6 432 789	197 452
CC	6 484 533	6 487 901	6 478 493	217 639

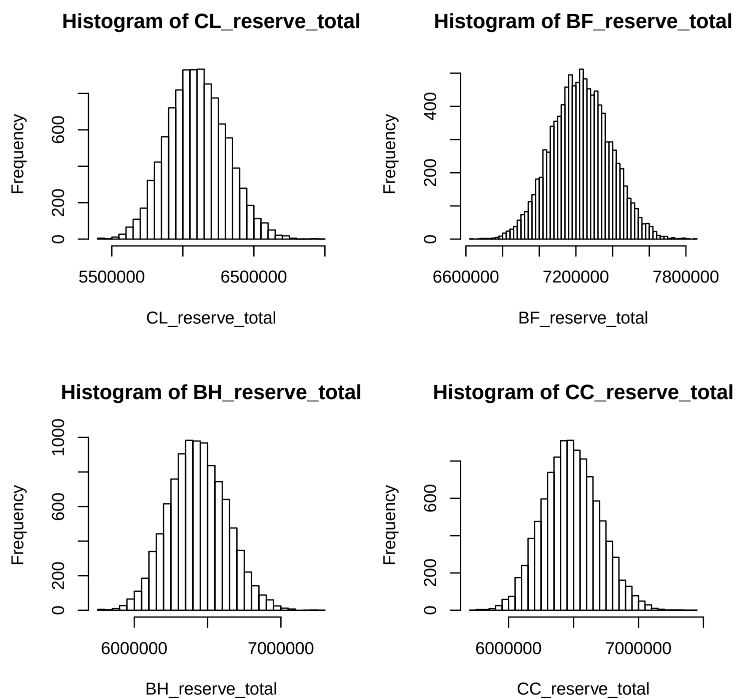
Table 6.5 shows us the bootstrap estimates of overall reserve R under the Autoregressive CL assumptions. For all bootstrap procedures we use 10 000 simulations. The average value of bootstrap reserves for Chain-ladder is above the calculated reserve estimate (Table 6.2). As mentioned, this is caused by the fact that

diagonal observations $C_{i,I-i}$ are rather bellow the expected value $C_{i,I-i-1}\hat{f}_{I-i-1}$. This phenomenon was not taken into account in Merz and Wüthrich [13], where only small negative bias of residuals was considered. Therefore we state analytic values, simulation values with zero residuals (upper triangle of expected values $C_{i,j} = C_{i,j-1}\hat{f}_{j-1}$) and bootstrap values of reserve R_i.

Residuals' negative bias causes that the bootstrap values are bellow zero-residual values.

The bootstrap distributions of $\hat{R}^{CL}, \hat{R}^{BF}, \hat{R}^{BH}, \hat{R}^{CC}$ under the Autoregressive Mack CL assumptions are shown in Figure 6.1.

Figure 6.1: Claims reserve bootstrap simulation histogram under AR CL



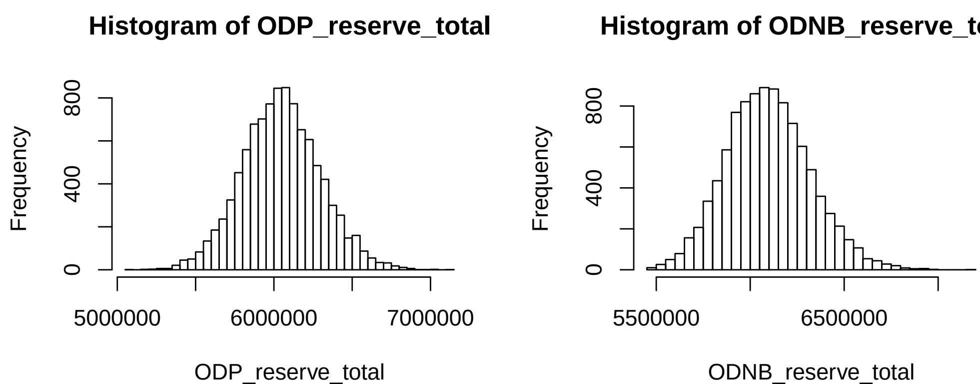
Next, we state the results of GLM Chain-ladder bootstrap following Section 5.2. GLM bootstrap reserve properties are shown in Table 6.6. We can see that the Over-dispersed Negative Binomial model shows due to its recursive property similar mean of estimated reserve as bootstrapped Chain-ladder (Table 6.5), but it has higher standard error than CL pseudo-observations.

