Verteilungsfreie Verfahren in der Schadenreservierung

KUMULATIVE DISSERTATION

Zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften der Universität Hamburg (Doctor Rerum Politicarum)

vorgelegt von

JOCHEN HEBERLE

geb. am 11.12.1978 in Tübingen

Hamburg, Februar 2014

Vorsitzender: Professor Dr. Bernhard Arnold Erstgutachter: Professor Dr. Michael Merz Zweitgutachter: Professor Dr. Dr. h.c. mult. Eberhard Schaich Datum der Disputation: 10.05.2013

Danksagung

Die vorliegende Dissertation ist zu einem Teil während meiner Zeit am Lehrstuhl für Statistik, Ökonometrie und Unternehmensforschung an der Eberhard Karls Universität Tübingen entstanden und zu einem weiteren Teil während meiner Zeit am Lehrstuhl für Mathematik und Statistik in den Wirtschaftswissenschaften an der Universität Hamburg. Ich möchte mich an dieser Stelle bei all jenen bedanken, die mich während dieser Zeit begleitet und auf verschiedene Weise zum Gelingen dieser Arbeit beigetragen haben:

Meinen wissenschaftlichen Lehrern, Herrn Professor Dr. Dr. h.c. mult. Eberhard Schaich sowie Herrn Professor Dr. Michael Merz, danke ich dafür, dass sie mir die Möglichkeit zur Anfertigung dieser Dissertation gegeben haben und mich dabei stets unterstützt haben. Herrn Merz gebührt ein besonderer Dank für seine Unterstützungen und Hilfestellungen, die weit über das Übliche hinaus gingen.

Dem Deutschen Verein für Versicherungswissenschaft e.V. danke ich für das mir zur Verfügung gestellte Promotionsstipendium, sowie für die Möglichkeit, auf verschiedenen Jahrestagungen meine zum jeweiligen Zeitpunkt aktuellen Forschungsergebnisse präsentieren zu können.

Herrn Professor Dr. Bernhard Arnold möchte ich für die Hilfestellung und Unterstützung zu Beginn meiner Lehrtätigkeit an der Universität Hamburg danken.

Für die Zusammenarbeit möchte ich mich sowohl bei meinen Tübinger als auch bei meinen Hamburger Kollegen bedanken. Ein besonderer Dank gilt hierbei Herrn Sebastian Happ, Herrn Dr. Luis Huergo und Frau Anne Thomas.

Zu guter Letzt möchte ich mich bei meinen Eltern und Geschwistern sowie bei meiner Lebensgefährtin für ihre Unterstützung recht herzlich bedanken.

Hamburg, Februar 2014

Jochen Heberle

Inhaltsverzeichnis

Ei	nleit	tung	8
1	Einf	ührung in die Schadenreservierung	11
	1.1	Abwicklungsdreiecke und Notation	12
	1.2	Das Chain-Ladder-Verfahren	14
	1.3	Beispiel	17
2	Boo	tstrapping	20
	2.1	Einführung	20
	2.2	Grundidee des Bootstrappings	21
	2.3	Mathematische Einführung	21
	2.4	Vorstellung verschiedener Bootstrappingverfahren	23
		2.4.1 Schätzung des Standardfehlers	23
		2.4.2 Ermittlung von Konfidenzintervallen mittels Bootstrappingverfahren .	24
	2.5	Beispiel	25
3	Der	Kalman-Filter	26
	3.1	Einführung	27
	3.2	Glättung, Filterung und Prognose	29
	3.3	Die Kalman-Rekursionsgleichungen	29
	3.4	Beispiel	32

Paper A: Bootstrapping the Chain-Ladder-Method for Several Correlated Run-Off Portfolios 38

1	Intr	oduction and motivation	38
	1.1	Predictive distribution of the claims reserves	38
	1.2	Claims reserving for several correlated run-off portfolios	40
	1.3	Bootstrapping in claims reserving	41
2	Cha	in-ladder method for several correlated run-off portfolios	41
	2.1	Notation and multivariate framework	42
	$\mathcal{D}\mathcal{D}$	Chain-ladder method and time-series model	43
	2.2		15
3	Der	ivation of the predictive distribution of the CL model for several corre-	15
3	Der late	ivation of the predictive distribution of the CL model for several corre- d run-off portfolios	47
3	Der late 3.1	ivation of the predictive distribution of the CL model for several corre- d run-off portfolios Step 1: Bootstrapping the Chain-ladder method for the derivation of the esti-	47
3	Der late 3.1	<pre>ivation of the predictive distribution of the CL model for several corre- d run-off portfolios Step 1: Bootstrapping the Chain-ladder method for the derivation of the esti- mation error</pre>	47
3	Der late 3.1	ivation of the predictive distribution of the CL model for several corre- d run-off portfolios Step 1: Bootstrapping the Chain-ladder method for the derivation of the esti- mation error	47 47 48
3	Der late 3.1	ivation of the predictive distribution of the CL model for several corredrun-off portfolios Step 1: Bootstrapping the Chain-ladder method for the derivation of the estimation error 3.1.1 Bootstrapping without considering correlations 3.1.2 Bootstrapping considering correlations	47 47 48 50
3	2.2 Der late 3.1 3.2	 ivation of the predictive distribution of the CL model for several corred run-off portfolios Step 1: Bootstrapping the Chain-ladder method for the derivation of the estimation error 3.1.1 Bootstrapping without considering correlations 3.1.2 Bootstrapping considering correlations Step 2: Sampling approach for including the process variance 	47 47 48 50 54

	3.2.2	Sampling approach considering correlations	57
4	Example		59
Re	eferences		62

Paper B: Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren für abhängige Run-Off-Portfolios 66

I	Emertung	00				
2	Schadenabwicklungsdreieck und Notation					
3	Multivariates Chain-Ladder-Verfahren					
4	Schadenabwicklungsergebnis für abhängige Run-Off-Portfolios	74				
5	Quantifizierung der Unsicherheit im Abwicklungsergebnis mehrerer abhän- giger Run-Off-Portfolios 5.1 Bedingter Prognosefehler für einzelne Anfalljahre 5.1.1 Bedingte Prozessvarianz 5.1.2 Bedingter Schätzfehler 5.2 Bedingter Prognosefehler für aggregierte Anfalljahre	78 78 79 80 84				
6	Parameterschätzung	87				
7	Beispiel	88				
Li	iteratur	89				

Paper C: Bootstrappingmethoden zur Bestimmung des einjährigen Reserverisikos für Solvenzbetrachtungen bei mehreren abhängigen Run-Off-Portfolios 92

1	Einl	eitung	92
	1.1	Die prädiktive Verteilung des einjährigen Reserverisikos	93
	1.2	Einjähriges Reserverisiko für korrelierte Run-Off Portfolios	94
	1.3	Bootstrappingverfahren in der Schadenreservierung	95
2	Mul	tivariates Chain-Ladder-Verfahren für korrelierte Run-Off Portfolios	96
	2.1	Schadenabwicklungsdreiecke und Notation	96
	2.2	Multivariates Chain-Ladder-Verfahren	97

3	3 Bestimmung einer prädiktiven Verteilung für mehrere korrelierte Run-Off Portfolios 100						
	 3.1 Bootstrappingmethode für das multivariate Chain-Ladder-Verfahren	101 102 104					
	 3.2 Zweistufiges Bootstrapverfahren zur Berücksichtigung der Prozessvarianz 3.2.1 Zweistufiger Bootstrap ohne Beachtung von Korrelationen 3.2.2 Zweistufiger Bootstrap mit Beachtung von Korrelationen 	106 108 109					
4	Beispiel	110					
Lit	eratur	112					
Ра	per D: Prediction of the ultimate claims based on cumulative clai payments and incurred losses data using Kalman-filter theory	ms 115					
1	Introduction 1.1 Motivation	115 115 116 118					
2	Kalman-filter theory on development triangles	119					
3	Mean square error of prediction3.1Single accident years3.2Aggregated accident years	123 123 124					
4	Estimation	125					
5	Example	125					
A	Appendix: ProofsA.1Proof of Theorem 2.5A.2Proof of Corollary 2.7	128 128 131					
Re	ferences	132					

Paper E: Bootstrapping the ultimate claims of cumulative claims payments and incurred losses data using Kalman-filter theory 135

1	Intro	oduction and motivation	135
	1.1	Predictive distribution of the claims reserves	136
	1.2	Claims reserving based on cumulative claims payments and incurred losses data	137
	1.3	Bootstrapping in claims reserving	137

2	Kalman-filter theory applied to development triangles2.1Notation	138 138
3	Derivation of the predictive (bootstrap) distribution for the claims reserve	of
	aggregated accident years	142
	3.1 Step 1: Bootstrapping the residuals for the derivation of the estimation error	144
	3.2 Step 2: Sampling approach for including the process variance	145
	3.3 Achieving the (bootstrap) distribution for the aggregated claims reserve	145
4	Example	147
Α	Appendix: Proof of Theorem 2.4	149
	A.1 Proof of the Kalman-smoothing formulas	150
	A.2 Proof of the Kalman-filtering formulas	153
Re	eferences	154

Einleitung

Die vorliegende Arbeit beschäftigt sich mit ausgewählten Problemen der Schadenreservierung, welche in Kapitel 1 näher erläutert werden. Hierbei liegt der Fokus auf verteilungsfreien Verfahren der Schadenreservierung zur Schätzung verschiedener Kenngrößen. Diese Kenngrößen sind unter anderem:

- Die Endschadenstände für einzelne sowie für aggregierte Anfalljahre.
- Die Einzelschadenreserven bzw. die Gesamtschadenreserve.
- Das Abwicklungsergebnis für die Periode (I, I + 1], welches die Differenz zwischen der Prognose für den Endschadenstand zu einem Zeitpunkt t = I und der Prognose für den Endschadenstand zum Zeitpunkt t = I + 1 darstellt.

Diese oben genannten Größen werden jeweils durch geeignete Prädiktoren bzw. Schätzer im Rahmen von stochastischen Modellen prognostiziert. Bei diesen Schätzungen der einzelnen Größen treten potenziell Prognosefehler auf, welche ebenfalls in den jeweiligen Abschnitten dieser Arbeit quantifiziert werden.

Die in dieser Arbeit behandelten Themen bzw. obig aufgezählte Punkte stellen zentrale Fragestellungen dar, welche Regulierungsbehörden bspw. in der Schweiz und in Europa mit dem Swiss Solvency Test (SST; siehe Bundesamt für Privatversicherungen [8]) bzw. mit Solvency II (siehe Europäisches Parlament und Rat der Europäischen Union [15]) ebenfalls als solche gekennzeichnet haben. Die in dieser Arbeit vorgestellten Verfahren besitzen in der Schadenreservierung eine breite Anwendbarkeit, die vor allem darin begründet liegt, dass diese Verfahren verteilungsfrei sind und damit geringere Anforderungen an die vorliegenden Daten gestellt werden als bspw. bei verteilungsgebundenen Verfahren. Verteilungsgebundene Verfahren finden sich in einer Vielzahl von Arbeiten, bspw. in Merz & Wüthrich [31], Gogol [18], Venter [42] oder Hertig [21].

Da die vorliegende Arbeit eine kumulative Dissertation darstellt, soll im weiteren Verlauf der Einleitung kurz auf die einzelnen Arbeiten eingegangen werden. Diese Arbeiten (Papers) sind:

- **Paper A**: "Bootstrapping the Chain-Ladder-Method for Several Correlated Run-Off Portfolios" in Zusammenarbeit mit Luis Huergo und Michael Merz.
- **Paper B**: "Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren für abhängige Run-Off-Portfolios" in Zusammenarbeit mit Luis Huergo und Michael Merz.

- **Paper C**: "Bootstrappingmethoden zur Bestimmung des einjährigen Reserverisikos für Solvenzbetrachtungen bei mehreren abhängigen Run-Off-Portfolios" in Zusammenarbeit mit Luis Huergo und Michael Merz.
- **Paper D**: "Prediction of the ultimate claims based on cumulative claims payments and incurred losses data using Kalman-filter theory" in alleiniger Autorenschaft.
- **Paper E**: "Bootstrapping the ultimate claims of cumulative claims payments and incurred losses data using Kalman-filter theory" in alleiniger Autorenschaft.

Bevor ein kurzer Überblick über den jeweiligen Inhalt gegeben wird, ist zu den einzelnen Arbeiten folgendes festzustellen: Paper B wurde auf der Jahrestagung 2008 des Deutschen Vereins für Versicherungswissenschaft in Dresden vorgestellt und ist in der Zeitschrift des Deutschen Vereins für Versicherungswissenschaften erschienen (siehe Heberle u. a. [20]). Auf der Jahrestagung 2009 des Deutschen Vereins für Versicherungswissenschaften wurde Paper C vorgestellt und Paper A ist 2010 in der Zeitschrift des Deutschen Vereins für Versicherungswissenschaft erschienen (siehe Heberle u. a. [19]). Die Arbeit D ist auf der IME¹-Konferenz 2012 in Hongkong präsentiert worden und Arbeit E wurde auf der ASTIN²-Konferenz 2012 in Mexiko-Stadt vorgestellt. Nachfolgend sollen die Inhalte der einzelnen Arbeiten kurz aufgezeigt werden.

Paper A beschäftigt sich, wie auch Paper C und Paper E, mit Bootstrappingmethoden. In Paper A wird die Gesamtreserve sowie der zugehörige Prognosefehler *N* abhängiger Abwicklungsdreiecke betrachtet. Der Fall *N* abhängiger Abwicklungsdreiecke ist dadurch motiviert, dass ein Abwicklungsdreieck oftmals in mehrere Teil-Dreiecke zerlegt wird, um "Homogenitätseigenschaften" zu erfüllen, welche in vielen Schadenreservierungsverfahren gefordert werden. Im abschließenden Beispiel werden verschiedene Risikokennzahlen für die Gesamtreserve, basierend auf der Bootstrapverteilung, ermittelt.

In Paper B werden explizite Formeln für die Prozessvarianz, den Schätzfehler und den Prognosefehler bei der Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren bestimmt. Dabei werden bereits bekannte Ergebnisse aus dem univariaten Chain-Ladder-Modell auf den Fall verallgemeinert, dass *N* abhängige Abwicklungsdreiecke vorliegen. Die Arbeit schließt mit einem Beispiel.

In Paper C wird im ersten Teil das einjährige Reserverisiko für das kommende Geschäftsjahr quantifiziert. Dieses einjährige Reserverisiko stellt die Unsicherheit im sogenannten Abwicklungsergebnis, welches als Differenz der "Best Estimate" Prognosen für den Endschadenstand am Anfang und am Ende des betrachteten Geschäftsjahres definiert ist, dar. Im zweiten Teil

¹Abkürzung für: Insurance: Mathematics and Economics

²Abkürzung für: Actuarial Studies in Non-Life Insurance

der Arbeit wird mittels Bootstrappingmethoden die Bootstrapverteilung dieses einjährigen Reserverisikos ermittelt. In beiden Teilen werden jeweils, wie bereits in den Arbeiten A und B, mehrere abhängige Abwicklungsdreiecke betrachtet. Auch diese Arbeit schließt mit einem Beispiel.

Während in den Arbeiten A, B und C das Chain-Ladder-Verfahren im Mittelpunkt steht, dient dieses Verfahren in Paper D und E lediglich zu Vergleichszwecken bzw. zur Parameterschätzung. Diese beiden Arbeiten beschäftigen sich mit dem Kalman-Filter und seiner Anwendung in der Schadenreservierung.

Hierbei wird von zwei unterschiedlichen Informationsquellen bzw. von zwei verschiedenen Abwicklungsdreiecken ausgegangen: dem sogenannten "paid"-Dreieck und dem "incurred"-Dreieck. Das paid-Dreieck enthält die geleisteten Zahlungen, während das incurred-Dreieck den Schadenaufwand enthält. Der Schadenaufwand stellt dabei die Summe aller geleisteten Zahlungen zuzüglich der Einzelschadenreserven dar. Es wird die Annahme getroffen, dass diese beiden beobachtbaren Abwicklungsdreiecke aus einem nicht beobachtbaren Abwicklungsdreieck zuzüglich einem Störterm generiert worden sind. Mittels Kalman-Filter-Theorie werden Rekursionen hergeleitet, mit denen der nicht beobachtbare Zustand, aus dem annahmegemäß die Beobachtbare Abwicklungsdreieck geschätzt wird. Hierbei wird nicht das komplette nicht beobachtbare Abwicklungsdreieck geschätzt, sondern lediglich der Teil, welcher für die Berechnung der Gesamtschadenreserve nötig ist. Die Arbeit wird ebenfalls mit einem Beispiel abgeschlossen.

Paper E baut auf Paper D auf, dabei sind die Modellannahmen sowie die beobachtbaren Informationsquellen identisch. Es wird mittels Bootstrappingverfahren eine simulierte Verteilung für die Gesamtschadenreserve ermittelt. Um das Bootstrappingverfahren anwenden zu können, ist die Bestimmung von Residuen unerlässlich. Hierfür ist, verglichen mit Paper D, die Schätzung des kompletten nicht beobachtbaren Abwicklungsdreiecks notwendig. Genauer gesagt ist für die Residuenbestimmung der obere linke Teil des unbeobachtbaren Abwicklungsdreiecks relevant, während für die Berechnung der Gesamtschadenreserve Kalman-Rekursionen für den unteren rechten Teil des Abwicklungsdreiecks nötig sind. Am Ende der Arbeit wird das aufgezeigte Verfahren anhand eines Beispiels dargestellt.

In den eben vorgestellten fünf Arbeiten kommen die folgenden Verfahren zum Einsatz, welche nachfolgend näher erläutert werden:

- Chain-Ladder-Verfahren
- Bootstrappingverfahren
- Kalman-Filter Algorithmus

1 Einführung in die Schadenreservierung

Die Schadenreservierung, welche von großer Bedeutung für ein Nicht-Leben Versicherungsunternehmen ist, hat als zentrale Aufgabe, für zukünftige versicherungstechnische Verpflichtungen, die aus bestehenden Verträgen resultieren, eine angemessene Reserve zu bilden und das damit verbundene Reserverisiko zu quantifizieren. Diese Frage wurde sehr lange durch einfache deterministische Verfahren beantwortet. Erst in der zweiten Hälfte des 20. Jahrhunderts haben *Aktuare* damit begonnen Schadenreservierungsmodelle in einem stochastischen Modellrahmen zu betrachten.

Ein in der Schadenreservierung relativ häufig anzutreffendes Problem ist das der sogenannten *Spätschäden.* Dieses ist darauf zurückzuführen, dass am Ende eines Geschäftsjahres noch nicht alle aufgetretenen und gemeldeten Schäden vollständig reguliert sind. Für diese Schäden hat das Versicherungsunternehmen eine ausreichende Reserve zu bilden. Bei den Spätschäden unterscheidet man grundsätzlich zweierlei Arten (vgl. Schmidt [37, S. 289]):

- Schäden, die bereits entstanden, jedoch dem Versicherungsunternehmen noch nicht gemeldet worden sind. Diese Schäden werden als IBNR³-Schäden bezeichnet.
- Schäden, die bereits gemeldet sind, deren spezifische Schadenhöhe jedoch noch nicht abschließend ermittelt bzw. zum derzeitigen Zeitpunkt noch nicht ermittelbar ist. Die für die Regulierung dieser Schäden gebildete Einzelschadenreserve ist möglicherweise nicht ausreichend. Diese Schäden werden als IBNER⁴-Schäden bezeichnet.

Ein beispielhafter Verlauf der Abwicklung eines Versicherungsanspruchs ist in Abbildung 1 dargestellt.



Abbildung 1: Typischer zeitlicher Verlauf eines Nicht-Leben Versicherungsanspruchs.

In Abbildung 1 ist dargestellt, dass der Schadeneintritt, die Schadenmeldung, die Schadenzahlungen etc. nicht notwendigerweise zum gleichen Zeitpunkt stattfinden müssen. Dies liegt

³Abkürzung für: incurred but not reported

⁴Abkürzung für: incurred but not enough reserved

unter Anderem an nachfolgend genannten Punkten (vgl. Wüthrich & Merz [45, S. 2]):

- Zwischen dem Eintreten eines Versicherungsfalls und der Meldung dieses Falls liegt immer eine gewisse Zeitspanne. Die zeitliche Differenz zwischen Schadeneintritt und -meldung kann dabei durchaus mehrere Jahre in Anspruch nehmen (bspw. könnte die Schadenmeldung an die Versicherung einer mit Asbest kontaminierten Wohnung durchaus mehrere Jahre bis Jahrzehnte nach dem Verbauen, und damit nach dem Schadeneintritt, stattfinden).
- Nachdem ein Versicherungsfall dem entsprechenden Versicherungsunternehmen mitgeteilt wurde, kann es mitunter mehrere Jahre dauern, bis der vollständige Umfang des eingetretenen Schadens bekannt ist. Bis zur endgültigen Regulierung des Schadens können somit einige Jahre vergehen.
- Nachdem ein Versicherungsfall abgeschlossen worden ist, kann es geschehen, dass der Versicherungsfall wiedereröffnet werden muss. Weitere Forderungen können auf das Versicherungsunternehmen zukommen, die möglicherweise beglichen werden müssen, bevor es zu einer erneuten Schließung des Versicherungsfalls kommen kann.

Einen Überblick über verschiedene Themengebiete der Schadenreservierung findet man unter anderem in Wüthrich & Merz [45], Taylor [40], Mack [29] oder in Kaas u. a. [22].

In den ersten drei Arbeiten A, B und C steht das in der Schadenreservierung sehr populäre *Chain-Ladder-Verfahren* im Mittelpunkt. Dieses Verfahren wird in Abschnitt 1.2 näher erläutert. Zuvor werden die hierfür benötigten Notationen und Begriffe eingeführt.

1.1 Abwicklungsdreiecke und Notation

Ein Abwicklungsdreieck stellt eine tabellarische Darstellung der Schadenzahlungen bzw. Schadenaufwendungen eines Versicherungsunternehmens in der Vergangenheit dar. In dieser tabellarischen Darstellung werden in den einzelnen Zeilen die (relativen) *Anfalljahre*, also die Jahre, in denen ein Schaden entstanden ist, dargestellt, wohingegen in den einzelnen Spalten die zugehörigen (relativen) *Abwicklungsjahre* aufgelistet sind. Die Regulierung eines in einem Anfalljahr aufgetretenen Schadens kann sich über mehrere Abwicklungsjahre erstrecken. In einem solchen Abwicklungsdreieck kann für jedes Anfalljahr und jedes Abwicklungsjahr

- die Anzahl der gemeldeten Schäden,
- die geleisteten Schadenzahlungen (*paid*), oder
- die Schadenaufwendungen (Summe aus Schadenzahlungen und Einzelfallreserven; *incurred*)

dargestellt werden (vgl. Schmidt [37, S. 290]).

Für nachfolgende Betrachtungen wird folgende Notation eingeführt (vgl. Abbildung 2):

- $i \in \{0, \dots, I\}$ bezeichnet das (relative) Anfalljahr.
- $j \in \{0, \dots, J\}$ be zeichnet das (relative) Abwicklungsjahr.
- $C_{i,j}$ ($i \in \{0, ..., I\}$ und $j \in \{0, ..., J\}$) stellt für das Anfalljahr i und das Abwicklungsjahr j die kumulierte Schadenzahlung dar.



Abbildung 2: Schematische Darstellung eines Abwicklungsdreiecks.

Bei Abwicklungsdreiecken, wie sie schematisch in Abbildung 2 dargestellt sind, werden, wie bereits erläutert, die einzelnen (relativen) Anfalljahre $i \in \{0, \ldots, J\}$ in den jeweiligen Zeilen und die (relativen) Abwicklungsjahre $j \in \{0, \ldots, J\}$ in den jeweiligen Spalten des Abwicklungsdreiecks abgetragen. Das in Abbildung 2 dargestellte Abwicklungsdreieck ist durch die Darstellung von relativen Abwicklungsjahren verglichen mit Kalenderjahren leichter zu interpretieren. Das heißt, die Abwicklungsjahre werden durch ihre Verzögerung in Bezug auf die Anfalljahre dargestellt. Die Anfalljahre werden letztendlich in gleicher Weise wie die Abwicklungsjahre dargestellt. Diese relative Notation vereinfacht die mathematische Darstellung (vgl. Schmidt [37, S. 291]). Durch diese Darstellung befindet sich die Zahlung des gleichen Kalenderjahres auf einer Diagonalen (von "unten links" nach "oben rechts").

Es wird angenommen, dass zum Zeitpunkt t = I die bereits geleisteten kumulierten Schadenzahlungen ("paid"-Dreieck) bzw. die getätigten kumulierten Schadenaufwendungen ("incurred"-Dreieck) im Anfalljahr $i \in \{0, ..., I\}$ und Abwicklungsjahr $j \in \{0, ..., J\}$ mit $i + j \leq I$ Realisationen einer Zufallsvariablen $C_{i,j}$ sind. Die Unterscheidung zwischen "paid"- und "incurred"-Dreiecken wird im Folgenden nicht weiter fortgeführt. Diese Unterscheidung spielt erst wieder in den Arbeiten D und E eine wichtige Rolle. Man stellt nun die Menge der beobachtbaren Realisierungen dar als:

$$\mathcal{D}_I = \{C_{i,j} \mid 0 \le i \le I, \ 0 \le j \le J \text{ und } i + j \le I\}$$

Der untere rechte Teil des Abwicklungsdreiecks enthält die zum Zeitpunkt t = I unbeobachtbaren Realisierungen der Zufallsvariablen $C_{i,j}$, i + j > I. Diese Menge lässt sich darstellen als:

$$\mathcal{D}_{I}^{c} = \{C_{i,j} \mid 0 \le i \le I, 0 \le j \le J \text{ und } i+j > I\}$$

Ein Ziel der Schadenreservierung besteht in der Prognose der noch ausstehenden Zahlungsverpflichtungen für einzelne Anfalljahre bzw. in der Prognose der insgesamt noch ausstehenden Zahlungsverpflichtungen. Diese beiden Größen werden für einzelne Anfalljahre $i \in \{1, ..., I\}$ mit R_i und für aggregierte Anfalljahre mit R bezeichnet. Es gilt folgender Zusammenhang:

$$R_i = C_{i,J} - C_{i,I-i}$$
 für $i = 1, ..., I$ (1.1)

$$R = \sum_{i=1}^{l} R_i \tag{1.2}$$

Da die Realisierungen der Zufallsvariablen $C_{i,j}$ zum Zeitpunkt I nicht beobachtbar sind, können auch die noch ausstehenden Zahlungsverpflichtungen R_i für $i \in \{1, ..., I\}$ bzw. die noch ausstehende Zahlungsverpflichtung R zum Zeitpunkt I noch nicht beobachtet werden (vgl. Wüthrich & Merz [45] oder Schmidt [37]).

1.2 Das Chain-Ladder-Verfahren

Das Chain-Ladder-Verfahren, welches vmtl. von Tarbell [39] erstmals beschrieben wurde, ist das populärste Verfahren in der Schadenreservierung. Es war zuerst ein rein heuristisches Verfahren, bevor dann stochastische Modelle entwickelt wurden, die dem Chain-Ladder-Verfahren zugrunde gelegt werden können. Das Chain-Ladder-Verfahren findet sich in einer Vielzahl von Arbeiten wieder, wie bspw. in Verrall [44], Mack [28], Ajne [1], Renshaw & Verrall [36], Quarg & Mack [35], Pröhl & Schmidt [34], Buchwalder u. a. [7], Gisler [17] oder in Verdonck u. a. [43].

Das Chain-Ladder-Verfahren besitzt eine Reihe von Vorteilen, wie bspw. seine Einfachheit, seine große Verbreitung oder die Tatsache, dass es mittlerweile sehr gut verstanden wird. Teilweise wird die Tatsache, dass das Chain-Ladder-Verfahren lediglich einen Datensatz benötigt, ebenfalls als Vorteil gewertet. Die Anwendbarkeit des Chain-Ladder-Verfahrens auf einen Datensatz ohne weitere Parameterannahmen treffen zu müssen wird allerdings oftmals auch als Nachteil gewertet, da keine weiteren Informationen über den datengenerierenden Prozess in das Verfahren eingehen. Ein weiterer oft angeführter Nachteil des Chain-Ladder-Verfahrens ist, dass es keine Kalenderjahr-Effekte wie bspw. Inflation berücksichtigt.

An dieser Stelle soll das verteilungsfreie Chain-Ladder-Verfahren (vgl. Mack [28]) vorgestellt werden, da dieses in den Arbeiten A, B und C eine zentrale Rolle spielt.

Für jedes Anfalljahr $i \in \{0, ..., I\}$ und jedes Abwicklungsjahr $j \in \{0, ..., J - 1\}$ wird durch

$$F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$$

der sogenannte *individuelle Abwicklungsfaktor* $F_{i,j}$ definiert. Dieser individuelle Abwicklungsfaktor stellt die relative Änderung der kumulierten Schäden innerhalb eines Anfalljahres in zwei aufeinanderfolgenden Abwicklungsjahren dar. Für i + j < I sind die Realisierungen dieser Zufallsvariablen zum Zeitpunkt t = I beobachtbar, während sie für $i + j \leq I$ unbeobachtbar sind.

Die zu den Zeitpunkten i + j > I unbeobachtbaren kumulierten Schadenzahlungen $C_{i,j}$ lassen sich wie folgt darstellen:

$$C_{i,j} = C_{i,I-i} \prod_{j=I-i}^{j-1} F_{i,j}$$
(1.3)

Die kumulierte Schadenzahlung $C_{i,I-i}$ stellt dabei die letzte beobachtbare kumulierte Schadenzahlung des Anfalljahres *i* zum Zeitpunkt t = I dar.

Für das Chain-Ladder-Verfahren werden nachfolgende Modellannahmen von Mack getroffen (vgl. Mack [28]).

Modellannahmen 1.1:

- 1. Die Mengen $\{C_{i,j} \mid 0 \le j \le J\}$ und $\{C_{k,j} \mid 0 \le j \le J\}$ werden für $i \ne k$ mit $i, k \in \{0, ..., I\}$ als unabhängig angenommen (Unabhängigkeit in den Anfalljahren).
- 2. Es existieren Chain-Ladder-Faktoren $f_0, \ldots, f_{J-1} > 0$, so dass für alle $i \in \{0, \ldots, I\}$ und $j \in \{0, \ldots, J-1\}$ gilt:

$$E[C_{i,j+1} | C_{i,0}, \dots, C_{i,j}] = E[C_{i,j+1} | C_{i,j}] = f_j C_{i,j}$$

3. Es existieren Varianzparameter $\sigma_0^2, \ldots, \sigma_{J-1}^2 > 0$, so dass für alle $i \in \{0, \ldots, I\}$ und $j \in \{0, \ldots, J-1\}$ gilt:

$$\operatorname{Var}\left(C_{i,j+1} \mid C_{i,j}\right) = \sigma_j^2 C_{i,j}$$

Da die Chain-Ladder-Faktoren f_0, \ldots, f_{J-1} unbekannt sind, wird im Folgenden das dem Chain-Ladder-Verfahren zugrundeliegende Standardverfahren zur Schätzung der Chain-Ladder-Faktoren vorgestellt.

Schätzer 1.2: Die Chain-Ladder-Faktoren f_j für $j \in \{0, ..., J - 1\}$ werden mittels der zum Zeitpunkt t = I zur Verfügung stehenden Beobachtungen \mathcal{D}_I geschätzt durch:

$$\hat{f}_{j} = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}$$
(1.4)

Der Chain-Ladder-Prädiktor für einzelne nicht beobachtbare kumulierte Schadenstände ist durch nachfolgenden Prädiktor gegeben (vgl. (1.3)).

Prädiktor 1.3: Der Chain-Ladder-Prädiktor $\widehat{C}_{i,j}$ für $\mathbb{E}[C_{i,j} | \mathcal{D}_I]$ mit i + j > I ist gegeben durch:

$$\widehat{C}_{i,j} = \widehat{E} \left[C_{i,j} \mid \mathcal{D}_I \right] = C_{i,I-i} \, \widehat{f}_{I-i} \, \cdots \, \widehat{f}_{j-1} \tag{1.5}$$

Die Modellannahmen 1.1 führen zu einer Reihe von Folgerungen, welche in Satz 1.4 zusammengefasst sind.

Satz 1.4: Unter den Modellannahmen 1.1 gilt:

- 1. $E[C_{i,J} | \mathcal{D}_I] = C_{i,I-i} f_{I-i} \cdots f_{J-1}$.
- 2. Für gegebenes \mathcal{B}_k , mit $\mathcal{B}_k = \{C_{i,j} \mid i+j \leq I \text{ und } 0 \leq j \leq k\}$ ist \hat{f}_j ein unverzerrter Schätzer für f_j , d.h. es gilt $\mathbb{E}[\hat{f}_j \mid \mathcal{B}_j] = f_j$.
- 3. Der Chain-Ladder-Schätzer \hat{f}_j ist unbedingt ein unverzerrter Schätzer für f_j , d.h. es gilt $E[\hat{f}_j] = f_j$.
- 4. Die Chain-Ladder-Schätzer $\hat{f}_0, \ldots, \hat{f}_{J-1}$ sind unkorreliert, d.h. es gilt

$$\mathbf{E}\left[\hat{f}_{0}\cdots\hat{f}_{j}\right]=\mathbf{E}\left[\hat{f}_{0}\right]\cdots\mathbf{E}\left[\hat{f}_{j}\right].$$

5. Für gegebenes $C_{i,I-i}$ ist $\widehat{C}_{i,J}$ ein unverzerrter Prädiktor für $E[C_{i,J} | C_{i,I-i}]$, d.h. es gilt

$$\mathbb{E}\left[\widehat{C}_{i,J} \mid C_{i,I-i}\right] = \mathbb{E}\left[C_{i,J} \mid C_{i,I-i}\right].$$

6. Der Prädiktor $\widehat{C}_{i,J}$ ist unbedingt ein unverzerrter Schätzer für $E[C_{i,J}]$, d.h. es gilt $E[\widehat{C}_{i,J}] = E[C_{i,J}]$.

7. Der Schätzer \hat{f}_j ist ein \mathcal{B}_{j+1} -messbarer unverzerrter Schätzer für f_j mit bedingt gegeben \mathcal{B}_j minimaler Varianz aus der Klasse der unverzerrten linearen Schätzer in $\{F_{i,j+1} \mid 0 \le i \le I - j - 1\}$ für f_j .

Beweis. Siehe Wüthrich & Merz [45, S. 17 ff.] bzw. Wüthrich & Merz [45, S. 38 ff.]

Unter der zusätzlichen Annahme

$$C_{i,j+1} = f_j C_{i,j} + \epsilon_{i,j+1}$$

und

$$\mathbf{E}\left[\epsilon_{i,j+1} \mid C_{i,j}\right] = 0$$

lässt sich leicht zeigen, dass der CL-Schätzer \hat{f}_j ein KQ-Schätzer für f_j ist. Es handelt sich insgesamt um eine gewichtete lineare Regression und man erhält die Schätzung \hat{f}_j durch Minimierung der Verlustfunktion

$$Q(f_j) = \sum_{i=0}^{I-j-1} \frac{(C_{i,j+1} - f_j C_{i,j})^2}{C_{i,j}}$$

Die erste und zweite Ableitung von Q sind gegeben durch

$$\frac{\partial Q(f_j)}{\partial f_j} = -2 \sum_{i=0}^{I-j-1} (C_{i,j+1} - f_j C_{i,j})$$
(1.6)

bzw.

$$\frac{\partial^2 Q(f_j)}{\partial f_j^2} = 2 \sum_{i=0}^{I-j-1} C_{i,j} > 0$$

und Nullsetzen von (1.6) liefert

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

1.3 Beispiel

Es soll abschließend das in Abschnitt 1.2 vorgestellte Chain-Ladder-Verfahren explizit auf ein Schadenabwicklungsdreieck (siehe Tabelle 1) angewendet werden.

Anfall-		Al	owicklu	ıngsjah	r j	
jahr <i>i</i>	0	1	2	3	4	5
0	1289	2400	3140	3875	4355	4565
1	1390	2630	3471	4272	4812	
2	1709	3278	4356	5358		
3	1950	3760	5110			
4	2150	4090				
5	2752					

Tabelle 1: Schadenabwicklungsdreieck mit jeweiligen Schadenständen für verschiedene Anfalljahre ($i \in \{0, ..., 5\}$) und Abwicklungsjahre ($j \in \{0, ..., 5\}$).

Zu Beginn werden die unbekannten Chain-Ladder-Faktoren f_j ($j \in \{0, ..., 5\}$) mittels (1.4) geschätzt. Man erhält Schätzungen \hat{f}_j ($j \in \{0, ..., 4\}$) gemäß Tabelle 2.

j	0	1	2	3	4
\hat{f}_j	1.904	1.332	1.231	1.125	1.048

Tabelle 2: Aus beobachteten Daten \mathcal{D}_I (siehe Tabelle 1) geschätzte Chain-Ladder-Faktoren \hat{f}_j $(j \in \{0, \dots, 4\}).$

Mit den in Tabelle 2 geschätzten Chain-Ladder-Faktoren \hat{f}_j $(j \in \{0, ..., 4\})$ lassen sich die noch ausstehenden Schadenstände nach (1.5) prognostizieren. Damit lassen sich dann auch die einzelnen noch ausstehenden Zahlungsverpflichtungen R_i $(i \in \{1, ..., 5\})$ sowie die insgesamt noch ausstehende Zahlungsverpflichtung R gemäß (1.1) und (1.2) prognostizieren. Die Prädiktoren $\widehat{C}_{i,j}$ $(i, j \in \{0, ..., 5\}$ mit i + j > 5) sowie \widehat{R}_i $(i \in \{1, ..., 5\})$ und \widehat{R} sind in Tabelle 3 zusammen mit den, gemäß Tabelle 1, bereits beobachteten Schadenständen $C_{i,j}$ $(i, j \in \{0, ..., 5\}$ mit $i + j \le 5$) aufgeführt.

Anfall-	Abwicklungsjahr <i>j</i>							
jahr <i>i</i>	0	1	2	3	4	5	\widehat{R}_i	
0	1289.00	2400.00	3140.00	3875.00	4355.00	4565.00		
1	1390.00	2630.00	3471.00	4272.00	4812.00	5044.03	232.04	
2	1709.00	3278.00	4356.00	5358.00	6028.82	6319.53	961.53	
3	1950.00	3760.00	5110.00	6292.56	7080.39	7421.81	2311.81	
4	2150.00	4090.00	5448.70	6709.65	7549.69	7913.74	3823.74	
5	2752.00	5238.79	6979.12	8594.23	9670.23	10136.53	7384.53	
Summe (\widehat{R})							14713.65	

Tabelle 3: Beobachtete und prognostizierte Schadenstände $C_{i,j}$ bzw. $\widehat{C}_{i,j}$ $(i, j \in \{0, ..., 5\})$ sowie prognostizierte noch ausstehende Zahlungsverpflichtungen \widehat{R}_i $(i \in \{1, ..., 5\})$ und prognostizierte kumulierte Zahlungsverpflichtung \widehat{R} .

2 Bootstrapping

2.1 Einführung

In diesem Abschnitt soll das Konzept des *Bootstrappings* vorgestellt werden, welches 1979 von Efron [11] eingeführt wurde. Ausführliche Einführungen in die Theorie des Bootstrappings findet man in Efron & Tibshirani [13], Efron [12] oder in Davison & Hinkley [9].

Bootstrappingverfahren wurden mittlerweile vielfach in der Schadenreservierung angewendet, bspw. in Lowe [27], England & Verrall [14], Pinheiro u. a. [33], Taylor & McGuire [41], Liu & Verrall [26], Peters u. a. [32] oder in Björkwall [4].

Bei Bootstrappingverfahren handelt es sich um eine statistische Methode, die als "Ziehen aus bereits beobachteten Daten" beschrieben werden kann. Dieses Ziehen aus vorhandenen Daten führt man durch, um bspw. den Erwartungswert, die Varianz oder die Verteilung einer zugrundeliegenden Zufallsvariablen zu schätzen. Bootstrappingverfahren werden vor allem dann verwendet, wenn die exakte Berechnung der gesuchten Größe analytisch nicht durchführbar ist.

Es existieren verschiedene Vor- und Nachteile bei der Anwendung von Bootstrappingverfahren. Vorteile sind:

- Bootstrappingverfahren sind einfach in ihrer Anwendung und Umsetzbarkeit.
- Die Schätzung von Standardfehlern oder Konfidenzintervallen von komplexen Parameterschätzern ist mittels Bootstrappingverfahren oftmals einfach möglich.
- Die Anwendung des Bootstraps ist ein anerkanntes Verfahren zur Beurteilung der Stabilität der erzielten Ergebnisse.
- In vielen Anwendungen ist es nicht möglich exakte Konfidenzintervalle zu bestimmen. Bootstrappingverfahren liefern hier vielfach asymptotisch exaktere Ergebnisse, als die Ermittlung von Standard-Konfidenzintervallen mit Hilfe von Standardfehlern und unter Verwendung der Normalverteilungsannahme.

Nachteile von Bootstrappingverfahren sind:

- Die Bootstrapverteilung eines Parameterschätzers kann für kleine Datensätze keine adäquate Schätzung der tatsächlichen Verteilung sein.
- Bootstrappingverfahren setzen eine Unabhängigkeitsannahme voraus, welche bei vielen Datensätzen im Voraus nicht gegeben ist (bspw. bei Zeitreihenmodellen). Dieses Problem kann durch geeignete Residuenbildung gelöst werden.

2.2 Grundidee des Bootstrappings

Eine zentrale Aufgabe in der Statistik ist es, aus einer gewonnenen Stichprobe Rückschlüsse auf die zugrundeliegende Grundgesamtheit zu ziehen. Um diese Rückschlüsse ziehen zu können wird oftmals mit Hilfe einer Schätzfunktion *T*, die auf den vorhandenen Beobachtungen basiert, gearbeitet. Da diese Schätzfunktion in der Regel unterschiedliche Resultate für unterschiedliche Stichproben derselben Grundgesamtheit liefert, ist man an der Verteilung dieser Resultate interessiert. Hierfür liefert das Bootstrapping eine Möglichkeit diese Verteilung zu approximieren.

Die Idee der Bootstrappingmethode ist, dass aus einer bereits vorhandenen Stichprobe der Größe $n \in \mathbb{N}$ mehrfach (*N*-mal) mit Zurücklegen Stichproben vom Umfang *n* gezogen werden. Dadurch erhält man $N \in \mathbb{N}$ Bootstrapstichproben der Größe *n*, wobei mit jeder dieser Bootstrapstichproben der entsprechende Schätzer *T* berechnet wird.

Damit resultiert eine empirische Verteilung des Schätzers *T*, woraus alle interessierenden Größen für den Schätzer *T* abgeleitet werden können. Dazu gehören bspw. Bootstrap-Konfidenzintervalle, welche aus einer sehr großen Anzahl *N* an Bootstrapstichproben derart berechnet werden, dass ein bestimmter Anteil $\frac{\alpha}{2}$ ($\alpha \in (0, 1)$; typischerweise $\alpha \in \{0.01, 0.02, 0.05\}$) am oberen bzw. unteren Ende von den, nach der Größe geordneten, Bootstrap-Stichprobenstatistiken abgeschnitten wird. Der Bereich, welcher nach diesen beiden Schnitten übrigbleibt, bildet dann das sogenannte $(1 - \alpha)$ -Bootstrap-Konfidenzintervall.

2.3 Mathematische Einführung

Gegeben sei eine stetige Zufallsvariable X mit Dichte f und Verteilungsfunktion F, d.h. $F(y) = P(X \le y) = \int_{-\infty}^{y} f(x) dx$. Gesucht sei nun eine Schätzung eines Parameter (-vektors) θ , welcher eine Funktion der interessierenden Verteilungsfunktion F ist, d.h. es gilt $\theta = g(F)$ für eine geeignete Funktion g. Basierend auf $n \in \mathbb{N}$ Realisierungen x_1, \ldots, x_n der Zufallsvariablen $X_i \stackrel{\text{i.i.d.}}{\sim} X$ wird, mit Hilfe einer Schätzfunktion

$$T_n: \mathbb{R}^n \to \mathbb{R}, \qquad \mathbf{x} = (x_1, \dots, x_n) \mapsto T_n(\mathbf{x})$$

ein Schätzwert $\hat{\theta}_n = T_n(\mathbf{x})$ für den unbekannten Parameter θ ermittelt.

Für die Beurteilung der Güte des Schätzers $\hat{\theta}_n = T_n(X_1, \dots, X_n)$ ist man an der Verteilung $F_{\hat{\theta}_n}$ von $\hat{\theta}_n$ interessiert. Aus dieser können dann in einem zweiten Schritt alle Kenngrößen abgeleitet werden, welche zur Quantifizierung der Schätzgüte benötigt werden.

Die Idee des Bootstrappingansatzes zur Ermittlung von $F_{\hat{\theta}_n}$ stellt sich nun wie folgt dar: Angenommen die Verteilung F der Zufallsvariable X sei bekannt. Es werden nun $N \in \mathbb{N}$ unabhängige (Bootstrap-) Kopien des Zufallsvektors (X_1, \ldots, X_n) betrachtet, wobei $X_i \stackrel{\text{i.i.d.}}{\sim} X$ gilt. Diese (Bootstrap-) Kopien werden als $X^k = (X_1^k, \ldots, X_n^k)$ mit $k \in \{1, \ldots, N\}$ bezeichnet. Für $k \in \{1, \ldots, N\}$ resultieren hieraus $T_n(X^k) \stackrel{\text{i.i.d.}}{\sim} F_{\hat{\theta}_n}$. Nach dem Satz von Glivenko-Cantelli (siehe bspw. Klenke [24, S. 117]) folgt

$$\hat{F}_N(T_n(X^1),\ldots,T_n(X^N)) \xrightarrow[schwach]{N \to \infty} F_{\hat{\theta}_n}, \qquad (2.1)$$

wobe
i \hat{F}_N die empirische Verteilungsfunktion darstellt.

Nach Gleichung (2.1) ist es also möglich, bei bekannter Verteilung F der Zufallsvariable X, durch unabhängiges Ziehen einer hinreichend großen Anzahl $N \in \mathbb{N}$ von i.i.d. Stichproben (X_1^k, \ldots, X_n^k) , mit $k \in \{1, \ldots, N\}$, die gesuchte Verteilung $F_{\hat{\theta}_n}$ des Schätzers $T_n(X_1, \ldots, X_n)$ beliebig genau zu approximieren. Das in der Praxis hierbei auftretende Problem ist jedoch, dass die Verteilung F von X unbekannt ist und somit eine Ziehung von N i.i.d. Stichproben (X_1^k, \ldots, X_n^k) , mit $k \in \{1, \ldots, N\}$, nicht möglich ist. Aus diesem Grund wird nicht aus der unbekannten Verteilung F der Zufallsvariablen X gezogen, sondern aus einer vorab ermittelten empirischen Verteilung \hat{F}_n , basierend auf den $n \in \mathbb{N}$ bekannten Realisierungen x_1, \ldots, x_n der Zufallsvariablen $X_i \overset{\text{i.i.d.}}{\sim} X$. Diese Approximation ist durch eine erneute Anwendung des Satzes von Glivenko-Cantelli motiviert:

$$\hat{F}_n \xrightarrow[schwach]{n \to \infty} F \tag{2.2}$$

Die Bootstrappingmethode besteht somit aus zwei Schritten:

Schritt 1: Ausgehend von *n* Realisierungen x_1, \ldots, x_n der Zufallsvariablen $X_i \stackrel{\text{i.i.d.}}{\sim} F$ wird die empirische Verteilung \hat{F}_n gebildet. Diese approximiert nach Gleichung (2.2) die Verteilung *F*. Hierbei kann eine bessere Approximation ausschließlich durch eine Erhöhung des Stichprobenumfangs *n* erzielt werden.

Schritt 2: Ausgehend von der empirischen Verteilung \hat{F}_n werden N unabhängige Stichproben $(\tilde{X}_1^k, \ldots, \tilde{X}_n^k)$ mit $k \in \{1, \ldots, N\}$ und $\tilde{X}_i^{k \text{ i.i.d.}}$ \hat{F}_n $(i \in \{1, \ldots, n\})$ gezogen. Aus diesen N unabhängigen Stichproben wird nun $\hat{F}_N(T_n(\tilde{X}^1), \ldots, T_n(\tilde{X}^N))$ bestimmt. Der in diesem Schritt entstehende Approximationsfehler kann durch Erhöhung von N beliebig verkleinert werden.

Zusammenfassend lässt sich für Schritt 1 und Schritt 2 festhalten, dass die Schwierigkeit beim Bootstrapping darin besteht, dass der Stichprobenumfang n oftmals nicht beliebig erhöht werden kann, wohingegen die N-fache Ziehung der Bootstrapstichproben sehr leicht erhöht werden kann.

In Abbildung 3 ist das grundsätzliche Vorgehen beim Bootstrapping dargestellt.



Abbildung 3: Schematische Darstellung des Bootstrappingkonzepts. Dabei wird aus einer unabhängigen Stichprobenrealisierung \mathbf{x} *n*-mal mit Zurücklegen gezogen. Dies wird *N*-mal durchgeführt. Hieraus erhält man die Bootstrapreplikationen $\tilde{\mathbf{x}}^1, \ldots, \tilde{\mathbf{x}}^N$. Mit diesen Bootstrapreplikationen und dem Schätzer T_n lassen sich nun Schätzungen für den unbekannten Parameter θ ermitteln.

2.4 Vorstellung verschiedener Bootstrappingverfahren

Es sollen im Folgenden zwei klassische Bootstrappingverfahren vorgestellt werden.

2.4.1 Schätzung des Standardfehlers

Es wird angenommen, dass ein unbekannter Parameter θ einer Grundgesamtheit zu schätzen ist. Mit $\hat{\theta}_n$ wird der Schätzer des Parameters θ bezeichnet, welcher auf einer Stichprobe der Größe *n* basiert. Es ist also $\hat{\theta}_n = T_n(x_1, \ldots, x_n)$ eine Funktion der Werte x_1, \ldots, x_n , welche Realisierungen der unabhängig und identisch verteilten Zufallsvariablen X_1, \ldots, X_n sind. Es wird nun der folgende Bootstrappingansatz durchgeführt:

- 1. Aus x_1, \ldots, x_n wird mit Zurücklegen *n*-mal gezogen und daraus wird $\tilde{\hat{\theta}}$ berechnet. Die Berechnung von $\hat{\hat{\theta}}$ geschieht nach derselben Methode/Funktionsvorschrift wie die Berechnung von $\hat{\theta}$.
- 2. Der 1. Schritt wird *N*-mal wiederholt, sodass dann die Bootstrapschätzungen $\hat{\tilde{\theta}}^1, \ldots, \hat{\tilde{\theta}}^N$ vorliegen.

3. In einem letzten Schritt schätzt man den Standardfehler von $\hat{\theta}$ durch:

$$\tilde{s}(\hat{\theta}) = \left(\frac{1}{N}\sum_{i=1}^{N} \left(\hat{\theta}^{i} - \hat{\theta}_{n}\right)^{2}\right)^{1/2}$$

2.4.2 Ermittlung von Konfidenzintervallen mittels Bootstrappingverfahren

Bei einem Konfidenzintervall für einen unbekannten Grundgesamtheitsparameter θ wird in einem ersten Schritt das Signifikanzniveau α für dieses Intervall im Voraus festgelegt (gebräuchliche Werte sind 95%, 99% oder 99.5%).

Konfidenzintervalle für einen unbekannten Grundgesamtheitsparameter θ sind durch die beiden Grenzen eines Intervalls gegeben, bspw. $[\hat{\theta}_I, \hat{\theta}_{II}]$. Dabei wird im Voraus die Wahrscheinlichkeit, dass der unbekannte Parameter θ innerhalb des Konfidenzintervalls $[\hat{\theta}_I, \hat{\theta}_{II}]$ liegt, vorgegeben (übliche Werte sind 95%, 99% oder 99.5%).

Üblicherweise basiert die Berechnung eines Konfidenzintervalls auf der Kenntnis der Verteilung der Schätzfunktion $\hat{\theta}$, oder zumindest auf der Kenntnis der asymptotischen Verteilung dieser Schätzfunktion. Diese Kenntnis wird beim nachfolgend vorgestellten Bootstrappingansatz nicht benötigt.

Angenommen es wurden N Bootstrap
replikationen durchgeführt, wobei aus jeder Replikation eine Bootstrap
schätzung $\hat{\theta}^i$ (i = 1, ..., N) für den unbekannten Grundgesamtheitsparameter
 θ resultiert. Ordnet man diese Werte der Größe nach, wobei diese mit
 $\hat{\theta}^{[i]}$ (i = 1, ..., N) bezeichnet werden, und wählt man diese Anordnung so, dass
 $\hat{\theta}^{[i]} \leq \hat{\theta}^{[i+1]}$ (i = 1, ..., N - 1) gilt, so ergibt sich daraus näherungsweise ein $(1 - \alpha)$ -Konfidenzintervall ($\alpha \in [0, 1]$) wie folgt:

$$\left[\hat{\hat{\theta}}^{\left[\left\lfloor\frac{\alpha}{2}N\right\rfloor\right]},\,\hat{\hat{\theta}}^{\left[\left\lceil\left(1-\frac{\alpha}{2}\right)N\right\rceil\right]}\right]$$
(2.3)

Mit "[]" bzw. mit "[]" wird die Gaußklammer (Abrundungsfunktion) bzw. die Aufrundungsfunktion bezeichnet.

Aus Formel (2.3) ergibt sich lediglich ein approximatives Konfidenzintervall, da lediglich eine endliche Anzahl an Werten $\hat{\theta}^{[i]}$ (i = 1, ..., N) vorhanden sind.

2.5 Beispiel

In einem abschließenden Beispiel soll anhand eines Datensatzes mittels Bootstrappingverfahren eine Approximation der Verteilung des Schätzers (arithmetisches Mittel)

$$T_{26}(X_1,\ldots,X_{26}) = \frac{1}{26} \sum_{i=1}^{26} X_i$$

mit $X_i \stackrel{\text{i.i.d.}}{\sim} X$ bestimmt werden. Der hier verwendete Datensatz mit n = 26 Beobachtungen ist in Tabelle 4 gegeben.

i	1	2	3	4	5	6	7	8	9	10	11	12	13
x_i	-0.6	7.6	47.4	16.6	-4.9	52.0	49.9	3.8	22.8	19.8	-7.9	7.3	40.1
i	14	15	16	17	18	19	20	21	22	23	24	25	26
x_i	23.9	14.8	29.7	25.3	24.4	-2.0	21.1	5.1	24.7	51.9	8.8	2.5	16.0

Tabelle 4: Uneingeschränkte Stichprobe der Größe n = 26 mit einem arithmetischen Mittel von $\bar{x} = 19.23462$ aus der zugehörigen Grundgesamtheit.

Führt man nun ein Bootstrappingverfahren mit N = 1000 zur Schätzung der unbekannten Verteilung F des arithmetischen Mittels der Grundgesamtheit durch, so erhält man als Histogramm Abbildung 4. Dieses Histogramm stellt dabei eine Näherung an die Dichtefunktion dar.



Abbildung 4: Histogramm der N = 1000 arithmetischen Mittel, berechnet aus den einzelnen Bootstrapreplikationen. Die gestrichelte Linie entspricht dem arithmetischen Mittel der Stichprobe ($\bar{x} = 19.23462$).

3 Der Kalman-Filter

In diesem Abschnitt soll eine Einführung bzw. ein kurzer Überblick über den Kalman-Filter (siehe Kálmán [23]) gegeben werden. Der Kalman-Filter (nach Rudolf E. Kálmán) stellt eine Methode dar, bei der aus vorliegenden, mit additivem Fehlerterm überlagerten, Beobachtungen Rückschlüsse auf den nicht messbaren bzw. nicht beobachtbaren "Zustand" getroffen werden können.

Der Kalman-Filter fand seinen ersten Einsatz in der Raumfahrttechnik, genauer im Apollo-Programm der NASA⁵. Dort kam er in Navigations- bzw. Leitsystemen zum Einsatz. Die NASA-Ingenieure hatten das Problem, aus mehreren optischen Beobachtungssystemen zur Positions- und Navigationsbestimmung der Mondkapsel, welche unterschiedliche Beobachtungen lieferten, die korrekte Position bzw. Navigation "herauszufiltern" (vgl. McGee & Schmidt [30]). Es wurde erkannt, dass mit Hilfe des Kalman-Filters dieses Positions-, Navigations- und Kontrollproblem der Mondkapsel in den Griff zu bekommen war. Anfänglich gab es zwar noch Schwierigkeiten beim Verständnis der Arbeit von Kalman, da ein relativ neuer Ansatz der Zustandsraumdarstellung unter Verwendung von Differenzengleichungen erster Ordnung verwendet wurde. Diese konnten jedoch im Laufe der Zeit beseitigt werden. Die Beseitigung dieser Schwierigkeiten führte dazu, dass der Kalman-Filter in abgewandelter Form als *Extended Kalman-Filter* (ursprünglich auch Kalman-Schmidt-Filter genannt) implementiert wurde (vgl. Baumann [3]).

Der Kalman-Filter wurde bereits mehrfach in der Schadenreservierung angewendet, unter anderem in de Jong & Zehnwirth [10], Verrall [44], Garrido & Romera [16], Kremer [25] oder in Atherino u. a. [2].

Das Grundkonzept des Kalman-Filters kann wie folgt beschrieben werden: Der Kalman-Filter ist ein rekursives Verfahren, welches in jedem Rechenschritt, basierend auf einer Beobachtung zum aktuellen Zeitpunkt und auf der zum vorherigen Zeitpunkt erfolgten Schätzung, den aktuellen Zustand schätzt.

Der Kalman-Filter unterscheidet also zwischen dem Zustand und der Beobachtung eines Systems. Im Falle der Positionsbestimmung der NASA-Mondkapsel haben verschiedene Messstationen unterschiedliche Beobachtungen zur aktuellen Position der Kapsel ermittelt. Aus diesen unterschiedlichen Beobachtungen war es nun möglich mit Hilfe des Kalman-Filters den Zustand, also die tatsächliche Position, zu bestimmen.

Bevor eine Einführung in den Kalman-Filter gegeben wird sollen noch Vor- und Nachteile dieses Verfahrens aufgeführt werden. Vorteile des Kalman-Filters sind:

 $^{^5\}mathrm{Abk}\ddot{\mathrm{u}}\mathrm{rzung}$ für: National Aeronautics and Space Administration

- Der Kalman-Filter ist in seiner Umsetzung sehr einfach.
- Der Kalman-Filter ist ein erwartungstreuer und konsistenter Schätzer.
- Innerhalb der Klasse der linearen Schätzer ist der Kalman-Filter jener mit minimaler Varianz (Minimum-Varianz-Schätzer).
- Der Kalman-Filter besitzt die Möglichkeit mit fehlenden Beobachtungen (missing-data-Problem) im Datensatz umzugehen.

Nachteile des Kalman-Filters sind:

- Der Kalman-Filter ist auf lineare Zustandsraummodelle beschränkt.
- Die Wahl der Startwerte sowie der Parameter sind vom Kalman-Filter nicht vorgegeben und können die Ergebnisse erheblich beeinflussen.
- Durch seinen iterativen Algorithmus können auftretende Rundungsfehler zur Divergenz des Kalman-Filters führen.

3.1 Einführung

Es wird im Folgenden von einem, möglicherweise multivariaten, Zeitreihenmodell ausgegangen. Eine Zeitreihe wird hierbei als Realisation eines stochastischen Prozesses aufgefasst. Die einzelnen Beobachtungen dieser Zeitreihe seien Realisationen der w-dimensionalen Zufallsvariablen Y_t (t = 1, 2, ...). Für ein Zustandsraummodell wird nun die Annahme getroffen, dass die einzelnen Beobachtungen Y_t sich als affin-lineare Funktion der v-dimensionalen Zustandsgrößen X_t (t = 1, 2, ...) darstellen lassen. Dies lässt sich durch

$$Y_t = G_t X_t + W_t \tag{3.1}$$

beschreiben, wobe
i $\boldsymbol{W}_t \sim WN(\boldsymbol{0}, R_t)$ einen White-Noise-Prozess mit Erwartungswertvektor Null und Kovarianz
matrix R_t beschreibt. Die Matrizen G_t
(t = 1, 2, ...) bilden eine Folge von $(w \times v)$ -Matrizen, welche exogen gegeben sind.

Für die Zustände X_t (t = 1, 2, ...) wird die stochastische Dynamik

$$\boldsymbol{X}_{t+1} = \boldsymbol{F}_t \boldsymbol{X}_t + \boldsymbol{V}_t \tag{3.2}$$

angenommen. Dabei sind F_t (t = 1, 2, ...) ebenfalls eine Folge von exogen gegebenen ($v \times v$)-Matrizen und $V_t \sim WN(0, Q_t)$ ein White-Noise-Prozess. Es wird weiterhin angenommen, dass V_t und W_s für alle t, s = 1, 2, ... unkorreliert sind, also $E[W_t V'_s] = 0$ gilt. Abschließend wird angenommen, dass der Initialisierungszustand X_1 unkorreliert von allen Störtermen V_t bzw. W_t ist.

Folgende Punkte sind hervorzuheben:

- Die hier vorgestellten Annahmen können auch dahingehend verallgemeinert werden, dass Korrelationen zwischen V_t und W_t möglich sind (vgl. Brockwell & Davis [6, S. 463 ff.]).
- Werden keine weiteren Annahmen getroffen, so ist es in praktischen Anwendungen oftmals schwierig die Matrizen F_t , G_t , Q_t und R_t (t = 1, 2, ...) zu schätzen. Aus diesem Grund wird häufig angenommen, dass diese Matrizen unabhängig von t und somit zeitinvariant sind.
- Aus den Gleichungen (3.1) und (3.2) folgt für t = 2, 3, ...

$$X_{t} = F_{t-1}X_{t-1} + V_{t-1}$$

= $F_{t-1} (F_{t-2}X_{t-2} + V_{t-2}) + V_{t-1}$
= $\dots = (F_{t-1} \cdots F_{1}) X_{1} + (F_{t-1} \cdots F_{2}) V_{1} + \dots + F_{t-1}V_{t-2} + V_{t-1}$
= $f_{t} (X_{1}, V_{1}, \dots, V_{t-1})$ (3.3)

und

$$Y_t = g_t (X_1, V_1, \dots, V_{t-1}, W_t).$$
(3.4)

Dabei stellen f_t und g_t (t = 2, 3, ...) geeignet spezifizierte Funktionen dar.

• Aus den Gleichungen (3.3) und (3.4) folgt, zusammen mit den Annahmen über die Störterme, dass

$$E [V_t X'_s] = 0$$
$$E [V_t Y'_s] = 0$$
$$E [W_t X'_s] = 0$$

für $1 \le s \le t$ und

$$\mathrm{E}\left[\boldsymbol{W}_{t}\boldsymbol{Y}_{s}^{\prime}\right]=\boldsymbol{0}$$

für $1 \le s < t$ gilt (vgl. Brockwell & Davis [5, S. 260 f.]).

3.2 Glättung, Filterung und Prognose

In diesem Abschnitt sollen drei Situationen aufgezeigt werden, in denen der Kalman-Filter angewendet wird. Dabei kommt jeweils der beste affin-lineare Prädiktor des Zustandsvektors X_t bei gegebenen Beobachtungen Y_1, Y_2, \ldots zum Einsatz. Zusätzlich wird angenommen, dass ein Zufallsvektor Y_0 existiert, welcher unkorreliert mit V_t und W_t (t > 0) ist.

Betrachtet man den Zustandsvektor X_t , bzw. soll dieser durch \hat{X}_t geschätzt werden, so unterscheidet man die folgenden Situationen:

- 1. Sind die Beobachtungsvektoren Y_0, \ldots, Y_{t-1} gegeben, so wird die Schätzung von X_t als *Prognoseproblem* bezeichnet.
- 2. Bei gegebenen Beobachtungsvektoren Y_0, \ldots, Y_t wird die Schätzung von X_t Filterproblem genannt.
- 3. Sind hingegen die Beobachtungsvektoren Y_0, \ldots, Y_k mit k > t gegeben, so wird die Schätzung von X_t als *Glättungsproblem* klassifiziert.

Nachfolgend soll nun noch das Konzept des besten affin-linearen unverzerrten Schätzers eingeführt werden.

Definition 3.1: Der beste affin-lineare Prädiktor für X_k basierend auf den Beobachtungen Y_0, \ldots, Y_t ist durch

$$\hat{X}_{k|t} = \operatorname{Pro}(X_k \mid \mathcal{A}_t)$$

mit

$$\mathcal{A}_t = \left\{ A \mid A = \boldsymbol{b} + \sum_{i=0}^t \boldsymbol{B}'_i \boldsymbol{Y}_i \text{ mit } \boldsymbol{b} \in \mathbb{R}^v \text{ und } \boldsymbol{B}_0, \dots, \boldsymbol{B}_t \in \mathbb{R}^{v \times w} \right\}$$

gegeben. "Pro" stellt dabei den Operator für die orthogonale Projektion auf den Unterraum A dar (siehe Brockwell & Davis [5, S. 282]).

Nachfolgend werden die Rekursionsgleichungen für das Prognose-, das Filter- sowie das Glättungsproblem vorgestellt. Dabei wird aus der Notation $\hat{X}_{k|t}$ ersichtlich, ob es sich bei der Prognose des Zustandsvektors X_k , bei gegebenen Beobachtungen Y_0, \ldots, Y_t um ein (1-Schritt) Prognoseproblem (k = t + 1), ein Filterproblem (k = t) oder um ein Glättungsproblem (k < t) handelt.

3.3 Die Kalman-Rekursionsgleichungen

Es wird nun zuerst die Situation betrachtet, in der die Beobachtungen Y_0, \ldots, Y_t gegeben sind und der Zustandsvektor X_{t+1} prognostiziert werden soll, d.h. es soll der Prädiktor $\hat{X}_{t+1|t}$ ermittelt werden. **Satz 3.2** (Kalman-Prognosealgorithmus): Für die in den Gleichungen (3.1) und (3.2) gegebenen Modellspezifikationen ist der 1-Schritt-Prädiktor $\hat{X}_{t+1|t}$ sowie die zugehörige Kovarianzmatrix $\Omega_{t+1|t}$ mit

$$\Omega_{t+1|t} = \mathbf{E} \left[\left(X_{t+1} - \hat{X}_{t+1|t} \right) \left(X_{t+1} - \hat{X}_{t+1|t} \right) \right]$$

bei gegebenen Startwerten

$$\hat{X}_{1|0}$$
 und $\Omega_{1|0} = \mathbb{E}\left[\left(X_1 - \hat{X}_{1|0}\right)\left(X_1 - \hat{X}_{1|0}\right)\right]$

für $t = 1, 2, \dots$ gegeben durch

$$\hat{X}_{t+1|t} = F_t \hat{X}_{t|t-1} + \Theta_t \Delta_t^{-1} \left(Y_t - G_t \hat{X}_{t|t-1} \right),$$

$$\Omega_{t+1|t} = F_t \Omega_{t|t-1} F_t' + Q_t - \Theta_t \Delta_t^{-1} \Theta_t',$$

wobei

$$\Delta_t = G_t \Omega_{t|t-1} G'_t + R_t,$$

$$\Theta_t = F_t \Omega_{t|t-1} G'_t$$

ist und Δ_t^{-1} eine verallgemeinerte Inverse von Δ_t bezeichnet.

Beweis. Siehe bspw. Brockwell & Davis [5, S. 273 f.].

Auf Basis des 1-Schritt-Prädiktors (siehe Satz 3.2) lässt sich nun der *h*-Schritt-Prädiktor angeben.

Korollar 3.3 (Kalman-*h*-Schritt-Prädiktor): Unter denselben Voraussetzungen wie in Satz 3.2 erhält man als h-Schritt Prädiktor für den Zustandsvektor X_{t+h} sowie für die zugehörige Kovarianzmatrix (h = 2, 3, ...):

$$\hat{X}_{t+h|t} = F_{t+h-1}\hat{X}_{t+h-1|t}$$
$$\Omega_{t+h|t} = F_{t+h-1}\Omega_{t+h-1|t}F'_{t+h-1} + Q_{t+h-1}$$

Beweis. Siehe bspw. Brockwell & Davis [5, S. 274].

Der noch folgende Kalman-Filteralgorithmus und Kalman-Glättungsalgorithmus basieren ebenfalls jeweils auf denselben Voraussetzungen wie Satz 3.2. Der Kalman-Filteralgorithmus ist nachfolgend beschrieben.

Satz 3.4 (Kalman-Filteralgorithmus): Unter denselben Voraussetzungen wie in Satz 3.2 erhält man als Prädiktor für X_t bei gegebenen Beobachtungen Y_0, \ldots, Y_t

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + \Omega_{t|t-1}G'_{t}\Delta_{t}^{-1}\left(Y_{t} - G_{t}\hat{X}_{t|t-1}\right)$$
$$\Omega_{t|t} = \Omega_{t|t-1} - \Omega_{t|t-1}G'_{t}\Delta_{t}^{-1}G_{t}\Omega'_{t|t-1}.$$

Beweis. Siehe bspw. Brockwell & Davis [5, S. 276].

Für den in Satz 3.5 beschriebenen Kalman-Glättungsalgorithmus wird noch die Bezeichnung

$$\Psi_{k|t} = \mathbf{E}\left[\left(\mathbf{X}_k - \hat{\mathbf{X}}_k\right)\left(\mathbf{X}_t - \hat{\mathbf{X}}_t\right)\right]$$

eingeführt.

Damit lautet der Kalman-Glättungsalgorithmus nun wie folgt.

Satz 3.5 (Kalman-Glättungsalgorithmus): Unter denselben Voraussetzungen wie in Satz 3.2 erhält man als Prädiktor für X_k bei gegebenen Beobachtungen Y_0, \ldots, Y_t mit k < t

$$\begin{split} \hat{X}_{k|t} &= \hat{X}_{k|t-1} + \Psi_{k|t} G_t' \Delta_t^{-1} \left(Y_t - G_t \hat{X}_{t|t-1} \right) \\ \Psi_{k|t+1} &= \Psi_{k|t} \left[F_t - \Theta_t \Delta_t^{-1} G_t \right]' \\ \Omega_{k|t} &= \Omega_{k|t-1} - \Psi_{k|t} G_t' \Delta_t^{-1} G_t \Psi_{k|t}'. \end{split}$$

Beweis. Siehe bspw. Brockwell & Davis [5, S. 277].

Zu den Sätzen 3.2, 3.4 und 3.5 sind folgende Anmerkungen zu machen:

- Sowohl der Kalman-Filteralgorithmus als auch der Kalman-Glättungsalgorithmus verwenden die aus dem Kalman-Prognosealgorithmus gewonnenen 1-Schritt-Prognosen $\hat{X}_{t|t-1}$. Aus diesem Grund findet immer zuerst der Kalman-Prognosealgorithmus Anwendung.
- Der Kalman-Glättungsalgorithmus stellt eine Rekursion dahingehend dar, dass für $\hat{X}_{k|t}$ die Prognose $\hat{X}_{k|t-1}$ benötigt wird, wobei hierfür $\hat{X}_{k|t-2}$ benötigt wird, usw. Dieser Rekursionsalgorithmus muss solange durchlaufen werden, bis $\hat{X}_{k|k-1}$ erreicht ist. Für $\hat{X}_{k|k-1}$ ist dann Satz 3.2 anzuwenden.

Abschließend soll nun noch ein Anwendungsbeispiel für die einzelnen Kalman-Algorithmen angeführt werden.

3.4 Beispiel

Um in diesem Beispiel, welches an Shumway & Stoffer [38, S. 331 ff.] angelehnt ist, einen Vergleich zwischen den aus den Kalman-Rekursionen gewonnenen Werten und den tatsächlichen Zuständen ziehen zu können, wird im Folgenden eine gegebene Zustandszeitreihe X_t (t = 1, ..., 100) durch einen stochastischen Störterm (weißes Rauschen) überlagert. Anschließend wird diese, durch den Störterm überlagerte Zeitreihe, als Beobachtungszeitreihe Y_t (t = 1, ..., 100) interpretiert und darauf die obig beschriebenen Kalman-Rekursionen angewendet.

Basierend auf diesen Annahmen, sowie auf weiteren Annahmen bzgl. den Startwerten $\hat{X}_{1|0}$ und $\Omega_{1|0}$, werden für t = 1, ..., 100 die Werte der Kalman-Prognose, Kalman-Filterung sowie der Kalman-Glättung berechnet. Die zugehörigen Varianzterme wurden ebenfalls für t = 1, ..., 100 berechnet. Die jeweiligen Ergebnisse sind in Abbildung 5 dargestellt.

In Abbildung 5 sind, unterteilt in drei Teilabbildungen, Kalman-Prognose, Kalman-Filterung und Kalman-Glättung dargestellt. In jeder Teilabbildung sind die einzelnen Beobachtungen durch ein "ד gekennzeichnet, während die geschätzten Kalman-Werte durch eine geradlinige Verbindung (dicke Linie) dargestellt sind. Die oberhalb bzw. unterhalb dieser Linie verlaufenden dünneren Linien sind die zugehörigen Konfidenzbänder. Vergleicht man die drei Teilabbildungen miteinander ist zu erkennen, dass die genannten Konfidenzbänder von der Prognose über die Filterung hin zur Glättung schmaler werden. Dies liegt darin begründet, dass in die Schätzung eines Kalman-Wertes zum Zeitpunkt $t \in \{1, ..., 100\}$ für die Prognose, die Filterung und die Glättung eine jeweils unterschiedliche Datenbasis zugrunde gelegt wird. So wird für die Prognose des Wertes zum Zeitpunkt t die Datenbasis bis zum Zeitpunkt t - 1, für die Filterung die Datenbasis bis zum Zeitpunkt t und für die Glättung die Datenbasis bis zum Endzeitpunkt benötigt. In die Kalman-Glättung gehen somit die meisten beobachteten Werte ein, während die Kalman-Prognose auf der kleinsten Datenbasis beruht. Diese Tatsache führt zu unterschiedliche breiten Konfidenzbändern.

Literatur

- [1] Ajne, B. (1994): Additivity of chain-ladder projections. *ASTIN Bulletin*, 24(2): 313–318.
- [2] Atherino, R., Pizzinga, A. & Fernandes, C. (2010): A Row-Wise Stacking of the Runoff Triangle: State Space Alternatives for IBNR Reserve Prediction. ASTIN Bulletin, 40: 917– 946.



Abbildung 5: Mittels Kalman-Prognose, Kalman-Filterung sowie Kalman-Glättung prognostizierte Zustände einer gegebenen Beobachtungszeitreihe, sowie zugehörige Standardabweichungsbänder (angelehnt an Shumway & Stoffer [38, S. 332]).

- [3] Baumann, C. (2010): Das Kalman Filter wird 50. Version 1.4. URL: http://www. convict.lu/htm/rob/Das%20Kalman%20Filter%20wird%2050_v_ 1.4.pdf (besucht am 01.05.2013).
- [4] Björkwall, S. (2011): "Stochastic claims reserving in non-life insurance: Bootstrap and smoothing models". Diss. Stockholm University, Department of Mathematics.
- [5] Brockwell, P. J. & Davis, R. A. (2002): Introduction to Time Series and Forecasting. Springer.
- [6] Brockwell, P. J. & Davis, R. A. (2006): *Time Series: Theory and Methods*. 2. Aufl. Springer series in statistics. New York, Heidelberg: Springer.
- [7] Buchwalder, M., Bühlmann, H., Merz, M. & Wüthrich, M. V. (2006): The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). *ASTIN Bulletin*, 36(2): 521–542.
- [8] Bundesamt für Privatversicherungen (2006): Swiss Solvency Test. Schweiz.
- [9] Davison, A. C. & Hinkley, D. V. (1997): *Bootstrap Methods and their Application*. Cambridge series on statistical and probabilistic mathematics. Cambridge [u.a.]: Cambridge Univ. Pr.
- [10] De Jong, P. & Zehnwirth, B. (1983): Claims reserving, state-space models and the Kalman filter. *Journal of the Institute of Actuaries*, 110: 157–181.
- [11] Efron, B. (1979): Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7: 1–26.
- [12] Efron, B. (1994): *The Jackknife, the Bootstrap and Other Resampling Plans.* 6. Aufl. Regional conference series in applied mathematics. Philadelphia, Pa.: Society for Industrial und Applied Mathematics.
- [13] Efron, B. & Tibshirani, R. J. (1993): *An Introduction to the Bootstrap.* 1. Aufl. Monographs on statistics and applied probability. Champam & Hall.
- [14] England, P. D. & Verrall, R. J. (1999): Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, 25(3): 281–293.
- [15] Europäisches Parlament und Rat der Europäischen Union (2009): Solvency II.
- [16] Garrido, J. & Romera, R. (1995): On credibility and robustness with the Kalman filter. *UC3M Working Papers. Statistics and Econometrics 1995-25-07*,
- [17] Gisler, A. (2006): The estimation error in the chain-ladder reserving method: a Bayesian approach. *ASTIN Bulletin*, 36: 554–565.
- [18] Gogol, D. (1993): Using expected loss ratios in reserving. *Insurance: Mathematics and Economics*, 12(3): 297–299.
- [19] Heberle, J., Huergo, L. & Merz, M. (2010): Bootstrapping the Chain-Ladder-Method of Several Correlated Run-Off Portfolios. *Zeitschrift für die gesamte Versicherungswissenschaft*, 98(5): 541–564.
- [20] Heberle, J., Huergo, L. & Merz, M. (2008): Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren für abhängige Run-Off-Portfolios. Zeitschrift für die gesamte Versicherungswissenschaft, 97(4): 439–461.
- [21] Hertig, J. (1985): A Statistical Approach to IBNR-Reserves in Marine Reinsurance. *ASTIN Bulletin*, 15(2): 171–183.

- [22] Kaas, R., Goovaerts, M., Dhaene, J. & Denuit, M. (2008): *Modern Actuarial Risk Theory: Using R.* 2. Aufl. Berlin: Springer.
- [23] Kálmán, R. E. (1960): A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1): 35–45.
- [24] Klenke, A. (2013): *Wahrscheinlichkeitstheorie*. 3., überarb. und erg. Aufl. Springer-Lehrbuch Masterclass. Berlin: Springer Spektrum.
- [25] Kremer, E. (1995): Robust Credibility via Robust Kalman Filtering,
- [26] Liu, H. & Verrall, R. (2010): Bootstrap Estimation of the Predictive Distributions of Reserves Using Paid and Incurred Claims. *Variance*, 4: 121–135.
- [27] Lowe, J. (1994): A practical guide to measuring reserve variability using: Bootstrapping, operational time and a distribution free approach. *Proceedings of the 1994 General Insurance Convention*.
- [28] Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2): 213–225.
- [29] Mack, T. (2002): *Schadenversicherungsmathematik*. 2., überarb. Aufl. Schriftenreihe angewandte Versicherungsmathematik. Karlsruhe: Verl. Versicherungswirtschaft.
- [30] McGee, L. A. & Schmidt, S. F. (1985): Discovery of the Kalman Filter as a Practical Tool for Aerospace and Industry. *NASA Technical Memorandum 86847*, 21.
- [31] Merz, M. & Wüthrich, M. V. (2010): Paid-incurred chain claims reserving method. *Insurance: Mathematics and Economics*, 46(3): 568–579.
- [32] Peters, G. W., Wüthrich, M. V. & Shevchenko, P. V. (2010): Chain ladder method: Bayesian bootstrap versus classical bootstrap. *Insurance: Mathematics and Economics*, 47(1): 36–51.
- [33] Pinheiro, P. J. R., Andrade e Silva, J. M. & de Lourdes Centeno, M. (2003): Bootstrap Methodology in Claim Reserving. *Journal of Risk and Insurance*, 70(4): 701–714.
- [34] Pröhl, C. & Schmidt, K. D. (2005): Multivariate chain-ladder. Dresdner Schriften zur Versicherungsmathematik, 3.
- [35] Quarg, G. & Mack, T. (2004): Munich Chain Ladder. *Blätter der DGVFM*, 26(4): 597–630.
- [36] Renshaw, A. E. & Verrall, R. J. (1998): A stochastic model underlying the chain-ladder technique. *British Actuarial Journal*, 4(4): 903–923.
- [37] Schmidt, K. D. (2009): *Versicherungsmathematik.* 3., überarb. und erw. Aufl. Springer-Lehrbuch. Berlin: Springer.
- [38] Shumway, R. H. & Stoffer, D. S. (2011): *Time Series Analysis and Its Applications: With R Examples.* 3. Aufl. Springer.
- [39] Tarbell, T. F. (1934): Incurred but not Reported Claim Reserves. *CAS 20*, 275–280.
- [40] Taylor, G. C. (2000): *Loss Reserving: An Actuarial Perspective*. Kluwer Academic Publishers, Boston.
- [41] Taylor, G. C. & McGuire, G. (2007): A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error. North American Actuarial Journal, 11(3): 70–88.

- [42] Venter, G. G. (2008): Distribution and Value of Reserves Using Paid and Incurred Triangles. *CAS E-Forum*, 348–375.
- [43] Verdonck, T., Van Wouwe, M. & Dhaene, J. (2009): A Robustification of the Chain-Ladder Method. *The North American Actuarial Journal*, 13: 280–298.
- [44] Verrall, R. J. (1989): A state space representation of the chain ladder linear model. *Journal of the Institute of Actuaries*, 116: 589–609.
- [45] Wüthrich, M. V. & Merz, M. (2008): *Stochastic Claims Reserving Methods in Insurance*. Chichester: Wiley Finance.


Bootstrapping the Chain-Ladder-Method for Several Correlated Run-Off Portfolios

Jochen Heberle, Luis Huergo, Michael Merz

Bootstrapping the Chain-Ladder-Method for Several Correlated Run-Off Portfolios

Jochen Heberle

Luis Huergo

Michael Merz

Abstract

The prediction of outstanding loss liabilities for a non-life run-off-portfolio as well as the quantification of the prediction error is one of the most important actuarial tasks in non-life insurance. In this paper we consider this prediction problem in a multivariate context. More precisely, we derive the predictive distribution of the claims reserves simultaneously for several correlated run-off-portfolios in the framework of the Chain-Ladder claims reserving method for several correlated run-off-portfolios.

1 Introduction and motivation

Claims reserving is one of the most important actuarial tasks in non-life insurance since

- often claims reserves are the *largest position* on the liability side of the balance sheet;
- the reserve risk is one of the *main risk driver* in non-life insurance and
- the total amount of the claims reserves has a *major impact* on income statements.

Therefore, given the available information about the past claims development, the prediction of adequate *claims reserves* for a non-life run-off portfolio as well as the quantification of the uncertainties in these reserves is a major task in actuarial practice and science (e.g. Rodermund [27], Teugels & Sundt [33] and Wüthrich & Merz [34]).

1.1 Predictive distribution of the claims reserves

In this paper we consider the claims reserving problem in a *multivariate context*. More precisely, we derive the *predictive distribution* of the claims reserves simultaneously for several correlated run-off portfolios in the framework of the Chain-ladder (CL) claims reserving method for several correlated run-off portfolios provided by Merz & Wüthrich [21].

The derivation of the predictive distribution of the claims reserves for one or several correlated run-off portfolios is the "holy grail" of stochastic claims reserving (cf. England & Verrall [10, p. 221]). The predictive distribution incorporates

- the *process variance* which quantifies the pure random part of our stochastic claims reserving problem and is caused by the fact that we deal with stochastic processes and
- the *estimation error* which quantifies the uncertainty from estimating the model parameters.

Unlike the process variance the estimation error decreases with the increase in the number of observations. In practical applications and solvency considerations, estimates for second moments such as the *mean square error (MSE)* and its both components process variance and estimation error are often sufficient or required by the regulator (e.g. Bundesamt für Privatversicherungen [5]). In these cases one often fits an analytic overall distribution (e.g. Lognormal distribution) to the total claims reserves by the method of moments using estimates of the expected claims reserves and the MSE as the first two moments of the fitted distribution. That is, one avoids the difficulties about marginal distributions, dependence structures and aggregation/convolution by simply making an overall choice for the resulting distribution (for an example to this approach we refer to the *Swiss Solvency Test*, see Bundesamt für Privatversicherungen [5]). Therefore, it is not surprising that the MSE is the most popular risk measure in actuarial practice and theory to quantify the uncertainties in claims reserves estimates (e.g. Wüthrich & Merz [34]).

However, besides best estimates of expected claims reserves and estimates of the MSE the predictive distribution for a portfolio of several correlated run-off portfolios provides us with additional information which is valuable for *solvency discussions, capital modelling* and *risk management considerations*. More precisely any measure on the predictive distribution can be evaluated such as

- 1. best estimates of the expected claims reserves;
- 2. estimates of the MSE;
- 3. estimates for prediction intervals, quantiles (e.g. Value-at-Risk) and more sophisticated risk measures such as the Expected-Shortfall.

Unfortunately, in most stochastic claims reserving models which have been proposed up to now one is not able to calculate the predictive distribution analytically since the distribution of the sum of random variables is required (see Wüthrich & Merz [34], Section 4.2.3 and 4.3.1 for two exceptions). Therefore, in many cases one is forced to adopt simulation techniques

such as *Bootstrapping techniques* or *Markov Chain Monte Carlo methods* to produce a simulated predictive distribution for the claims reserves.

1.2 Claims reserving for several correlated run-off portfolios

The simultaneous study of several individual run-off portfolios by means of claims reserving methods suitable for correlated run-off portfolios is motivated by several facts:

- 1. In practice it is quite natural to subdivide a non-life run-off portfolio into several correlated subportfolios, such that each subportfolio satisfies the homogeneity properties of a certain stochastic claims reserving methods (e.g. CL method or additive claims reserving method);
- 2. It addresses the problem of dependence between run-off portfolios of different lines of business (e.g. bodily injury claims in auto liability and in general liability business);
- 3. The multivariate approach has the advantage that by observing one run-off portfolio we learn about the behaviour of the other run-off portfolios (e.g. portfolio of small claims versus portfolio of large claims);
- 4. It resolves the problem of additivity. This means the estimators of the ultimate claims for the whole portfolio are obtained by summation over the estimators of the ultimate claims for the individual run-off portfolios (e.g. Ajne [1] and Klemmt [17]);
- 5. It allows for an unified approach for estimating the resulting MSE, Value-at-Risk, Expected-Shortfall etc. for the total portfolio.