Bootstrap distributions of Over-dispersed Poisson and Over-dispersed Negative Binomial model reserve are shown in Figure 6.2.

Table 6.6: GLM Chain-ladder bootstrap results

	R_i^{CL}	mean (R_i^B)		sd(R_i^B)	
acc	CL	ODP	ODNB	ODP	ODNB
0	0	0	0	0	0
1	15 126	15 156	15 116	11 734	13 449
2	26 257	26 213	26 029	13 800	15 560
3	34 538	34 447	35 100	13 772	15 942
4	85 302	85 156	84 385	17 724	19 851
5	156 494	156 344	162 433	21 635	25 513
6	286 121	286 132	283 874	26 564	28 775
7	449 167	449 441	462 303	31 459	34 491
8	1 043 242	1 043 624	1 076 624	53 043	53 793
9	3 950 815	3 950 845	3 950 151	184 799	98 363
agg	6 047 064	6 047 358	6 096 015	249 305	223 554

Figure 6.2: ODP & ODNB claims reserve bootstrap histogram



Estimates of Mack Distribution-free BF parameters from Section 4.5 are given in Table 6. Since we based prior estimates $\hat{\mu}_i$ on the premium adjusted be constant loss ratio $\hat{\mu}_i/\Pi_i = 0.75$, the Mack's estimator for ultimate claim variance (4.64) is replaced with coefficient of variation (4.85).

We can see, that the BF MSEP results clearly show higher values than previous methods. Such large numbers in model's parameter and process error are mainly caused by high variance parameters $\hat{\sigma}_j^2$ for first development periods.

Table 6.7: Mack Distribution-free Bornhuetter-Ferguson

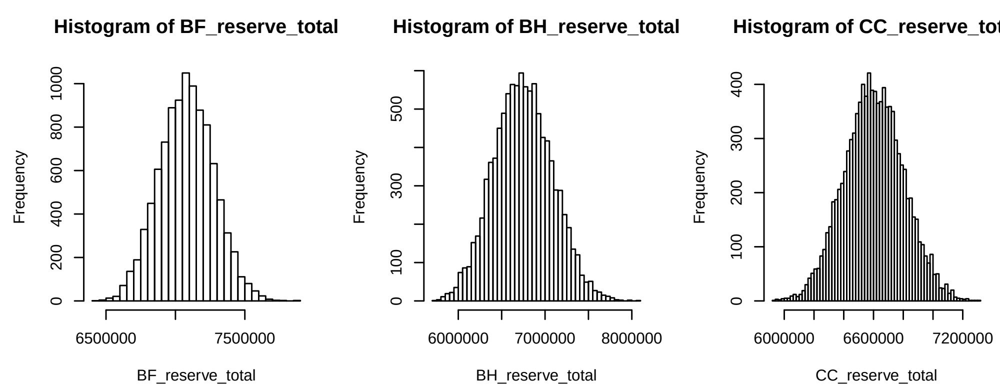
acc	$\hat{\beta}_{I-i}^{Mack}$	$\hat{R}_i^{BF}$	$\hat{R}_i^{BFMack}$	$\widehat{\mathsf{Var}}(C_{i,J} \mid D_I)^{\frac{1}{2}}$	$\lvert \hat{C}_{i,J} - \mathsf{E}[C_{i,J} \mid D_I] \rvert$	$(\widehat{\mathsf{msep}})^{1/2}$
0	1.0000	1.0000	0	0	0	0
1	0.9986	0.9985	16 931	2 754	22 356	22 525
2	0.9975	0.9974	28 520	10 365	27 955	29 815
3	0.9965	0.9962	39 934	10 664	36 084	37 627
4	0.9914	0.9908	100 447	35 477	55 436	65 816
5	0.9845	0.9835	185 440	46 717	78 602	91 437
6	0.9701	0.9685	350 282	79 164	104 291	130 933
7	0.9484	0.9461	587 826	94 253	141 352	169 894
8	0.8800	0.8768	1 336 741	182 640	260 365	318 037
9	0.5896	0.5862	4 720 547	550 499	454 072	713 604
agg	-	-	7 366 670	596 009	583 224	833 893