In the literature claims reserving methods for portfolios of several correlated run-off portfolios have been studied by Braun [2], Schmidt [29, 28], Pröhl & Schmidt [26], Mildenhall [24], Hess et al. [12], Hürlimann [14] and Merz & Wüthrich [21, 23, 22, 20]. Simulation based approaches which extend the bootstrapping technique from a single run-off portfolio to several correlated run-off portfolios are given by Brehm [3], Kirschner et al. [16] and Taylor & McGuire [32, 31]. The derivation of the predictive distribution of the claims reserves simultaneously for several correlated run-off portfolios using bootstrapping techniques is more sophisticated than the derivation of the predictive distribution only for a single run-off portfolio. In the following we show how bootstrapping can be used to capture correlated run-off portfolios provided by Merz & Wüthrich [21] and compare our results to the analytical ones in Merz & Wüthrich [23]. The derivation of the predictive distribution for the claims reserves for several correlated run-off portfolios in the CL method for several correlated run-off portfolios provided by Merz & Wüthrich [21] and compare our results to the analytical ones in Merz & Wüthrich [23]. The derivation of the predictive distribution for the claims reserves for several correlated run-off portfolios in the CL method probably to the most popular claims reserving technique in theory and practise.

1.3 Bootstrapping in claims reserving

The bootstrap technique introduced by Efron & Tibshirani [7] is a powerful technique for obtaining information about an aggregate distribution from single samples of data. In the claims reserving literature bootstrap methods appear, for example, in Brehm [3], Taylor [30], Wüthrich & Merz [34], Kirschner et al. [16], Taylor & McGuire [32, 31], England & Verrall [9, 10], England [8], Lowe [18] or Pinheiro et al. [25] and have become very popular in actuarial practice since bootstrapping

- 1. is as well established statistical technique;
- 2. can easily be applied;
- 3. can often be derived in a spreadsheet and
- 4. combined with a simulation for the process variance the predictive distribution can be obtained.

For applying the bootstrap technique to claims reserving we only need a model structure which allows to resample observations (pseudo observations) using the samples of data. Bootstrapping can be described as "simulating from an estimated model" by resampling new pseudo observations from the data themselves and re-fitting the model to the pseudo observations. In the following we consider the CL claims reserving model suitable for several correlated run-off portfolios provided by Merz & Wüthrich [21] and use bootstrapping techniques to derive the predictive distribution of the claims reserves. For this purpose we proceed in the following two steps:

- 1. In the first step we use bootstrapping techniques to obtain the estimation error of the CL model.
- 2. In the second step we supplement the bootstrapping by a simulation approach to incorporate the process variance, giving the full predictive distribution of the claims reserves for the whole portfolio of several correlated run-off portfolios (cf. Figure 1).

2 Chain-ladder method for several correlated run-off portfolios

In the following we consider the CL claims reserving model provided by Merz & Wüthrich [21]. This claims reserving model is suitable for several correlated run-off portfolios considering development year-based correlations.



Figure 1: Obtaining the predictive distribution using bootstrapping techniques.

2.1 Notation and multivariate framework

For ease of exposition, we assume that the data of the run-off portfolios consist of $N \ge 1$ *run-off triangles* of observations of the same size. However, the following described CL model can also be applied to other shapes of data (e.g. run-off trapezoids). In these N triangles the indices

 $n, 1 \le n \le N,$ refer to run-off portfolios (run-off triangles), $i, 0 \le i \le I,$ refer to accident years (rows), $j, 0 \le j \le J = I,$ refer to development years (columns).

The cumulative claims (i.e. cumulative payments, claims incurred or total number of reported claims) are given by

$$C_{i,j}^{(n)}$$
 for $n = 1, ..., N$ and $i, j = 0, ..., I$.

This means that $C_{i,j}^{(n)}$ denotes the cumulative claim of all claims up to development year *j* occurred in the accident year *i* in the *n*-th run-off portfolio.

Usually, at time *I* (*calendar year I*), we have a total of observations over all portfolios given by

$$\mathcal{D}_{I}^{N} = \left\{ C_{i,j}^{(n)} \mid i \leq I, \ i+j \leq I, \ 1 \leq n \leq N \right\}$$

and we need to predict the random variables in its complement

$$\mathcal{D}_I^{N,c} = \left\{ C_{i,j}^{(n)} \mid i \leq I, \ i+j > I, \ 1 \leq n \leq N \right\}.$$

Figure 2 shows the claims data structure for the N run-off portfolios described above.



Figure 2: Claims development of run-off triangle $n \in \{1, ..., N\}$.

2.2 Chain-ladder method and time-series model

Often the CL method is understood as a purely mechanical algorithm to derive predictions $\widehat{C}_{i,J}^{(n)}$ for the ultimate claims $C_{i,J}^{(n)}$. For run-off triangle $n \in \{1, \ldots, N\}$ the algorithmic definition of the CL method at time *I* reads as follows:

1. There are constants $f_i^{(n)}$ (j = 0, ..., J - 1) so that for all *i* and $k \ge I - i$

$$\widehat{C}_{i,k}^{(n)} = C_{i,I-i}^{(n)} f_{I-i}^{(n)} f_{I-i+1}^{(n)} \dots f_{k-1}^{(n)}$$

with $\widehat{C}_{i,l-i}^{(n)} = C_{i,l-i}^{(n)}$ is an appropriate predictor for $C_{i,k}^{(n)}$.

2. The Chain-ladder factors (age-to-age factors) $f_i^{(n)}$ are estimated by

$$\widehat{f}_{j}^{(n)} = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}} F_{i,j}^{(n)},$$
(2.1)

where

$$F_{i,j}^{(n)} = \frac{C_{i,j+1}^{(n)}}{C_{i,j}^{(n)}}.$$

The quotients $F_{i,j}^{(n)}$ are called *individual development factors*. It is Mack's merit (cf. Mack [19]) that he first gave a *distribution-free* stochastic model which satisfies the CL algorithm, and for which he estimated the MSE for a single run-off portfolio.

The *claims reserves* for single accident years $i \in \{1, ..., I\}$ are estimated by

$$\widehat{R}_{i}^{(n)} = \widehat{C}_{i,J}^{(n)} - C_{i,I-i}$$
 and $\widehat{R}_{i} = \sum_{n=1}^{N} \widehat{R}_{i}^{(n)}$

for single run-off portfolios $n \in \{1, ..., N\}$ and the total run-off portfolio, respectively. The claims reserves estimates for aggregated accident years are given by

$$\widehat{R}^{(n)} = \sum_{i=1}^{I} \widehat{R}_{i}^{(n)}$$
 and $\widehat{R} = \sum_{n=1}^{N} \widehat{R}^{(n)}$

for single run-off portfolios $n \in \{1, ..., N\}$ and the total run-off portfolio, respectively.

In this paper, we consider the following CL consistent time series framework for *N* correlated run-off triangles provided by Merz & Wüthrich [21]:

Model Assumptions 2.1 (CL consistent time series model for N triangles):

- Cumulative claims of different accident years *i* are independent.
- There exist constants $f_j^{(n)}, \sigma_j^{(n)} > 0$ and random variables $\epsilon_{i,j+1}^{(n)}$ such that for all $i \in \{0, \ldots, I\}, j \in \{0, \ldots, J-1\}$ and $n = 1, \ldots, N$ we have

$$C_{i,j+1}^{(n)} = f_j^{(n)} C_{i,j}^{(n)} + \sigma_j^{(n)} \sqrt{C_{i,j}^{(n)}} \epsilon_{i,j+1}^{(n)}.$$
(2.2)

• The random variables $\epsilon_{i,j+1}^{(n)}$ and $\epsilon_{k,l+1}^{(m)}$ are independent if $i \neq k$ or $j \neq l$ and it holds

$$\mathbb{E}\left[\left.\epsilon_{i,j}^{(n)} \mid \mathcal{B}_{0}^{N}\right] = 0, \qquad \text{Var}\left(\left.\epsilon_{i,j}^{(n)} \mid \mathcal{B}_{0}^{N}\right) = 1 \qquad \text{and} \qquad P\left(C_{i,j+1}^{(n)} > 0 \mid \mathcal{B}_{0}^{N}\right) = 1$$

with $\mathcal{B}_0^N = \left\{ C_{0,0}^{(n)}, \dots, C_{I,0}^{(n)} \mid n = 1, \dots, N \right\}$ for all $i \in \{0, \dots, I\}, j \in \{0, \dots, J-1\}$ and $n = 1, \dots, N$.

• The *N*-dimensional random variables $\boldsymbol{\epsilon}_{i,j} = \left(\epsilon_{i,j}^{(1)}, \dots, \epsilon_{i,j}^{(N)}\right)^T$ have the (conditionally)

correlation-matrices

$$\Sigma_{j} = \operatorname{Corr}\left(\boldsymbol{\epsilon}_{i,j} \mid \mathcal{B}_{0}^{N}\right) = \begin{pmatrix} 1 & \rho_{j}^{(1,2)} & \cdots & \rho_{j}^{(1,N)} \\ \rho_{j}^{(2,1)} & 1 & \cdots & \rho_{j}^{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{j}^{(N,1)} & \rho_{j}^{(N,2)} & \cdots & 1 \end{pmatrix},$$
(2.3)

where $\rho_j^{(n,m)} \in (-1,1)$ for $n, m \in \{1, ..., N\}$ and $n \neq m$. Since Var $\left(\epsilon_{i,j}^{(n)} \mid \mathcal{B}_0^N\right) = 1$ holds true the correlation-matrices Σ_j are also covariance-matrices of $\epsilon_{i,j}$, i.e.:

$$\Sigma_{j} = \operatorname{Corr}\left(\boldsymbol{\epsilon}_{i,j} \mid \mathcal{B}_{0}^{N}\right) = \operatorname{Cov}\left(\boldsymbol{\epsilon}_{i,j} \mid \mathcal{B}_{0}^{N}\right)$$
(2.4)

Remarks 2.2:

- The Time Series Model 2.1 defines autoregressive processes. It dictates the definition used for the residuals in the bootstrap and it provides a mechanism for generating sets of "pseudo" observations by means of these residuals.
- The random variables $\epsilon_{i,j+1}$ are defined conditionally, given \mathcal{B}_0^N , in order to ensure that the cumulative claims $C_{i,j+1}^{(n)}$ stay positive, $P[\cdot | \mathcal{B}_0^N]$ -a.s. This implies that all the derivations and statements are done under the conditional probability measure $P[\cdot | \mathcal{B}_0^N]$. In order to not overload the notation this conditional probability measure is abbreviated by P.
- Within the CL framework for several correlated run-off portfolios Braun [2] and Merz & Wüthrich [21, 23] proposed the development year-based correlations given by (2.3). Often correlations between different run-off triangles are attributed to claims inflation. Under this point of view it may seem more reasonable to allow for correlation between the cumulative claims of the same calender year (diagonals of the run-off triangles). This would introduce dependencies between accident years and, therefore, contradict the CL model assumption of independent accident years. That is, all calender year-based dependencies should be removed from the data before calculating the claims reserves with the CL method. After correcting the data for the calender year-based correlations, further direct and indirect sources for correlations between different run-off triangles of a portfolio exist and should be taken into account (cf. Houltram [13]).

We immediately obtain the following corollary:

Corollary 2.3: Under Model Assumptions 2.1, $(C_{i,j}^{(n)})_{i\geq 1}$ are Markov chains with

- 1. different accident years of one or of distinct triangles are independent,
- 2. E $\left[C_{i,j}^{(n)} \mid C_{i,j-1}^{(n)}\right] = f_{j-1}^{(n)} C_{i,j-1}^{(n)},$

3. Var $\left[C_{i,j}^{(n)} \mid C_{i,j-1}^{(n)}\right] = \left(\sigma_{j-1}^{(n)}\right)^2 C_{i,j-1}^{(n)}$ 4. Cov $\left[C_{i,j}^{(n)}, C_{k,j}^{(m)} \mid C_{i,j-1}^{(n)}, C_{k,j-1}^{(m)}\right] = \sigma_{j-1}^{(n)} \sigma_{j-1}^{(m)} \sqrt{C_{i,j-1}^{(n)} C_{i,j-1}^{(m)}} \rho_j^{(n,m)} \mathbf{1}_{\{i=k\}}.$

Observe the properties 1.-3. are the classical Mack [19] conditions extended to several run-off triangles.

Lemma 2.4: Under Model Assumptions 2.1 we have that $\widehat{f}_l^{(n)}$ is an unbiased estimator for $f_l^{(n)}$. Moreover $\widehat{f}_l^{(n)}$ and $\widehat{f}_k^{(m)}$ are uncorrelated for $l \neq k$, and given $C_{i,l-i}^{(n)}$

$$\widehat{C}_{i,J}^{(n)} = C_{i,I-i}^{(n)} \ \widehat{f}_{I-i}^{(n)} \ \widehat{f}_{I-i+1}^{(n)} \ \dots \ \widehat{f}_{J-1}^{(n)}$$

is an unbiased estimator for $E\left[C_{i,J}^{(n)} \mid C_{i,I-i}^{(n)}\right]$.

Proof. The proof is similar to the one of Theorems 1 and 2 in Mack [19]. \Box

Appropriate estimators for the variance and correlation parameters $\sigma_j^{(n)}$ and $\rho_j^{(n,m)}$ are given by

$$\widehat{\sigma}_{j}^{(n)} = \sqrt{\frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j}^{(n)} \left(F_{i,j}^{(n)} - \widehat{f}_{j}^{(n)}\right)^{2}}$$
(2.5)

and

$$\widehat{\rho}_{j}^{(n,m)} = c_{j+1} \sum_{i=0}^{I-j-1} \frac{\sqrt{C_{i,j}^{(n)} C_{i,j}^{(m)}}}{\widehat{\sigma}_{j}^{(n)} \widehat{\sigma}_{j}^{(m)}} \left(F_{i,j}^{(n)} - \widehat{f}_{j}^{(n)}\right) \left(F_{i,j}^{(n)} - \widehat{f}_{j-1}^{(m)}\right)$$
(2.6)

for $j \in \{0, \ldots, J-1\}$ and $n, m \in \{1, \ldots, N\}$ with $n \neq m$ and $\widehat{\rho}_j^{(n,m)} = 1$ for n = m, where

$$c_{j+1} = \frac{1}{I - j - 2 + w_{j+1}^2} \quad \text{and} \quad w_{j+1}^2 = \frac{\left(\sum\limits_{i=0}^{I - j - 1} \sqrt{C_{i,j}^{(n)} C_{i,j}^{(m)}}\right)^2}{\sum\limits_{i=0}^{I - j - 1} C_{i,j}^{(n)} \sum\limits_{i=0}^{I - j - 1} C_{i,j}^{(m)}}.$$

For the properties of these estimators see Wüthrich & Merz [34]. However, we are only able to estimate $\sigma_{J-1}^{(n)}$ and $\rho_{J-1}^{(n,m)}$ for $n \neq m$ from the data with formula (2.5) and (2.6), respectively, if I > J. This is due to the lack of data in the tails. In the case I = J we simply extrapolate the last parameters $\hat{\sigma}_{J-1}^{(n)}$ and $\hat{\rho}_{J-1}^{(n,m)}$ for $n \neq m$ by exponentially decreasing series (cf. Mack [19])

$$\widehat{\sigma}_{J-1}^{(n)} = \min\left\{ \left(\widehat{\sigma}_{J-2}^{(n)} \right)^2 \middle| \widehat{\sigma}_{J-3}^{(n)} , \, \widehat{\sigma}_{J-3}^{(n)} , \, \widehat{\sigma}_{J-2}^{(n)} \right\}$$
(2.7)

and

$$\widehat{\rho}_{J-1}^{(n,m)} = \frac{\min\left\{\frac{\left(\widehat{\rho}_{J-2}^{(n,m)} \ \widehat{\sigma}_{J-2}^{(n)} \ \widehat{\sigma}_{J-2}^{(m)}\right)^{2}}{\left|\widehat{\rho}_{J-3}^{(n,m)} \ \widehat{\sigma}_{J-3}^{(m)} \ \widehat{\sigma}_{J-3}^{(m)}\right|}, \ \left|\widehat{\rho}_{J-2}^{(n,m)} \ \widehat{\sigma}_{J-2}^{(m)} \ \widehat{\sigma}_{J-2}^{(m)} \ \widehat{\sigma}_{J-2}^{(m)} \ \widehat{\sigma}_{J-3}^{(m)} \ \widehat{\sigma}_{J-3}^{(m)} \ \widehat{\sigma}_{J-3}^{(m)}\right|}\right\}}{\widehat{\sigma}_{J-1}^{(n)} \ \widehat{\sigma}_{J-1}^{(m)}} .$$
(2.8)

For our purpose, i.e. to demonstrate the bootstrapping technique for several correlated run-off triangles, these estimators are sufficient. However, in practice it is a serious issue choosing appropriate tail parameters, which needs a deep discussion that we do not want to enter here.

3 Derivation of the predictive distribution of the CL model for several correlated run-off portfolios

In this section we carry out in two steps the derivation of the predictive distribution for several correlated run-off portfolios in the framework of the CL consistent Time Series Model 2.1.

3.1 Step 1: Bootstrapping the Chain-ladder method for the derivation of the estimation error

In this subsection we carry out step 1 of the derivation of the predictive distribution (cf. Figure 1). We derive the estimation error in the CL method for several correlated run-off portfolios by bootstrapping.

In the application of bootstrap techniques it is essential that the observations are the realizations of independent and identically distributed random variables. However, it is obvious that this does not hold for the claims sizes in a run-off triangle. Therefore, in the following we does not bootstrap the claims data themselves but residuals which are consistent with the model being fitted (cf. England & Verrall [10]). The autoregressive structure (2.2) dictates the definition of residuals which are consistent with the CL model. Together with the parameter estimators (2.1), (2.5)-(2.7) and using the data in the original run-off triangles we obtain for every individual run-off triangle $n \in \{1, ..., N\}$ the $K = \frac{1}{2}I(I + 1)$ residuals given by

$$r_{i,j+1}^{(n)} = \frac{\sqrt{C_{i,j}^{(n)}} \left(F_{i,j}^{(n)} - \hat{f}_j^{(n)}\right)}{\widehat{\sigma}_i^{(n)}}$$
(3.1)

for $i \in \{0, ..., I-1\}$ and $j \in \{0, ..., J-1\}$ with $i + j \le I - 1$. Note that these residuals $r_{i,j+1}^{(n)}$ are observable, whereas the random variables $\epsilon_{i,j+1}^{(n)}$ are not observable for unknown CL factors $f_i^{(n)}$. The scaled residuals are then computed as

$$\overline{r}_{i,j+1}^{(n)} = rac{r_{i,j+1}^{(n)} - \mu_{j+1}}{\sigma_r},$$

where

$$\mu_{j+1} = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} r_{i,j+1}^{(n)}$$
 for $j = 0, \dots, J-1$

and

$$\sigma_r = \sqrt{\frac{1}{K-1} \sum_{i=0}^{I-1} \sum_{j=0}^{J-2} \left(r_{i,j+1}^{(n)} - \mu_{j+1}\right)^2}.$$

These residuals are known as *scaled Pearson residuals* and for a fixed $n \in \{1, ..., N\}$ the set

$$\overline{r}^{(n)} = \left\{ \overline{r}_{i,j+1}^{(n)} \mid i = 0, \dots, I-1 \text{ and } j = 0, \dots, J-1 \text{ with } i+j \le I-1 \right\}$$

of all scaled Pearson residuals constitutes the *bootstrap distribution* $G^{(n)}$ for the run-off triangle $n \in \{1, ..., N\}$. We expect that approximately 95% of the scaled Pearson residuals from $\overline{r}^{(n)}$ lie in the interval (-2, 2).

3.1.1 Bootstrapping without considering correlations

Given the set $\overline{r}^{(n)}$ of the *K* scaled Pearson residuals $\overline{r}_{i,j+1}^{(n)}$ for every run-off triangle $n \in \{1, ..., N\}$ we resample *K* residuals by sampling with replacement from $\overline{r}^{(n)}$ for every run-off triangle *n*. Since we do not consider correlations at the moment we do this independent from the other run-off triangles $k \neq n$, i.e. we consider the random variables

$$\widetilde{\overline{r}}_{i,j+1}^{(n)} \sim G^{(n)} \quad \text{for } n \in \{1, \dots, N\},\tag{3.2}$$

where $\tilde{r}_{i,j+1}^{(n)}$ and $\tilde{r}_{k,l+1}^{(k)}$ are independent for $n \neq k$. From (3.1) and (3.2) it is straightforward to generate *pseudo/bootstrap observations* for the individual development factors $F_{i,j}^{(n)}$ in run-off triangle $n \in \{1, \ldots, N\}$, which are consistent with the CL model, i.e.

$$\widetilde{F}_{i,j}^{(n)} = \widehat{f}_j^{(n)} + \frac{\widehat{\sigma}_j^{(n)}}{\sqrt{C_{i,j}^{(n)}}} \, \overline{\widetilde{r}}_{i,j+1}^{(n)} \tag{3.3}$$

for $i \in \{0, ..., I - 1\}$ and $j \in \{0, ..., J - 1\}$ with $i + j \leq I - 1$. It is important to note that, given the total of observation \mathcal{D}_{I}^{N} , the bootstrap individual development factors $\widetilde{F}_{i,j}^{(n)}$ and $\widetilde{F}_{i,j}^{(k)}$ are independent for $n \neq k$. Using (2.1), these pseudo observations $\widetilde{F}_{i,j}^{(n)}$ lead to bootstrap CL estimates

$$\widetilde{f}_{j}^{(n)} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}^{(n)}}{\sum\limits_{i=0}^{I-j-1} C_{i,j}^{(n)}} \widetilde{F}_{i,j}^{(n)},$$
(3.4)

of the CL factors $f_i^{(n)}$ and the bootstrap CL estimates for the cumulative claims are given by

$$\widetilde{C}_{i,j}^{(n)} = C_{i,I-i}^{(n)} \, \widetilde{f}_{I-i}^{(n)} \, \dots \, \widetilde{f}_{j-1}^{(n)} \quad \text{for } I - i \le j \le J$$
(3.5)

with $\widetilde{C}_{i,I-i}^{(n)} = C_{i,I-i}^{(n)}$. Finally, we obtain for a individual run-off triangle $n \in \{1, ..., N\}$ the bootstrap claims reserves

$$\widetilde{R}_{i}^{(n)} = \widetilde{C}_{i,J}^{(n)} - C_{i,I-i}^{(n)}$$
$$\widetilde{R}^{(n)} = \sum_{i=1}^{I} \widetilde{R}_{i}^{(n)}$$

for single and aggregated accident years, respectively. The bootstrap claims reserves for the entire run-off portfolio without considering correlations are given by

$$\widetilde{R}_{i} = \sum_{n=1}^{N} \widetilde{R}_{i}^{(n)}$$
$$\widetilde{R} = \sum_{i=1}^{I} \widetilde{R}_{i}.$$

Repeating this bootstrap sample many times we obtain the bootstrap distribution of the claims reserves for single and aggregated accident years as well as for single run-off portfolios $n \in \{1, ..., N\}$ and the whole run-off portfolio. Note, that the bootstrap distribution for the claims reserves captures only the estimation error but not the process variance.

Remarks 3.1:

• The described bootstrapping approach is equivalent to construct the *n*-th run-off triangle of pseudo cumulative claims by

$$\widetilde{C}_{i,0} = C_{i,0} \quad \text{for } i = 0, \dots, I$$
$$\widetilde{C}_{i,j+1}^{(n)} = \widehat{f}_j^{(n)} C_{i,j}^{(n)} + \widehat{\sigma}_j^{(n)} \sqrt{C_{i,j}^{(n)}} \widetilde{\overline{r}}_{i,j+1}^{(n)}$$

for $0 \le j \le J - 1$ with $0 \le i + j \le I - 1$, where the actual cumulative claims $C_{i,j}^{(n)}$ act as volume measure, and to use instead of (3.4) the bootstrapped CL estimates

$$\widetilde{f}_{j}^{(n)} = \frac{\sum_{i=0}^{I-j-1} \widetilde{C}_{i,j+1}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}} \quad \text{for } j = 0, \dots, J-1.$$

This approach corresponds to the conditional resampling approach described in Buchwalder et al. [4] and Merz & Wüthrich [21].

• The derived bootstrap distribution of the parameter estimates and cumulative claims can be used to obtain a bootstrap estimate of the standard error of the parameters (i.e. calculating the standard error of the bootstrap distribution of each parameter in turn), a bootstrap estimate of the covariance matrix of the parameters and a bootstrap estimate of the standard error of the fitted values such as ultimate claims and claims reserves.

3.1.2 Bootstrapping considering correlations

In the following we respect the development year-based correlation structure (2.3) between the N run-off portfolios. Therefore, we consider the N-dimensional random variables

$$\widetilde{\overline{\mathbf{r}}}_{i,j} = \left(\widetilde{\overline{r}}_{i,j}^{(1)}, \dots, \widetilde{\overline{r}}_{i,j}^{(N)}\right)^T \sim \mathbf{G}(\mathbf{0}, \widehat{\Sigma}_{i,j}^*)$$

for $i \in \{0, ..., I - 1\}$ and $j \in \{1, ..., J\}$ with $i + j \leq I$, where $\tilde{\overline{r}}_{i,j}^{(n)}$ is given by (3.2). As seen in Section 3.1.1 the scaled Pearson residuals $\tilde{\overline{r}}_{i,j}^{(1)}, ..., \tilde{\overline{r}}_{i,j}^{(N)}$ are, except for sample effects $\hat{\Sigma}_{i,j}^* = \operatorname{Corr}\left(\tilde{\overline{r}}_{i,j} \mid \mathcal{B}_0^N\right)$ uncorrelated with expectation 0 and variance 1. Thereby, $\hat{\Sigma}_{i,j}^* = \operatorname{Corr}\left(\tilde{\overline{r}}_{i,j} \mid \mathcal{B}_0^N\right) = \operatorname{Cov}\left(\tilde{\overline{r}}_{i,j} \mid \mathcal{B}_0^N\right)$ holds true. In order to induce the desired correlation structure Σ_j , we transform the residuals $\tilde{\overline{r}}_{i,j}$ (possibly allowing for an actual sample covariance matrix $\hat{\Sigma}_{i,j}^*$) by appropriate $N \times N$ -matrices $M_{i,j}$ so that we obtain

$$\operatorname{Corr}\left(M_{i,j}\,\widetilde{\widetilde{\mathbf{r}}}_{i,j}\right) = \Sigma_{j} = \begin{pmatrix} 1 & \rho_{j}^{(1,2)} & \cdots & \rho_{j}^{(1,N)} \\ \rho_{j}^{(2,1)} & 1 & \cdots & \rho_{j}^{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{j}^{(N,1)} & \rho_{j}^{(N,2)} & \cdots & 1 \end{pmatrix}$$

for $i \in \{0, ..., I-1\}$ and $j \in \{1, ..., J\}$ with $i + j \le I$, where $\rho_j^{(n,m)} \in (-1, 1)$ for $n, m \in \{1, ..., N\}$ and $n \ne m$. Theses matrices $M_{i,j}$ are given by (cf. Heberle et al. [11])

$$M_{i,j} = B_j D_j^{1/2} D_{i,j}^{* - 1/2} B_{i,j}^{* T}$$
(3.6)

where

- D_i diagonal matrix of eigenvalues of Σ_i ,
- $D_{i,j}^*$ diagonal matrix of eigenvalues of $\hat{\Sigma}_{i,j}^*$,
- B_j matrix of orthogonal and normalized eigenvectors of Σ_j ,
- $\mathbf{B}^*_{i,j}$ matrix of orthogonal and normalized eigenvectors of $\hat{\Sigma}^*_{i,j}$.

This means we consider the transformed N-dimensional random variables

$$\begin{aligned} \mathbf{Y}_{i,j} &= \mathbf{M}_{i,j} \, \widetilde{\mathbf{\bar{r}}}_{i,j} \\ &= \mathbf{B}_j \, \mathbf{D}_j^{1/2} \, \mathbf{D}_{i,j}^{* \ -1/2} \, \mathbf{B}_{i,j}^{* \ T} \, \widetilde{\mathbf{\bar{r}}}_{i,j} \end{aligned}$$

and obtain

$$Cov \left(\mathbf{Y}_{i,j} \mid \mathcal{B}_{0}^{N}\right) = \mathbf{M}_{i,j} \ Cov \left(\widetilde{\mathbf{r}}_{i,j} \mid \mathcal{B}_{0}^{N}\right) \ \mathbf{M}_{i,j}^{T}$$

$$= \mathbf{M}_{i,j} \ \widehat{\boldsymbol{\Sigma}}_{i,j}^{*} \ \mathbf{M}_{i,j}^{T}$$

$$= \mathbf{B}_{j} \ \mathbf{D}_{j}^{1/2} \ \mathbf{D}_{i,j}^{* - 1/2} \ \mathbf{B}_{i,j}^{* T} \ \widehat{\boldsymbol{\Sigma}}_{i,j}^{*} \ \left(\mathbf{B}_{j} \ \mathbf{D}_{j}^{1/2} \ \mathbf{D}_{i,j}^{* - 1/2} \ \mathbf{B}_{i,j}^{* T}\right)^{T}$$

$$= \mathbf{B}_{j} \ \mathbf{D}_{j}^{1/2} \ \mathbf{D}_{i,j}^{* - 1/2} \ \mathbf{B}_{i,j}^{* T} \ \widehat{\boldsymbol{\Sigma}}_{i,j}^{*} \ \mathbf{B}_{i,j}^{*} \ \mathbf{D}_{i,j}^{* - 1/2} \ \mathbf{D}_{j}^{1/2} \ \mathbf{B}_{j}^{T}.$$

By means of the well-known spectral decomposition of a positive definite symmetric matrix (cf. Zurmühl & Falk [35, p. 192]) we obtain

$$\mathbf{B}_{i,j}^{* T} \, \widehat{\boldsymbol{\Sigma}}_{i,j}^{*} \, \mathbf{B}_{i,j}^{*} = \mathbf{D}_{i,j}^{*}$$
$$\mathbf{B}_{j} \, \mathbf{D}_{j} \, \mathbf{B}_{j}^{T} = \boldsymbol{\Sigma}_{j}.$$

This implies

$$\operatorname{Cov}\left(\mathbf{Y}_{i,j} \mid \mathcal{B}_0^N\right) = \Sigma_j$$

and for this reason (cf. (2.4))

$$\operatorname{Corr}\left(\mathbf{Y}_{i,j} \mid \mathcal{B}_0^N\right) = \Sigma_j.$$

This means the transformed residuals $Y_{i,j} = M_{i,j} \tilde{\overline{r}}_{i,j}$ feature the desired development year-based correlation structure Σ_j (cf. (2.4)).

The columnwise transformation procedure for two development triangles is displayed in Figure 3 (the bar and tilde symbols have been omitted in order not to overload the graphic). For example the pseudo-observations $\tilde{\vec{r}}_{i,t}^{(1)}$ and $\tilde{\vec{r}}_{i,t}^{(2)}$ for the development years t = 1, 2 have the target covariance matrices

$$\Sigma_1 = \begin{bmatrix} 1 & \rho_1^{(1,2)} \\ \rho_1^{(2,1)} & 1 \end{bmatrix}$$

and

$$\Sigma_2 = \begin{bmatrix} 1 & \rho_2^{(1,2)} \\ \rho_2^{(2,1)} & 1 \end{bmatrix},$$

respectively. These matrices can be estimated with formulas (2.7)-(2.8).



Figure 3: Correlated pseudo-observations for the development years t = 1 and t = 2.

Computational issues: Theoretically, the procedure described above should be carried out for every column $j \in \{1, ..., J\}$. However, due to some nontrivial computational issues related to (2.7)-(2.8), we assume that the correlation matrices Σ_j are the same for all development years j. This simplifying assumption has many advantages: On the one hand, it is possible to gather all the pseudo-observations for a development triangle in one column vector $\tilde{\vec{r}}^{(n)}$ for $n \in \{1, ..., N\}$ and all the vectors in a matrix R which can be transformed by (3.6). On the other hand, the correlation parameter $\rho^{(n,m)}$ can be estimated by the sample correlation coefficient of the Pearson residuals without the extrapolation methods related to small sample sizes. The simplified Matrix of residuals is displayed in Figure 4.

From a theoretical point of view, however, the column wise transformation of the pseudoobservations is preferable since it completely mirrors the CL-consistent theoretical model. First development triangle

Second development triangle



Figure 4: Simplified matrix of the resampled residuals.

It rests to check whether such a simplifying assumption is able to mimick the results of the analytical formula for the MSE derived in Merz & Wüthrich [21]. In order to decide whether this simplifying assumption is tenable, extensive simulations have been carried out. The surprising result of these simulations show that, within the expectable accuracy, both methods deliver comparable results and succeed in mimicking the results of the analytical formula for the MSE derived in Merz & Wüthrich [21]. Since the computational issues related to the more elaborated column wise transformation of the pseudo-observations are non trivial, we will describe the steps of the simplified bootstrap.

For the sake of completeness, it must be pointed out that alternative methods to induce a desired correlation structure of the resampled residuals have been suggested by the literature. The most important of them are

 Restricted Pairing: This method was originally proposed by Iman & Conover [15] for the generation of artificial samples with a desired (rank) correlation structure and is often related to the method of Latin Hypercube sampling (see Dandekar et al. [6]). This method bases on the columnwise rearrangement of the observations, such that the new pairs match a desired *rank* correlation. A modification of the method was proposed by Kirschner et al. [16] for bootstrap purposes in claims reserving. 2. Sampling across triangles: Instead of sampling residuals, their positions are sampled and the elements of every development triangle corresponding to the sampled positions are used to create new triangles. Since the corresponding elements of every triangle are kept together, the original correlation structure should remain unchanged throughout the replications.

The results of our simulations show that, for actual sets of data, the direct transformation of the residuals performs as well as the sampling across triangles (method 2.) and considerably better than the rank method (method 1.).

In spite of the simplicity of the sampling across triangles method, the direct transformation of the residuals has the advantage that it can be used with no modifications for the simulation of the process variance. For this reason we concentrate on the latter method.

Implementation of the algorithm:

- 1. Collect in a vector **a** the positions of the observations, except for the first column¹. Repeat this for each of the N run-off triangles.
- 2. Resample the vector of positions **a** independently for every run-off triangle.
- 3. Generate a vector with the residuals corresponding to the resampled positions (vector $\tilde{\mathbf{r}}^{(n)}$). Repeat this for each of the *N* run-off triangles.
- 4. Collect the *N* vectors $\tilde{\vec{r}}^{(n)}$ in a matrix of residuals denoted by R. The dimension of the matrix R is given by *dimension of vector* dim(a) × *N*.
- 5. Transform the matrix R according to transformation (3.6).
- 6. By means of vector **a** construct new run-off triangles with the transformed residuals.
- 7. Construct pseudo-observations $\widetilde{C}_{i,j}^{(n)}$ according to (3.3)-(3.5).

3.2 Step 2: Sampling approach for including the process variance

If we are only interested in the standard deviation of the CL estimates $\widehat{f}_{j}^{(n)}$ and in the estimation error in the best estimates for the ultimate claims and claims reserves we could stop now and calculate the standard deviation of the bootstrap CL estimates $\widetilde{f}_{j}^{(n)}$ and project forward from the latest observed cumulative claims $C_{i,I-i}^{(n)}$ to the ultimate claims $\widetilde{C}_{i,J}^{(n)}$ using (3.5). However, our aim is a little bit more ambitious and we want to obtain a predictive distribution for the claims reserves. Therefore, in this subsection we carry out Step 2 of the derivation of the predictive

¹The first development year has per definition no residuals.

distribution (cf. Figure 1). This means that we combine the first step with an additional step to incorporate the process error in the CL method for several run-off portfolios. Analogously to Step 1, thereby we do this without and with considering correlations. The resulting predictive distribution can be used to forecast into the future.

England & Verrall [9, 10], England [8] and Pinheiro et al. [25] used the two-step bootstrap technique described above to derive the predictive distribution of the claims reserves for a individual run-off portfolio in the framework of different univariate claims reserving methods such as the over-dispersed Poisson model, the Lognormal model, the Gamma model, generalized linear models and the univariate CL model (see Wüthrich & Merz [34] for a comprehensive description of these models).

Analogously to Step 1 the autoregressive structure (2.2) dictates how the forecast for the not observable cumulative claims in the lower run-off triangles can be obtained for each run-off triangle $n \in \{1, ..., N\}$ and bootstrap iteration. By applying the bootstrap CL estimates $\tilde{f}_j^{(n)}$ and the variance parameter estimates (2.5) to the latest observed cumulative claims on the diagonal, $C_{i,I-i}^{(n)}$, and the previous simulated forecast, $\hat{C}_{i,j}^{*(n)}$, we obtain the one-step-ahead forecasts

$$\widehat{C}_{i,I-i+1}^{*(n)} = \widetilde{f}_{I-i}^{(n)} C_{i,I-i}^{(n)} + \widehat{\sigma}_{I-i}^{(n)} \sqrt{C_{i,I-i}^{(n)}} \epsilon_{i,I-i+1}^{*(n)}$$
(3.7)

for i = 1, ..., I as well as the two-step-ahead forecasts and beyond

$$\widehat{C}_{i,j+1}^{*(n)} = \widetilde{f}_{j}^{(n)} \ \widehat{C}_{i,j}^{*(n)} + \widehat{\sigma}_{j}^{(n)} \ \sqrt{\widehat{C}_{i,j}^{*(n)}} \ \epsilon_{i,j+1}^{*(n)}$$
(3.8)

for i = 2, ..., I and j = I - i + 1, ..., J - 1. Analogously to Time Series Model 2.1, we assume for the random variables $\epsilon_{i,j}^{*(n)}$:

 ε^{*(n)}_{i,j} and ε^{*(n)}_{k,l} are independent if i ≠ k or j ≠ l,
 E [ε^{*(n)}_{i,j}] = 0 and Var (ε^{*(n)}_{i,j}) = 1,
 ε^{*}_{i,j} = (ε^{*(1)}_{i,j},...,ε^{*(N)}_{i,j}) and ε_{i,j} = (ε⁽¹⁾_{i,j},...,ε^(N)_{i,j}) have the same correlation matrix Σ_j, i.e. Corr (ε^{*}_{i,j},ε^{*}_{i,j}) = Σ_j

for i = 0, ..., I and j = 1, ..., J (cf. (2.3)).

We obtain the following corollary which is the analogon of Corollary 2.3:

Corollary 3.2: Under assumptions 1.-3. and the assumption of known parameter estimates $\tilde{f}_j^{(n)}$ and $\hat{\sigma}_j^{(n)}$, we have

$$1. \quad \mathbb{E} \left[\widehat{C}_{i,l-i+1}^{*(n)} \mid C_{i,l-i}^{(n)} \right] = \widetilde{f}_{l-i}^{(n)} C_{i,l-i}^{(n)} \text{ and } \mathbb{E} \left[\widehat{C}_{i,j+1}^{*(n)} \mid \widehat{C}_{i,j}^{*(n)} \right] = \widetilde{f}_{j}^{(n)} \widehat{C}_{i,j}^{*(n)}$$

$$2. \quad \text{Var} \left[\widehat{C}_{i,l-i+1}^{*(n)} \mid C_{i,l-1}^{(n)} \right] = \left(\widehat{\sigma}_{l-i}^{(n)} \right)^2 C_{i,l-i}^{(n)} \text{ and } \text{Var} \left[\widehat{C}_{i,j+1}^{*(n)} \mid \widehat{C}_{i,j}^{*(n)} \right] = \left(\widehat{\sigma}_{j}^{(n)} \right)^2 \widehat{C}_{i,j}^{*(n)}$$

$$3. \quad \text{Cov} \left[\widehat{C}_{i,l-i+1}^{*(n)}, \widehat{C}_{k,l-i+1}^{*(m)} \mid C_{i,l-i}^{(n)}, C_{k,l-i}^{(m)} \right] = \widehat{\sigma}_{l-i}^{(n)} \widehat{\sigma}_{l-i}^{(m)} \sqrt{C_{i,l-i}^{(n)} C_{i,l-i}^{(m)}} \rho_{l-i+1}^{(n,m)} \mathbb{1}_{\{i=k\}}$$

$$4. \quad \text{Cov} \left[\widehat{C}_{i,j+1}^{*(n)}, \widehat{C}_{k,j+1}^{*(m)} \mid \widehat{C}_{i,j}^{*(n)}, \widehat{C}_{k,j}^{*(m)} \right] = \widehat{\sigma}_{j}^{(n)} \widehat{\sigma}_{j}^{(m)} \sqrt{\widehat{C}_{i,j}^{*(n)} \widehat{C}_{i,j}^{*(m)}} \rho_{j+1}^{(n,m)} \mathbb{1}_{\{i=k\}}$$

For the derivation of the predictive distribution we have to make a specific assumption on the distribution function of the random variables $\epsilon_{i,j}^{*(n)}$. There is a number of different ways in which this can be done such as Normal distribution, shifted Gamma distribution, shifted Lognormal distribution, Uniform distribution etc. In the following example we use a Normal distribution $\mathcal{N}(0, 1)$.

Remarks 3.3:

- It holds $\mathbb{E}\left[\epsilon_{i,j}^{*(n)}\right] = 0$ and $\operatorname{Var}\left(\epsilon_{i,j}^{*(n)}\right) = 1$. However, the model is now not distribution-free anymore.
- The use of the Normal distribution $\mathcal{N}(0,1)$ allows the simulation of negative cumulative claims. In the following we consider this problem as a pure mathematical one. Where this is likely to occur, this can be avoided by defining the distribution of $\epsilon_{i,j}^{*(n)}$ conditionally, given \mathcal{B}_0^N (cf. Model Assumptions 2.1).

3.2.1 Sampling approach without considering correlations

The sampling approach described above does not consider correlations if we set $\rho_j^{(n,m)} = 0$ for all $n \neq m$ and j = 1, ..., J. In this case the random variables $\epsilon_{i,j}^{*(n)}, \epsilon_{i,j}^{*(m)} \sim \mathcal{N}(0, 1)$ are uncorrelated for $n \neq m$ and the correlation-matrix for $\epsilon_{i,j}^{*}$ is given by the $N \times N$ -identity matrix

$$\Sigma_j = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

for j = 1, ..., J.

Finally, for an individual run-off triangle $n \in \{1, ..., N\}$ the claims reserves incorporating the estimation error and process variance are given by

$$\widehat{R}_{i}^{*(n)} = \widehat{C}_{i,J}^{*(n)} - C_{i,I-i}^{(n)}$$

$$\widehat{R}^{*(n)} = \sum_{i=1}^{I} \widehat{R}_{i}^{*(n)}$$
(3.9)

for single and aggregated accident years, respectively. The claims reserves for the entire run-off portfolio without considering correlations are given by

$$\widehat{R}_{i}^{*} = \sum_{n=1}^{N} \widehat{R}_{i}^{*(n)}$$

$$\widehat{R}^{*} = \sum_{i=1}^{I} \widehat{R}_{i}^{*}.$$
(3.10)

Repeating this calculation for every bootstrap sample we obtain the predictive distribution of the claims reserves for single and aggregated accident years as well as for single run-off portfolios $n \in \{1, ..., N\}$ and the whole run-off portfolio. The predictive distribution for the claims reserves captures the estimation error as well as the process variance.

3.2.2 Sampling approach considering correlations

For a single run-off triangle, the random variable $\epsilon_{i,j}^*$ to be sampled reflect the uncertainty of the forecast $\widehat{C}_{i,j}^*$. When *N* correlated run-off triangles are considered, however, the correlation structure of the $\epsilon_{i,j}^{*(1)}, \ldots, \epsilon_{i,j}^{*(N)}$ should be taken into account. Therefore in the following we assume that $\epsilon_{i,j}^* = (\epsilon_{i,j}^{*(1)}, \ldots, \epsilon_{i,j}^{*(N)})$ is multivariate normally distributed with expectation $\mu = \mathbf{0}$ and correlation matrix Σ_j given by (2.3).

In analogy to Section 3.1.2 we make the simplifying assumption that the column wise correlation matrices Σ_j^* are the same for all j and the $\rho^{n,m}$ for $n, m \in \{1, ..., N\}$ can be approximated by the sample correlation coefficients of the Pearson residuals. The uncorrelated draws for all run-off triangles can thus be gathered in a matrix as depicted in Figure 5 for the special case of two run-off triangles and transformed with the method described in Section 3.1.2.

Computational steps: In order to include the process variance a sample of size equal to the number of unobserved ultimate claims in the run-off triangle, i.e. $K = \frac{1}{2}I(I + 1)$, is drawn from a standard-normal distribution for every run-off triangle. The *K* samples are then collected in a matrix and transformed according to (3.6). Subsequently, the elements are plugged into the corresponding positions of the run-off triangle.

First development triangle

Second development triangle

$ \begin{array}{c} \epsilon_{0,1}^{(1)} \\ \epsilon_{1,1}^{(1)} \\ \epsilon_{2,1}^{(1)} \\ \epsilon_{3,1}^{(1)} \\ \epsilon_{4,1}^{(1)} \\ \epsilon_{5,1}^{(1)} \end{array} $	<i>j</i> = 1	$\left\{\begin{array}{c} \epsilon_{0,1}^{(2)} \\ \epsilon_{1,1}^{(2)} \\ \epsilon_{2,1}^{(2)} \\ \epsilon_{2,1}^{(2)} \\ \epsilon_{3,1}^{(2)} \\ \epsilon_{4,1}^{(2)} \\ \epsilon_{5,1}^{(2)} \end{array}\right.$
	÷	
$egin{array}{c} \epsilon_{0,4}^{(1)} \ \epsilon_{1,4}^{(1)} \ \epsilon_{2,4}^{(1)} \end{array} ight brace$	<i>j</i> = 4	$\left\{egin{array}{c} \epsilon_{0,4}^{(2)} \ \epsilon_{1,4}^{(2)} \ \epsilon_{2,4}^{(2)} \end{array} ight.$
$\left. egin{array}{c c} \epsilon_{0,5}^{(1)} \\ \epsilon_{1,5}^{(1)} \end{array} ight angle$	<i>j</i> = 5	$\left\{ egin{array}{c} \epsilon_{0,5}^{(2)} \ \epsilon_{1,5}^{(2)} \end{array} ight.$
$\epsilon_{0,J}^{(1)} ight\}$	j = J	$\left\{ egin{array}{c} \epsilon^{(2)}_{0,J} \end{array} ight.$

Figure 5: Simplified matrix of the sampled errors (the asterisk has been omitted in order not to overload the graphic).

Implementation of the algorithm:

- 1. Collect in a vector **b** of dimension *K* the positions of the unobserved ultimate claims. Repeat this for each of the *N* run-off triangles.
- 2. Draw a sample of size K from a standard-normal distribution. Repeat this for each of the N run-off triangles.
- 3. Collect these *N* samples in a matrix denoted by Q. The dimension of the matrix Q is given by $K \times N$.
- 4. Transform the matrix Q according to transformation (3.6).
- 5. By means of vector **b** plug the transformed drawings into the run-off triangle. Repeat this for each of the *N* run-off triangles.
- 6. Construct the pseudo-observations according to equations (3.7)-(3.8).
- 7. Compute the claims reserves according to equations (3.9)-(3.10).

This procedure completes one bootstrap resampling. Although it is often the case that applications of the bootstrap procedure for the calculation of central moments get acceptable results with as few repetitions as 200, the use of the bootstrap for risk management demands a considerably larger number of replications. The reason for that is that we are not interested in the mean value of the simulated bootstrap distribution but in its higher percentiles. A large number of bootstrap replications is needed for the tails of the distribution to have the chance to get filled in. An insufficient number of replications results in an underestimation of risk measures such as Value-at-Risk or Expected-Shortfall. For our simulations we have used 9999 bootstrap replications.

4 Example

For our simulation we have used the two run-off triangles from Braun [2] (cf. Table 1 and 2). After constructing the Pearson residuals the resulting sample correlation coefficient $\hat{\rho}$ is 0.43. The two triangles are thus positively correlated. For that reason we expect the variance of the aggregated reserves of the two triangles to be larger than the sum of the variances of the single run-off triangles.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	59,966	163,152	254,512	349,524	433,265	475,778	513,660	520,309	527,978	539,039	537,301	540,873	547,696	549,589
1	49,685	153,344	272,936	383,349	458,791	503,358	532,615	551,437	555,792	556,671	560,844	563,571	562,795	
2	51,914	170,048	319,204	425,029	503,999	544,769	559,475	577,425	588,342	590,985	601,296	602,710		
3	84,937	273,183	407,318	547,288	621,738	687,139	736,304	757,440	758,036	782,084	784,632			
4	98,921	278,329	448,530	561,691	641,332	721,696	742,110	752,434	768,638	768,373				
5	71,708	245,587	416,882	560,958	654,652	726,813	768,358	793,603	811,100					
6	92,350	285,507	466,214	620,030	741,226	827,979	873,526	896,728						
7	95,731	313,144	553,702	755,978	857,859	962,825	1,022,241							
8	97,518	343,218	575,441	769,017	934,103	1,019,303								
9	173,686	459,416	722,336	955,335	1,141,750									
10	139,821	436,958	809,926	1,174,196										
11	154,965	528,080	1,032,684											
12	196,124	772,971												
13	204,325													

Table 1: Observed cumulative claims $C_{i,j}^{(1)}$ in Portfolio A (in Mio. US\$).

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	114,423	247,961	312,982	344,340	371,479	371,102	380,991	385,468	385,152	392,260	391,225	391,328	391,537	391,428
1	152,296	305,175	376,613	418,299	440,308	465,623	473,584	478,427	478,314	479,907	480,755	485,138	483,974	
2	144,325	307,244	413,609	464,041	519,265	527,216	535,450	536,859	538,920	539,589	539,765	540,742		
3	145,904	307,636	387,094	433,736	463,120	478,931	482,529	488,056	485,572	486,034	485,016			
4	170,333	341,501	434,102	470,329	482,201	500,961	504,141	507,679	508,627	507,752				
5	189,643	361,123	446,857	508,083	526,562	540,118	547,641	549,605	549,693					
6	179,022	396,224	497,304	553,487	581,849	611,640	622,884	635,452						
7	205,908	416,047	520,444	565,721	600,609	630,802	648,365							
8	210,951	426,429	525,047	587,893	640,328	663,152								
9	213,426	509,222	649,433	731,692	790,901									
10	249,508	580,010	722,136	844,159										
11	258,425	686,012	915,109											
12	368,762	909,066												
13	394,997													

Table 2: Observed cumulative claims $C_{i,j}^{(2)}$ in Portfolio B (in Mio. US\$).

The simplified method described in Sections 3.1 and 3.2 was used with and without considering correlations. For the simulation 9999 bootstrap replications were run. Tables 3 and 4 display

the percental deviations of the mean values and standard deviations of the simulated claims reserves for single accident years in the run-off triangles A and B from the results calculated by the analytical formulas for the mean value and the root mean square error of prediction (RMSE) in Merz & Wüthrich [21]. It is important to remark that the analytical results used as a benchmark do consider the correlation structure, which is estimated with formula (3.20) in Merz & Wüthrich [21]. The benchmark represents thus the "best" analytical estimation of the root mean square error of prediction (RMSE) available in a multivariate setting.

The results depicted in Table 3 correspond to the simulation taking the correlation between the two run-off triangles into account, whereas Table 4 shows the results of the simulation without considering the correlation. It is evident, that both methods are able to approximate the results calculated by the analytical formulas for the first two moments of the distributions of the reserves, since the marginal distributions are not supposed to be affected by the presence of a correlation structure.

	Portfoli percental de	o A eviation	Portfoli percental de	o B eviation
i	mean value	RMSE	mean value	RMSE
1	0.662	-0.395	-6.075	0.299
2	-0.166	1.477	-4.416	0.314
3	-0.258	-1.414	3.744	-1.221
4	0.196	-0.204	3.802	-0.019
5	-0.622	-0.397	2.551	0.018
6	0.205	-0.277	6.265	0.700
7	-0.590	0.043	3.137	-0.454
8	-0.246	-1.062	1.731	1.467
9	-0.417	0.156	0.928	0.243
10	-0.215	-1.309	0.827	-1.304
11	-0.184	-0.054	0.455	-0.212
12	-0.221	0.127	0.401	0.451
13	-0.221	-1.076	0.647	-0.672

Table 3: Deviation in % of the bootstrapping results from the analytical results for single accident years. Method considering correlation.

The story changes when we move to aggregated run-off portfolios. The simulated RMSE for the bootstrap simulation considering correlation is 505644.8 and thus deviates from the RMSE calculated by the analytical formula derived in Merz & Wüthrich [21] only by -0.041%. The corresponding percental deviation between analytical formula for the root process variance (RPV) derived in Merz & Wüthrich [21] and simulated RPV is -1.164% and the percental deviation between analytical formula for the root estimation error (REE) derived in [21] and simulated REE is 1.790%.

	Portfoli percental de	o A eviation	Portfolio B percental deviation		
i	mean value	RMSE	mean value	RMSE	
1	0.239	-0.229	3.984	-1.704	
2	-1.147	-0.555	6.507	-1.777	
3	-0.741	-1.444	-2.938	0.394	
4	-0.834	-0.575	-6.119	-0.366	
5	-1.642	-0.536	-4.479	-0.352	
6	-0.708	0.144	-2.816	-1.580	
7	-0.621	-0.734	-0.685	-0.075	
8	-0.110	-0.849	-0.330	-1.363	
9	-0.208	-1.820	-0.116	-0.360	
10	-0.297	-0.735	-0.337	-0.486	
11	-0.048	-0.439	-0.117	-1.861	
12	-0.062	-1.379	-0.091	-1.160	
13	-0.133	-0.447	-0.219	-1.209	

Table 4: Deviation in % of the bootstrapping results from the analytical results for single accident years. Method without considering correlation.

On the other hand, the simulated RMSE for the bootstrap simulation without considering correlation is 449557.7 corresponding to a percental deviation from the analytical RMSE of 11.056%. The corresponding percental deviation between theoretical and simulated root process variance is 10.407% and the percental deviation between theoretical and simulated root estimation error is 12.108%. This large deviation is consistent with the difference between the results of the analytical RMSE in Merz & Wüthrich [21]. Since the latter is a univariate approach, it cannot take the correlation structure into account.

The situation becomes even worse when we calculate risk measures such as Value-at-Risk and Expected-Shortfall. Tables 5 and 6 display the Value-at-Risk and Expected-Shortfall corresponding to the 90*th*, 95*th* and 99*th* percentiles for the bootstrap simulations with and without considering correlation.

		Percentile	
	90%	95%	99%
Value-at-Risk Expected-Shortfall	8,868,117 9,119,350	9,056,291 9,287,647	9,449,510 9,607,772

Table 5: Risk measures Value-at-Risk and Expected-Shortfall for the predictive bootstrap distribution of the total reserves. Method considering correlation.

		Percentile	
	90%	95%	99%
Value-at-Risk	8,787,663	8,950,341	9,280,683
Expected-Shortfall	9,007,405	9,149,286	9,438,070

Table 6: Risk measures Value-at-Risk and Expected-Shortfall for the predictive bootstrap distri-
bution of the total reserves. Method without considering correlation.

References

- [1] Ajne, B. (1994): Additivity of chain-ladder projections. ASTIN Bulletin, 24(2): 313–318.
- [2] Braun, C. (2004): The prediction error of the chain ladder method applied to correlated run off triangles. *ASTIN Bulletin*, 34(2): 399–423.
- [3] Brehm, P. J. (2002): Correlation and the aggregation of unpaid loss distributions. *CAS E-Forum*, 1–24.
- [4] Buchwalder, M., Bühlmann, H., Merz, M. & Wüthrich, M. V. (2006): The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). *ASTIN Bulletin*, 36(2): 521–542.
- [5] Bundesamt für Privatversicherungen (2006): Swiss Solvency Test. Schweiz.
- [6] Dandekar, R. A., Cohen, M. & Kirkendall, N. (2001): Applicability of Latin Hypercube Sampling to Create Multi Variate Synthetic Micro Data. *ETK-NTTS 2001: European Communities Luxembourg*: 839–847.
- [7] Efron, B. & Tibshirani, R. J. (1993): *An Introduction to the Bootstrap*. 1st ed. Monographs on statistics and applied probability. Champam & Hall.
- [8] England, P. D. (2002): Addendum to "Analytic and bootstrap estimates of prediction errors in claims reserving". *Insurance: Mathematics and Economics*, 31(3): 461–466.
- [9] England, P. D. & Verrall, R. J. (1999): Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, 25(3): 281–293.
- [10] England, P. D. & Verrall, R. J. (2007): Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1(2): 221–270.
- [11] Heberle, J., Huergo, L. & Merz, M. (2009): Bootstrappingmethoden zur Bestimmung des einjährigen Reserverisikos für Solvenzbetrachtungen bei mehreren abhängigen Run-Off-Portfolios.
- [12] Hess, K. T., Schmidt, K. D. & Zocher, M. (2006): Multivariate loss prediction in the multivariate additive model. *Insurance: Mathematics and Economics*, 39(2): 185–191.
- [13] Houltram, A. (2003): Reserving Judgement Considerations Relevant to Insurance Liability Assessment. *XIV General Insurance Seminar, Institute of Actuaries of Australia.*
- [14] Hürlimann, W. (2005): Approximate bounds for the IBNR claims reserves based on the bivariate chain-ladder model. *Belgian Actuarial Bulletin*, 5(1): 46–51.

- [15] Iman, R. L. & Conover, W. J. (1982): A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics – Simulation and Computation*, 11(3): 311–334.
- [16] Kirschner, G. S., Kerley, C. & Isaacs, B. (2008): Two approaches to calculating correlated reserve indications across multiple lines of business. *Variance*, 2: 15–38.
- [17] Klemmt, H. (2004): Trennung von Schadenarten und Additivität bei Chain Ladder Prognosen. *Fall meeting of the German ASTIN Group, Munich*.
- [18] Lowe, J. (1994): A practical guide to measuring reserve variability using: Bootstrapping, operational time and a distribution free approach. *Proceedings of the 1994 General Insurance Convention*.
- [19] Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2): 213–225.
- [20] Merz, M. & Wüthrich, M. V. (2009): Combining Chain-Ladder and Additive Loss Reserving Methods for Dependent Lines of Business. *Variance*, 3: 270–291.
- [21] Merz, M. & Wüthrich, M. V. (2008): Prediction error of the chain ladder reserving method applied to correlated run off triangles. *Annals of Actuarial Science*, 2(1): 25–50.
- [22] Merz, M. & Wüthrich, M. V. (2009): Prediction error of the multivariate additive loss reserving method for dependent lines of business. *Variance*, 3: 131–151.
- [23] Merz, M. & Wüthrich, M. V. (2008): Prediction error of the multivariate chain ladder reserving method. *North American Actuarial Journal*, 12(2): 175–197.
- [24] Mildenhall, S. J. (2006): A multivariate Bayesian claim count development model with closed form posterior and predictive distributions. *CAS E-Forum*, 451–493.
- [25] Pinheiro, P. J. R., Andrade e Silva, J. M. & de Lourdes Centeno, M. (2003): Bootstrap Methodology in Claim Reserving. *Journal of Risk and Insurance*, 70(4): 701–714.
- [26] Pröhl, C. & Schmidt, K. D. (2005): Multivariate chain-ladder. Dresdner Schriften zur Versicherungsmathematik, 3.
- [27] Rodermund, M. (1990): *Foundations of Casualty Actuarial Science*. Casualty Actuarial Society (CAS).
- [28] Schmidt, K. D. (2006): Methods and models of loss reserving based on run-off triangles: a unifying survey. *CAS E-Forum*, 269–317.
- [29] Schmidt, K. D. (2006): Optimal and additive loss reserving for dependent lines of business. *CAS E-Forum*, 319–351.
- [30] Taylor, G. C. (2000): *Loss Reserving: An Actuarial Perspective*. Kluwer Academic Publishers, Boston.
- [31] Taylor, G. C. & McGuire, G. (2007): A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error. *North American Actuarial Journal*, 11(3): 70–88.
- [32] Taylor, G. C. & McGuire, G. (2005): Synchronous bootstrapping of seemingly unrelated regressions.
- [33] Teugels, J. L. & Sundt, B. (2004): *Encyclopedia of Actuarial Science*. Vol. 1. Wiley & Sons Ltd.

- [34] Wüthrich, M. V. & Merz, M. (2008): *Stochastic Claims Reserving Methods in Insurance*. Chichester: Wiley Finance.
- [35] Zurmühl, R. & Falk, S. (1984): *Matrizen und ihre Anwendungen, Teil 1: Grundlagen*. Berlin: Springer.

Paper **B**

Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren für abhängige Run-Off-Portfolios

Jochen Heberle, Luis Huergo, Michael Merz

Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren für abhängige Run-Off-Portfolios

Jochen Heberle

Luis Huergo

Michael Merz

Zusammenfassung

Wir zeigen in diesem Artikel, wie die beiden Schätzer für das Abwicklungsergebnis und den Prognosefehler aus Merz & Wüthrich [6] auf den Fall *N* korrelierter Run-Off-Portfolios verallgemeinert werden können. Die simultane Betrachtung mehrerer korrelierter Run-Off-Portfolios ist dabei vor allem dadurch motiviert, dass es in der Praxis der Schadenreservierung sehr häufig notwendig ist, ein Run-Off-Portfolio erst in mehrere (Sub-) Run-Off-Portfolios zu zerlegen, bevor die für eine sinnvolle Anwendung spezieller Schadenreservierungsverfahren notwendigen Homogenitätseigenschaften erfüllt sind. Das Ergebnis ist jeweils eine explizite Formel für die Prozessvarianz, den Schätzfehler und den Prognosefehler bei der Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren. Die Ergebnisse werden anhand eines Beispiels dargestellt.

1 Einleitung

In der Nichtleben-Versicherung stellen die Schadenreserven für IBNR- und RBNS-Schäden in der Regel den größten versicherungstechnischen Passivposten in der Bilanz dar. Die Bestimmung von Schadenreserven ist somit für jeden Nichtleben-Versicherer von großer wirtschaftlicher Bedeutung. Die Berechnung der angemessenen Schadenreserve in zwei aufeinanderfolgenden Zeitpunkten t = I und t = I + 1 führt zu einem sogenannten *Abwicklungsergebnis* im Kalenderjahr (I, I + 1]. Das Abwicklungsergebnis ist dabei definiert als die Differenz zwischen der Prognose für den Endschadenstand zum Zeitpunkt t = I und der Prognose für den Endschadenstand zum Zeitpunkt t = I und der Prognose für den Endschadenstand erwartungstreu, so ist das erwartete Abwicklungsergebnis gleich Null. Da jedoch bei der Schadenreservierung zukünftige Zahlungen (d.h. Zufallsvariablen) prognostiziert werden, weicht das Abwicklungsergebnis in der Regel stark von dem erwarteten Wert ab. In diesem Beitrag stellen wir einen mathematischen Ansatz vor, wie diese Volatilität für *N* korrelierte Run-Off-Portfolios im Rahmen des multivariaten Chain-Ladder-Reservierungsverfahrens von Merz & Wüthrich [7] geschätzt werden kann. Das heißt wir zeigen wie die Konzepte und Resultate aus

Merz & Wüthrich [6] für ein Run-Off-Portfolio auf den Fall *N* korrelierter Run-Off-Portfolios verallgemeinert werden können.

Die Betrachtung des Abwicklungsergebnisses anstelle wie üblich des Endschadenstandes bzw. der Schadenreserve trägt dem Einjahreshorizont neuer risikobasierter Solvenzanforderungen wie z.B. Solvency II und Swiss Solvency Test Rechnung (vgl. Sandström [9] und Bundesamt für Privatversicherungen [2]). Somit steht in dieser Arbeit, im Unterschied zu den meisten Veröffentlichungen zur Schadenreservierung, nicht die Volatilität der Schadenreserve oder des Endschadenstands im Mittelpunkt, sondern die Volatilität des Abwicklungsergebnisses im Kalenderjahr (I, I + 1]. Diese Sichtweise liefert wichtige Informationen über die kurzfristige Perfomance des Versicherers und ist damit für anstehende Managemententscheidungen, Investoren, Aufsichtsämter und Analysten von großer Bedeutung. Darüber hinaus ist sie eine wichtige Voraussetzung für eine längerfristige positive wirtschaftliche Entwicklung des Versicherers.

Die simultane Betrachtung mehrerer abhängiger Run-Off-Portfolios ist dabei vor allem dadurch motiviert, dass es in der Praxis der Schadenreservierung sehr häufig notwendig ist, ein Run-Off-Portfolio erst in mehrere (Sub-) Run-Off-Portfolios zu zerlegen, bevor die für die sinnvolle Anwendung eines Schadenreservierungsverfahrens notwendigen Homogenitätseigenschaften erfüllt sind (in dieser Arbeit die Annahmen des Chain-Ladder-Reservierungsverfahrens). Ferner können auf diese Weise zum einen existierende Abhängigkeiten zwischen verschiedenen Versicherungsbranchen berücksichtigt werden und zum anderen kann aus der Schadenabwicklung anderer Run-Off-Portfolios gelernt werden (vgl. Merz & Wüthrich [7]).

2 Schadenabwicklungsdreieck und Notation

Im Folgenden betrachten wir simultan $N \ge 1$ Run-Off-Portfolios, die alle die gleiche Dimension haben. Wir nehmen an, dass $C_{i,j}^{(n)}$ die kumulierten Schadenzahlungen für das Anfalljahr $i \in \{0, \ldots, I\}$ bis einschließlich zum Abwicklungsjahr $j \in \{0, \ldots, J\}$ für das Run-Off-Portfolio $n \in \{1, \ldots, N\}$ bezeichnet. Die Bezeichnung Anfalljahr bezieht sich dabei auf dasjenige Jahr, in dem der Schaden eingetreten ist. Um die Notation übersichtlich zu halten, nehmen wir wie allgemein üblich I = J an. Alle Formeln behalten jedoch in entsprechender Form auch für den allgemeineren Fall $I \ge J$ Gültigkeit. Das heißt wir betrachten Abwicklungsdreiecke zum Zeitpunkt I und Abwicklungstrapeze zum Zeitpunkt I + 1 (vgl. Abbildung 1). Ferner nehmen wir an, dass alle Anfalljahre nach J Abwicklungsjahren vollständig abgewickelt sind, es gelte also $C_{i,J+k}^{(n)} = C_{i,J}^{(n)}$ für alle $i = 0, \ldots, I, k \in \mathbb{N}$ und $n = 1, \ldots, N$.

Anfall-	Abwicklungsjahr j	Anjall-	Abwicklungsjahr j
jahr i	$0 \dots j \dots J$	jahr i	0 j J
0		0	
÷	$\mathcal{D}_{I}^{(n)}$:	$\mathcal{D}_{I+1}^{(n)}$
I - j		I-j	
:		:	
Ι		I	

Abbildung 1: *n*-tes Abwicklungsdreieck und Abwicklungstrapez zu den Zeitpunkten t = I bzw. t = I + 1.

Zum Zeitpunkt I liegen für das n-te Abwicklungsdreieck ($n \in \{1, ..., N\}$) insgesamt die Beobachtungen

$$\mathcal{D}_{I}^{(n)} = \left\{ C_{i,j}^{(n)} \mid 0 \le i + j \le I \text{ und } 0 \le i \le I \right\}$$

vor (vgl. Abbildung 1, links). Die Menge der Beobachtungen für alle N Abwicklungsdreiecke zum Zeitpunkt I ist somit gegeben durch

$$\mathcal{D}_I^N = \bigcup_{n=1}^N \mathcal{D}_I^{(n)}.$$

Die Beobachtungen zum Zeitpunkt I + 1 für das *n*-te Abwicklungsdreieck und für alle N Abwicklungsdreiecke sind dagegen gegeben durch

$$\mathcal{D}_{I+1}^{(n)} = \left\{ C_{i,j}^{(n)} \mid 0 \le i+j \le I+1 \text{ und } 0 \le i \le I \right\} = \mathcal{D}_{I}^{(n)} \cup \left\{ C_{i,I-i+1}^{(n)} \mid 1 \le i \le I \right\}$$

bzw.

$$\mathcal{D}_{I+1}^N = \bigcup_{n=1}^N \mathcal{D}_{I+1}^{(n)}.$$

Für die folgende Betrachtung ist es hilfreich für die N Abwicklungsdreiecke durch

$$C_{i,j} = \left(C_{i,j}^{(1)}, \dots, C_{i,j}^{(N)}\right)'$$

die kumulierten Schadenzahlungen des Anfalljahres $i \in \{0, ..., I\}$ und des Abwicklungsjahres $j \in \{0, ..., J\}$ zu Vektoren zusammenzufassen. Ferner definieren wir

$$\mathcal{B}_{k}^{N} = \left\{ C_{i,j}^{(n)} \in \mathcal{D}_{I}^{N} \mid 0 \leq j \leq k \right\} \subseteq \mathcal{D}_{I}^{N}$$

für $k \in \{0, ..., J\}$ und den N-dimensionalen Spaltenvektor $\mathbb{1} = (1, ..., 1)'$ bestehend aus Einsen.

Das klassische Ziel der (multivariaten) Schadenreservierung ist die Prognose des Endschadenstandes $C_{i,J}$ für alle $i \in \{1, \ldots, I\}$ mit Hilfe der Beobachtungen \mathcal{D}_I^N . Jedoch eine Periode später, d.h. zum Zeitpunkt I + 1, erhalten wir neue Beobachtungen auf den Diagonalen der N Abwicklungsdreiecke. Wir sind somit an der Prognose des Endschadenstandes $C_{i,J}$ mit Hilfe der aktualisierten Beobachtungen \mathcal{D}_{I+1}^N interessiert. Im Folgenden analysieren wir die Fluktuationen dieser aktualisierten Prognosen für die Endschadenstände der N korrelierten Run-Off-Portfolios.

3 Multivariates Chain-Ladder-Verfahren

Im Folgenden verwenden wir zur Prognose des Endschadenstandes $C_{i,J}$ im Zeitpunkt I und I + 1 das multivariate Chain-Ladder-Verfahren von Pröhl & Schmidt [8] und Schmidt [10] in der Zeitreihendarstellung von Merz & Wüthrich [7].

Das Chain-Ladder-Verfahren basiert auf dem Studium der individuellen Abwicklungsfaktoren

$$F_{i,j}^{(n)} = \frac{C_{i,j}^{(n)}}{C_{i,j-1}^{(n)}} \quad \text{und} \quad F_{i,j} = \left(F_{i,j}^{(1)}, \dots, F_{i,j}^{(N)}\right)'$$

für $i \in \{0, ..., I\}, j \in \{1, ..., J\}$ und $n \in \{1, ..., N\}$. Ferner bezeichnen wir mit

$$D(\boldsymbol{a}) = \begin{pmatrix} a_1 & 0 \\ & \ddots & \\ 0 & & a_N \end{pmatrix} \quad \text{und} \quad D(\boldsymbol{a})^b = \begin{pmatrix} a_1^b & 0 \\ & \ddots & \\ 0 & & a_N^b \end{pmatrix}$$

die $N \times N$ -Diagonalmatrizen des N-dimensionalen Vektors $\boldsymbol{a} = (a_1, \ldots, a_N)' \in \mathbb{R}^N$ bzw. $(a_1^b, \ldots, a_N^b)' \in \mathbb{R}^N$ für ein $b \in \mathbb{R}$. Damit erhalten wir für $C_{i,j}$ die Darstellung

$$C_{i,j} = D(C_{i,j-1})F_{i,j} = D(F_{i,j})C_{i,j-1}$$

für alle j = 1, ..., J und i = 0, ..., I.

Wir definieren nun das multivariate Chain-Ladder-Zeitreihenmodell wie folgt:

Modellannahmen 3.1 (Multivariates Chain-Ladder-Zeitreihenmodell):

- Kumulierte Schadenzahlungen $C_{i,j}$ aus verschiedenen Anfalljahren i sind unabhängig.
- Es existieren N-dimensionale Vektoren

$$f_j = \left(f_j^{(1)}, \dots, f_j^{(N)}\right)'$$
 und $\sigma_j = \left(\sigma_j^{(1)}, \dots, \sigma_j^{(N)}\right)'$

mit $f_j^{(n)} > 0$, $\sigma_j^{(n)} > 0$ für alle $n \in \{1, ..., N\}$ und N-dimensionale Zufallsvariablen

$$\boldsymbol{\epsilon}_{i,j+1} = \left(\epsilon_{i,j+1}^{(1)},\ldots,\epsilon_{i,j+1}^{(N)}\right)',$$

so dass für alle $i \in \{0, ..., I\}$ und $j \in \{0, ..., J - 1\}$

$$C_{i,j+1} = \mathrm{D}(f_j)C_{i,j} + \mathrm{D}(C_{i,j})^{1/2} \mathrm{D}(\boldsymbol{\epsilon}_{i,j+1})\boldsymbol{\sigma}_j$$

gilt.

• Die Zufallsvariablen $\epsilon_{i,j+1}$ sind unabhängig mit

$$\mathbb{E}\left[\boldsymbol{\epsilon}_{i,j+1} \mid \boldsymbol{\mathcal{B}}_{0}^{N}\right] = \mathbf{0}, \qquad P\left(C_{i,j}^{(n)} > 0 \mid \boldsymbol{\mathcal{B}}_{0}^{N}\right) = 1$$

und

$$\operatorname{Cov}\left(\boldsymbol{\epsilon}_{i,j+1}, \boldsymbol{\epsilon}_{i,j+1} \mid \mathcal{B}_{0}^{N}\right) = \operatorname{E}\left[\boldsymbol{\epsilon}_{i,j+1}\boldsymbol{\epsilon}_{i,j+1}^{\prime} \mid \mathcal{B}_{0}^{N}\right] = \begin{pmatrix} 1 & \rho_{j}^{(1,2)} & \cdots & \rho_{j}^{(1,N)} \\ \rho_{j}^{(2,1)} & 1 & \cdots & \rho_{j}^{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{j}^{(N,1)} & \rho_{j}^{(N,2)} & \cdots & 1 \end{pmatrix}.$$

Dabei ist
$$\rho_j^{(n,m)} \in (-1,1)$$
 für alle $n,m \in \{1,\ldots,N\}$ mit $n \neq m$.

Dieses multivariate Chain-Ladder-Zeitreihenmodell wurde erstmals in Merz & Wüthrich [5, 7] analysiert und ist ein Spezialfall des Chain-Ladder-Modells von Pröhl & Schmidt [8]. Das Zeitreihenmodell 3.1 besitzt den Vorteil, dass es angibt, wie neue Beobachtungen generiert werden können. Dies ist für die analytische Herleitung eines Schätzers für den bedingten Schätzfehler sehr hilfreich. Die Zufallsvariablen $\epsilon_{i,j+1}$ sind dabei bedingt gegeben \mathcal{B}_0^N definiert. Dies stellt sicher, dass die kumulierten Schadenzahlungen $C_{i,j+1}$ positiv sind.

Mit Hilfe der Identität

$$D(b)ac' D(d) = D(a)bd' D(c)$$
(3.1)

für alle *N*-dimensionale Vektoren *a*, *b*, *c* und *d* folgt

$$Cov (C_{i,j}, C_{i,j} | C_{i,j-1})$$

= $D(C_{i,j-1})^{1/2} E [D(\epsilon_{i,j})\sigma_{j-1}\sigma'_{j-1}D(\epsilon_{i,j})] D(C_{i,j-1})^{1/2}$
= $D(C_{i,j-1})^{1/2} D(\sigma_{j-1}) E [\epsilon_{i,j}\epsilon'_{i,j}] D(\sigma_{j-1}) D(C_{i,j-1})^{1/2}$.

Bemerkungen 3.2:

- Der Parameter f_j wird als *N*-dimensionaler Chain-Ladder-Faktor für das *j*-te Abwicklungsjahr bezeichnet. Er gibt an wie sich $C_{i,j+1}$ aus $C_{i,j}$ ergibt.
- Die Modellannahmen 3.1 implizieren

$$E[C_{i,j} | C_{i,j-1}] = D(f_{j-1})C_{i,j-1}$$
(3.2)

$$Cov (C_{i,j}, C_{i,j} | C_{i,j-1}) = D(C_{i,j-1})^{1/2} \Sigma_{j-1} D(C_{i,j-1})^{1/2}$$
(3.3)

mit

$$\Sigma_{j-1} = D(\boldsymbol{\sigma}_{j-1}) E[\boldsymbol{\epsilon}_{i,j} \boldsymbol{\epsilon}'_{i,j}] D(\boldsymbol{\sigma}_{j-1})$$

$$= \begin{pmatrix} (\sigma_{j-1}^{(1)})^2 & \sigma_{j-1}^{(1)} \sigma_{j-1}^{(2)} \rho_{j-1}^{(1,2)} & \cdots & \sigma_{j-1}^{(1)} \sigma_{j-1}^{(N)} \rho_{j-1}^{(1,N)} \\ \sigma_{j-1}^{(2)} \sigma_{j-1}^{(1)} \rho_{j-1}^{(2,1)} & (\sigma_{j-1}^{(2)})^2 & \cdots & \sigma_{j-1}^{(2)} \sigma_{j-1}^{(N)} \rho_{j-1}^{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{j-1}^{(N)} \sigma_{j-1}^{(1)} \rho_{j-1}^{(N,1)} & \sigma_{j-1}^{(N)} \sigma_{j-1}^{(2)} \rho_{j-1}^{(N,2)} & \cdots & (\sigma_{j-1}^{(N)})^2 \end{pmatrix}$$

$$(3.4)$$

für alle $0 \le i \le I$ und $1 \le j \le J$. Zusammen mit der Unabhängigkeitsannahme für verschiedene Anfalljahre stellen die Gleichungen (3.2) und (3.3) das multivariate Analogon des univariaten Chain-Ladder-Modells von Mack [3] dar.

- Die Korrelationen zwischen den N
 Run-Off-Portfolios werden durch die Korrelationsmatrizen
 Σ_{j-1} erfasst.
- Im Zeitreihenmodell 3.1 werden strengere Annahmen getroffen als im Chain-Ladder-Modell von Pröhl & Schmidt [8] bzw. Schmidt [10]. Im Unterschied zum Modell in diesen beiden Arbeiten werden im Modell 3.1 unabhängige Anfalljahre, deterministische Chain-Ladder-Faktoren f_j und deterministische $N \times N$ -Korrelationsmatrizen Σ_j unterstellt.
- Für eine Motivation der im Modell 3.1 unterstellten Abwicklungsjahr basierten Korrelationen siehe die Arbeiten von Braun [1] sowie Merz & Wüthrich [5, 7].

Die Modellannahmen 3.1 implizieren unmittelbar

$$\mathbb{E}\left[C_{i,J} \mid \mathcal{D}_{I}^{N}\right] = \mathbb{E}\left[C_{i,J} \mid C_{i,I-i}\right] = \prod_{j=I-i}^{J-1} \mathbb{D}(f_{j})C_{i,I-i}$$

und

$$\mathbb{E}\left[C_{i,J} \mid \mathcal{D}_{I+1}^{N}\right] = \mathbb{E}\left[C_{i,J} \mid C_{i,I-i+1}\right] = \prod_{j=I-i+1}^{J-1} \mathbb{D}(f_{j})C_{i,I-i+1}.$$

Das heißt sobald die Chain-Ladder-Faktoren f_j bekannt sind, kann der erwartete Endschaden $C_{i,J}$, gegeben die Beobachtungen \mathcal{D}_I bzw. \mathcal{D}_{I+1} , berechnet werden. Allerdings sind in den meisten praktischen Anwendungen die Chain-Ladder-Faktoren f_j nicht bekannt und müssen daher geschätzt werden. Im multivariaten Chain-Ladder-Modell werden die Chain-Ladder-Faktoren f_j ($j = 0, \ldots, J - 1$) wie folgt geschätzt (vgl. Pröhl & Schmidt [8] und Schmidt [10]):

1. Zum Zeitpunkt t = I durch

$$\hat{f}_{j}^{I} = \left(\hat{f}_{j}^{I,(1)}, \dots, \hat{f}_{j}^{I,(N)}\right)' = (S_{j}^{I})^{-1} \sum_{i=0}^{I-j-1} \mathcal{D}(C_{i,j})^{1/2} \Sigma_{j}^{-1} \mathcal{D}(C_{i,j})^{1/2} F_{i,j+1},$$
(3.5)

mit

$$S_j^I = \sum_{i=0}^{I-j-1} D(C_{i,j})^{1/2} \Sigma_j^{-1} D(C_{i,j})^{1/2}.$$
 (3.6)

2. Zum Zeitpunkt t = I + 1 durch

$$\hat{f}_{j}^{I+1} = \left(\hat{f}_{j}^{I+1,(1)}, \dots, \hat{f}_{j}^{I+1,(N)}\right)' = (S_{j}^{I+1})^{-1} \sum_{i=0}^{I-j} D(C_{i,j})^{1/2} \Sigma_{j}^{-1} D(C_{i,j})^{1/2} F_{i,j+1}, \quad (3.7)$$

mit

$$S_j^{I+1} = \sum_{i=0}^{I-j} D(C_{i,j})^{1/2} \Sigma_j^{-1} D(C_{i,j})^{1/2}.$$
 (3.8)

Der Chain-Ladder-Schätzer \hat{f}_{j}^{I+1} zum Zeitpunkt I + 1 nutzt die Beobachtungen, die im Kalenderjahr (I, I + 1] neu hinzukommen. Die Schätzer (3.5) und (3.7) sind erwartungstreue Gauss-Markov-Schätzer für f_{j} (vgl. Merz & Wüthrich [7]). Ferner sind \hat{f}_{l}^{m} und \hat{f}_{j}^{m} (für m = Iund m = I + 1) mit $j \neq l$ unkorreliert (vgl. Lemma 3.5 in Merz & Wüthrich [7]). Dies impliziert,
bedingt gegeben $C_{i,I-i}$, ist der Chain-Ladder-Prädiktor

$$\widehat{C_{i,j}^{I}} = \left(\widehat{C_{i,j}^{I,(1)}}, \dots, \widehat{C_{i,j}^{I,(N)}}\right)' = \prod_{l=I-i}^{j-1} \mathcal{D}(\widehat{f}_{l}^{I})C_{i,I-i}$$
(3.9)

ein erwartungstreuer Prädiktor für E [$C_{i,j} \mid \mathcal{D}_I^N$] mit $j \ge I - i$ und, bedingt gegeben $C_{i,I-i+1}$, ist der Chain-Ladder-Prädiktor

$$\widehat{C_{i,j}^{l+1}} = \left(\widehat{C_{i,j}^{l+1,(1)}}, \dots, \widehat{C_{i,j}^{l+1,(N)}}\right)' = \prod_{l=l-i+1}^{j-1} \mathcal{D}(\widehat{f}_l^{l+1}) C_{i,l-i+1}$$
(3.10)

ein erwartungstreuer Prädiktor für E [$C_{i,j} \mid \mathcal{D}_{I+1}^N$] mit $j \ge I - i + 1$ (vgl. Merz & Wüthrich [6] für den univariaten Fall).

Bemerkungen 3.3:

- Die Realisierungen der Schätzer $\hat{f}_0^I, \ldots, \hat{f}_{J-1}^I$ sind zum Zeitpunkt I bekannt. Da jedoch die kumulierten Schadenzahlungen $C_{I,1}, \ldots, C_{I-J+1,J}$ zum Zeitpunkt I + 1 unbekannt sind, gilt dies auch für die Realisierungen von $\hat{f}_0^{I+1}, \ldots, \hat{f}_{J-1}^{I+1}$.
- Im Fall N = 1 (d.h. lediglich ein Run-Off-Portfolio bzw. Abwicklungsdreieck) erhalten wir aus (3.5) und (3.7) die klassischen univariaten Chain-Ladder-Schätzer

$$\hat{f}_{j}^{I} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum\limits_{k=0}^{I-j-1} C_{k,j}} F_{i,j+1} \quad \text{und} \quad \hat{f}_{j}^{I+1} = \sum_{i=0}^{I-j} \frac{C_{i,j}}{\sum\limits_{k=0}^{I-j} C_{k,j}} F_{i,j+1}. \quad (3.11)$$

• In Braun [1] und Merz & Wüthrich [5] basieren die Prognosen für die N korrelierten Run-Off-Portfolios auf den univariaten Chain-Ladder Schätzern (3.11). Es werden also die Korrelationen zwischen den verschiedenen Run-Off-Portfolios nicht berücksichtigt. In der hier verwendeten Matrix-Notation würde dies bedeuten, dass die multivariaten Chain-Ladder-Faktoren f_i zum Zeitpunkt I und I + 1 durch

$$\hat{f}_{j}^{I,(0)} = \left(\sum_{i=0}^{I-j-1} D(C_{i,j})\right)^{-1} \sum_{i=0}^{I-j-1} D(C_{i,j}) F_{i,j+1}$$
(3.12)

bzw.

$$\hat{f}_{j}^{I+1,(0)} = \left(\sum_{i=0}^{I-j} D(C_{ij})\right)^{-1} \sum_{i=0}^{I-j} D(C_{ij})F_{ij+1}$$

geschätzt werden. Beide Chain-Ladder-Schätzer \hat{f}_j^m und $\hat{f}_j^{m,(0)}$ (für m = I und m = I + 1) sind erwartungstreue Schätzer für f_j . Der Chain-Ladder-Schätzer \hat{f}_j^m besitzt jedoch die kleinere erwartete quadratische Abweichung von f_j . Ist die $N \times N$ Kovarianzmatrix Σ_j diagonal, so stimmen \hat{f}_j^m und $\hat{f}_j^{m,(0)}$ überein.

• Die Prädiktoren $\widehat{C_{i,j}^{I+1}}$ basieren auf den Chain-Ladder-Schätzern zum Zeitpunkt t = I + 1und nutzen damit die im Kalenderjahr (I, I + 1] neu hinzukommenden Beobachtungen.