Another sign, that model does not fit well is quite significant bias in model residuals (mean = -0.44). This effect is reflected in bootstrap simulation, where lower reserve values are therefore expected (see Table 6). Bootstrap distributions for BF, BH, CC reserves are shown in Figure 6.3.

Table 6.8: BF, BH, CC bootstrapping under Mack Distribution-free BF

method	analytic $\hat{R}_i$	mean($\hat{R}_i^B$)	sd($\hat{R}_i^B$)
BF	7 228 192	7 084 962	195 388
BH	6 385 797	6 746 509	336 001
CC	6 484 533	6 605 771	199 332

Figure 6.3: Claims reserve bootstrap histogram under Mack Distribution-free BH

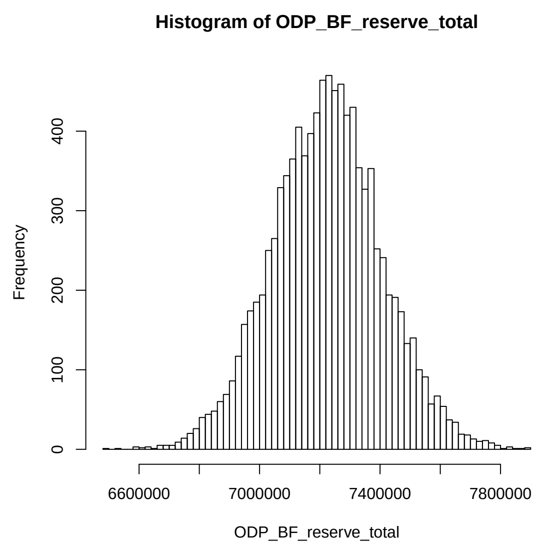


Values of parameter process and estimation error for ODP Bornhuetter-Ferguson can be found in Alai et al. [1], since we use the same data. Table 6 shows the mean and standard deviation of bootstrap BF reserve under ODP model assumptions.

Table 6.9: Over-dispersed Poisson BF bootstrapping

method	$\hat{R}_i$	mean($\hat{R}_i^B$)	sd($\hat{R}_i^B$)
BF	7 228 192	7 224 318	181 298

Figure 6.4: Claims reserve bootstrap simulation histogram



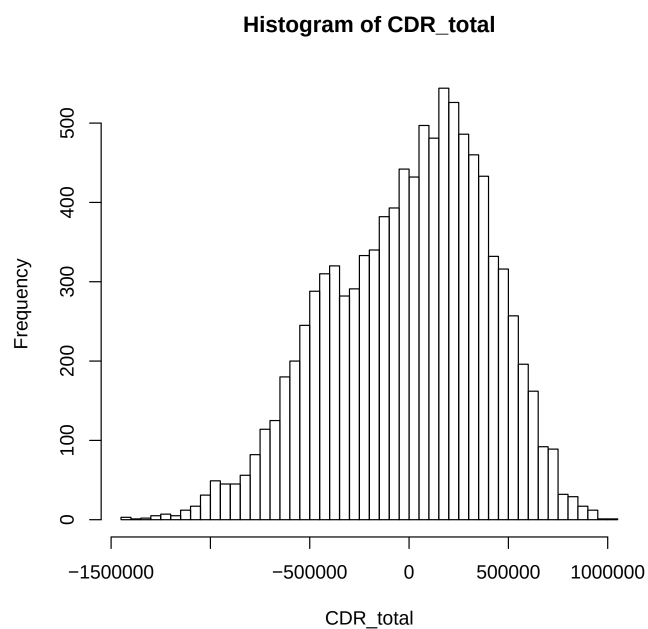
Finally, in Table 6.10 we present Merz-Wüthrich estimates for CDR MSEP and bootstrap results. We can see, that the standard deviation of CDR pseudo-observations coincides with Merz-Wüthrich variance formula (5.16).