Wir erhalten folgendes Lemma:

Lemma 3.4: Unter den Modellannahmen 3.1 gilt

1.
$$C_{i,I-i+1}, \hat{f}_{I-i+1}^{I+1}, \dots, \hat{f}_{J-1}^{I+1}$$
 sind, bedingt gegeben \mathcal{D}_{I}^{N} , unabhängig.
2. $\mathbb{E}\left[\left.\hat{f}_{l}^{I+1} \right| \left.\mathcal{D}_{I}^{N}\right] = (S_{l}^{I+1})^{-1}S_{l}^{I}\hat{f}_{l}^{I} + (S_{l}^{I+1})^{-1}\mathbb{D}(C_{I-l,l})^{1/2}\Sigma_{l}^{-1}\mathbb{D}(C_{I-l,l})^{1/2}f_{l},$
3. $\mathbb{E}\left[\left.\widehat{C_{i,j}^{I+1}} \right| \left.\mathcal{D}_{I}^{N}\right] = \prod_{l=I-i+1}^{j-1}\mathbb{E}\left[\left.\mathbb{D}(\hat{f}_{l}^{I+1}) \right| \left.\mathcal{D}_{I}^{N}\right]\mathbb{D}(f_{I-i})C_{i,I-i}.$

Beweis. Der Beweis ist ähnlich zum Beweis des Lemmas 2.4 in Merz und Wüthrich [6] (univariates Chain-Ladder-Verfahren).

4 Schadenabwicklungsergebnis für abhängige Run-Off-Portfolios

Die noch ausstehenden Zahlungsverpflichtungen für das *i*-te Anfalljahr ($i \in \{1, ..., I\}$) zu den Zeitpunkten t = I und t = I + 1 sind definiert durch

$$\mathbf{R}_{i}^{I} = \left(R_{i}^{I,(1)}, \dots, R_{i}^{I,(N)}\right)' = C_{i,J} - C_{i,I-i}$$

bzw.

$$\boldsymbol{R}_{i}^{I+1} = \left(R_{i}^{I+1,(1)}, \ldots, R_{i}^{I+1,(N)}\right)' = \boldsymbol{C}_{i,J} - \boldsymbol{C}_{i,I-i+1}.$$

Das Abwicklungsergebnis (AWE) ist dann wie folgt definiert:

Definition 4.1 (AWE für einzelne und aggregierte Anfalljahre): Das Abwicklungsergebnis für das *i*-te Anfalljahr ($i \in \{1, ..., I\}$) im Kalenderjahr (I, I + 1] ist definiert durch

$$AWE_{i}(I+1) = \left(AWE_{i}(I+1)^{(1)}, \dots, AWE_{i}(I+1)^{(N)}\right)'$$

= E $\left[R_{i}^{I} \mid \mathcal{D}_{I}^{N}\right] - \left(X_{i,I-i+1} + E\left[R_{i}^{I+1} \mid \mathcal{D}_{I+1}^{N}\right]\right).$ (4.1)

Dabei bezeichnet $X_{i,I-i+1} = C_{i,I-i+1} - C_{i,I-i}$ die Schadenzahlungen im Kalenderjahr (I,I + 1] und Anfalljahr i. Das Abwicklungsergebnis für aggregierte Anfalljahre ist gegeben durch

$$AWE(I + 1) = \left(AWE(I + 1)^{(1)}, \dots, AWE(I + 1)^{(N)}\right)'$$
$$= \sum_{i=1}^{I} AWE_i(I + 1).$$
(4.2)

Das Abwicklungsergebnis (4.1) und (4.2) ist erst zum Zeitpunkt I + 1 bekannt und damit zum Zeitpunkt I eine Zufallsvariable. Für das Abwicklungsergebnis im Anfalljahr i gilt

$$\mathbf{AWE}_{i}(I+1) = \mathbf{E} \left[C_{i,J} \mid \mathcal{D}_{I}^{N} \right] - \mathbf{E} \left[C_{i,J} \mid \mathcal{D}_{I+1}^{N} \right].$$
(4.3)

Aus (4.3) folgt unmittelbar

$$\mathbb{E}\left[\mathbf{AWE}_{i}(I+1) \mid \mathcal{D}_{I}^{N}\right] = \mathbf{0} \qquad \text{bzw.} \qquad \mathbb{E}\left[\sum_{n=1}^{N} \mathbf{AWE}_{i}(I+1)^{(n)} \mid \mathcal{D}_{I}^{N}\right] = \mathbf{0}$$
(4.4)

und

$$\mathbb{E}\left[\mathbf{AWE}(I+1) \mid \mathcal{D}_{I}^{N}\right] = \mathbf{0} \qquad \text{bzw.} \qquad \mathbb{E}\left[\sum_{i=1}^{I}\sum_{n=1}^{N}\mathbf{AWE}_{i}(I+1)^{(n)} \mid \mathcal{D}_{I}^{N}\right] = \mathbf{0}.$$
 (4.5)

Das Abwicklungsergebnis spiegelt die zeitliche Entwicklung der Prognose für den Endschadenstand im Kalenderjahr (I, I + 1] wieder. Ist z.B. zum Zeitpunkt I + 1 ein Versicherer der Meinung, dass die Schadenrückstellungen reduziert werden können, so ist das Abwicklungsergebnis im Zeitpunkt I + 1 positiv. Auf der anderen Seite erhalten wir ein negatives Abwicklungsergebnis, falls der Versicherer zum Zeitpunkt I + 1 der Meinung ist, dass die Schadenrückstellungen erhöht werden müssen.

Aus (4.4) und (4.5) folgt, dass für bekannte Chain-Ladder-Faktoren f_j das zum Zeitpunkt I erwartete Abwicklungsergebnis gleich Null ist. In den meisten praktischen Anwendungen sind jedoch die Chain-Ladder-Faktoren f_j unbekannt und müssen daher durch \hat{f}_j^I bzw. \hat{f}_j^{I+1}

geschätzt werden. Bedingt gegeben $C_{i,I-i}$ ($1 \le i \le I$) gilt, dass der Prädiktor

$$\widehat{R_i^I} = \left(\widehat{R_i^{I,(1)}}, \dots, \widehat{R_i^{I,(N)}}\right)' = \widehat{C_{i,J}^I} - C_{i,I-i}$$
(4.6)

ein erwartungstreuer Prädiktor für E $[\mathbf{R}_i^I \mid \mathcal{D}_I^N]$ ist. Analog gilt, dass bedingt gegeben $C_{i,I-i+1}$

$$\widehat{R_{i}^{I+1}} = \left(\widehat{R_{i}^{I+1,(1)}}, \dots, \widehat{R_{i}^{I+1,(N)}}\right)' = \widehat{C_{i,J}^{I+1}} - C_{i,I-i+1}$$
(4.7)

ein erwartungstreuer Prädiktor für $E[\mathbf{R}_{i}^{I+1} \mid \mathcal{D}_{I+1}^{N}]$ ist.

Ersetzen wir in (4.1) E $[\mathbf{R}_{i}^{I} | \mathcal{D}_{I}^{N}]$ und E $[\mathbf{R}_{i}^{I+1} | \mathcal{D}_{I+1}^{N}]$ durch ihre Prädiktoren (4.6) und (4.7), so erhalten wir das zum Zeitpunkt I + 1 beobachtbare Abwicklungsergebnis im Kalenderjahr (I, I + 1]:

Schätzer 4.2 (Beobachtbares AWE für einzelne und aggregierte Anfalljahre): Das beobachtbare Abwicklungsergebnis für das Anfalljahr $i \in \{1, ..., I\}$ im Kalenderjahr (I, I + 1] ist definiert durch

$$\widehat{\mathbf{AWE}}_{i}(I+1) = \left(\widehat{\mathbf{AWE}}_{i}(I+1)^{(1)}, \dots, \widehat{\mathbf{AWE}}_{i}(I+1)^{(N)}\right)'$$
$$= \widehat{\mathbf{R}}_{i}^{I} - \left(\mathbf{X}_{i,I-i+1} + \widehat{\mathbf{R}}_{i}^{I+1}\right) = \widehat{C}_{i,J}^{I} - \widehat{C}_{i,J}^{I+1}.$$
(4.8)

Das aggregierte beobachtbare Abwicklungsergebnis ist gegeben durch

$$\widehat{AWE}(I+1) = \left(\widehat{AWE}(I+1)^{(1)}, \dots, \widehat{AWE}(I+1)^{(N)}\right)'$$
$$= \sum_{i=1}^{I} \widehat{AWE}_i(I+1).$$
(4.9)

Das beobachtbare Abwicklungsergebnis ist das zum Zeitpunkt I + 1 resultierende Abwicklungsergebnis, wenn die ausstehenden Schadenverpflichtungen für die N Run-Off-Portfolios zu den Zeitpunkten I und I + 1 mit Hilfe des multivariaten Chain-Ladder-Verfahren prognostiziert werden. Dabei ist jedoch zu beachten, dass (4.8) und (4.9) keine "echten" Prädiktoren für $AWE_i(I + 1)$ bzw. AWE(I + 1) sind, da sie zum Zeitpunkt I nicht beobachtbar sind. Das heißt wir sind nicht in der Lage mit Hilfe der bis zum Zeitpunkt I vorhanden Beobachtungen \mathcal{D}_I^N in den N Run-Off-Portfolios (4.8) und (4.9) auszuwerten. Dies ist erst nach dem Kalenderjahr (I, I + 1] möglich, wenn die Beobachtungen \mathcal{D}_{I+1}^N vorliegen. Daher wird das beobachtbare Abwicklungsergebnis durch seine bedingte Erwartung, gegeben die Beobachtungen \mathcal{D}_I^N zum Zeitpunkt I, approximiert. **Schätzer 4.3** (Schätzung des AWE für einzelne und aggregierte Anfalljahre): Der Prädiktor des Abwicklungsergebnisses für das Anfalljahr $i \in \{1, ..., I\}$ im Kalenderjahr (I, I + 1] ist definiert durch

$$\widehat{\widehat{AWE}}_i(I+1) = \left(\widehat{\widehat{AWE}}_i(I+1)^{(1)}, \dots, \widehat{\widehat{AWE}}_i(I+1)^{(N)}\right)'$$
$$= \mathbb{E}\left[\left.\widehat{AWE}_i(I+1) \right| \mathcal{D}_I^N\right].$$

Der Prädiktor des aggregierten Abwicklungsergebnisses ist gegeben durch

$$\widehat{\widehat{AWE}}(I+1) = \left(\widehat{\widehat{AWE}}(I+1)^{(1)}, \dots, \widehat{\widehat{AWE}}(I+1)^{(N)}\right)^{I}$$
$$= \sum_{i=1}^{I} \widehat{\widehat{AWE}}_{i}(I+1).$$

Wir erhalten das folgende Lemma:

Lemma 4.4: Mit den Modellannahmen 3.1 gilt für den Schätzer des AWE im Anfalljahr $i \in \{1, ..., I\}$

$$\widehat{\widehat{AWE}}_{i}(I+1) = \left(\prod_{l=I-i}^{J-1} \mathcal{D}(\widehat{f}_{l}^{I}) - \prod_{l=I-i+1}^{J-1} \mathcal{D}(\widehat{h}_{l}) \mathcal{D}(f_{I-i})\right) C_{i,I-i}$$
(4.10)

mit

$$\hat{\boldsymbol{h}}_{l} = \left(\hat{\boldsymbol{h}}_{l}^{(1)}, \dots, \hat{\boldsymbol{h}}_{l}^{(N)}\right)' \\ = \mathbb{E}\left[\left.\hat{\boldsymbol{f}}_{l}^{I+1} \right| \mathcal{D}_{l}^{N}\right] = (S_{l}^{I+1})^{-1} S_{l}^{I} \hat{\boldsymbol{f}}_{l}^{I} + (S_{l}^{I+1})^{-1} (S_{l}^{I+1} \boldsymbol{f}_{l} - S_{l}^{I} \boldsymbol{f}_{l}).$$
(4.11)

Beweis. Der Beweis folgt leicht aus Gleichung (3.9) und (3.10) und Lemma 3.4 (Teil 2. und 3.). $\hfill \square$

Beachte, die rechte Seite von (4.10) enthält die unbekannten Parameter f_l (in \hat{h}_l) und f_{I-i} . Wenn wir diese Chain-Ladder-Faktoren durch die Chain-Ladder-Schätzer \hat{f}_l^I und \hat{f}_{I-i}^I im Zeitpunkt I ersetzen, dann resultiert **0**. In diesem Sinne haben wir beste Schätzer für die Schadenreserve.

5 Quantifizierung der Unsicherheit im Abwicklungsergebnis mehrerer abhängiger Run-Off-Portfolios

In diesem Abschnitt quantifizieren wir die Volatilität der Prädiktoren für das Abwicklungsergebnis

$$\sum_{n=1}^{N} \widehat{\widehat{AWE}}_{i}(I+1)^{(n)} = \mathbb{1}' \widehat{\widehat{AWE}}_{i}(I+1)$$
(5.1)

und

$$\sum_{i=1}^{I} \sum_{n=1}^{N} \widehat{\widehat{AWE}}_i (I+1)^{(n)} = \mathbb{1}' \widehat{\widehat{AWE}} (I+1)$$
(5.2)

mit Hilfe des bedingten Prognosefehlers gegeben die Beobachtungen bis zum Zeitpunkt I.

5.1 Bedingter Prognosefehler für einzelne Anfalljahre

In diesem Abschnitt ermitteln wir analytisch einen Schätzer für den bedingten Prognosefehler für ein einzelnes Anfalljahr.

Definition 5.1 (Bedingter Prognosefehler für einzelne Anfalljahre): Der bedingte Prognosefehler für Anfalljahr $i \in \{1, ..., I\}$ ist definiert durch

$$MSE\left(\sum_{n=1}^{N}\widehat{AWE}_{i}(I+1)^{(n)}\right)$$

= $E\left[\left(\sum_{n=1}^{N}\widehat{AWE}_{i}(I+1)^{(n)} - \sum_{n=1}^{N}AWE_{i}(I+1)^{(n)}\right)^{2} \middle| \mathcal{D}_{I}^{N}\right].$ (5.3)

Wie üblich zerlegen wir den bedingten Prognosefehler in seine beiden Bestandteile bedingte

Prozessvarianz und bedingter Schätzfehler. Mit (5.1) erhalten wir für (5.3)

$$MSE\left(\sum_{n=1}^{N}\widehat{AWE}_{i}(I+1)^{(n)}\right) = \underbrace{\mathbb{1}' \operatorname{Var}\left(AWE_{i}(I+1) \mid \mathcal{D}_{I}^{N}\right)\mathbb{1}}_{\text{bedingte Prozessvarianz}} + \underbrace{\mathbb{1}'\widehat{AWE}_{i}(I+1)\widehat{AWE}_{i}(I+1)'\mathbb{1}}_{\text{bedingter Schätzfehler}}.$$

Bemerkungen 5.2:

- Die bedingte Prozessvarianz beschreibt die Varianz innerhalb des stochastischen Modells, d.h. den reinen Zufall, welcher nicht eliminiert werden kann. Der bedingte Schätzfehler reflektiert dagegen die Unsicherheit in der Schätzung der Chain-Ladder-Faktoren durch \hat{f}_{j}^{I} und \hat{f}_{j}^{I+1} zum Zeitpunkt I bzw. I + 1.
- Es ist wichtig festzuhalten, dass eine Solvenzbetrachtung, welche auf der Unsicherheit des Prädiktors für das Abwicklungsergebnis basiert, lediglich die Unsicherheit in den Schadenreserven für das Kalenderjahr (I, I + 1] erfasst. Diese einjährige Betrachtung ist jedoch wichtig für einen längerfristigen Erfolg und für unmittelbar anstehende Managemententscheidungen. Ebenso ist sie für Investoren, Analysten, Aufsichtsämter usw. von großer Bedeutung. Um die Unsicherheit des vollständigen Run-Offs zu erfassen, kann z.B. ein Kapitalkostenzuschlag zum verbleibenden Run-Off nach Kalenderjahr (I, I + 1] addiert werden, welcher auf der Unsicherheit des Prädiktors für das Abwicklungsergebnis basiert (vgl. Bundesamt für Privatversicherungen [2]). Ein solcher Kapitalkostenzuschlag stellt dann sicher, dass das zukünftig benötigte Solvenzkapital finanziert werden kann.

Im Folgenden ermitteln wir jeweils einen Schätzer für die bedingte Prozessvarianz und für den bedingten Schätzfehler in einem einzelnen Anfalljahr. Für weitere Informationen zum bedingten und unbedingten Prognosefehler siehe z.B. Kapitel 3 in Wüthrich & Merz [11].

5.1.1 Bedingte Prozessvarianz

Für die bedingte Prozessvarianz im Anfalljahr $i \in \{1, ..., I\}$ gilt:

Lemma 5.3 (Bedingte Prozessvarianz für einzelne Anfalljahre): Unter den Modellannahmen 3.1 ist die bedingte Prozessvarianz des Prädiktors für das Abwicklungsergebnis im Anfalljahr $i \in \{1, ..., I\}$ für das Kalenderjahr (I, I + 1] gegeben durch

$$\mathbb{1}' \operatorname{Var} \left(\operatorname{AWE}_{i}(I+1) \mid \mathcal{D}_{I}^{N} \right) \mathbb{1} = \operatorname{E} \left[C_{iJ} \mid \mathcal{D}_{I}^{N} \right] \operatorname{D}(C_{i,I-i})^{-1/2} \operatorname{D}(f_{I-i})^{-1} \Sigma_{I-i} \operatorname{D}(f_{I-i})^{-1} \times \operatorname{D}(C_{i,I-i})^{-1/2} \operatorname{E} \left[C_{iJ} \mid \mathcal{D}_{I}^{N} \right].$$
(5.4)

Beweis. Der Beweis ist ähnlich zum Beweis des Lemmas 4.2 in Merz & Wüthrich [6] (univariates Chain-Ladder-Verfahren).

Bemerkungen 5.4:

• Wir erhalten einen Schätzer für die bedingte Prozessvarianz im Anfalljahr $i \in \{1, ..., I\}$, wenn wir auf der rechten Seite von (5.4) die Parameter f_{I-i} und Σ_{I-i} durch ihre Schätzer zum Zeitpunkt I (siehe Abschnitt 6) sowie den Endschadenstand $C_{i,J}$ durch seinen Chain-Ladder-Prädiktor $\widehat{C_{i,I}^{I}}$ zum Zeitpunkt I ersetzen:

$$\mathbb{1}' \widehat{\operatorname{Var}} \left(\operatorname{AWE}_{i}(I+1) \mid \mathcal{D}_{I}^{N} \right) \mathbb{1} = \widehat{C_{i,J}^{I}} \operatorname{D}(C_{i,I-i})^{-1/2} \operatorname{D}(\widehat{f}_{I-i})^{-1} \widehat{\Sigma}_{I-i} \\ \times \operatorname{D}(\widehat{f}_{I-i})^{-1} \operatorname{D}(C_{i,I-i})^{-1/2} \widehat{C_{i,J}^{I}}$$

• Der Ausdruck (5.4) kann auch als die bedingte Prozessvarianz für das Abwicklungsjahr I - i + 1 im *i*-ten Anfalljahr aufgefasst werden. Durch Addition dieser Ausdrücke über die Abwicklungsjahre I - i + 1, ..., J hinweg (für festes Anfalljahr *i*), erhalten wir die bedingte Prozessvarianz des Prädiktors für den Endschadenstand im Anfalljahr *i* im multivariaten Chain-Ladder-Verfahren (vgl. Lemma 4.1 in Merz & Wüthrich [7]). Für den Fall N = 1 erhalten wir aus (5.4) die bedingte Prozessvarianz für ein einzelnes Anfalljahr *i* und ein einzelnes Run-Off-Portfolio (vgl. Merz & Wüthrich [6]).

5.1.2 Bedingter Schätzfehler

In diesem Abschnitt untersuchen wir den Schätzfehler für ein einzelnes Anfalljahr $i \in \{1, ..., I\}$, d.h.:

$$\mathbb{1}'\widehat{\widehat{\mathrm{AWE}}}_i(I+1)\widehat{\widehat{\mathrm{AWE}}}_i(I+1)'\,\mathbb{1}$$

Wir sind also an einer Schätzung der Präzision der Chain-Ladder-Schätzer $\hat{f}_{I-i}^{I}, \ldots, \hat{f}_{J-1}^{I}$ im Zeitpunkt *I* interessiert. Mit Hilfe von Lemma 4.4 erhält man für den bedingten Schätzfehler

$$\begin{split} \widehat{\mathbf{1}'\widehat{\mathbf{AWE}}}_{i}(I+1)\widehat{\widehat{\mathbf{AWE}}}_{i}(I+1)'\,\mathbb{1} &= \mathbb{1}'\left(\prod_{l=I-i}^{J-1}\,\mathrm{D}(\widehat{f}_{l}^{I}) - \prod_{l=I-i+1}^{J-1}\,\mathrm{D}(\widehat{h}_{l})\,\mathrm{D}(f_{I-i})\right)C_{i,I-i} \\ &\times C'_{i,I-i}\left(\prod_{l=I-i}^{J-1}\,\mathrm{D}(\widehat{f}_{l}^{I}) - \prod_{l=I-i+1}^{J-1}\,\mathrm{D}(\widehat{h}_{l})\,\mathrm{D}(f_{I-i})\right)\,\mathbb{1}, \end{split}$$

wobei \hat{h}_l durch (4.11) definiert ist. Mit (3.1) erhalten wir daraus weiter

$$\widehat{\mathbb{1}'\widehat{AWE}}_{i}(I+1)\widehat{\widehat{AWE}}_{i}(I+1)' \mathbb{1}$$

= $\mathbb{1}' D(C_{i,I-i}) (\hat{g}_{i|J} - D(f_{I-i})\hat{g}_{i|J}^{*}) (\hat{g}_{i|J} - D(f_{I-i})\hat{g}_{i|J}^{*})' D(C_{i,I-i}) \mathbb{1}.$ (5.5)

Dabei sind $\hat{g}_{i|j}$ und $\hat{g}_{i|j}^*$ für $j = I - i + 1, \dots, J$ durch

$$\hat{\boldsymbol{g}}_{i|j} = \mathbf{D}(\hat{\boldsymbol{f}}_{I-i}^{I}) \cdot \ldots \cdot \mathbf{D}(\hat{\boldsymbol{f}}_{j-1}^{I}) \mathbb{1}$$
$$\hat{\boldsymbol{g}}_{i|j}^{*} = \mathbf{D}(\hat{\boldsymbol{h}}_{I-i+1}) \cdot \ldots \cdot \mathbf{D}(\hat{\boldsymbol{h}}_{j-1}) \mathbb{1}$$

definiert. Aus (5.5) wird ersichtlich, dass das Hauptproblem in der Quantifizierung der Volatilität der Chain-Ladder-Schätzer und somit in der Berechnung der $N \times N$ -Matrix

$$\hat{\boldsymbol{g}}_{i|J}\hat{\boldsymbol{g}}_{i|J}' = \left(\prod_{j=I-i}^{J-1} \hat{f}_{j}^{I,(n)} \hat{f}_{j}^{I,(m)}\right)_{1 \le n,m \le N}$$
(5.6)

liegt (anlog $\hat{g}_{i|J}(\hat{g}_{i|J}^*)'$, $\hat{g}_{i|J}^*\hat{g}_{i|J}'$ und $\hat{g}_{i|J}^*(\hat{g}_{i|J}^*)'$). Dabei ist zu beachten, dass die Chain-Ladder-Schätzer \hat{f}_k^I und \hat{f}_l^I für $k \neq l$ zwar unkorreliert (vgl. Lemma 3.5 in Merz & Wüthrich [7]) aber nicht stochastisch unabhängig sind (vgl. Mack u. a. [4] und Wüthrich u. a. [12]). Daraus folgt, dass der Erwartungswert von (5.6) nicht berechnet werden kann, da im Allgemeinen

$$\mathbf{E}\left[\prod_{j=I-i}^{J-1} \hat{f}_j^{I,(n)} \hat{f}_j^{I,(m)}\right] \neq \prod_{j=I-i}^{J-1} \mathbf{E}\left[\hat{f}_j^{I,(n)} \hat{f}_j^{I,(m)}\right]$$

gilt. Wir approximieren deshalb die rechte Seite von (5.5) mit Hilfe des (multivariaten) bedingten Resampling-Ansatzes in Merz & Wüthrich [7] und berechnen dazu die Erwartungswerte

$$\prod_{j=I-i}^{J-1} \mathbb{E}\left[\left.\hat{f}_{j}^{I,(n)}\hat{f}_{j}^{I,(m)}\right| \ \mathcal{B}_{j}^{N}\right].$$

Hierzu benötigen wir die Volatilität von \hat{f}_j^I um den "wahren" Wert f_j , da die Schätzer \hat{f}_j^I der Chain-Ladder-Faktoren f_j nun als Zufallsvariablen aufgefasst werden. Im bedingten Resampling-Ansatz werden deshalb, gegeben $C_{i,j}$, durch

$$\tilde{C}_{i,j+1} = D(f_j)C_{i,j} + D(C_{i,j})^{1/2}D(\tilde{\epsilon}_{i,j+1})\sigma_j.$$
(5.7)

für alle $i \in \{0, ..., I\}$ und $j \in \{0, ..., J - 1\}$ Pseudo-Beobachtungen generiert, wobei

$$\tilde{\boldsymbol{\epsilon}}_{i,j+1}$$
 und $\boldsymbol{\epsilon}_{i,j+1}$ unabhängig und identisch verteilt sind. (5.8)

Für die so generierten Pseudo-Beobachtungen gilt $\tilde{C}_{i,j+1} \stackrel{(d)}{=} C_{i,j+1}$, bedingt gegeben $C_{i,j}$. Dies führt schließlich zu den folgenden Pseudo-Chain-Ladder-Schätzern

$$\hat{f}_{j}^{I} = \left(\sum_{i=0}^{I-j-1} D(C_{i,j})^{1/2} \Sigma_{j}^{-1} D(C_{i,j})\right)^{-1} \sum_{i=0}^{I-j-1} D(C_{i,j})^{1/2} \Sigma_{j}^{-1} D(C_{i,j}) D(C_{i,j})^{-1} \tilde{C}_{i,j+1}$$
$$= f_{j} + \left(D(C_{i,j})^{1/2} \Sigma_{j}^{-1} D(C_{i,j})\right)^{-1} \sum_{i=0}^{I-j-1} D(C_{i,j})^{1/2} \Sigma_{j}^{-1} D(\tilde{\epsilon}_{i,j+1}) \sigma_{j}.$$
(5.9)

Bemerkungen 5.5:

- In Gleichung (5.9) und im Folgenden verwenden wir für die Pseudo-Chain-Ladder-Schätzer ebenfalls die Notation \hat{f}_{j}^{I} .
- Das Ziel ist nun, die Volatilität der Pseudo-Chain-Ladder-Schätzer (5.9), gegeben \mathcal{D}_{I}^{N} , zu quantifizieren. Dabei ist zu beachten, dass die Pseudo-Beobachtungen { $\tilde{C}_{i,j} \mid i+j \leq I$ } und die Pseudo-Chain-Ladder-Schätzer $\hat{f}_{0}^{I}, \ldots, \hat{f}_{J-1}^{I}$ Zufallsvariablen sind, bedingt gegeben die Informationen \mathcal{D}_{I}^{N} . Außerdem sind die Beobachtungen $C_{i,j}$ und die neuen Residuen $\tilde{\epsilon}_{i,j}$ stochastisch unabhängig. Dies führt zu einer multiplikativen Struktur und erlaubt die analytische Herleitung eines Schätzers für den bedingten Schätzfehler.

Wir erhalten das folgende Lemma:

Lemma 5.6: Unter den Modellannahmen 3.1 und den Annahmen (5.7) und (5.8) gilt

- 1. Die Schätzer $\hat{f}_0^I, \ldots, \hat{f}_{J-1}^I$ sind, bedingt gegeben \mathcal{D}_I^N , unabhängig.
- 2. $\mathbb{E}\left[\left.\hat{f}_{j}^{I}\right| \mathcal{D}_{I}^{N}\right] = f_{j} \, f \ddot{u} r \, 0 \leq j \leq J 1.$
- 3. E $\left[\hat{f}_{j}^{I,(n)}\hat{f}_{j}^{I,(m)} \mid \mathcal{D}_{I}^{N}\right] = f_{j}^{(n)}f_{j}^{(m)} + W_{j}(n,m), wobei W_{j}(n,m) der(n,m) te Eintrag der N \times N-Matrix$

$$W_j = \left(\sum_{k=0}^{I-j-1} D(C_{k,j})^{1/2} \Sigma_j^{-1} D(C_{k,j})^{1/2}\right)^{-1}$$
(5.10)

ist.

- 4. E $\left[\hat{\boldsymbol{h}}_{j} \mid \mathcal{D}_{I}^{N}\right] = \boldsymbol{f}_{j} f \ddot{\boldsymbol{u}} r \, 0 \leq j \leq J 1.$
- 5. Die Schätzer $\hat{h}_0, \ldots, \hat{h}_{J-1}$ sind, bedingt gegeben \mathcal{D}_I^N , unabhängig.
- 6. Die Schätzer \hat{f}_{j}^{J} und \hat{h}_{k} sind für $j \neq k$, bedingt gegeben \mathcal{D}_{I}^{N} , unabhängig.

7. E $\left[\hat{h}_{j}^{(n)}\hat{h}_{j}^{(m)} \mid \mathcal{D}_{I}^{N}\right] = h_{j}^{(n)}h_{j}^{(m)} + H_{j}(n,m)$, wobei $H_{j}(n,m)$ der (n,m)-te Eintrag der $N \times N$ -Matrix H_{j} ist. Die Matrix H_{j} ist gegeben durch:

$$H_j = (S_j^{I+1})^{-1} S_j^{I} W_j S_j^{I} (S_j^{I+1})^{-1}$$
(5.11)

8. E $\left[\hat{f}_{j}^{I,(n)}\hat{h}_{j}^{(m)} \mid \mathcal{D}_{I}^{N}\right] = f_{j}^{(n)}h_{j}^{(m)} + K_{j}(n,m), \text{ wobei } K_{j}(n,m) \text{ der } (n,m) \text{ -te Eintrag der } N \times N - Matrix$

$$K_j = W_j S_j^I (S_j^{I+1})^{-1}.$$
(5.12)

ist.

Beweis. Der Beweis ist eine einfache Übungsaufgabe. Siehe auch den Beweis des Lemmas 4.6 in Merz & Wüthrich [7]. $\hfill \Box$

Anwendung des Erwartungswert
operators E [· | \mathcal{D}_I^N] auf die rechte Seite von (5.5) liefert mit Lemma 5.6 den folgenden Ausdruck für den bedingten Schätzfehler

$$\mathbb{1}' D(C_{i,I-i}) E\left[\left(\hat{g}_{i|J} - D(f_{I-i}) \hat{g}_{i|J}^{*} \right) \left(\hat{g}_{i|J} - D(f_{I-i}) \hat{g}_{i|J}^{*} \right)' \middle| \mathcal{D}_{I}^{N} \right] D(C_{i,I-i}) \mathbb{1}$$

$$= \mathbb{1}' D(C_{i,I-i}) \left(E\left[\hat{g}_{i|J} \hat{g}_{i|J}' \middle| \mathcal{D}_{I}^{N} \right] + D(f_{I-i}) E\left[\hat{g}_{i|J}^{*} (\hat{g}_{i|J}^{*})' \middle| \mathcal{D}_{I}^{N} \right] D(f_{I-i})$$

$$- E\left[\hat{g}_{i|J} (\hat{g}_{i|J}^{*})' \middle| \mathcal{D}_{I}^{N} \right] D(f_{I-i}) - D(f_{I-i}) E\left[\hat{g}_{i|J}^{*} \hat{g}_{i|J}' \middle| \mathcal{D}_{I}^{N} \right] \right) D(C_{i,I-i}) \mathbb{1}$$

$$= \mathbb{1}' D(C_{i,I-i}) \left(\Omega_{i,J}^{(n,m)} \right)_{1 \le n,m \le N} D(C_{i,I-i}) \mathbb{1},$$

$$(5.13)$$

wobei die $N \times N$ -Matrix $\left(\Omega_{i,J}^{(n,m)}\right)_{1 \le n,m \le N}$ durch

$$\begin{aligned} \left(\Omega_{iJ}^{(n,m)}\right)_{1 \le n,m \le N} &= \mathbb{E}\left[\hat{\boldsymbol{g}}_{i|J}\hat{\boldsymbol{g}}_{i|J}' \mid \mathcal{D}_{I}^{N}\right] + \mathbb{D}(\boldsymbol{f}_{I-i}) \mathbb{E}\left[\hat{\boldsymbol{g}}_{i|J}^{*}(\hat{\boldsymbol{g}}_{i|J}^{*})' \mid \mathcal{D}_{I}^{N}\right] \mathbb{D}(\boldsymbol{f}_{I-i}) \\ &- \mathbb{E}\left[\hat{\boldsymbol{g}}_{i|J}(\hat{\boldsymbol{g}}_{i|J}^{*})' \mid \mathcal{D}_{I}^{N}\right] \mathbb{D}(\boldsymbol{f}_{I-i}) - \mathbb{D}(\boldsymbol{f}_{I-i}) \mathbb{E}\left[\hat{\boldsymbol{g}}_{i|J}^{*}\hat{\boldsymbol{g}}_{i|J}' \mid \mathcal{D}_{I}^{N}\right] \end{aligned}$$

definiert ist. Die einzelnen Elemente dieser Matrix sind definiert durch

$$\Omega_{i,J}^{(n,m)} = \prod_{j=I-i}^{J-1} \left(f_j^{(n)} f_j^{(m)} + W_j(n,m) \right) + f_{I-i}^{(n)} f_{I-i}^{(m)} \prod_{j=I-i+1}^{J-1} \left(h_j^{(n)} h_j^{(m)} + H_j(n,m) \right) - f_{I-i}^{(n)} f_{I-i}^{(m)} \left[\prod_{j=I-i+1}^{J-1} \left(f_j^{(n)} h_j^{(m)} + K_j(n,m) \right) + \prod_{j=I-i+1}^{J-1} \left(f_j^{(m)} h_j^{(n)} + K_j(n,m) \right) \right].$$
(5.14)

Ersetzt man nun die Parameter auf der rechten Seite von (5.4) und (5.14) durch ihre Schätzer (siehe Abschnitt 6), so erhalten wir den folgenden Schätzer für den bedingten Prognosefehler eines Anfalljahres:

Ergebnis 5.7 (Bedingter Prognosefehler für einzelne Anfalljahre): Unter den Modellannahmen 3.1 erhält man für den bedingten Prognosefehler im Anfalljahr $i \in \{1, ..., I\}$ den Schätzer

$$\widehat{\text{MSE}}\left(\sum_{n=1}^{N}\widehat{\widehat{\text{AWE}}}_{i}(I+1)^{(n)}\right)$$

$$=\underbrace{\widehat{C}_{i,J}^{I} D(C_{i,I-i})^{-1/2} D(\widehat{f}_{I-i}^{I})^{-1} \widehat{\Sigma}_{I-i} D(\widehat{f}_{I-i}^{I})^{-1} D(C_{i,I-i})^{-1/2} \widehat{C}_{i,J}^{I}}_{Schätzer für die bedingte Prozeßvarianz}$$

$$+\underbrace{\mathbb{1}' D(C_{i,I-i}) \left(\widehat{\Omega}_{i,J}^{(n,m)}\right)_{1 \le n,m \le N} D(C_{i,I-i}) \mathbb{1}}_{Schätzer für den bedingten Schätzfehler}.$$
(5.15)

Dabei gilt

$$\hat{\Omega}_{i,J}^{(n,m)} = \prod_{j=I-i}^{J-1} \left(\hat{f}_{j}^{I,(n)} \hat{f}_{j}^{I,(m)} + \hat{W}_{j}(n,m) \right) + \hat{f}_{I-i}^{I,(n)} \hat{f}_{I-i}^{I,(m)} \prod_{j=I-i+1}^{J-1} \left(\hat{h}_{j}^{I,(n)} \hat{h}_{j}^{I,(m)} + \hat{H}_{j}(n,m) \right) \\ - \hat{f}_{I-i}^{I,(n)} \hat{f}_{I-i}^{I,(m)} \left[\prod_{j=I-i+1}^{J-1} \left(\hat{f}_{j}^{I,(n)} \hat{h}_{j}^{I,(m)} + \hat{K}_{j}(n,m) \right) + \prod_{j=I-i+1}^{J-1} \left(\hat{f}_{j}^{I,(m)} \hat{h}_{j}^{I,(n)} + \hat{K}_{j}(n,m) \right) \right], \quad (5.16)$$

wobei $\hat{W}_j(n,m)$, $\hat{H}_j(n,m)$ und $\hat{K}_j(n,m)$ die Elemente der Schätzungen \hat{W}_j , \hat{H}_j und \hat{K}_j der $N \times N$ -Matrizen W_j , H_j und K_j sind (siehe Abschnitt 6).

Bemerkung 5.8: Im Fall N = 1 erhalten wir aus (5.15) und (5.16) den bedingten Schätzfehler bzw. den Schätzer für den bedingten Prognosefehler des Anfalljahres *i* für ein einzelnes Run-Off-Portfolio (vgl. Formeln (4.15) und (4.29) in Merz & Wüthrich [6]).

5.2 Bedingter Prognosefehler für aggregierte Anfalljahre

In diesem Abschnitt ermitteln wir analytisch einen Schätzer für den bedingten Prognosefehler für aggregierte Anfalljahre.

Definition 5.9 (Bedingter Prognosefehler für aggregierte Anfalljahre): Der bedingte Prognosefehler für aggregierte Anfalljahr ist definiert durch

$$MSE\left(\sum_{i=1}^{I}\sum_{n=1}^{N}\widehat{\widehat{AWE}}_{i}(I+1)^{(n)}\right)$$

= $E\left[\left(\sum_{i=1}^{I}\sum_{n=1}^{N}\widehat{\widehat{AWE}}_{i}(I+1)^{(n)} - \sum_{i=1}^{I}\sum_{n=1}^{N}AWE_{i}(I+1)^{(n)}\right)^{2} \middle| \mathcal{D}_{I}^{N}\right].$ (5.17)

Mit (5.2) erhalten wir für (5.17) die Darstellung

$$MSE\left(\sum_{i=1}^{I}\sum_{n=1}^{N}\widehat{\widehat{AWE}}_{i}(I+1)^{(n)}\right)$$

= $\mathbb{1}'E\left[\left(\widehat{\widehat{AWE}}(I+1) - AWE(I+1)\right)\left(\widehat{\widehat{AWE}}(I+1) - AWE(I+1)\right)'\right| \mathcal{D}_{I}^{N}\right] \mathbb{1}.$

Bei der Berechnung eines Schätzers für den bedingten Prognosefehler für aggregierte Anfalljahre ist zu beachten, dass $\widehat{AWE}_i(I+1)$ und $\widehat{AWE}_k(I+1)$ zum Teil auf den gleichen Beobachtungen basieren und daher abhängig sind. Wir betrachten zuerst lediglich zwei Anfalljahre *i* und *k* mit $1 \le i < k \le I$ und erhalten

$$MSE\left(\sum_{n=1}^{N}\widehat{AWE}_{i}(I+1)^{(n)} + \sum_{n=1}^{N}\widehat{AWE}_{k}(I+1)^{(n)}\right)$$

= 1' Var $\left(AWE_{i}(I+1) + AWE_{k}(I+1) \mid \mathcal{D}_{I}^{N}\right)$ 1
+ 1' $\left(\widehat{AWE}_{i}(I+1) + \widehat{AWE}_{k}(I+1)\right) \left(\widehat{AWE}_{i}(I+1) + \widehat{AWE}_{k}(I+1)\right)'$ 1.

Aufgrund der Unabhängigkeit verschiedener Anfalljahre (vgl. Modellannahmen 3.1) erhalten wir für die bedingte Prozessvarianz

$$\mathbb{1}' \operatorname{Var} \left(\operatorname{AWE}_{i}(I+1) + \operatorname{AWE}_{k}(I+1) \mid \mathcal{D}_{I}^{N} \right) \mathbb{1}$$

= $\mathbb{1}' \operatorname{Var} \left(\operatorname{AWE}_{i}(I+1) \mid \mathcal{D}_{I}^{N} \right) \mathbb{1} + \mathbb{1}' \operatorname{Var} \left(\operatorname{AWE}_{k}(I+1) \mid \mathcal{D}_{I}^{N} \right) \mathbb{1}.$

Für den bedingten Schätzfehler erhalten wir

$$\mathbb{1}' \left(\widehat{\widehat{AWE}}_i(I+1) + \widehat{\widehat{AWE}}_k(I+1) \right) \left(\widehat{\widehat{AWE}}_i(I+1) + \widehat{\widehat{AWE}}_k(I+1) \right)' \mathbb{1}$$

$$= \mathbb{1}' \widehat{\widehat{AWE}}_i(I+1) \widehat{\widehat{AWE}}_i(I+1)' \mathbb{1} + \mathbb{1}' \widehat{\widehat{AWE}}_k(I+1) \widehat{\widehat{AWE}}_k(I+1)' \mathbb{1}$$

$$+ 2 \mathbb{1}' \widehat{\widehat{AWE}}_i(I+1) \widehat{\widehat{AWE}}_k(I+1)' \mathbb{1}.$$

Wir erhalten somit für den bedingten Prognosefehler zweier aggregierter Anfalljahrei und k

$$MSE\left(\sum_{n=1}^{N}\widehat{\widehat{AWE}}_{i}(I+1)^{(n)} + \sum_{n=1}^{N}\widehat{\widehat{AWE}}_{k}(I+1)^{(n)}\right)$$
$$= MSE\left(\sum_{n=1}^{N}\widehat{\widehat{AWE}}_{i}(I+1)^{(n)}\right) + MSE\left(\sum_{n=1}^{N}\widehat{\widehat{AWE}}_{k}(I+1)^{(n)}\right)$$
$$+ 2\,\mathbb{1}'\widehat{\widehat{AWE}}_{i}(I+1)\widehat{\widehat{AWE}}_{k}(I+1)'\,\mathbb{1}.$$

Zusätzlich zu den beiden bedingten Prognosefehler einzelner Anfalljahre erhalten wir zweimal den Mischterm

$$\widehat{1'\widehat{AWE}}_{i}(I+1)\widehat{\widehat{AWE}}_{k}(I+1) \,\mathbb{1}$$

= $\mathbb{1}' D(C_{i,I-i}) \left(\hat{g}_{i|J} - D(f_{I-i}) \hat{g}_{i|J}^{*} \right) \left(\hat{g}_{k|J} - D(f_{I-k}) \hat{g}_{k|J}^{*} \right)' D(C_{k,I-k}) \,\mathbb{1}.$ (5.18)

Anwendung des Erwartungswertoperators E [$\cdot \mid \mathcal{D}_I^N$] auf die rechte Seite von (5.18) und erneute Anwendung des im letzten Abschnitt beschriebenen bedingten Resampling-Ansatz liefert (vgl. (5.13) für einzelne Anfalljahre)

$$\begin{split} & \mathbb{1}' \operatorname{D}(C_{i,I-i}) \operatorname{E} \left[(\hat{g}_{i|J} - \operatorname{D}(f_{I-i}) \hat{g}_{i|J}^{*}) (\hat{g}_{k|J} - \operatorname{D}(f_{I-k}) \hat{g}_{k|J}^{*}) \mid \mathcal{D}_{I}^{N} \right] \mathbb{1} \\ &= \mathbb{1}' \operatorname{D}(C_{i,I-i}) \left(\operatorname{E} \left[\hat{g}_{i|J} \hat{g}_{k|J}^{\prime} \mid \mathcal{D}_{I}^{N} \right] + \operatorname{D}(f_{I-i}) \operatorname{E} \left[\hat{g}_{i|J}^{*} (\hat{g}_{k|J}^{*})^{\prime} \mid \mathcal{D}_{I}^{N} \right] \operatorname{D}(f_{I-k}) \\ &- \operatorname{E} \left[\hat{g}_{i|J} (\hat{g}_{k|J}^{*})^{\prime} \mid \mathcal{D}_{I}^{N} \right] \operatorname{D}(f_{I-k}) - \operatorname{D}(f_{I-i}) \operatorname{E} \left[\hat{g}_{i|J}^{*} \hat{g}_{k|J}^{\prime} \mid \mathcal{D}_{I}^{N} \right] \right) \operatorname{D}(C_{k,I-k}) \mathbb{1} \\ &= \mathbb{1}' \operatorname{D}(C_{i,I-i}) \left(\operatorname{E} \left[\hat{g}_{i|J} \hat{g}_{i|J}^{\prime} \mid \mathcal{D}_{I}^{N} \right] + \operatorname{D}(f_{I-i}) \operatorname{E} \left[\hat{g}_{i|J}^{*} (\hat{g}_{i|J}^{*})^{\prime} \mid \mathcal{D}_{I}^{N} \right] \operatorname{D}(f_{I-i}) \\ &- \operatorname{E} \left[\hat{g}_{i|J} (\hat{g}_{i|J}^{*})^{\prime} \operatorname{D}(\hat{h}_{I-i}) \mid \mathcal{D}_{I}^{N} \right] - \operatorname{D}(f_{I-i}) \operatorname{E} \left[\hat{g}_{i|J}^{*} \hat{g}_{i|J}^{\prime} \mid \mathcal{D}_{I}^{N} \right] \right) \prod_{j=I-k}^{I-i-1} \operatorname{D}(f_{j}) \operatorname{D}(C_{k,I-k}) \mathbb{1} \\ &= \mathbb{1}' \operatorname{D}(C_{i,I-i}) \left(\Theta_{i,J}^{(n,m)} \right)_{1 \le n,m \le N} \prod_{j=I-k}^{I-i-1} \operatorname{D}(f_{j}) \operatorname{D}(C_{k,I-k}) \mathbb{1}, \end{split}$$

wobei die $N \times N$ -Matrix $\left(\Theta_{i,J}^{(n,m)}\right)_{1 \le n,m \le N}$ definiert ist durch

$$\begin{split} \left(\Theta_{i,J}^{(n,m)} \right)_{1 \le n,m \le N} &= \mathbb{E} \left[\hat{\boldsymbol{g}}_{i|J} \hat{\boldsymbol{g}}_{i|J}' \mid \mathcal{D}_{I}^{N} \right] + \mathbb{D}(\boldsymbol{f}_{I-i}) \mathbb{E} \left[\hat{\boldsymbol{g}}_{i|J}^{*} (\hat{\boldsymbol{g}}_{i|J}^{*})' \mid \mathcal{D}_{I}^{N} \right] \mathbb{D}(\boldsymbol{f}_{I-i}) \\ &- \mathbb{E} \left[\hat{\boldsymbol{g}}_{i|J} (\hat{\boldsymbol{g}}_{i|J}^{*})' \mathbb{D}(\hat{\boldsymbol{h}}_{I-i}) \mid \mathcal{D}_{I}^{N} \right] - \mathbb{D}(\boldsymbol{f}_{I-i}) \mathbb{E} \left[\hat{\boldsymbol{g}}_{i|J}^{*} \hat{\boldsymbol{g}}_{i|J}' \mid \mathcal{D}_{I}^{N} \right]. \end{aligned}$$

Die einzelnen Einträge dieser Matrix sind gegeben durch

$$\begin{split} \Theta_{i,J}^{(n,m)} &= \prod_{j=I-i}^{J-1} \left(f_j^{(n)} f_j^{(m)} + W_j(n,m) \right) + f_{I-i}^{(n)} f_{I-i}^{(m)} \prod_{j=I-i+1}^{J-1} \left(h_j^{(n)} h_j^{(m)} + H_j(n,m) \right) \\ &- \prod_{j=I-i}^{J-1} \left(f_j^{(n)} h_j^{(m)} + K_j(n,m) \right) - f_{I-i}^{(n)} f_{I-i}^{(m)} \prod_{j=I-i+1}^{J-1} \left(f_j^{(m)} h_j^{(n)} + K_j(n,m) \right) . \end{split}$$

Nach Ersetzen der Parameter durch die entsprechenden Parameterschätzungen (siehe Abschnitt 6) erhalten wir schließlich den folgenden Schätzer für den bedingten Prognosefehler für aggregierte Anfalljahre:

Ergebnis 5.10 (Bedingter Prognosefehler für aggregierte Anfalljahre): Unter den Modellannahmen 3.1 erhält man für den bedingten Prognosefehler für aggregierte Anfalljahre den Schätzer

$$\widehat{\mathrm{MSE}}\left(\sum_{i=1}^{I}\sum_{n=1}^{N}\widehat{\widehat{\mathrm{AWE}}}_{i}(I+1)^{(n)}\right) = \sum_{i=1}^{I}\widehat{\mathrm{MSE}}\left(\sum_{n=1}^{N}\widehat{\widehat{\mathrm{AWE}}}_{i}(I+1)^{(n)}\right)$$
$$+ 2\sum_{1\leq i< k\leq I}\mathbbm{1}' \mathrm{D}(C_{i,I-i})\left(\widehat{\Theta}_{i,J}^{(n,m)}\right)_{1\leq n,m\leq N}\prod_{j=I-k}^{I-i-1}\mathrm{D}(\widehat{f}_{j}^{I}) \mathrm{D}(C_{k,I-k})\mathbbm{1}.$$

Dabei gilt

$$\hat{\Theta}_{i,J}^{(n,m)} = \prod_{j=I-i}^{J-1} \left(\hat{f}_j^{I,(n)} \hat{f}_j^{I,(m)} + \hat{W}_j(n,m) \right) + \hat{f}_{I-i}^{I,(n)} \hat{f}_{I-i}^{I,(m)} \prod_{j=I-i+1}^{J-1} \left(\hat{h}_j^{I,(n)} \hat{h}_j^{I,(m)} + \hat{H}_j(n,m) \right) \\ - \prod_{j=I-i}^{J-1} \left(\hat{f}_j^{I,(n)} \hat{h}_j^{I,(m)} + \hat{K}_j(n,m) \right) - \hat{f}_{I-i}^{I,(n)} \hat{f}_{I-i}^{I,(m)} \prod_{j=I-i+1}^{J-1} \left(\hat{f}_j^{I,(m)} \hat{h}_j^{I,(n)} + \hat{K}_j(n,m) \right)$$

wobei $\hat{W}_j(n,m)$, $\hat{H}_j(n,m)$ und $\hat{K}_j(n,m)$ die Elemente der Schätzungen \hat{W}_j , \hat{H}_j und \hat{K}_j der $N \times N$ -Matrizen W_j , H_j und K_j sind (siehe Abschnitt 6).

6 Parameterschätzung

Die Schätzer für die Chain-Ladder-Faktoren f_j sind in (3.5) gegeben. Jedoch im Gegensatz zum univariaten Chain-Ladder-Verfahren sind dies nur implizite Schätzer. Denn die multivariaten Chain-Ladder-Schätzer \hat{f}_j^I hängen von den Parametern σ_j und E [$\epsilon_{i,j+1}\epsilon'_{i,j+1}$] ab, welche ihrerseits wieder von den multivariaten Chain-Ladder-Schätzern \hat{f}_j^I abhängen (siehe (3.5) und (3.4)). Merz & Wüthrich [7] schlagen deshalb eine iterative Schätzung dieser Parameter vor. Ausgangspunkt sind die Startwerte $\hat{f}_j^{I,(0)}$ (siehe (3.12)), welche beste erwartungstreue Schätzer für f_j darstellen, falls die *N* Run-Off Portfolios unkorreliert sind. Im Falle korrelierter Run-Off Portfolios ist dies jedoch nicht mehr der Fall (siehe Bemerkungen 3.3). Mit den Startwerten $\hat{f}_j^{I,(0)}$ erhalten wir erste Schätzungen $\hat{\sigma}_j^{(1)}$ und $\hat{E} [\epsilon_{i,j+1} \epsilon'_{i,j+1}]^{(1)}$ für σ_j bzw. $E [\epsilon_{i,j+1} \epsilon'_{i,j+1}]$. Diese werden verwendet um eine verbesserte Schätzung $\hat{f}_j^{I,(1)}$ für f_j zu erhalten. Der neue Schätzer $\hat{f}_j^{I,(1)}$ wird wiederum herangezogen, um verbesserte Schätzungen $\hat{\sigma}_j^{(2)}$ und $\hat{E} [\epsilon_{i,j+1} \epsilon'_{i,j+1}]^{(2)}$ für σ_j bzw. $E [\epsilon_{i,j+1} \epsilon'_{i,j+1}]$ zu berechnen. Diese iterative Schätzprozedur wird solange fortgeführt, bis die Schätzer $\hat{f}_j^{(k-1)}$, $\hat{\sigma}_j^{(k)}$ und $\hat{E} [\epsilon_{i,j+1} \epsilon'_{i,j+1}]^{(k)}$ ein vorgegebenes Abbruchkriterium erfüllen (vgl. Merz & Wüthrich [7]).

Motiviert durch (3.4), (3.6), (3.8), (5.10), (5.11) und (5.12) werden diese Parameterschätzungen schließlich zur Schätzung von Σ_j , S_j^I , S_j^{I+1} , W_j , H_j und K_j verwendet:

$$\begin{split} \hat{\Sigma}_{j} &= D(\hat{\sigma}_{j}) \, \hat{E} \left[\boldsymbol{\epsilon}_{i,j+1} \boldsymbol{\epsilon}_{i,j+1}' \right] D(\hat{\sigma}_{j}), \\ \hat{S}_{j}^{I} &= \sum_{i=0}^{I-j-1} D(C_{i,j})^{1/2} (\hat{\Sigma}_{j})^{-1} D(C_{i,j})^{1/2}, \\ \hat{S}_{j}^{I+1} &= \sum_{i=0}^{I-j} D(C_{i,j})^{1/2} (\hat{\Sigma}_{j})^{-1} D(C_{i,j})^{1/2}, \\ \hat{W}_{j} &= \left(\sum_{k=0}^{I-j-1} D(C_{k,j})^{1/2} (\hat{\Sigma}_{j})^{-1} D(C_{k,j})^{1/2} \right)^{-1} = \left(\hat{S}_{j}^{I} \right)^{-1}, \\ \hat{H}_{j} &= \left(\hat{S}_{j}^{I+1} \right)^{-1} \hat{S}_{j}^{I} \hat{W}_{j} \hat{S}_{j}^{I} \left(\hat{S}_{j}^{I+1} \right)^{-1} = \left(\hat{S}_{j}^{I+1} \right)^{-1} \hat{S}_{j}^{I} \left(\hat{S}_{j}^{I+1} \right)^{-1}, \\ \hat{K}_{j} &= \hat{W}_{j} \hat{S}_{j}^{I} \left(\hat{S}_{j}^{I+1} \right)^{-1} = \left(\hat{S}_{j}^{I+1} \right)^{-1}. \end{split}$$

7 Beispiel

Wir verwenden die Daten aus Braun [1] und Merz & Wüthrich [5, 7]. In diesen Arbeiten werden zwei Run-Off-Portfolios A (allgemeine Haftpflicht) und B (Autohaftpflicht) betrachtet. Somit gilt N = 2 und die Abwicklungsdaten sind in den Tabellen 1 und 2 in Braun [1] enthalten. Als Parameterschätzungen $\hat{f}_{j}^{(k-1)}$, $\hat{\sigma}_{j}^{(k)}$ und $\hat{\rho}_{j}^{(1,2)(k)}$ für k = 1,2,3 verwenden wir die Werte in Table 3 in Merz & Wüthrich [7].

Tabelle 1 enthält die über die Anfalljahre aggregierte Schadenreserve und die Schätzungen für die bedingte Prozessvarianz, den bedingten Schätzfehler und den bedingten Prognosefehler des Prädiktors für den über die Anfalljahre und die Run-Off-Portfolios aggregierten Endschadenstand $\sum_{i=1}^{I} \sum_{n=1}^{2} \widehat{C_{i,J}^{(n)}}$ (vgl. Merz & Wüthrich [7]). Das heißt diese Werte quantifizieren die klassische langfristige Unsicherheit bis alle Schäden vollständig abgewickelt sind.

	Run-Off Subportfolio				Run-Off Portfolio A+B						
	А		В		Iteration $k = 1$		Iteration $k = 2$		Iteration $k = 3$		
Geschätzte Reserven	6,155,261		2,063,612		8,218,874		8,215,227		8,215,350		
√Prozessvarianz	330,485	5.4%	134,676	6.5%	396,731	4.8%	396,799	4.8%	396,805	4.8%	
√Schätzfehler	270,878	5.4%	91,599	4.4%	313,122	3.8%	313,071	3.8%	313,074	3.8%	
\sqrt{MSE}	427,311	6.9%	162,874	7.9%	505,412	6.1%	505,433	6.2%	505,440	6.2%	

Tabelle 1: Schadenreserve und Schätzung für den bedingten Prognosefehler des Prädiktors für den Endschadenstand bei aggregierten Anfalljahren.

Tabelle 2 enthält das über die Anfalljahre aggregierte Abwicklungsergebnis und die Schätzungen für die bedingte Prozessvarianz, den bedingten Schätzfehler und den bedingten Prognosefehler des Prädiktors für das über die Anfalljahre aggregierte Abwicklungsergebnis. Somit quantifizieren diese Werte das einjährige Risiko für das Kalenderjahr (I, I + 1].

	Ru	Run-Off Subportfolio				Run-Off Portfolio A+B						
	A		В		Iteration $k = 1$		Iteration $k = 2$		Iteration $k = 3$			
erwartetes AWE	0		0		0		0		0			
$\sqrt{\text{Prozessvarianz}}$	264,002	4.3%	115,609	5.6%	317,374	3.9%	317,298	3.9%	317,298	3.9%		
√Schätzfehler	134,779	2.2%	50,367	2.4%	156,923	1.9%	156,883	1.9%	156,883	1.9%		
$\sqrt{\text{MSE}}$	296,416	4.8%	126,104	6.1%	354,049	4.3%	353,964	4.3%	353,964	4.3%		

Tabelle 2: Abwicklungsergebnis und Schätzung für den bedingten Prognosefehler des Prädiktors für das Abwicklungsergebnis bei aggregierten Anfalljahren.

Literatur

- [1] Braun, C. (2004): The prediction error of the chain ladder method applied to correlated run off triangles. *ASTIN Bulletin*, 34(2): 399–423.
- [2] Bundesamt für Privatversicherungen (2006): Swiss Solvency Test. Schweiz.
- [3] Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2): 213–225.
- [4] Mack, T., Quarg, G. & Braun, C. (2006): The mean square error of prediction in the chain ladder reserving method a comment. *ASTIN Bulletin*, 36(2): 543–552.
- [5] Merz, M. & Wüthrich, M. V. (2008): Prediction error of the chain ladder reserving method applied to correlated run off triangles. *Annals of Actuarial Science*, 2(1): 25–50.
- [6] Merz, M. & Wüthrich, M. V. (2007): Prediction error of the expected claims development result in the chain ladder method. *Bulletin of Swiss Association of Actuaries*, 1: 117–137.
- [7] Merz, M. & Wüthrich, M. V. (2008): Prediction error of the multivariate chain ladder reserving method. *North American Actuarial Journal*, 12(2): 175–197.

- [8] Pröhl, C. & Schmidt, K. D. (2005): Multivariate chain-ladder. Dresdner Schriften zur Versicherungsmathematik, 3.
- [9] Sandström, A. (2007): Solvency, a historical review and some pragmatic solutions. *Bulletin* of Swiss Association of Actuaries, 1: 11–34.
- [10] Schmidt, K. D. (2006): Optimal and additive loss reserving for dependent lines of business. *CAS E-Forum*, 319–351.
- [11] Wüthrich, M. V. & Merz, M. (2008): *Stochastic Claims Reserving Methods in Insurance*. Chichester: Wiley Finance.
- [12] Wüthrich, M. V., Merz, M. & Bühlmann, H. (2008): Bounds on the estimation error in the chain ladder method. *Scandinavian Actuarial Journal*, 2008: 283–300.

Paper C

Bootstrappingmethoden zur Bestimmung des einjährigen Reserverisikos für Solvenzbetrachtungen bei mehreren abhängigen Run-Off-Portfolios

Jochen Heberle, Luis Huergo, Michael Merz

Bootstrappingmethoden zur Bestimmung des einjährigen Reserverisikos für Solvenzbetrachtungen bei mehreren abhängigen Run-Off-Portfolios

Jochen Heberle

Luis Huergo

Michael Merz

Zusammenfassung

In diesem Artikel steht nicht die Schadenreserve und ihre Unsicherheit im Mittelpunkt, sondern die Quantifizierung der Unsicherheit im sog. Abwicklungsergebnis für das nächste Geschäftsjahr, das als die Differenz der "Best Estimate" Prognosen für den Endschadenstand am Anfang und am Ende des betrachteten Geschäftsjahres definiert ist. Das heißt es steht die Quantifizierung des einjährigen Reserverisikos für das kommende Geschäftsjahr im Mittelpunkt. Ein zweiter Aspekt des Artikels ist es, dass gezeigt werden soll, wie das einjährige Reserverisikos mittels moderner Bootstrapping-Methoden für mehrere abhängige Run-Off-Portfolios quantifiziert werden kann. Es wird gezeigt wie die vollständige Verteilung des Abwicklungsergebnisses mehrerer abhängiger Versicherungsportfolios simuliert werden kann.

1 Einleitung

In der Nichtleben-Versicherung stellen die Schadenrückstellungen für IBNR- und RBNS-Schäden i.d.R. den größten versicherungstechnischen Passivposten in der Bilanz dar. Die Bewertung von Schadenrückstellungen und die Messung des Reserverisikos sind somit von großer wirtschaftlicher Bedeutung für jeden Nichtleben-Versicherer und es ist nicht verwunderlich, dass in den letzten 15 Jahren sehr viele verschiedene (stochastische) Schadenreservierungsverfahren entwickelt worden sind (vgl. z.B. Wüthrich & Merz [25]).

Im Gegensatz zu den meisten anderen Artikeln über das Reserverisiko steht jedoch in dieser Arbeit nicht die Schadenreserve und ihre Unsicherheit im Mittelpunkt, sondern die Quantifizierung der Unsicherheit im sog. *Abwicklungsergebnis* für das nächste Geschäftsjahr. Somit ist das Ziel die Quantifizierung des *einjährigen Reserverisikos* für das kommende Geschäftsjahr. Die Messung des einjährigen Reserverisikos anstelle des *vollständigen Reserverisikos* bis alle Schäden vollständig abgewickelt sind, ist aufgrund neuer risikobasierter Solvenzanforderungen mit einjährigem Zeithorizont, wie z.B. Solvency II in der EU oder Swiss Solvency Test (SST) in der Schweiz, ein wichtiges und aktuelles Thema für alle Nichtleben-Versicherer (vgl. z.B. Bundesamt für Privatversicherungen [5], Sandström [22] und Wüthrich u. a. [26]). Dies wird auch durch die aktuelle Studie "AISAM-ACME study on non-life long tail liabilities: Reserve risk and risk margin assessment under Solvency II (2007)" belegt (siehe AISAM-ACME [1]).

Klassische Formeln zur Messung des Reserverisikos sind nicht konsistent mit einem Einjahres-Horizont wie er in Solvency II und dem SST betrachtet wird. Sie quantifizieren das vollständige Reserverisiko für alle zukünftigen Kalenderjahre und überschätzen daher das einjährige Reserverisiko und somit auch das erforderliche Solvenzkapital für das kommende Kalenderjahr (vgl. AISAM-ACME [1]).

Das zweite Ziel dieses Artikels ist es, zu zeigen, wie das einjährige Reserverisiko mittels *Bootstrapping-Methoden* für mehrere korrelierte Run-Off Portfolios quantifiziert werden kann. Genauer soll gezeigt werden, wie die *prädiktive Verteilung* des Abwicklungsergebnisses des kommenden Kalenderjahres für mehrere abhängige Versicherungsportfolios simuliert werden kann. Dies erlaubt neben der Schätzung des Erwartungswerts und der Varianz des Abwicklungsergebnisses u.a. auch die Schätzung von Quantilen und Risikomaßen. Solche statistischen "Kennziffern" sind für moderne risikobasierte Solvenzvorschriften sowie die wert- und risikobasierte Steuerung von Versicherungsunternehmen von hoher Bedeutung (vgl. z.B. Nguyen [19]).

1.1 Die prädiktive Verteilung des einjährigen Reserverisikos

Im Folgenden wird gezeigt wie mit Hilfe von Bootstrapping-Methoden im Rahmen des in Heberle u. a. [10] betrachteten multivariaten *Chain-Ladder-Verfahrens* und Modellrahmens für mehrere korrelierte Run-Off Portfolios die prädiktive Verteilung des einjährigen Reserverisikos für das kommende Kalenderjahr ermittelt werden kann.

Die prädiktive Verteilung setzt sich aus den beiden Unsicherheitsquellen

- 1. Prozessvarianz und
- 2. Schätzfehler

zusammen. Die Prozessvarianz beschreibt dabei die Varianz innerhalb des zu Grunde gelegten stochastischen Modells und der Schätzfehler resultiert aus der Schätzung der Modellparameter.

Die prädiktive Verteilung des einjährigen Reserverisikos ermöglicht neben der (approximativen) Berechnung des Erwartungswertes und der Varianz auch die Bestimmung von

- Quantilen (z.B. Value-at-Risk) und
- Risikomaßen (z.B. Expected-Shortfall)

für das Abwicklungsergebnis im kommenden Kalenderjahr. Da jedoch in den meisten stochastischen Schadenreservierungsmodellen die Berechnung der prädiktiven Verteilung analytisch nicht möglich ist (vgl. Wüthrich & Merz [25]), ist man i.d.R. bei der Bestimmung der prädiktiven Verteilung auf den Einsatz von Simulationsmethoden, wie z.B. das Bootstrappingverfahren, angewiesen.

1.2 Einjähriges Reserverisiko für korrelierte Run-Off Portfolios

Die Betrachtung des Abwicklungsergebnisses für mehrere korrelierte Run-Off Portfolios ist motiviert durch:

- Bei der Anwendung eines speziellen Schadenreservierungsverfahrens in der Praxis (z.B. Chain-Ladder-Verfahren) muss zuvor sehr häufig das Portfolio in mehrere Subportfolios zerlegt werden, bevor die für eine sinnvolle Anwendung des Verfahrens notwendigen Homogenitätseigenschaften erfüllt sind. Dabei entstehen jedoch mehrere – i.d.R. stark korrelierte – Schadenabwicklungsdreiecke.
- 2. Ein multivariater Ansatz ermöglicht einen einheitlichen Ansatz zur Schätzung des Erwartungswerts, der Varianz, des Value-at-Risk, des Expected-Shortfall, des Abwicklungsergebnisses usw. für ein Gesamtportfolios bestehend aus mehreren korrelierten Subportfolios.
- Beobachtungen aus einem Run-Off Portfolio ermöglichen Rückschlüsse auf das Abwicklungsverhalten anderer – korrelierter – Run-Off Portfolios (z.B. Run-Off Portfolios von Klein- bzw. Großschäden).
- 4. Ein multivariater Ansatz erlaubt die Berücksichtigung der Abhängigkeiten zwischen den Run-Off Portfolios verschiedener Versicherungsbranchen (z.B. Körperschäden in der Autohaftpflicht- und der allgemeinen Haftpflichtversicherung).

Das klassische Problem der Schadenreservierung – d.h. die Bestimmung einer Prognose für den Endschadenstand bzw. die Schadenreserve sowie die Schätzung des damit verbundenen Prognosefehlers (vollständiges Reserverisiko) – für mehrere korrelierte Run-Off Portfolios wurde bereits in einer ganzen Reihe von Arbeiten analytisch untersucht (vgl. z.B. Braun [2], Schmidt [23], Pröhl & Schmidt [21], Hess u. a. [11] und Merz & Wüthrich [18, 16, 17]). In Brehm [3], Kirschner u. a. [13] und Taylor & McGuire [24] wird dagegen gezeigt, wie das Bootstrappingverfahren zur Prognose des Endschadenstands und zur Schätzung des Prognosefehlers für mehrere korrelierte Run-Off Portfolios herangezogen werden kann.

Im Folgenden soll gezeigt werden, wie im Rahmen des in Heberle u. a. [10] betrachteten multivariaten Chain-Ladder-Verfahrens und Modellrahmens bei der Quantifizierung der Unsicherheit im Abwicklungsergebnis des kommenden Kalenderjahrs (d.h. des einjährigen Reserverisikos) für ein ganzes Portfolio aus mehreren korrelierten Run-Off Portfolios Korrelationen zwischen den verschiedenen Run-Off Portfolios berücksichtigt werden können. Während jedoch in Heberle u. a. [10] analytisch ein Schätzer für den bedingten Prognosefehler bei der Prognose des Abwicklungsergebnisses ermittelt wird, erfolgt in dieser Arbeit die Quantifizierung der Unsicherheit im Abwicklungsergebnis des kommenden Kalenderjahrs mittels Bootstrappingverfahren. Dieser simulative Ansatz ermöglicht die Bestimmung einer prädiktiven Verteilung für das Abwicklungsergebnis im kommenden Kalenderjahr.

1.3 Bootstrappingverfahren in der Schadenreservierung

Das Bootstrappingverfahren in seiner ursprünglichen Form geht auf Efron & Tibshirani [6] zurück und ist eine flexible sowie verteilungsfreie Simulationsmethode zur Erzeugung einer Verteilung aus einer gegebenen Stichprobe. In der Schadenreservierung wurden Bootstrappingverfahren unter anderem bereits von Brehm [3], Taylor & McGuire [24], Kirschner u. a. [13], England & Verrall [9, 8], Lowe [14] und Pinheiro u. a. [20] eingesetzt. Der Einsatz von Bootstrappingverfahren ist jedoch auch in der aktuariellen Praxis in den letzten Jahren sehr beliebt geworden, da Bootstrappingverfahren

- 1. ein etabliertes statistisches Verfahren darstellen,
- 2. leicht angewendet und implementiert werden können und
- 3. durch die Kombination mit einer Simulation für die Prozessvarianz die Berechnung der prädiktiven Verteilung ermöglichen.

Um das Bootstrappingverfahren in der Schadenreservierung anwenden zu können, wird lediglich ein Schadenreservierungsmodell benötigt, welches die Möglichkeit bietet aus gegebenen Beobachtungen neue *Pseudo-Beobachtungen* zu generieren. Bootstrapping kann dabei allgemein als "Simulation aus einem geschätzten Schadenreservierungsmodell" beschrieben werden, bei der Pseudo-Beobachtungen erzeugt werden und anschließend das Schadenreservierungsmodell an die Pseudo-Beobachtungen anpasst wird.

Im Folgenden betrachten wir das multivariate Chain-Ladder Modell (vgl. Heberle u. a. [10]) für N korrelierte Run-Off Portfolios und ermitteln mittels Bootstrapping die prädiktive Verteilung

des Abwicklungsergebnisses im nächsten Kalenderjahr. Wir gehen dazu in zwei Schritten vor:

- 1. Berechnung des Schätzfehlers für das multivariate Chain-Ladder Modell.
- 2. Simulation der Prozessvarianz und Aggregation mit dem Schätzfehler zur prädiktiven Verteilung des Abwicklungsergebnisses im nächsten Kalenderjahr.

2 Multivariates Chain-Ladder-Verfahren für korrelierte Run-Off Portfolios

In diesem Abschnitt wird kurz das multivariate Chain-Ladder Modell aus Merz & Wüthrich [16] vorgestellt. Dieses multivariate Schadenreservierungsmodell ist für N korrelierte Run-Off-Portfolios geeignet, wobei die Korrelationen durch die Abwicklungsjahreffekte induziert werden.

2.1 Schadenabwicklungsdreiecke und Notation

Im Folgenden betrachten wir simultan $N \ge 1$ Run-Off Portfolios gleicher Dimension. Wir nehmen an, dass $C_{i,j}^{(n)}$ die kumulierten Schadenzahlungen für das Anfalljahr $i \in \{0, ..., I\}$ bis einschließlich zum relativen Abwicklungsjahr $j \in \{0, ..., J\}$ im Run-Off-Portfolio $n \in \{1, ..., N\}$ bezeichnet. Die Bezeichnung Anfalljahr bezieht sich dabei auf dasjenige Jahr, in dem der Schaden eingetreten ist. Um die Notation übersichtlich zu halten, nehmen wir (wie allgemein üblich) I = J an. Alle Formeln behalten jedoch in entsprechender Form auch für den allgemeineren Fall I > J Gültigkeit. Wir betrachten also N Abwicklungsdreiecke zum Zeitpunkt I und NAbwicklungstrapeze eine Periode später, d.h. zum Zeitpunkt I + 1 (vgl. Abbildung 1).

Anfall-	Abwicklungsjahr j					Anjall-	Abwicklungsjahr j					
jahr i	0		j		J	jahr i	0		j		J	
0						0						
÷		$\mathcal{D}_{I}^{(n)}$)			÷		$\mathcal{D}_{I+}^{(n)}$	ı) 1			
I - j		-				I - j		_	-			
:						÷						
Ι						Ι				1		

Abbildung 1: Abwicklungsdreieck $n \in \{1, ..., N\}$ bzw. Abwicklungstrapez $n \in \{1, ..., N\}$ zu den Zeitpunkten t = I bzw. t = I + 1.

Ferner nehmen wir an, dass alle Schäden spätestens nach J Jahren vollständig abgewickelt sind. Es gelte somit $C_{i,J+k}^{(n)} = C_{i,J}^{(n)}$ für alle $i = 0, ..., I, k \in \mathbb{N}$ und n = 1, ..., N.

Zum Zeitpunkt I liegen für das Abwicklungsdreieck $n \in \{1, \dots, N\}$ insgesamt die Beobachtungen

$$\mathcal{D}_{I}^{(n)} := \left\{ C_{i,j}^{(n)} \mid 0 \le i+j \le I \text{ und } 0 \le i \le I \right\}$$

vor (vgl. Abbildung 1, links). Die Menge der Beobachtungen für alle N Abwicklungsdreiecke zum Zeitpunkt I ist somit gegeben durch

$$\mathcal{D}_I^N := \bigcup_{n=1}^N \mathcal{D}_I^{(n)}$$

Die Beobachtungen zum Zeitpunkt I + 1 für das *n*-te Abwicklungsdreieck und für alle N Abwicklungsdreiecke sind dagegen gegeben durch

$$\mathcal{D}_{I+1}^{(n)} \coloneqq \mathcal{D}_{I}^{(n)} \cup \left\{ C_{i,I-i+1}^{(n)} \mid 1 \le i \le I \right\}$$

bzw.

$$\mathcal{D}_{I+1}^N := \bigcup_{n=1}^N \mathcal{D}_{I+1}^{(n)}.$$

Ferner definieren wir

$$\mathcal{B}_{k}^{N} := \left\{ C_{i,j}^{(n)} \in \mathcal{D}_{I}^{N} \mid 0 \leq j \leq k \right\} \subseteq \mathcal{D}_{I}^{N}$$

für $k \in \{0, ..., J\}$. Dies ist die Menge aller Beobachtungen zum Zeitpunkt *I* bis zum relativen Abwicklungsjahr *k*.

2.2 Multivariates Chain-Ladder-Verfahren

Oft wird der Chain-Ladder Algorithmus als ein rein mechanischer Algorithmus zur Berechnung von Prognosen $\widehat{C}_{i,J}^{(n)}$ für die Endschadenstände $C_{i,J}^{(n)}$ verstanden. Für das *n*-te Abwicklungsdreieck zum Zeitpunkt I lässt sich der Chain-Ladder Algorithmus wie folgt darstellen:

1. Es existieren sogenannte Chain-Ladder Faktoren $f_j^{(n)} > 0$ für j = 0, ..., J - 1, so dass für alle $i \in \{0, ..., I\}$ und $I - i + 1 \le k \le J$

$$C_{i,k}^{(n)} = C_{i,I-i}^{(n)} \cdot f_{I-i}^{(n)} \cdot f_{I-i+1}^{(n)} \cdot \dots \cdot f_{k-1}^{(n)}$$
(2.1)

gilt.

2. Die Chain-Ladder Faktoren $f_i^{(n)}$ in (2.1) werden durch die sog. Chain-Ladder Schätzer

$$\widehat{f}_{j}^{(n)} := \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}} \cdot F_{i,j}^{(n)}$$
(2.2)

geschätzt. Dabei ist

$$F_{i,j}^{(n)} := \frac{C_{i,j+1}^{(n)}}{C_{i,j}^{(n)}}$$

der *individuelle Abwicklungsfaktor* für $i \in \{0, ..., I\}, j \in \{0, ..., J-1\}$ und $n \in \{1, ..., N\}$.

3. Der Prädiktor

$$\hat{C}_{i,J}^{(n)} := C_{i,I-i}^{(n)} \cdot \hat{f}_{I-i}^{(n)} \cdot \hat{f}_{I-i+1}^{(n)} \cdot \ldots \cdot \hat{f}_{J-1}^{(n)}$$

wird als *Chain-Ladder Prädiktor* für $C_{i,J}^{(n)}$ bezeichnet.

Im Folgenden betrachteten wir das Zeitreihenmodell von Merz & Wüthrich [16], welches dem Chain-Ladder-Verfahren bei *N* korrelierten Run-Off Portfolios zu Grunde gelegt werden kann:

Modellannahmen 2.1 (Multivariates Chain-Ladder Zeitreihenmodell):

- Kumulierte Schadenzahlungen aus verschiedenen Anfalljahren i sind unabhängig.
- Es existieren Konstanten $f_j^{(n)} > 0$, $\sigma_j^{(n)} > 0$ und Zufallsvariablen $\epsilon_{i,j+1}^{(n)}$ so, dass für alle $i \in \{0, \ldots, I\}, j \in \{0, \ldots, J-1\}$ und $n \in \{1, \ldots, N\}$ gilt:

$$C_{i,j+1}^{(n)} = f_j^{(n)} \cdot C_{i,j}^{(n)} + \sigma_j^{(n)} \cdot \sqrt{C_{i,j}^{(n)}} \cdot \epsilon_{i,j+1}^{(n)}.$$
(2.3)

• Die Zufallsvariablen $\epsilon_{i,j+1}^{(n)}$ und $\epsilon_{k,l+1}^{(m)}$ sind für $i \neq k$ und $j \neq l$ unabhängig und es gilt

$$\mathbb{E}\left[\left.\epsilon_{i,j}^{(n)} \mid \mathcal{B}_{0}^{N}\right] = 0, \quad \text{Var}\left(\left.\epsilon_{i,j}^{(n)} \mid \mathcal{B}_{0}^{N}\right) = 1 \quad und \quad P\left(C_{i,j+1}^{(n)} > 0 \mid \mathcal{B}_{0}^{N}\right) = 1$$

für alle $i \in \{0, ..., I\}, j \in \{0, ..., J\}$ und $n \in \{1, ..., N\}$.

• Die N-dimensionale Zufallsvariable $\epsilon_{i,j} := (\epsilon_{i,j}^{(1)}, \dots, \epsilon_{i,j}^{(N)})'$ besitzt die bedingte Korrelationsmatrix

$$\Sigma_{j} := \operatorname{Corr} \left(\boldsymbol{\epsilon}_{i,j} \mid \mathcal{B}_{0}^{N} \right) := \begin{pmatrix} 1 & \rho_{j}^{(1,2)} & \cdots & \rho_{j}^{(1,N)} \\ \rho_{j}^{(2,1)} & 1 & \cdots & \rho_{j}^{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{j}^{(N,1)} & \rho_{j}^{(N,2)} & \cdots & 1 \end{pmatrix},$$
(2.4)

mit $\rho_j^{(n,m)} \in (-1,1)$ für alle $n,m \in \{1,\ldots,N\}$ und $n \neq m$. Da Var $\left(\epsilon_{i,j}^{(n)} \mid \mathcal{B}_0^N\right) = 1$ gilt folgt, dass die Korrelationsmatrix Σ_j gleichzeitig die Kovarianzmatrix von $\epsilon_{i,j}$ ist, d.h. es gilt:

$$\Sigma_{j} = \operatorname{Corr}\left(\boldsymbol{\epsilon}_{i,j} \mid \mathcal{B}_{0}^{N}\right) = \operatorname{Cov}\left(\boldsymbol{\epsilon}_{i,j} \mid \mathcal{B}_{0}^{N}\right)$$
(2.5)

Dieses multivariate Chain-Ladder Zeitreihenmodell wurde erstmals in Merz & Wüthrich [16] analysiert und ist ein Spezialfall des Chain-Ladder Modells von Pröhl & Schmidt [21]. In diesem Modell wird durch (2.4) eine Abwicklungsjahr-basierte Korrelationsstruktur unterstellt. Diese Annahme ist zwar auf der einen Seite sehr restriktiv, auf der anderen Seite ist sie jedoch für das Chain-Ladder-Verfahren auch die natürlichste Korrelationsannahme (cf. Braun [2]). In mancher Hinsicht erscheint es zuerst vielleicht intuitiver eine Kalenderjahr-basierte Korrelationsstruktur zu unterstellen (z.B. aufgrund von Inflation). Dies würde jedoch Abhängigkeiten zwischen den verschiedenen Anfalljahren induzieren und damit der für das Chain-Ladder-Verfahren klassischen Annahme unabhängiger Anfalljahre widersprechen. Es sollten daher generell vor der Anwendung des Chain-Ladder-Verfahrens zuvor alle Kalenderjahr-basierten Abhängigkeiten aus den Daten entfernt werden. Nach Entfernung der Kalenderjahr-basierten Korrelationen existieren jedoch noch weitere (Abwicklungsjahr-basierte) Korrelationen, die durch ein geeignetes multivariates Chain-Ladder Modell berücksichtigt werden sollten (cf. Houltram [12]).

Es gilt:

Lemma 2.2: Unter den Modellannahmen 2.1 ist $(C_{i,j}^{(n)})_{j\geq 1}$ für i = 0, ..., I eine Markovkette und es gilt:

- 1. Die kumulierten Schadenzahlungen zweier verschiedener Anfalljahre i und k gleicher oder verschiedener Run-Off Portfolios sind unabhängig,
- 2. E $\left[C_{i,j}^{(n)} \mid C_{i,j-1}^{(n)} \right] = f_{j-1}^{(n)} \cdot C_{i,j-1}^{(n)}$
- 3. Die Chain-Ladder Schätzer $\hat{f}_i^{(n)}$ sind erwartungstreue Schätzer für $f_i^{(n)}$,
- 4. Die Chain-Ladder Schätzer $\hat{f}_i^{(n)}$ und $\hat{f}_l^{(m)}$ sind für $j \neq l$ unkorreliert,
- 5. Var $\left(C_{i,j}^{(n)} \mid C_{i,j-1}^{(n)}\right) = \left(\sigma_{j-1}^{(n)}\right)^2 \cdot C_{i,j-1}^{(n)}$ und
- 6. Cov $\left(C_{i,j}^{(n)}, C_{k,j}^{(m)} \mid C_{i,j-1}^{(n)}, C_{k,j-1}^{(m)}\right) = \sigma_{j-1}^{(n)} \cdot \sigma_{j-1}^{(m)} \cdot \sqrt{C_{i,j-1}^{(n)} \cdot C_{k,j-1}^{(m)}} \cdot \rho_j^{(n,m)} \cdot \mathbf{1}_{\{i=k\}},$
- 7. Bedingt gegeben $C_{i,I-i}^{(n)}$, ist der Chain-Ladder Prädiktor $\hat{C}_{i,J}^{(n)}$ für $C_{i,J}^{(n)}$ erwartungstreu, d.h. es gilt $\mathbb{E}\left[\hat{C}_{i,J}^{(n)} \mid C_{i,I-i}^{(n)}\right] = C_{i,J}^{(n)}$.

Beweis. Siehe Merz & Wüthrich [16].

Ein erwartungstreuer Schätzer für den Varianzparameter $\sigma_i^{(n)}$ ist gegeben durch

$$\hat{\sigma}_{j}^{(n)} := \sqrt{\frac{1}{I-j-1} \cdot \sum_{i=0}^{I-j-1} C_{ij}^{(n)} \cdot \left(F_{i,j}^{(n)} - \hat{f}_{j}^{(n)}\right)^{2}}$$
(2.6)

für $j \in \{0, ..., J-1\}$ und $n \in \{1, ..., N\}$. Dabei ist jedoch zu beachten, dass (2.6) zur Schätzung von $\sigma_{J-1}^{(n)}$ nur im Falle von I > J herangezogen werden kann. Im Falle I = J kann jedoch eine Schätzung für $\sigma_{J-1}^{(n)}$ z.B. durch die Extrapolation

$$\hat{\sigma}_{J-1}^{(n)} \coloneqq \min\left\{ \left(\hat{\sigma}_{J-2}^{(n)} \right)^2 \middle/ \hat{\sigma}_{J-3}^{(n)} , \, \hat{\sigma}_{J-3}^{(n)} , \, \hat{\sigma}_{J-2}^{(n)} \right\}$$
(2.7)

ermittelt werden (siehe Mack [15]).

3 Bestimmung einer prädiktiven Verteilung für mehrere korrelierte Run-Off Portfolios

In diesem Abschnitt wird mit Hilfe der Bootstrapmethode in zwei Schritten die *prädiktive Verteilung* für das Abwicklungsergebnis im kommenden Kalenderjahr ermittelt. Im ersten Schritt wird der resultierende *Schätzfehler* bei der Schätzung der Chain-Ladder Faktoren quantifiziert und in einem zweiten Schritt die *Prozessvarianz* berücksichtigt. Dabei unterscheiden wir jeweils zwischen den beiden Fällen a) Nichtberücksichtigung und b) Berücksichtigung einer eventuell vorhandenen Abwicklungsjahr-basierten Korrelationsstruktur zwischen den Daten in den verschiedenen Run-Off Portfolios. Das Ziel ist somit eine prädiktive Verteilung für das Abwicklungsergebnis im kommenden Kalenderjahr unter Berücksichtigung der durch die Modellannahmen 2.1 beschriebenen Korrelationsstruktur (2.4) zwischen den *N* verschiedenen Abwicklungsdreiecken. Dieses zweistufige Bootstrapverfahren wurde bereits von England & Verrall [8, 9], England [7] und Pinheiro u. a. [20] im Falle eines einzelnen Run-Off Portfolios zur Ermittlung der prädiktiven Verteilung für den Endschadenstand und die Schadenreserve verwendet.