An interesting information is provided by CDR bootstrap histogram, since it is, unlike the reserve distributions, rather asymmetrical. Therefore other risk factors like Value at Risk or Expected Shortfall could be in place.

Table 6.10: Merz-Wüthrich CDR

acc	$\left(\mathsf{msep}(\widehat{CDR}(I+1))\right)^{\frac{1}{2}}$	$\left(\mathsf{msep}_{CDR}(0)\right)^{\frac{1}{2}}$	$\left(\widehat{\mathsf{Var}}(CDR(I+1))\right)^{\frac{1}{2}}$	$\mathrm{sd}(\widehat{CDR}^B))$
0	0	0	0	0
1	268	187	191	189
2	885	518	717	710
3	2 949	1 440	2 573	2 544
4	7 018	3 128	6 283	6 210
5	32 470	13 156	29 685	29 292
6	66 178	24 645	61 418	60 853
7	50 296	18 729	46 679	46 269
8	104 311	34 121	98 572	97 316
9	385 773	121 417	366 168	364 838
agg	420 221	137 303	388 168	395 936

Figure 6.5: CDR bootstrap simulation histogram



Bibliography

[1] Alai D., Merz M., Wuthrich M.: *Mean Square Error of Prediction in the Bornhuetter-Ferguson Claims Reserving Method*, Annals of Actuarial Science, 2009

[2] Bornhuetter R. L. , Ferguson R. E.: *The actuary and IBNR*, Proceedings of the Casualty Actuarial Society, LIX, pp. 181–195, 1972

[3] Buchwalder M., Buhlman H., Merz M., Wuthrich M.: *The mean square error of prediction in chain ladder reserving method (Mack and Murphy revisited)*, Astin Bulletin, 2006

[4] England P.D., Verral R.J.: *Stochastic claims reserving in general insurance*, Bell & Bain Ltd., Glasgow, 2002

[5] England P.D., Verral R.J.: *Predictive distributions of outstanding liabilities in general insurance*, Annals of Actuarial Science 1, pp. 221–270., 2007

[6] Gogol D.: *Using expected loss ratios in reserving*, Insurance: Mathematics and Economics 12, pp. 297–299, 1993

[7] Mack, T. *Distribution-free calculation of the standard error of chain ladder reserve estimates*, ASTIN Bulletin, 1993

[8] Mack, T. *The Prediction Error of Bornhuetter-Ferguson*, ASTIN Bulletin 38, pp. 87–103, 2008

[9] Mack, T. *Credible claims reserves: the Benktander method*, ASTIN Bulletin 30/2, pp. 333–347, 2000

[10] Mack, T. *Parameter Estimation for Bornhuetter-Ferguson*, Casualty Actuarial Society Forum, 2006

[11] McCullagh, P., Nelder J. A. *Generalized linear models (second edition)*, Chapman and Hall, London, 1989

[12] Merz M., Wüthrich M.: *Modelling The Claims Development Result For Solvency Purposes*, Casualty Actuarial Society E-Forum, 2008

[13] Merz M. , Wüthrich M.: *Stochastic Claims Reserving Methods in Insurance*, United Kingdom, Wiley Finance, 2008

[14] Renshaw A.E., Verrall R.J. *A Stochastic model underlying the Chain-ladder technique*, Insurance: Mathematics and Economics 26 (2000), pp. 91–99, 1998

[15] Verrall R.J. *An investigation into stochastic claims reserving models and the chain-ladder technique*, Insurance: Mathematics and Economics 26, 2000