Das Abwicklungsergebnis eines einzelnen Abwicklungsdreiecks ist dabei wie folgt definiert:

Definition 3.1: Das Abwicklungsergebnis für das n-te Abwicklungsdreieck im i-ten Anfalljahr $(i \in \{1, ..., I\})$ im Kalenderjahr (I, I + 1] ist definiert durch

$$AWE_{i}(I+1)^{(n)} := E\left[R_{i}^{(n),I} \mid \mathcal{D}_{I}^{(n)}\right] - \left(X_{i,I-i+1}^{(n)} + E\left[R_{i}^{(n),I+1} \mid \mathcal{D}_{I+1}^{(n)}\right]\right).$$

Dabei bezeichnet $X_{i,I-i+1}^{(n)} = C_{i,I-i+1}^{(n)} - c_{i,I-i}^{(n)}$ die Schadenzahlung im Kalenderjahr (I, I + 1] und Anfalljahr i. Außerdem wird mit $R_i^{(n),I} = C_{i,J}^{(n)} - C_{i,I-i}^{(n)}$ die noch ausstehende Zahlungsverpflichtung für das i-te Anfalljahr ($i \in \{1, ..., I\}$) und n-te Abwicklungsdreieck ($n \in \{1, ..., N\}$) zum Zeitpunkt t = I definiert. Ebenso ist $R_i^{(n),I+1} = C_{i,J}^{(n)} - C_{i,I-i+1}^{(n)}$ die noch ausstehende Zahlungsverpflichtung für das i-te Anfalljahr und n-te Abwicklungsjahr zum Zeitpunkt t = I + 1.

Aus Definition 3.1 erhält man für den Schätzer des Abwicklungsergebnisses für das *n*-te Abwicklungsdreieck im *i*-ten Anfalljahr ($i \in \{1, ..., I\}$) im Kalenderjahr (I, I + 1] unmittelbar

$$\widehat{\text{AWE}}_{i}(I+1)^{(n)} = \hat{C}_{i,J}^{(n),I} - \hat{C}_{i,J}^{(n),I+1}.$$

Für aggregierte Anfalljahre erhält man

$$\widehat{AWE}(I+1)^{(n)} = \sum_{i=1}^{I} \widehat{AWE}_i(I+1)^{(n)}.$$

3.1 Bootstrappingmethode für das multivariate Chain-Ladder-Verfahren

Für die qualifizierte Anwendung der Bootstrappingmethode müssen die Beobachtungen Realisierungen identisch-verteilter Zufallsvariablen sein. Diese Voraussetzung ist jedoch bei den Daten eines Abwicklungsdreiecks nicht erfüllt. Deshalb wird die Bootstrappingmethode nicht direkt auf die Schadenzahlungen $C_{i,j}^{(n)} \in \mathcal{D}_I^{(n)}$ für $n = 1, \ldots, N$ angewendet, sondern auf die Residuen (cf. England & Verrall [9]). Die rekursive Struktur (2.3) gibt dabei die Definition der Residuen vor, die mit den Annahmen des Chain-Ladder Modells 2.1 konsistent sind.

Mit den Parameterschätzern (2.2) und (2.6)-(2.7) erhält man aus (2.3) für das Abwicklungsdreieck $n \in \{1, ..., N\}$ die $K := \frac{1}{2} \cdot I \cdot (I + 1)$ Pearson-Residuen

$$r_{i,j+1}^{(n)} := \frac{\sqrt{C_{i,j}^{(n)}} \cdot \left(F_{i,j}^{(n)} - \hat{f}_j^{(n)}\right)}{\hat{\sigma}_j^{(n)}}$$

für i = 0, ..., I - 1, j = 0, ..., J - 1 und $i + j \le I - 1$. Dabei ist zu beachten, dass die Residuen $r_{i,j+1}^{(n)}$ zum Zeitpunkt t = I beobachtbar sind, während die Störvariablen $\epsilon_{i,j+1}^{(n)}$ für unbekannte Chain-Ladder Faktoren $f_i^{(n)}$ zum Zeitpunkt t = I nicht beobachtbar sind.

Die skalierten Residuen $\bar{r}_{i,j+1}^{(n)}$ sind gegeben durch

$$\bar{r}_{i,j+1}^{(n)} := rac{r_{i,j+1}^{(n)} - \mu_{j+1}^{(n)}}{\sigma^{(n)}}$$

mit

$$\mu_{j+1}^{(n)} := \frac{1}{I-j-1} \cdot \sum_{i=0}^{I-j-1} r_{i,j+1}^{(n)} \quad \text{für } j = 0, \dots, J-1$$

und

$$\sigma^{(n)} := \sqrt{\frac{1}{K-1} \cdot \sum_{i=0}^{I-1} \sum_{j=0}^{J-2} \left(r_{i,j+1}^{(n)} - \mu_{j+1}^{(n)} \right)^2}.$$

Die Werte $\bar{r}_{i,j+1}^{(n)}$ werden als *skalierte Pearson-Residuen* bezeichnet. Für ein gegebenes Run-Off Portfolio $n \in \{1, ..., N\}$ ist

$$\bar{r}^{(n)} := \left\{ \bar{r}_{i,j+1}^{(n)} \mid i = 0, \dots, I-1; \ j = 0, \dots, J-1; \ i+j \le J-1 \right\}$$
(3.1)

die Menge aller skalierten Pearson-Residuen. Durch (3.1) ist die *Bootstrapverteilung* $G^{(n)}$ für das *n*-te Run-Off Portfolio definiert, aus der die Stichproben für das *n*-te Run-Off Portfolio gezogen werden.

3.1.1 Bootstrapping ohne Berücksichtigung von Korrelationen

Aus der Menge $\bar{r}^{(n)}$ der *K* skalierten Pearson-Residuen $\bar{r}_{i,j+1}^{(n)}$ für das Abwicklungsdreieck $n \in \{1, \ldots, N\}$, werden nun Residuen mit Zurücklegen gezogen. Da keine Abhängigkeiten zwischen den Run-Off Portfolios berücksichtigt werden, gilt für die durch die Ziehungen induzierte Zufallsvariable $\tilde{r}_{i,j+1}^{(n)}$

$$\tilde{\bar{r}}_{i,j+1}^{(n)} \sim G^{(n)} \quad \text{für } n \in \{1, \dots, N\},$$
(3.2)

und die Zufallsvariablen $\tilde{r}_{i,j+1}^{(n)}$ und $\tilde{r}_{i,j+1}^{(k)}$ sind – bis auf minimale Abweichungen, welche auf die Zufälligkeit der Stichprobenziehung (sogenannte Stichprobeneffekte) zurückzuführen sind – für $n \neq k$ unkorreliert. Mit Hilfe von (3.2) lassen sich neue *Pseudo-Beobachtungen* für die individuellen Abwicklungsfaktoren $F_{i,j}^{(n)}$ des Abwicklungsdreiecks $n \in \{1, \ldots, N\}$ durch

$$\tilde{F}_{i,j}^{(n)} := \hat{f}_j^{(n)} + \frac{\hat{\sigma}_j^{(n)}}{\sqrt{C_{i,j}^{(n)}}} \cdot \tilde{r}_{i,j+1}^{(n)}$$

für $i \in \{0, ..., I - 1\}, j \in \{0, ..., J - 1\}$ und $i + j \le I - 1$ generieren. Zusammen mit (2.2) lassen sich daraus *Bootstrap Chain-Ladder Faktoren*

$$\tilde{f}_{j}^{(n)} := \sum_{i=0}^{I-j-1} \frac{C_{i,j}^{(n)}}{\sum\limits_{i=0}^{I-j-1} C_{i,j}^{(n)}} \cdot \tilde{F}_{i,j}^{(n)}$$
(3.3)

für $j \in \{0, ..., J-1\}$ und $n \in \{1, ..., N\}$ erzeugen. Die Nichtberücksichtigung von Korrelationen bedeutet, dass die in (2.4) angegebene Matrix Σ_j durch die Einheitsmatrix gegeben ist. Mit den in (3.3) bestimmten Bootstrap Chain-Ladder Faktoren lassen sich nun durch

$$\tilde{C}_{i,j}^{(n)} := C_{i,I-i}^{(n)} \cdot \tilde{f}_{I-i}^{(n)} \cdot \ldots \cdot \tilde{f}_{j-1}^{(n)} \quad \text{für } I - i + 1 \le j \le J$$
(3.4)

für das Abwicklungsdreieck $n \in \{1, ..., N\}$ Pseudo-Beobachtungen für die kumulierten Schadenzahlungen im unteren Abwicklungsdreieck generieren. Wird dieses Vorgehen hinreichend oft wiederholt, kann eine Bootstrapverteilung (d.h. "simulierte" empirische Verteilung) für das Abwicklungsergebnis im kommenden Kalenderjahr ermittelt werden. Dabei ist zu beachten, dass durch die oben beschriebenen Schritte lediglich der Schätzfehler erfasst wird. Die zusätzliche Berücksichtigung der Prozessvarianz erfolgt in Abschnitt 3.2.

Bemerkungen 3.2:

• Die Pseudo-Beobachtungen für die kumulierten Schadenzahlungen des *n*-ten Run-Off Portfolios lassen sich auch durch

$$\tilde{C}_{i,0}^{(n)} := C_{i,0}^{(n)} \quad \text{für } i = 0, \dots, I$$

$$\tilde{C}_{i,j+1}^{(n)} := \hat{f}_{j}^{(n)} \cdot C_{i,j}^{(n)} + \hat{\sigma}_{j}^{(n)} \cdot \sqrt{C_{i,j}^{(n)}} \cdot \tilde{r}_{i,j+1}^{(n)}$$
(3.5)

für $0 \le j \le J - 1$ und $0 \le i + j \le I - 1$ generieren und statt (3.3) können die Bootstrap Chain-Ladder Faktoren durch

$$\tilde{f}_{j}^{(n)} := \frac{\sum_{i=0}^{I-j-1} \tilde{C}_{i,j+1}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}}$$

für $j \in \{0, ..., J - 1\}$ berechnet werden. Dieser Ansatz entspricht dem in Buchwalder u. a. [4] und Merz & Wüthrich [16] betrachteten *bedingten Resampling-Ansatz*.

• Die ermittelte Bootstrapverteilung für die Parameter und die kumulierten Schadenzahlungen kann z.B. zur Ermittlung einer Bootstrapschätzung für die Standardfehler der Parameter herangezogen werden. Ferner können mit Hilfe der Bootstrapverteilung Bootstrapschätzungen für die Kovarianzmatrix der Parameter sowie die Standardfehler der kumulierten Schadenzahlungen und des Abwicklungsergebnisses im kommenden Kalenderjahr ermittelt werden.

3.1.2 Bootstrapping unter Berücksichtigung von Korrelationen

Im Gegensatz zum vorhergehenden Abschnitt 3.1.1 soll nun die Abwicklungsjahr-basierte Korrelationsstruktur (2.4) zwischen den einzelnen Run-Off Portfolios berücksichtigt werden. Dazu betrachten wir den *N*-dimensionalen Zufallsvektor

$$\widetilde{\overline{\mathbf{r}}}_{i,j+1} := \left(\widetilde{\overline{r}}_{i,j+1}^{(1)}, \dots, \widetilde{\overline{r}}_{i,j+1}^{(N)}\right)' \sim G(\mathbf{0}, \widehat{\Sigma}_{i,j}^*)$$

für $i \in \{0, ..., I\}$ und $j \in \{0, ..., J - 1\}$ bestehend aus den skalierten Pearson-Residuen (3.2). Wie oben bereits beschrieben sind die Residuen $\widetilde{r}_{i,j+1}^{(1)}, ..., \widetilde{r}_{i,j+1}^{(N)}$ – bis auf Stichprobeneffekte – unkorreliert mit Erwartungswert 0. Die möglichen Stichprobeneffekte zwischen den Residuen $\widetilde{r}_{i,j+1}^{(1)}, ..., \widetilde{r}_{i,j+1}^{(N)}$ werden im Folgenden durch die Kovarianzmatrix $\widehat{\Sigma}_{i,j}^*$ berücksichtigt. Das heißt es wird

$$\operatorname{Cov}\left(\widetilde{\overline{\mathbf{r}}}_{i,j+1} \mid \mathcal{B}_{0}^{N}\right) = \widehat{\Sigma}_{i,j}^{*}$$

angenommen. Das Ziel ist es nun den Zufallsvektor $\tilde{\overline{\mathbf{r}}}_{i,j+1}$ durch geeignete $N \times N$ -Matrizen $M_{i,j}$ so zu transformieren, dass der neue Zufallsvektor $M_{i,j}$ $\tilde{\overline{\mathbf{r}}}_{i,j+1}$ die Abwicklungsjahr-basierte Kovarianzstruktur (2.5) besitzt. Damit ist die Matrix $M_{i,j}$ so zu bestimmen, dass

$$\operatorname{Cov}\left(M_{i,j}\,\widetilde{\widetilde{\mathbf{r}}}_{i,j+1}\,\middle|\,\,\mathcal{B}_{0}^{N}\right) = \Sigma_{j} \tag{3.6}$$

für $i \in \{0, ..., I\}$ und $j \in \{0, ..., J - 1\}$ gilt.

Zur Generierung der durch (3.6) beschriebenen Kovarianzstruktur verwenden wir nachfolgenden wohlbekannten Satz aus der Matrizentheorie:

Satz 3.3: Es sei A eine symmetrische $N \times N$ Matrix mit rg(A) = N. Dann existiert eine $N \times N$ Matrix B, so dass

$$A = BDB^{T}$$

mit

$$D := \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_N \end{pmatrix}$$

gilt. Dabei sind $\lambda_1, \ldots, \lambda_N$ die von Null verschiedenen Eigenwerte von A. Die Spaltenvektoren von B entsprechen den (paarweise orthonormalen) zugehörigen Eigenvektoren. Insbesondere gilt $BB^T = B^T B = I$, d.h. $B^{-1} = B^T$.

Beweis. Siehe z.B. Zurmühl & Falk [27, S. 192].

Um nun die gewünschte Abwicklungsjahr-basierte Kovarianzstruktur (3.6) zu erhalten, wird der Zufallsvektor $\tilde{\vec{r}}_{i,j+1}$ (für $i \in \{0, ..., I\}$ und $j \in \{0, ..., J-1\}$) mit der $N \times N$ -Matrix

$$M_{i,j} := B_j D_j^{1/2} D_{i,j}^{* - 1/2} B_{i,j}^{* T}$$

multipliziert. Es wird dann der N-dimensionale Zufallsvektor

$$\mathbf{Y}_{i,j} := M_{i,j} \, \widetilde{\mathbf{r}}_{i,j+1} = B_j \, D_j^{1/2} \, D_{i,j}^{* - 1/2} \, B_{i,j}^{* T} \, \widetilde{\mathbf{r}}_{i,j+1}$$
(3.7)

mit

- D_j ist die Diagonalmatrix der Eigenwerte von Σ_j ;
- $D^*_{i,j}$ ist die Diagonalmatrix der Eigenwerte von $\widehat{\Sigma}^*_{i,j};$
- B_i ist die Matrix der orthonormalen Eigenvektoren (spaltenweise) von Σ_i und
- $B^*_{i,j}$ ist die Matrix der orthonormalen Eigenvektoren (spaltenweise) von $\widehat{\Sigma}^*_{i,j}$.

betrachtet. Wir erhalten mit Lemma 3.3 für den Zufallsvektor $Y_{i,j}$:

$$Cov \left(\mathbf{Y}_{i,j} \mid \mathcal{B}_{0}^{N} \right) = M_{i,j} Cov \left(\widetilde{\mathbf{r}}_{i,j+1} \mid \mathcal{B}_{0}^{N} \right) M_{i,j}^{T} = M_{i,j} \widehat{\Sigma}_{i,j}^{*} M_{i,j}^{T} = B_{j} D_{j}^{1/2} D_{i,j}^{* - 1/2} B_{i,j}^{* T} \widehat{\Sigma}_{i,j}^{*} \left(B_{j} D_{j}^{1/2} D_{i,j}^{* - 1/2} B_{i,j}^{* T} \right)^{T} = B_{j} D_{j}^{1/2} D_{i,j}^{* - 1/2} B_{i,j}^{* T} \widehat{\Sigma}_{i,j}^{*} B_{i,j}^{*} D_{i,j}^{* - 1/2} D_{j}^{1/2} B_{j}^{T}.$$

Wegen

$$(B_{i,j}^*)^T \ \widehat{\Sigma}_{i,j}^* \ B_{i,j}^* = D_{i,j}^*$$

folgt daraus weiter

$$\operatorname{Cov}\left(\mathbf{Y}_{i,j} \mid \mathcal{B}_{0}^{N}\right) = B_{j} D_{j} B_{j}^{T} = \Sigma_{j}.$$

Das heißt die neuen transformierten Residuen $Y_{i,j}$ weisen im Gegensatz zu den skalierten Pearson-Residuen $\tilde{\bar{r}}_{i,j+1}$ die gewünschte Abwicklungsjahr-basierte Kovarianzstruktur (2.5) – und damit die gewünschte Korrelationsstruktur – auf.

Für die Berücksichtigung der Prozessvarianz bei der Quantifizierung der Unsicherheit im Abwicklungsergebnis des kommenden Kalenderjahrs ist ein zweiter Schritt notwendig.

3.2 Zweistufiges Bootstrapverfahren zur Berücksichtigung der Prozessvarianz

Interessiert man sich ausschließlich für die Standardabweichung der Chain-Ladder Schätzungen $\hat{f}_{j}^{(n)}$ für die Chain-Ladder Faktoren $f_{j}^{(n)}$ oder für den Schätzfehler bei der Prognose des Endschadenstands, der Schadenreserve und des Abwicklungsergebnisses im kommenden Kalenderjahr, dann muss lediglich noch die Standardabweichung in den Bootstrap Chain-Ladder Faktoren $\tilde{f}_{j}^{(n)}$ berechnet und mit Hilfe von (3.4) die Pseudo-Beobachtungen für die Endschadenstände, die Schadenreserve bzw. das Abwicklungsergebnis im kommenden Kalenderjahr ermittelt werden. Da wir jedoch eine prädiktive Verteilung für das Abwicklungsergebnis im kommenden Kalenderjahr ermittelt werden die Prozessvarianz bei der Anwendung der Chain-Ladder Methode auf korrelierte Run-Off Portfolios berücksichtigt wird. Die so gewonnene prädiktive Verteilung kann dann für Prognosen zukünftiger Werte verwendet werden. Analog zum ersten Schritt unterscheiden wir dabei wieder zwischen den beiden Fällen a) Nichtberücksichtigung und b) Berücksichtigung einer eventuell vorhandenen Abwicklungsjahr-basierten Korrelationsstruktur.

Analog zum ersten Schritt gibt uns die autoregressive Struktur (2.3) vor, wie die nicht beobachteten kumulierten Schadenzahlungen im unteren Abwicklungsdreieck für jedes einzelne Abwicklungsdreieck $n \in \{1, ..., N\}$ und jede einzelne Bootstrapiteration ermittelt werden können. Wir erhalten dann mit den Bootstrap Chain-Ladder Faktoren $\tilde{f}_j^{(n)}$ und den Schätzungen (2.6)-(2.7) für die Varianzparameter $\sigma_j^{(n)}$ die Ein-Schritt Prognose für die kumulierte Schadenzahlung

$$\hat{C}_{i,I-i+1}^{*(n)} = \tilde{f}_{I-i}^{(n)} \cdot C_{i,I-i}^{(n)} + \hat{\sigma}_{I-i}^{(n)} \cdot \sqrt{C_{i,I-i}^{(n)}} \cdot \epsilon_{i,I-i+1}^{*(n)}$$
(3.8)

für i = 1, ..., I, sowie als Zwei-Schritt Prognose

$$\hat{C}_{i,j+1}^{*(n)} = \tilde{f}_{j}^{(n)} \cdot \hat{C}_{i,j}^{*(n)} + \hat{\sigma}_{j}^{(n)} \cdot \sqrt{\hat{C}_{i,j}^{*(n)}} \cdot \epsilon_{i,j+1}^{*(n)}$$
(3.9)

für i = 2, ..., I und j = I - i + 1, ..., J - 1. Dabei werden, analog zu den Modellannahmen 2.1, für die Zufallsvariablen $\epsilon_{i,j+1}^{*(n)}$ folgende Annahmen getroffen:

- 1. $\epsilon_{i,j+1}^{*(n)}$ und $\epsilon_{k,l+1}^{*(n)}$ sind unabhängig, falls $i \neq k$ oder $j \neq l$.
- 2. Es gilt E $\left[\epsilon_{i,j}^{*(n)}\right] = 0$ und Var $\left(\epsilon_{i,j}^{*(n)}\right) = 1$.
- 3. Der *N*-dimensionale Zufallsvektor $\epsilon_{i,j}^* = (\epsilon_{i,j}^{*(1)}, \dots, \epsilon_{i,j}^{*(N)})'$ hat dieselbe Varianz-Ko-

varianzmatrix wie $\boldsymbol{\epsilon}_{i,j} = \left(\epsilon_{i,j}^{(1)}, \ldots, \epsilon_{i,j}^{(N)}\right)'$, d.h.

$$\Sigma_j^* = \operatorname{Cov}\left(\boldsymbol{\epsilon}_{i,j}^* \mid \mathcal{B}_0^N\right) = \Sigma_j.$$

Korollar 3.4: Unter Verwendung von 1.-3. erhält man (für i = 2, ..., I und j = I - i + 1, ..., J - 1)

1.

$$\mathbb{E}\left[\hat{C}_{i,I-i+1}^{*(n)} \mid C_{i,I-i}^{(n)}\right] = \tilde{f}_{I-i}^{(n)} \cdot C_{i,I-i}^{(n)}$$

und

$$\mathbb{E}\left[\hat{C}_{i,j+1}^{*(n)} \middle| \hat{C}_{i,j}^{*(n)}\right] = \tilde{f}_{j}^{(n)} \cdot \hat{C}_{i,j}^{*(n)}$$

2.

Var
$$(\hat{C}_{i,I-i+1}^{*(n)} | C_{i,I-i}^{(n)}) = (\hat{\sigma}_{I-i}^{(n)})^2 \cdot C_{i,I-i}^{(n)}$$

und

Var
$$(\hat{C}_{i,j+1}^{*(n)} \mid \hat{C}_{i,j}^{*(n)}) = (\hat{\sigma}_{j}^{(n)})^2 \cdot \hat{C}_{i,j}^{*(n)}$$

3.

$$\operatorname{Cov}\left(\hat{C}_{i,I-i+1}^{*(n)}, \hat{C}_{k,I-i+1}^{*(m)} \middle| C_{i,I-i}^{(n)}, C_{k,I-i}^{(m)}\right) = \hat{\sigma}_{I-i}^{(n)} \cdot \hat{\sigma}_{I-i}^{(m)} \cdot \sqrt{C_{i,I-i}^{(n)} \cdot C_{k,I-i}^{(m)}} \cdot \rho_{I-i+1}^{(n,m)} \cdot \mathbf{1}_{\{i=k\}}$$

und

$$\operatorname{Cov}\left(\hat{C}_{i,j+1}^{*(n)}, \hat{C}_{k,j+1}^{*(m)} \middle| \hat{C}_{i,j}^{*(n)}, \hat{C}_{k,j}^{*(m)}\right) = \hat{\sigma}_{j}^{(n)} \cdot \hat{\sigma}_{j}^{(m)} \cdot \sqrt{\hat{C}_{i,j}^{*(n)} \cdot \hat{C}_{k,j}^{*(m)}} \cdot \rho_{j+1}^{(n,m)} \cdot \mathbf{1}_{\{i=k\}}.$$

Um die prädiktive Verteilung nun herleiten zu können muss eine Annahme über die Verteilung der $\epsilon_{i,j+1}^{*(n)}$ getroffen werden. Wir werden im folgenden annehmen, dass die $\epsilon_{i,j+1}^{*(n)}$ standardnormalverteilt sind, also $\epsilon_{i,j+1}^{*(n)} \sim \mathcal{N}(0,1)$.

Bemerkungen 3.5:

• Es gilt E $\left[\epsilon_{i,j}^{*(n)}\right] = 0$ und Var $\left(\epsilon_{i,j}^{*(n)}\right) = 1$ für $\epsilon_{i,j}^{*(n)} \sim \mathcal{N}(0,1)$. Dies bedeutet, dass das vorliegende Modell nicht mehr verteilungsfrei ist.

• Die Verwendung der Standardnormalverteilung führt dazu, dass bei der Simulation auch negative kumulierte Schadenzahlungen resultieren können. Dieses Problem kann dadurch behoben werden, dass man die $\epsilon_{i,j}^{*(n)}$ durch eine bedingte Verteilung, gegeben \mathcal{B}_0^N , definiert (vgl. Modellannahmen 2.1).

3.2.1 Zweistufiger Bootstrap ohne Beachtung von Korrelationen

Der in Abschnitt 3.2 beschriebene Ansatz lässt Korrelationen völlig außer Acht, wenn wir $\rho_j^{(n,m)} = 0$ für alle $n \neq m$ und $j = 1, \ldots, J$ setzen. In diesem Fall sind die Zufallsvariablen $\epsilon_{i,j}^{*(n)}, \epsilon_{i,j}^{*(m)} \sim \mathcal{N}(0,1)$ für $n \neq m$ unkorreliert und man erhält als Korrelationsmatrix

$$\Sigma_j^* = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

für j = 0, ..., J - 1. Damit und mit (3.8) und (3.9) lassen sich dann die unteren Einträge der N Abwicklungsdreiecke für jede einzelne Bootstrap-Iteration zum Zeitpunkt t = I bestimmen. Aus diesen gebootstrapten Abwicklungsdreiecken lassen sich dann die $\hat{C}_{i,I}^{(n),I}$ ablesen.

Für die Schätzung der der Endschadenstände $\hat{C}_{i,J}^{(n),I+1}$, welche für die Schätzung des Abwicklungsergebnisses benötigt werden, muss eine weitere Diagonale in den Bootstrapdreiecken durch eine weitere Ziehung skalierter Pearson-Residuen gebildet werden. Es müssen also zusätzlich zu den *K* bereits gezogenen Residuen weitere I - 1 Residuen gezogen werden. Diese I - 1 neuen Residuen werden, mittels Formel (3.5), für die Berechnung der $\tilde{C}_{i,I-i+1}^{(n)}$ (i = 1, ..., I) verwendet.

Für diese neuen Bootstrapdreiecke ergeben sich nach (3.3) neue Bootstrap Chain-Ladder Faktoren, welche analog für die Berechnung der Ein-Schritt bzw. Zwei-Schritt Prognose nach (3.8) und (3.9) verwendet werden. Als Ein-Schritt bzw. Zwei-Schritt Prognose für die kumulierte Schadenzahlung erhält man damit

$$\hat{C}_{i,j+1}^{*(n)} = \tilde{f}_{j}^{(n)} \cdot \hat{C}_{i,j}^{*(n)} + \hat{\sigma}_{j}^{(n)} \cdot \sqrt{\hat{C}_{i,j}^{*(n)}} \cdot \epsilon_{i,j+1}^{*(n)}$$
(3.10)

für i = 2, ..., I und j = I - i + 1, ..., J - 1. Dabei ist zu beachten, dass die Bootstrap Chain-Ladder Faktoren $\tilde{f}_{i}^{(n)}$ aus den um eine weitere Diagonale erweiterten Dreiecken berechnet werden
müssen. Somit gilt

$$\tilde{f}_{j}^{(n)} = \tilde{f}_{j}^{(n),l+1} = \frac{\sum_{i=0}^{l-j} \tilde{C}_{i,j+1}^{(n)}}{\sum_{i=0}^{l-j} C_{i,j}^{(n)}}.$$

Aus diesen gebootstrapten Abwicklungsdreiecken lassen sich somit Schätzungen für die Endschadenstände $\hat{C}_{i,J}^{(n),I+1}$ zum Zeitpunkt t = I + 1 ermitteln. Damit und mit obigen Ergebnissen lassen sich somit Schätzungen für das Abwicklungsergebnis bestimmen.

3.2.2 Zweistufiger Bootstrap mit Beachtung von Korrelationen

Bei Anwendung des zweistufigen Bootstrap-Verfahrens werden, analog zu Paragraph 3.2.1, in einem ersten Schritt die bisher erzeugten Bootstrapdreiecke um eine weitere Diagonale erweitert (siehe Formel (3.5)).

Im zweiten Schritt werden dann $N \cdot (I \cdot J - K) \mathcal{N}(0, 1)$ -verteilte Zufallszahlen gezogen, welche mit Hilfe der in 3.1.2 ermittelten Matrizen $M_{i,j} = B_j D_j^{1/2} D_{i,j}^{* - 1/2} B_{i,j}^{* T}$ auf die gewünschte Korrelationsstruktur gebracht werden. Man berechnet also (siehe (3.7))

$$\mathbf{Y}_{i,j}^* := M_{i,j} \boldsymbol{\epsilon}_{i,j+1}^*.$$

Daraus folgt

Corr
$$\left(\mathbf{Y}_{i,j}^* \mid \mathcal{B}_0^N\right) = \Sigma_j.$$

Das heißt die neuen transformierten Zufallszahlen $\mathbf{Y}_{i,j}^*$ weisen im Gegensatz zu den unkorrelierten Zufallszahlen $\boldsymbol{\epsilon}_{i,j+1}^*$ die gewünschte Abwicklungsjahr-basierte Korrelationsstruktur (2.5) auf.

Damit lassen sich nun die kumulierten Schadenzahlungen $\hat{C}_{i,j}^{*(n)}$ im unteren *n*-ten Abwicklungsdreieck mit Hilfe der Ein-Schritt bzw. Zwei-Schritt Prognose zu den Zeitpunkten t = I und t = I + 1 bestimmen (siehe (3.8), (3.9) und (3.10)). Hierfür muss lediglich das entsprechende $\epsilon_{i,i+1}^{*(n)}$ durch den entsprechenden Eintrag aus $Y_{i,j}^{*}$ ersetzt werden.

Hieraus lässt sich dann für die *N* korrelierten Abwicklungsdreiecke eine Schätzung für das Abwicklungsergebnis ermitteln:

$$\widehat{AWE}(I+1) = \sum_{n=1}^{N} \widehat{AWE}(I+1)^{(n)}$$
(3.11)

Da hierfür die transformierten $Y_{i,j}^*$ verwendet wurden, wird in Formel (3.11) die Korrelationsstruktur der *N* korrelierten Abwicklungsdreiecke beachtet.

4 Beispiel

Wir verwenden die beiden Abwicklungsdreiecke aus Braun [2] bzw. Merz & Wüthrich [16, 18]. In diesen Arbeiten werden zwei Run-Off Portfolios A (allgemeine Haftpflicht) und B (Kfz-Haftpflicht) betrachtet (siehe Tabellen 1 und 2).

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	59966	163152	254512	349524	433265	475778	513660	520309	527978	539039	537301	540873	547696	549589
1	49685	153344	272936	383349	458791	503358	532615	551437	555792	556671	560844	563571	562795	
2	51914	170048	319204	425029	503999	544769	559475	577425	588342	590985	601296	602710		
3	84937	273183	407318	547288	621738	687139	736304	757440	758036	782084	784632			
4	98921	278329	448530	561691	641332	721696	742110	752434	768638	768373				
5	71708	245587	416882	560958	654652	726813	768358	793603	811100					
6	92350	285507	466214	620030	741226	827979	873526	896728						
7	95731	313144	553702	755978	857859	962825	1022241							
8	97518	343218	575441	769017	934103	1019303								
9	173686	459416	722336	955335	1141750									
10	139821	436958	809926	1174196										
11	154965	528080	1032684											
12	196124	772971												
13	204325													

Tabelle 1: Kumulierte Schadenzahlungen $C_{i,j}^{(1)}$ in Portfolio A (in Mio. US\$).

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	114423	247961	312982	344340	371479	371102	380991	385468	385152	392260	391225	391328	391537	391428
1	152296	305175	376613	418299	440308	465623	473584	478427	478314	479907	480755	485138	483974	
2	144325	307244	413609	464041	519265	527216	535450	536859	538920	539589	539765	540742		
3	145904	307636	387094	433736	463120	478931	482529	488056	485572	486034	485016			
4	170333	341501	434102	470329	482201	500961	504141	507679	508627	507752				
5	189643	361123	446857	508083	526562	540118	547641	549605	549693					
6	179022	396224	497304	553487	581849	611640	622884	635452						
7	205908	416047	520444	565721	600609	630802	648365							
8	210951	426429	525047	587893	640328	663152								
9	213426	509222	649433	731692	790901									
10	249508	580010	722136	844159										
11	258425	686012	915109											
12	368762	909066												
13	394997													

Tabelle 2: Kumulierte Schadenzahlungen $C_{i,j}^{(2)}$ in Portfolio B (in Mio. US\$).

Bei der Ermittlung der prädiktiven Verteilung des einjährigen Reserverisikos wurde für die Abwicklungsjahr-basierten Korrelationsmatrizen $\Sigma_j = \Sigma$ unterstellt, d.h. es wurde davon ausgegangen, dass die Korrelationsstruktur für verschiedene Abwicklungsjahre *j* konstant ist. Für das vorliegende Beispiel wurden die Korrelationen zwischen den Pearson-Residuen $\bar{r}^{(n)}$ (für n = 1, ..., N) berechnet. Daraus erhält man für den vorliegenden Fall (N = 2) eine

 2×2 -Korrelationsmatrix $\Sigma,$ welche die folgenden Einträge besitzt:

$$\Sigma = \begin{pmatrix} 1 & 0.431 \\ 0.431 & 1 \end{pmatrix}$$

Man erhält also einen Korrelationskoeffizienten zwischen den beiden Abwicklungsdreiecken A und B von $\rho^{(1,2)} = 0.431$ und somit eine positive Korrelation. Führt man nun die hier vorgestellte Bootstrapmethode, jedoch ohne Beachtung jeglicher Korrelationsstruktur, durch, so erhält man Schätzungen, die sehr nahe bei den theoretischen Ergebnissen in Heberle u. a. [10] sind. Wird nun die oben beschriebene Korrelationsstruktur beachtet, so erhält man die Ergebnisse, welche in Tabelle 3 dargestellt sind. Hierfür wurden 10,000 Bootstraps durchgeführt. Die Anzahl durchzuführender Bootstraps ist vor allem deshalb relativ groß zu wählen, da wir uns für den Tail der prädiktiven Verteilung interessieren und deshalb eine solide Datenbasis benötigen. Aus Tabelle 3 ist ersichtlich, dass die Differenz der ersten beiden Momente der prädiktiven Verteilung des einjährigen Reserverisikos zu den entsprechenden theoretischen Momenten sehr klein ist. Diese gute Anpassung soll als Rechtfertigung dienen, um nun aus der prädiktiven Verteilung weitere Risikomaße zu berechnen. In Tabelle 4 sind deshalb der Value-at-Risk und der Expected-Shortfall zu verschiedenen Sicherheitsniveaus angegeben.

	theoret mit k = 1	theoretische Ergebnisse mit k Iterationen k = 1 $k = 2$ $k = 3$				
einjähriges Reserverisiko $\sqrt{\text{MSE}}$	0	0	0	4,087		
	354,049	353,964	353,964	355,571		

Tabelle 3: Vergleich der theoretischen Ergebnisse für das einjährige Reserverisiko mit den Ergebnissen aus dem Bootstrap.

	90%	95%	99%
Value-at-Risk	453,364	580,795	843,460
Expected-Shortfall	623,869	736,511	956,908

Tabelle 4: Value-at-Risk und Expected-Shortfall für die ermittelte prädiktive Verteilung.

In Abbildung 2 ist die prädiktive Verteilung des einjährigen Reserverisikos dargestellt.



Abbildung 2: Prädiktive Verteilung des einjährigen Reserverisikos.

Literatur

- [1] AISAM-ACME (2007): AISAM-ACME study on non-life long-tail liabilities. Reserve risk and risk margin assessment under Solvency II. *Insurance: Mathematics and Economics*,
- [2] Braun, C. (2004): The prediction error of the chain ladder method applied to correlated run off triangles. *ASTIN Bulletin*, 34(2): 399–423.
- [3] Brehm, P. J. (2002): Correlation and the aggregation of unpaid loss distributions. *CAS E-Forum*, 1–24.
- [4] Buchwalder, M., Bühlmann, H., Merz, M. & Wüthrich, M. V. (2006): The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). *ASTIN Bulletin*, 36(2): 521–542.
- [5] Bundesamt für Privatversicherungen (2006): Swiss Solvency Test. Schweiz.
- [6] Efron, B. & Tibshirani, R. J. (1993): *An Introduction to the Bootstrap.* 1. Aufl. Monographs on statistics and applied probability. Champam & Hall.
- [7] England, P. D. (2002): Addendum to "Analytic and bootstrap estimates of prediction errors in claims reserving". *Insurance: Mathematics and Economics*, 31(3): 461–466.
- [8] England, P. D. & Verrall, R. J. (1999): Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, 25(3): 281–293.
- [9] England, P. D. & Verrall, R. J. (2007): Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1(2): 221–270.
- [10] Heberle, J., Huergo, L. & Merz, M. (2008): Prognose des Abwicklungsergebnisses mittels Chain-Ladder-Verfahren für abhängige Run-Off-Portfolios. *Zeitschrift für die gesamte Versicherungswissenschaft*, 97(4): 439–461.

- [11] Hess, K. T., Schmidt, K. D. & Zocher, M. (2006): Multivariate loss prediction in the multivariate additive model. *Insurance: Mathematics and Economics*, 39(2): 185–191.
- [12] Houltram, A. (2003): Reserving Judgement Considerations Relevant to Insurance Liability Assessment. *XIV General Insurance Seminar, Institute of Actuaries of Australia.*
- [13] Kirschner, G. S., Kerley, C. & Isaacs, B. (2008): Two approaches to calculating correlated reserve indications across multiple lines of business. *Variance*, 2: 15–38.
- [14] Lowe, J. (1994): A practical guide to measuring reserve variability using: Bootstrapping, operational time and a distribution free approach. *Proceedings of the 1994 General Insurance Convention*.
- [15] Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2): 213–225.
- [16] Merz, M. & Wüthrich, M. V. (2008): Prediction error of the chain ladder reserving method applied to correlated run off triangles. *Annals of Actuarial Science*, 2(1): 25–50.
- [17] Merz, M. & Wüthrich, M. V. (2009): Prediction error of the multivariate additive loss reserving method for dependent lines of business. *Variance*, 3: 131–151.
- [18] Merz, M. & Wüthrich, M. V. (2008): Prediction error of the multivariate chain ladder reserving method. *North American Actuarial Journal*, 12(2): 175–197.
- [19] Nguyen, T. (2008): Handbuch der wert- und risikoorientierten Steuerung von Versicherungsunternehmen. Karlsruhe: Verlag Versicherungswirtschaft.
- [20] Pinheiro, P. J. R., Andrade e Silva, J. M. & de Lourdes Centeno, M. (2003): Bootstrap Methodology in Claim Reserving. *Journal of Risk and Insurance*, 70(4): 701–714.
- [21] Pröhl, C. & Schmidt, K. D. (2005): Multivariate chain-ladder. Dresdner Schriften zur Versicherungsmathematik, 3.
- [22] Sandström, A. (2007): Solvency, a historical review and some pragmatic solutions. *Bulletin* of Swiss Association of Actuaries, 1: 11–34.
- [23] Schmidt, K. D. (2006): Methods and models of loss reserving based on run-off triangles: a unifying survey. *CAS E-Forum*, 269–317.
- [24] Taylor, G. C. & McGuire, G. (2007): A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error. North American Actuarial Journal, 11(3): 70–88.
- [25] Wüthrich, M. V. & Merz, M. (2008): *Stochastic Claims Reserving Methods in Insurance*. Chichester: Wiley Finance.
- [26] Wüthrich, M. V., Merz, M. & Lysenko, N. (2009): Uncertainty in the claims development result in the chain ladder method. *Scandinavian Actuarial Journal*, 1: 63–84.
- [27] Zurmühl, R. & Falk, S. (1984): *Matrizen und ihre Anwendungen, Teil 1: Grundlagen*. Berlin: Springer.

Paper D

Prediction of the ultimate claims based on cumulative claims payments and incurred losses data using Kalman-filter theory

Jochen Heberle

Prediction of the ultimate claims based on cumulative claims payments and incurred losses data using Kalman-filter theory

Jochen Heberle

Abstract

One of the main tasks in non-life insurance is the prediction of outstanding loss liabilities for run-off portfolios. Additionally, the quantification of the prediction uncertainty is also of great interest. In this paper we look at this actuarial problem in a bivariate framework, i.e. we assume that we have two different sources of information, namely a paid- and an incurred run-off portfolio. Based on these two information sources we derive a formula for the ultimate claim and for the mean-square-error of prediction (MSEP) using Kalman-filter techniques. Finally, an example for a given triangle pair (paid-/incurred-triangle data) is presented.

1 Introduction

1.1 Motivation

One of the main tasks of reserving actuaries is to predict outstanding loss liabilities. Usually, this prediction is based on observed data and only one source of information is used. For example the classical Chain-Ladder method relies only on cumulative claims payments in the past.

In this paper we address the problem linked to the incorporation of two different sources of observations, namely cumulative claims payments and incurred losses data, for the prediction of the ultimate claims. Probably one of the first papers which discussed this problem is Halliwell [6] who combined claims payments and incurred losses data for claims reserving from a statistical point of view. The analysis of Halliwell [6] is based on a regression framework.

Another well-known article describes the Munich-Chain-Ladder method (cf. Quarg & Mack [10]) which also tries to predict the ultimate claims with combined claims payments and

incurred losses data. The aim of the Munich-Chain-Ladder method is to reduce the gap between two different Chain-Ladder predictions, the Chain-Ladder prediction based on cumulative claims payments and incurred losses data. The results are two (different) predictions for the same ultimate claim, which are not necessarily identical. This problem is also discussed in Dahms [3] and Merz & Wüthrich [9]. In Dahms [3] the model combines two projections of claims payments and incurred losses data, such that both projections lead to the same prediction of the ultimate claims. In Merz & Wüthrich [9] the authors use a Bayesian log-normal claims reserving model which is based on claims payments and incurred losses data. This approach also leads to one unique ultimate claim.

The goal of this paper is to derive an alternative method to predict the (unique) ultimate claims based on cumulative claims payments and incurred losses data using Kalman-filter theory. Therefore, we consider two observed triangles, one with cumulative claims payments data and the other one with incurred losses data. For both triangles we assume that they are the result of a linear transformation with noise of an unobservable (state-) triangle (cf. Figure 1 and Model Assumptions 2.1).

Kalman-filter theory was already used in claims reserving by de Jong & Zehnwirth [5], Verrall [11], Wright [12] and Atherino et al. [1]. De Jong & Zehnwirth [5] adapted the Kalman-filter method to one single development triangle and estimated the parameters using maximum likelihood methods. The idea of de Jong & Zehnwirth [5] was adopted by Verrall [11] who used a similar approach but estimated the parameters using Bayesian methods. In Atherino et al. [1] the rows of the development triangle are stacked, resulting in a univariate time series with several missing values. In Wüthrich [13] the problem of missing values is solved without the use of Kalman-filter theory, they present a direct method for solving the problem of missing values.

1.2 The development triangles

In the following $C_{i,j}^{Pa}$ denotes the cumulative claims payments in accident year $i \in \{0, ..., I\}$ up to development year $j \in \{0, ..., J\}$ and $C_{i,j}^{In}$ denotes the cumulative incurred losses data for accident year $i \in \{0, ..., I\}$ until development year $j \in \{0, ..., J\}$. To simplify notation, we assume that I = J. Of course, all formulas can be generalized to the case I > J (development trapezoids).

We assume that we are at time t = I. Therefore, we have the set of observed cumulative claims payments data

$$\mathcal{D}_{I}^{\mathrm{Pa}} = \left\{ C_{i,j}^{\mathrm{Pa}} \mid 0 \le i \le I, 0 \le j \le J \text{ with } i+j \le I \right\}$$



Figure 1: At the top is the unobservable (state-) triangle, while the two triangles in the middle are the observed triangles with cumulative payments and claims incurred data. At the bottom is the predicted triangle using Kalman-filter theory.

and the set of incurred losses data

$$\mathcal{D}_{I}^{\mathrm{In}} = \left\{ C_{i,j}^{\mathrm{In}} \mid 0 \le i \le I, 0 \le j \le J \text{ with } i+j \le I \right\}.$$

Combining $\mathcal{D}_{I}^{\mathrm{Pa}}$ and $\mathcal{D}_{I}^{\mathrm{In}}$ leads to the set

$$\mathcal{D}_I = \mathcal{D}_I^{\operatorname{Pa}} \cup \mathcal{D}_I^{\operatorname{In}}.$$

Our goal is to predict the ultimate claims $C_{i,J}$ in the unobservable (state-) triangle by means of \mathcal{D}_I . In the following $C_{i,j}$ (i = 0, ..., I and j = 0, ..., J) denote the cumulative claims data in the unobservable (state-) triangle.

1.3 Idea of state-space models

Since the Kalman-filter is a recursion algorithm associated to state-space models we give a short introduction to these models. For a more detailed introduction see for example Brockwell & Davis [2, pp. 259 sqq.] or Hamilton [7, pp. 372 sqq.].

A state-space model for a time series, denoted with Y_t ($t \in \mathbb{N}$), can be displayed using two equations:

1. The "observation equation" (cf. equation (1.1)), which expresses the observations Y_t as an affine-linear function of a state variable X_t ($t \in \mathbb{N}$) plus a noise term W_t ($t \in \mathbb{N}$):

$$Y_t = G_t X_t + W_t \tag{1.1}$$

2. The "state-space equation" (cf. equation (1.2)), which describes the stochastic dynamics of the state X_t ($t \in \mathbb{N}$) at some time t + 1 as an affine-linear function of the previous state X_t ($t \in \mathbb{N}$) plus noise term V_t ($t \in \mathbb{N}$):

$$\boldsymbol{X}_{t+1} = \boldsymbol{F}_t \boldsymbol{X}_t + \boldsymbol{V}_t \tag{1.2}$$

Thereby, G_t and F_t ($t \in \mathbb{N}$) are two sequences of matrices and W_t and V_t ($t \in \mathbb{N}$) are two white noise processes. for details on the model assumptions see Brockwell & Davis [2, pp. 260 sq.].

2 Kalman-filter theory on development triangles

The idea of Kalman-filter theory is to express a dynamical stochastic linear system in a particular form called the *state-space representation* and the Kalman-filter algorithm provides sequentially updated linear projections for the state of the dynamic system (cf. Hamilton [7]).

In the following we denote predicted values in the unobservable (state-) triangle with $\widehat{C}_{i,j}^{I}$ $(i \in \{0, \ldots, I\} \text{ and } j \in \{0, \ldots, J\})$, while the "*I*" denotes the given observation set, i.e. \mathcal{D}_{I} .

The set of two-dimensional observations $\{(C_{i,j}^{\text{Pa}}, C_{i,j}^{\text{In}}) \mid i + j \leq I\}$ is used to derive:

- 1. Kalman-predicted values $\widehat{C}_{i,j}^{I}$ with i+j > I and $i \leq I$ to get predictions for the unobservable ultimate claims $C_{i,J}$ for i = 1, ..., I (see Theorem 2.5).
- 2. Kalman-error-variances $(\hat{\sigma}_{i,j}^{I})^{2}$ (i = 0, ..., I and j = 0, ..., J) for predicted values to obtain an estimator for the mean square error of prediction. This is provided in Section 3.

There is also the possibility to calculate so called "smoothed" and "filtered" values for the upper left part of the unobservable (state-) triangle. Since we are only interested in the prediction of ultimate claims this is omitted.

To apply the results from Kalman-filter theory to the present claims reserving problem, we work with the observed claims data, respectively with the cumulative payments and claims incurred data, simultaneously.

Therefore, we make the following model assumptions:

Model Assumptions 2.1 (state-space representation of claims payments):

(1) There exist vectors $h_j = (h_j^{\text{Pa}}, h_j^{\text{In}})' \in \mathbb{R}^2$ and positive semidefinite 2 × 2-matrices $T_{i,j} \in \mathbb{R}^{2 \times 2}$, such that

$$\boldsymbol{C}_{i,j}^{\text{Pa,In}} = \begin{pmatrix} C_{i,j}^{\text{Pa}} \\ C_{i,j}^{\text{In}} \end{pmatrix} = \begin{pmatrix} h_j^{\text{Pa}} \\ h_j^{\text{In}} \end{pmatrix} \boldsymbol{C}_{i,j} + \begin{pmatrix} \boldsymbol{w}_{i,j}^{\text{Pa}} \\ \boldsymbol{w}_{i,j}^{\text{In}} \end{pmatrix} = \boldsymbol{h}_j \boldsymbol{C}_{i,j} + \boldsymbol{w}_{i,j}$$

for i = 0, ..., I and j = 0, ..., J, where $(\mathbf{w}'_{i,j})_{j=0,...,J}^{i=0,...,I} = (w_{i,j}^{\text{Pa}}, w_{i,j}^{\text{In}})_{j=0,...,J}^{i=0,...,I}$ constitutes a twodimensional white-noise-process with

$$E[\boldsymbol{w}_{i,j}] = \boldsymbol{0} \quad \text{and} \quad E[\boldsymbol{w}_{i,j}\boldsymbol{w}'_{k,l}] = \begin{cases} T_{i,j} & \text{if } i = k \text{ and } j = l \\ \boldsymbol{O} & \text{otherwise} \end{cases}$$

The 2 \times 2-matrix **O** contains zeros only.

(2) There exist parameters $g_j \in \mathbb{R}$ and $q_{i,j} \in \mathbb{R}$, such that

$$C_{i,j+1} = g_j C_{i,j} + v_{i,j}$$

for i = 0, ..., I and j = 0, ..., J - 1, where $(v_{i,j})_{j=0,...,J-1}^{i=0,...,I}$ constitutes a one-dimensional white-noise-process with

$$E[v_{i,j}] = 0 \quad \text{and} \quad E[v_{i,j}v_{k,l}] = \begin{cases} q_{i,j} & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise} \end{cases}$$

(3) The white-noise-processes $(\mathbf{w}_{i,j})_{j=0,\dots,J}^{i=0,\dots,I}$ and $(\upsilon_{i,j})_{j=0,\dots,J-1}^{i=0,\dots,I}$ are uncorrelated, i.e.

$$\mathrm{E}\left[\upsilon_{i,j}\boldsymbol{w}_{k,l}\right] = \mathbf{0}$$

for i, k = 0, ..., I, j = 0, ..., J - 1 and l = 0, ..., J.

(4) Cumulative claims $C_{i,j}$ of different accident years i = 0, ..., I are independent.

Remarks 2.2:

- If we were interested in a prediction of the whole (state-) triangle we would have to predict smoothed values ($\widehat{C}_{i,j}^{I}$ for i + j < I), filtered values ($\widehat{C}_{i,j}^{I}$ for i + j = I) and predicted values ($\widehat{C}_{i,j}^{I}$ for i + j > I) from the two-dimensional observations ($C_{i,j}^{\text{Pa}}, C_{i,j}^{\text{In}}$) with $i + j \leq I$ since the claims $C_{i,j}$ in the (state-) triangle are unobservable for all i, j.
- Model Assumption 2.1 (3) is not necessary (cf. Brockwell & Davis [2]), i.e. the Kalmanfilter algorithm can also be derived in the case of correlated noise, but this assumption makes some formulas easier to handle.
- Obviously, $C_{i,j}$ and $C_{i,j}^{\text{Pa,In}}$ for i = 0, ..., I and j = 0, ..., J can be written as

$$C_{i,j} = g_{j-1}C_{i,j-1} + v_{i,j-1} = g_{j-1}(g_{j-2}C_{i,j-2} + v_{i,j-2}) + v_{i,j-1}$$

= ... = $p_{i,j}(C_{i,0}, g_0, \dots, g_{j-1}, v_{i,0}, \dots, v_{i,j-1})$ (2.1)

and

$$C_{i,j}^{\text{Pa,In}} = r_{i,j}(C_{i,0}, \boldsymbol{h}_j, \boldsymbol{w}_{i,j}, g_0, \dots, g_{j-1}, \upsilon_{i,0}, \dots, \upsilon_{i,j-1}), \qquad (2.2)$$

respectively, where $p_{i,j}$ and $r_{i,j}$ for i = 0, ..., I and j = 0, ..., J are some appropriate functions.

• From equations (2.1) and (2.2) and the assumptions on the noise terms, it follows that

$$E[C_{i,j}v_{i,k}] = 0$$
 and $E[C_{i,j}w_{i,k}] = 0$

holds true for i = 0, ..., I and j, k = 0, ..., J with $j \le k$.

We make the following definition for the Kalman predictors.

Definition 2.3 (Best affine-linear predictors): In the Kalman-filter framework in business period I the predictors $\widehat{C}_{i,j}^{I}$ for the unobservable claims amounts $C_{i,j}$ are defined by

$$\widehat{C}_{i,j}^{I} = \operatorname{Pro}(C_{i,j} \mid \mathcal{A}^{I})$$

(i = 0, ..., I and j = 0, ..., J) where "Pro" is the orthogonal projection operator on the subspace \mathcal{A}^{I} (cf. Brockwell & Davis [2, p. 272]). The set \mathcal{A}^{I} is given by

$$\mathcal{A}^{I} = \left\{ A \mid A = a + \sum_{l=0}^{I-i} a'_{l} C^{\mathrm{Pa,In}}_{i,l} \text{ with } a \in \mathbb{R}, a_{0}, \ldots, a_{I-i} \in \mathbb{R}^{2} \text{ and } i \in \{0, \ldots, I\} \right\}.$$

This is the subspace of all affine-linear predictors.

Remarks 2.4:

• Definition 2.3 for the projection operator is equivalent to:

$$\widehat{C}_{i,j}^{I} = \underset{A \in \mathcal{A}^{I}}{\arg\min} \mathbb{E}\left[(C_{i,j} - A)^{2} \right]$$

- Definition 2.3 implies that the predictors $\widehat{C}_{i,j}^{I}$ are best (w.r.t. the MSEP) affine-linear predictors out of \mathcal{A}^{I} .
- For i + j < I the predictor $\widehat{C}_{i,j}^{I}$ is called Kalman-smoothed predictor of $C_{i,j}$ since the "prediction-time" i + j is in the past.
- For i + j = I the predictor is called Kalman-filtered predictor of $C_{i,j}$ since at time I this is the predictor for the current calendar year.
- For i + j > I the predictor $\widehat{C}_{i,j}^{I}$ is called Kalman-prediction.
 - In the case i + j = I + 1 the Kalman-prediction is a 1-step-prediction.
 - In the case i + j = I + h with h > 1 the Kalman-prediction is a *h*-step-prediction.

In the following we focus only on Kalman-predicted values, since we are only interested in the prediction of ultimate claims $C_{i,J}$ (i = 1, ..., I). Therefore, the Kalman-predicted values are denoted by $\widehat{C}_{i,j}^{I}$.

For achieving an estimator for the mean square error of prediction we have to make a definition for the error-variances of $\widehat{C}_{i,j}^{l}$. These error-variances are defined by

$$\left(\hat{\sigma}_{i,j}^{I}\right)^{2} = \mathbf{E}\left[\left(C_{i,j} - \widehat{C}_{i,j}^{I}\right)^{2}\right]$$
(2.3)

for i = 0, ..., I and j = 0, ..., J.

The following theorem describes the computation of $\widehat{C}_{i,j}^{I}$ as well as the estimation of the errorvariances $(\widehat{\sigma}_{i,j}^{I})^{2}$. The estimator for the error-variances is used in Section 3 where we derive an estimator for the mean square error of prediction of the ultimate claim predictors $\widehat{C}_{i,J}^{I}$ for $i = 1, \ldots, I$.

Theorem 2.5 (Kalman-prediction): Under Model Assumptions 2.1 and the initial values $\widehat{C}_{i,0}^{I}$ and $\left(\widehat{\sigma}_{i,0}^{I}\right)^{2}$ the (one-step) Kalman-predicted values are given by

$$\widehat{C}_{i,j+1}^{i+j} = g_j \widehat{C}_{i,j}^{i+j-1} + \gamma'_{i,j} \Delta_{i,j}^{-1} \left(C_{i,j}^{\operatorname{Pa,In}} - \boldsymbol{h}_j \widehat{C}_{i,j}^{i+j-1} \right)$$
(2.4)

and

$$\left(\hat{\sigma}_{i,j+1}^{i+j}\right)^2 = g_j^2 \left(\hat{\sigma}_{i,j}^{i+j-1}\right)^2 + q_{i,j} - \gamma_{i,j}' \Delta_{i,j}^{-1} \gamma_{i,j}$$
(2.5)

with

$$\boldsymbol{\gamma}_{i,j} = g_j \left(\hat{\sigma}_{i,j}^{i+j-1} \right)^2 \boldsymbol{h}_j$$
$$\boldsymbol{\Delta}_{i,j} = \boldsymbol{h}_j \left(\hat{\sigma}_{i,j}^{i+j-1} \right)^2 \boldsymbol{h}'_j + T_{i,j}$$

for i = 0, ..., I and j = 0, ..., J - 1.

Proof. The proof is given in Appendix A.1.

Remarks 2.6:

- The term $\Delta_{i,j}^{-1}$ in (2.4) and (2.5) is interpreted as generalized inverse of $\Delta_{i,j}$.
- The "prediction" formula (2.4) will only be used for predicting the unobservable claims for the next calendar year, i.e. the claims $C_{i,I-i+1}$ for i = 1, ..., I. This means that we are at time t = I and that we want to predict the values at time t = I + 1.

- Equation (2.4) only depends on the previous calculated value and the so called innovation $L_{i,j} = C_{i,j}^{\text{Pa,In}} h_j \widehat{C}_{i,j}^{i+j-1}$ (cf. equation (A.4)). I.e. the previous observations $C_{i,0}^{\text{Pa,In}}, \ldots, C_{i,j-1}^{\text{Pa,In}}$ are not directly needed for the calculation of $\widehat{C}_{i,j+1}^{i+j}$ since they are already contained in $\widehat{C}_{i,j}^{i+j-1}$.
- The parameters h_j , $T_{i,j}$, g_j and $q_{i,j}$ (i = 0, ..., I, j = 0, ..., J 1) have to be estimated from the given data or they have to be given for example by experts.

At time t = I we are only able to predict the claims for the next calendar year t = I + 1 by means of Theorem 2.5, i.e. the main part in the lower right (state-) triangle can not be filled. Therefore, in the following corollary we provide the Kalman algorithm for *h*-step prediction.

Corollary 2.7 (*h*-step Kalman-prediction): Under the assumptions of Theorem 2.5 the *h*-step Kalman predictions are given by

$$\widehat{C}_{i,I-i+h+1}^{I} = g_{I-i+h} \widehat{C}_{i,I-i+h}^{I} = \dots = \widehat{C}_{i,I-i+1}^{I} \prod_{l=1}^{h} g_{I-i+l}$$
(2.6)

and

$$\left(\hat{\sigma}_{i,I-i+h+1}^{I}\right)^{2} = g_{I-i+h}^{2} \left(\hat{\sigma}_{i,I-i+h}^{I}\right)^{2} + q_{i,I-i+h}, \qquad (2.7)$$

for $i \leq I$ and h = 1, ..., J - I + i - 1.

Proof. The proof is given in Appendix A.2.

3 Mean square error of prediction

3.1 Single accident years

With Theorem 2.5 and Corollary 2.7 we are able to predict the ultimate claims $C_{i,J}$ (i = 1, ..., I). In this section we want to measure the quality of these predictions by means of second moments. This means that we calculate the so-called mean square error of prediction (MSEP). At time t = I the MSEP of the ultimate claim predictors $\widehat{C}_{i,J}^{I}$ (i = 1, ..., I) is defined by

MSEP
$$\left(\widehat{C}_{i,J}^{I}\right) = \mathbb{E}\left[\left(C_{i,J} - \widehat{C}_{i,J}^{I}\right)^{2}\right].$$

The MSEP for a given accident year is equal to the corresponding error-variance (cf. equation (2.3)), i.e.:

MSEP
$$\left(\widehat{C}_{i,J}^{I}\right) = \left(\widehat{\sigma}_{i,J}^{I}\right)^{2}$$

Remarks 3.1:

- For accident year i = 1 the MSEP is given through the one-step Kalman-prediction of the error-variance given in Theorem 2.5.
- For accident years *i* > 1 the MSEPs are given through the *h*-step Kalman-predictions of the error-variances given in Corollary 2.7.

3.2 Aggregated accident years

In this section we derive an estimator for the MSEP of aggregated accident years. Therefore, we start with the MSEP for two accident years and, in a second step, extend the result to more than two accident years.

The MSEP for two accident years $i, k \in \{1, ..., I\}$ with i < k is given by:

$$MSEP\left(\widehat{C}_{i,J}^{I} + \widehat{C}_{k,J}^{I}\right) = E\left[\left(C_{i,J} - \widehat{C}_{i,J}^{I} + C_{k,J} - \widehat{C}_{k,J}^{I}\right)^{2}\right]$$

Using $E[(X + Y)^2] = E[X^2] + E[Y^2] + 2E[XY]$, where $X = C_{i,J} - \widehat{C}_{i,J}^I$ and $Y = C_{k,J} - \widehat{C}_{k,J}^I$, we achieve for the MSEP of the sum of two accident years:

$$MSEP\left(\widehat{C}_{i,J}^{Kal,P} + \widehat{C}_{k,J}^{Kal,P}\right) = E\left[\left(C_{i,J} - \widehat{C}_{i,J}^{I}\right)^{2}\right] + E\left[\left(C_{k,J} - \widehat{C}_{k,J}^{I}\right)^{2}\right] + 2 E\left[\left(C_{i,J} - \widehat{C}_{i,J}^{I}\right)\left(C_{k,J} - \widehat{C}_{k,J}^{I}\right)\right] = MSEP\left(\widehat{C}_{i,J}^{Kal,P}\right) + MSEP\left(\widehat{C}_{k,J}^{Kal,P}\right) + 2 E\left[\left(C_{i,J} - \widehat{C}_{i,J}^{I}\right)\left(C_{k,J} - \widehat{C}_{k,J}^{I}\right)\right]$$
(3.1)

Obviously, the MSEP for two accident years is the sum of the MSEPs for single accident years plus the mixed term containing both accident years. Using the independence of different accident years (cf. Model Assumption 2.1 (4)) the last term in equation (3.1) can be written as:

$$\mathbb{E}\left[\left(C_{iJ} - \widehat{C}_{iJ}^{I}\right)\left(C_{kJ} - \widehat{C}_{kJ}^{I}\right)\right] = \mathbb{E}\left[C_{iJ} - \widehat{C}_{iJ}^{I}\right] \mathbb{E}\left[C_{kJ} - \widehat{C}_{kJ}^{I}\right]$$
$$= \left(\mathbb{E}\left[C_{iJ}\right] - \mathbb{E}\left[\widehat{C}_{iJ}^{I}\right]\right)\left(\mathbb{E}\left[C_{kJ}\right] - \mathbb{E}\left[\widehat{C}_{kJ}^{I}\right]\right)$$

Since $\widehat{C}_{i,J}^{I}$ is an unbiased estimator for $\mathbb{E}[C_{i,J}]$, the last term in equation (3.1) is equal to zero. Therefore, the MSEP over all accident years is given by:

$$MSEP\left(\sum_{i=1}^{I} \widehat{C}_{i,J}^{I}\right) = \sum_{i=1}^{I} MSEP\left(\widehat{C}_{i,J}^{I}\right)$$

4 Estimation

In Model Assumptions 2.1 all cumulative payments $C_{i,j}$ (i = 0, ..., I and j = 1, ..., J) are random variables, so we have to predict $C_{i,J}$ with the given data \mathcal{D}_I . The initial values $C_{i,0}$ (i = 0, ..., I) have to be given in the model, i.e. we have to make an assumption on the initial values $\widehat{C}_{i,0}^I$. Using Kalman-filter theory, the best predictor – in terms of the MSEP – for the ultimate claims $C_{i,J}$ (i = 1, ..., I) is given through Theorem 2.5 and Corollary 2.7.

The following steps have to be realized to compute the predictions of the ultimate claims $C_{i,J}$ (i = 1, ..., I) as well as the mean square error of prediction for single accident years MSEP ($\widehat{C}_{i,J}^{I}$) and for aggregated accident years MSEP ($\sum_{i=1}^{I} \widehat{C}_{i,J}^{I}$).

- 1. The model parameters h_j , $T_{i,j}$ (i = 0, ..., I and j = 0, ..., J), g_j and $q_{i,j}$ (i = 0, ..., I and j = 0, ..., J 1) have to be estimated through the given data \mathcal{D}_I or they have to be determined in advance for example through experts.
- 2. The model parameters h_j , $T_{i,j}$, g_j and $q_{i,j}$ and the observed data \mathcal{D}_I are used to compute (one-step) predicted Kalman-values using Theorem 2.5. The error-variances $(\hat{\sigma}_{i,j}^I)^2$ are also estimated in this step, using the second part of Theorem 2.5.
- 3. The one-step predicted Kalman-values are used to compute *h*-step predicted Kalman-values using Corollary 2.7. The *h*-step error-variances $(\hat{\sigma}_{i,I-i+h+1}^{I})^2$, $h = 1, \ldots, J I + i 1$, are also estimated in this step.
- 4. In the last step estimators for the mean square error of prediction for single accident years and for aggregated accident years are computed.

5 Example

For our example we use two run-off triangles from Dahms [3] (cf. Table 1 and Table 2).

	0	1	2	3	4	5	6	7	8	9
0	1,216,632	1,347,072	1,786,877	2,281,606	2,656,224	2,909,307	3,283,388	3,587,549	3,754,403	3,921,258
1	798,924	1,051,912	1,215,785	1,349,939	1,655,312	1,926,210	2,132,833	2,287,311	2,567,056	
2	1,115,636	1,387,387	1,930,867	2,177,002	2,513,171	2,931,930	3,047,368	3,182,511		
3	1,052,161	1,321,206	1,700,132	1,971,303	2,298,349	2,645,113	3,003,425			
4	808,864	1,029,523	1,229,626	1,590,338	1,842,662	2,150,351				
5	1,016,862	1,251,420	1,698,052	2,105,143	2,385,339					
6	948,312	1,108,791	1,315,524	1,487,577						
7	917,530	1,082,426	1,484,405							
8	1,001,238	1,376,124								
9	841,930									

Table 1: Observed	cumulative	claims $C_{i,i}^{\text{Pa}}$	(triangle A).
-------------------	------------	------------------------------	---------------

	0	1	2	3	4	5	6	7	8	9
0	3,362,115	5,217,243	4,754,900	4,381,677	4,136,883	4,094,140	4,018,736	3,971,591	3,941,391	3,921,258
1	2,640,443	4,643,860	3,869,954	3,248,558	3,102,002	3,019,980	2,976,064	2,946,941	2,919,955	
2	2,879,697	4,785,531	4,045,448	3,467,822	3,377,540	3,341,934	3,283,928	3,257,827		
3	2,933,345	5,299,146	4,451,963	3,700,809	3,553,391	3,469,505	3,413,921			
4	2,768,181	4,658,933	3,936,455	3,512,735	3,385,129	3,298,998				
5	3,228,439	5,271,304	4,484,946	3,798,384	3,702,427					
6	2,927,033	5,067,768	4,066,526	3,704,113						
7	3,083,429	4,790,944	4,408,097							
8	2,761,163	4,132,757								
9	3,045,376									

Table 2: Observed cumulative claims $C_{i,j}^{\text{In}}$ (triangle B).

For predicting the unobservable state-triangle we have to make assumptions about the initial values $\widehat{C}_{i,0}^{I}$ and the initial error-variances $(\widehat{\sigma}_{i,0}^{I})^{2}$ (i = 0, ..., I). For the Kalman-parameters h_{j} , $T_{i,j}, g_{j}$ and $q_{i,j}$ (i = 0, ..., I and j = 0, ..., J - 1) assumptions have to be made as well. To avoid overparametrization we set $T = T_{i,j}$ and $q = q_{i,j}$. So the Kalman-parameters $T_{i,j}$ and $q_{i,j}$ will be constant for all accident years and all development years.

We set

$$\boldsymbol{h}_{j} = \begin{pmatrix} h_{j}^{\mathrm{Pa}} \\ h_{j}^{\mathrm{In}} \end{pmatrix} = \begin{pmatrix} \frac{j}{2J} + \frac{1}{2} \\ -\frac{j}{2J} + \frac{3}{2} \end{pmatrix}$$

to weight the incurred-data at the beginning stronger than the paid-data. For j = J both weights are equal. Additionally, the values g_i (j = 0, ..., J - 1) are set to:

$$g_j = \sqrt{\hat{f}_j^{\text{CL,Pa}} \hat{f}_j^{\text{CL,In}}}$$

The initial values $\widehat{C}_{i,0}^{I}$ as well as $(\widehat{\sigma}_{i,0}^{I})^{2}$ (i = 0, ..., I) are set to $(\alpha = 0.5)$:

$$\begin{split} \widehat{C}_{i,0}^{I} &= \alpha C_{i,0}^{\text{Pa}} + (1-\alpha) C_{i,0}^{\text{In}} \\ \widehat{\sigma}_{i,0}^{I} &= \alpha \widehat{\sigma}_{0}^{\text{CL,Pa}} + (1-\alpha) \widehat{\sigma}_{0}^{\text{CL,Ir}} \end{split}$$

Thereby, $\hat{f}_{j}^{\text{CL,Pa}}$ (respectively $\hat{f}_{j}^{\text{CL,In}}$) as well as $\hat{\sigma}_{0}^{\text{CL,Pa}}$ (respectively $\hat{\sigma}_{0}^{\text{CL,In}}$) are the well known Chain-Ladder estimators (cf. Mack [8]).

With these parameters and initial values we are able to predict the unobservable lower right part of the development triangle using Theorem 2.5 and Corollary 2.7. The estimated values are listed in Table 3. There are also so called Kalman-smoothed and Kalman-filtered values listed in Table 3 (upper left part of the estimated development triangle). These estimates are only needed if one is interested in the whole estimation of the unobservable triangle (the formulas for Kalman-smoothing and Kalman-filtering are not given in this paper).

In Table 4 the reserves in the Kalman-filter method are compared with the reserves given in

	0	1	2	3	4	5	6	7	8	9
0	1,999,733	2,853,907	2,999,620	3,057,545	3,233,783	3,428,580	3,567,794	3,672,410	3,793,668	3,867,137
1	1,471,078	2,099,445	2,206,638	2,249,252	2,378,896	2,522,192	2,624,599	2,701,555	2,790,752	2,844,797
2	1,759,518	2,511,095	2,639,304	2,690,275	2,845,337	3,016,727	3,139,214	3,231,259	3,337,945	3,402,588
3	1,738,793	2,481,527	2,608,230	2,658,607	2,811,844	2,981,213	3,102,252	3,193,210	3,298,640	3,362,521
4	1,524,015	2,175,012	2,286,065	2,330,220	2,464,528	2,612,974	2,719,061	2,798,784	2,891,191	2,947,182
5	1,824,088	2,603,265	2,736,183	2,789,029	2,949,778	3,127,449	3,254,423	3,349,843	3,460,445	3,527,460
6	1,670,545	2,384,139	2,505,869	2,554,265	2,701,480	2,864,195	2,980,482	3,067,869	3,169,161	3,230,535
7	1,708,083	2,437,716	2,562,179	2,611,659	2,762,182	2,928,554	3,047,453	3,136,805	3,240,373	3,303,126
8	1,643,611	2,345,701	2,465,463	2,513,076	2,657,917	2,818,009	2,932,420	3,018,398	3,118,057	3,178,441
9	1,646,188	2,349,377	2,469,326	2,517,014	2,662,082	2,822,425	2,937,015	3,023,128	3,122,943	3,183,422

Table 3: Estimated cumulative claims $\widehat{C}_{i,j}^{I}$.

Dahms et al. [4]. The Kalman-filter reserves always lie between the Chain-Ladder paid and the Chain-Ladder incurred reserves and also between the Munich Chain-Ladder paid and the Munich Chain-Ladder incurred reserves.

accident			res	erves		
year <i>i</i>	CL paid	CL incurred	MCL paid	MCL incurred	CLRM	Kalman
1	114,086	337,984	104,606	338,200	314,902	277,742
2	394,121	31,884	457,484	30,850	66,994	220,077
3	608,750	331,436	664,871	330,205	359,384	359,097
4	697,742	1,018,350	615,436	1,021,361	981,883	796,831
5	1,234,157	1,103,928	1,271,110	1,102,396	1,115,768	1,142,121
6	1,138,623	1,868,664	919,102	1,894,861	1,786,947	1,742,958
7	1,638,793	1,997,651	1,498,163	2,020,310	1,942,518	1,818,721
8	2,359,939	1,418,779	3,181,319	1,320,492	1,569,657	1,802,317
9	1,979,401	2,556,612	1,602,089	2,703,242	2,590,718	2,341,492
Σ	10,165,612	10,665,287	10,314,181	10,761,918	10,728,771	10,501,355

Table 4: Estimated claims reserves for the Chain-Ladder (CL) method, the Munich Chain-Ladder (MCL) method, the complementary loss ratio method (CLRM) and the Kalman-filter method.

The square roots of the MSEPs for aggregated accident years for the Kalman-filter method as well as for the Chain-Ladder and the complementary loss ratio method are given in Table 5.

method	$\widehat{\text{MSEP}}^{1/2}$
CL paid	1,517,480
CL incurred	455,794
CLRM paid	467,814
CLRM incurred	471,873
Kalman	443,149

Table 5: MSEPs for the ultimate claim prediction $\widehat{C}_{i,J}$ for $C_{i,J}$ from the Chain-Ladder (CL) method, the complementary loss ratio method (CLRM) and the Kalman-filter method.

A Appendix: Proofs

In Section A.1 we provide the proof of Theorem 2.5 and in Section A.2 the proof of Corollary 2.7 will be given.

A.1 Proof of Theorem 2.5

Proof. The following equation describes the best affine-linear predictor of $C_{i,j}$ in terms of all observations \mathcal{D}_I :

$$\operatorname{Pro}_{I}(C_{i,j}) = \operatorname{Pro}(C_{i,j} \mid A^{I}) = \underset{A \in \mathcal{A}^{I}}{\operatorname{arg\,min}} \operatorname{E}\left[(C_{i,j} - A)^{2}\right]$$
(A.1)

with

$$\mathcal{R}^{I} = \left\{ A \mid A = a + \sum_{l=0}^{I-i} \boldsymbol{a}'_{l} \boldsymbol{C}^{\mathrm{Pa,In}}_{i,l} \text{ with } a \in \mathbb{R}, \boldsymbol{a}_{0}, \dots, \boldsymbol{a}_{I-i} \in \mathbb{R}^{2} \text{ and } i \in \{0, \dots, I\} \right\}.$$

We start the proof with the innovation $L_{i,j}$ defined by $L_{i,0} = C_{i,0}^{\text{Pa,In}}$ and

$$L_{i,j} = C_{i,j}^{\text{Pa,In}} - \text{Pro}_{i+j-1}(C_{i,j}^{\text{Pa,In}})$$
(A.2)

for i = 0, ..., I and j = 1, ..., J. $Pro_{i+j-1}(C_{i,j}^{Pa,In})$ has to be interpreted as:

$$\operatorname{Pro}_{i+j-1}(C_{i,j}^{\operatorname{Pa,In}}) = \begin{pmatrix} \operatorname{Pro}_{i+j-1}(C_{i,j}^{\operatorname{Pa}}) \\ \operatorname{Pro}_{i+j-1}(C_{i,j}^{\operatorname{In}}) \end{pmatrix}$$
(A.3)

Using Model Assumptions 2.1 (1), Definition 2.3 and equation (A.3) formula (A.2) can be written as

$$\boldsymbol{L}_{i,j} = \boldsymbol{C}_{i,j}^{\text{Pa,In}} - \boldsymbol{h}_{j} \widehat{C}_{i,j}^{i+j-1} = \boldsymbol{h}_{j} C_{i,j} + \boldsymbol{w}_{i,j} - \boldsymbol{h}_{j} \widehat{C}_{i,j}^{i+j-1} = \boldsymbol{h}_{j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) + \boldsymbol{w}_{i,j}$$
(A.4)

for i = 0, ..., I and j = 1, ..., J.

In the Kalman-filter framework, every $y \in \mathbb{R}$ of the form

$$y = a_0 C_{i,0} + \ldots + a_k C_{i,k}$$

can be expressed as

$$y = b_0 C_{i,0} + \ldots + b_{k-1} C_{i,k-1} + d'_k L_{i,k}$$

where $b_0, \ldots, b_{k-1} \in \mathbb{R}$ and $d_k \in \mathbb{R}^2$ only depends on $a_0, \ldots, a_k \in \mathbb{R}$. Together with $\mathbb{E}[C_{i,j}L_{i,k}] = \mathbf{0}$ for all j < k it follows that:

$$\operatorname{Pro}_{i+k}(\cdot) = \operatorname{Pro}_{i+k-1}(\cdot) + \operatorname{Pro}(\cdot \mid L_{i,k})$$

The expression $Pro(\cdot | L_{i,k})$ is meant as best affine-linear predictor only depending on $L_{i,k}$ (cf. (A.1)). This leads to:

$$\widehat{C}_{i,j+1}^{i+j} = \operatorname{Pro}_{i+j}(C_{i,j+1}) = \operatorname{Pro}_{i+j-1}(C_{i,j+1}) + \operatorname{Pro}(C_{i,j+1} \mid L_{i,j})$$
(A.5)

Furthermore, we define

$$\boldsymbol{\gamma}_{i,j} = \mathbf{E}\left[C_{i,j+1}\boldsymbol{L}_{i,j}\right] \tag{A.6}$$

and

$$\Delta_{i,j} = \mathbb{E}\left[L_{i,j}L'_{i,j}\right] \tag{A.7}$$

for i = 0, ..., I and j = 0, ..., J.

Equation (A.9) as well as (A.10) holds true, since (cf. last point in Remarks 2.2):

$$\mathbb{E}\left[C_{i,j}v_{i,j}\right] = 0 \qquad \text{and} \qquad \mathbb{E}\left[C_{i,j}\boldsymbol{w}_{i,j}\right] = \boldsymbol{0} \qquad (A.8)$$

Using Model Assumptions 2.1 (2) and (3) and equations (A.4) and (A.8) $\gamma_{i,j}$ can be written as:

$$\begin{aligned} \boldsymbol{\gamma}_{i,j} &= \mathbb{E}\left[\left(g_j C_{i,j} + v_{i,j} \right) \left(\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \boldsymbol{h}_j + \boldsymbol{w}_{i,j} \right) \right] \\ &= g_j \mathbb{E}\left[C_{i,j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \right] \boldsymbol{h}_j + g_j \mathbb{E}\left[C_{i,j} \boldsymbol{w}_{i,j} \right] \\ &+ \mathbb{E}\left[v_{i,j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \right] \boldsymbol{h}_j + \mathbb{E}\left[v_{i,j} \boldsymbol{w}_{i,j} \right] \\ &= g_j \left(\widehat{\sigma}_{i,j}^{i+j-1} \right)^2 \boldsymbol{h}_j \end{aligned}$$
(A.9)

With (A.4) and (A.8) we get for $\Delta_{i,j}$:

$$\Delta_{i,j} = \mathbb{E} \left[\left(\boldsymbol{h}_{j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) + \boldsymbol{w}_{i,j} \right) \left(\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \boldsymbol{h}_{j}' + \boldsymbol{w}_{i,j}' \right) \right] \\ = \boldsymbol{h}_{j} \mathbb{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right)^{2} \right] \boldsymbol{h}_{j}' + \boldsymbol{h}_{j} \mathbb{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \boldsymbol{w}_{i,j}' \right] \right] \\ + \mathbb{E} \left[\boldsymbol{w}_{i,j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \right] \boldsymbol{h}_{j}' + \mathbb{E} \left[\boldsymbol{w}_{i,j} \boldsymbol{w}_{i,j}' \right] \\ = \boldsymbol{h}_{j} \mathbb{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right)^{2} \right] \boldsymbol{h}_{j}' + \mathbb{E} \left[\boldsymbol{w}_{i,j} \boldsymbol{w}_{i,j}' \right] \\ = \boldsymbol{h}_{j} \left(\widehat{\sigma}_{i,j}^{i+j-1} \right)^{2} \boldsymbol{h}_{j}' + \boldsymbol{T}_{i,j}$$
(A.10)

With $\gamma_{i,j}$ and $\Delta_{i,j}$, i.e. with equation (A.9) and equation (A.10), and with the fact that $Pro(C_{i,j+1} | L_{i,j})$ can be displayed as

$$Pro(C_{i,j+1} \mid L_{i,j}) = \boldsymbol{m}L_{i,j}$$

where the two-dimensional vector \boldsymbol{m} is given through

$$m = E [C_{i,j+1}L'_{i,j}](E [L_{i,j}L'_{i,j}])^{-1}$$

and $(E[L_{i,j}L'_{i,j}])^{-1}$ is any generalized inverse of $E[L_{i,j}L'_{i,j}]$ (cf. Brockwell & Davis [2, p. 272]), equation (A.5) can be written as (using Model Assumptions 2.1 (2) and equations (A.6) and (A.7)):

$$\widehat{C}_{i,j+1}^{i+j} = \operatorname{Pro}_{i+j-1}(g_j C_{i,j} + v_{i,j}) + \operatorname{E}[C_{i,j+1}L'_{i,j}](\operatorname{E}[L_{i,j}L'_{i,j}])^{-1}L_{i,j}
= g_j \widehat{C}_{i,j}^{i+j-1} + \gamma'_{i,j} \Delta_{i,j}^{-1}L_{i,j}$$
(A.11)

Using $L_{i,j} = C_{i,j}^{\text{Pa,In}} - h_j \widehat{C}_{i,j}^{i+j-1}$ (cf. equation (A.4)) equation (A.11) leads to formula (2.4):

$$\widehat{C}_{i,j+1}^{i+j} = g_j \widehat{C}_{i,j}^{i+j-1} + \boldsymbol{\gamma}'_{i,j} \boldsymbol{\Delta}_{i,j}^{-1} \left(C_{i,j}^{\text{Pa,In}} - \boldsymbol{h}_j \widehat{C}_{i,j}^{i+j-1} \right)$$

The next step is to verify the formula for the error-variances (cf. equation (2.5)). We observe that (cf. (A.8))

$$\left(\hat{\sigma}_{i,j+1}^{i+j} \right)^2 = \mathbf{E} \left[\left(C_{i,j+1} - \widehat{C}_{i,j+1}^{i+j} \right)^2 \right]$$

= $\mathbf{E} \left[C_{i,j+1}^2 \right] - 2 \mathbf{E} \left[C_{i,j+1} \widehat{C}_{i,j+1}^{i+j} \right] + \mathbf{E} \left[\left(\widehat{C}_{i,j+1}^{i+j} \right)^2 \right]$
= $\mathbf{E} \left[C_{i,j+1}^2 \right] - \mathbf{E} \left[\left(\widehat{C}_{i,j+1}^{i+j} \right)^2 \right]$ (A.12)

holds true for $i = 0, \ldots, I$ and $j = 0, \ldots, J - 1$.

The first term of equation (A.12) can be displayed as follows:

$$E[C_{i,j+1}^{2}] = E[(g_{j}C_{i,j} + v_{i,j})^{2}] = g_{j}^{2} E[C_{i,j}^{2}] + 2g_{j} E[C_{i,j}v_{i,j}] + E[v_{i,j}^{2}]$$

= $g_{j}^{2} E[C_{i,j}^{2}] + q_{i,j}$ (A.13)

The second part of equation (A.12) together with equation (A.11) leads to:

$$E\left[\left(\widehat{C}_{i,j+1}^{i+j}\right)^{2}\right] = E\left[\left(g_{j}\widehat{C}_{i,j}^{i+j-1} + \boldsymbol{\gamma}_{i,j}^{\prime}\boldsymbol{\Delta}_{i,j}^{-1}\boldsymbol{L}_{i,j}\right)^{2}\right]$$
$$= g_{j}^{2} E\left[\left(\widehat{C}_{i,j}^{i+j-1}\right)^{2}\right]$$
$$+ 2g_{j}\widehat{C}_{i,j}^{i+j-1}\boldsymbol{\gamma}_{i,j}^{\prime}\boldsymbol{\Delta}_{i,j}^{-1} E\left[\boldsymbol{L}_{i,j}\right]$$
$$+ \boldsymbol{\gamma}_{i,j}^{\prime}\boldsymbol{\Delta}_{i,j}^{-1} E\left[\boldsymbol{L}_{i,j}L_{i,j}^{\prime}\right](\boldsymbol{\Delta}_{i,j}^{-1})^{\prime}\boldsymbol{\gamma}_{i,j}$$
(A.14)

 $E[L_{i,j}]$ can be written as (cf. Model Assumptions 2.1 (1)):

$$\mathbf{E}\left[\boldsymbol{L}_{i,j}\right] = \mathbf{E}\left[\boldsymbol{C}_{i,j}^{\mathrm{Pa,In}} - \boldsymbol{h}_{j}\widehat{\boldsymbol{C}}_{i,j}^{i+j-1}\right] = \boldsymbol{h}_{j}\mathbf{E}\left[\boldsymbol{C}_{i,j} - \widehat{\boldsymbol{C}}_{i,j}^{i+j-1}\right] + \mathbf{E}\left[\boldsymbol{w}_{i,j}\right] = \boldsymbol{0}$$

Therefore, since $(\Delta_{i,j}^{-1})' = (\Delta_{i,j}')^{-1} = \Delta_{i,j}^{-1}$ holds true, equation (A.14) leads to:

$$E\left[\left(\widehat{C}_{i,j+1}^{i+j}\right)^{2}\right] = g_{j}^{2} \operatorname{E}\left[\left(\widehat{C}_{i,j}^{i+j-1}\right)^{2}\right] + \boldsymbol{\gamma}_{i,j}^{\prime} \boldsymbol{\Delta}_{i,j}^{-1} \boldsymbol{\gamma}_{i,j}$$
(A.15)

Using (A.13) and (A.15) we obtain for (A.12):

$$\begin{pmatrix} \hat{\sigma}_{i,j+1}^{i+j} \end{pmatrix}^2 = g_j^2 \operatorname{E} [C_{i,j}^2] + q_{i,j} - g_j^2 \operatorname{E} \left[\left(\widehat{C}_{i,j}^{i+j-1} \right)^2 \right] - \gamma'_{i,j} \Delta_{i,j}^{-1} \gamma_{i,j}$$

$$= g_j^2 \operatorname{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right)^2 \right] + q_{i,j} - \gamma'_{i,j} \Delta_{i,j}^{-1} \gamma_{i,j}$$

$$= g_j^2 \left(\widehat{\sigma}_{i,j}^{i+j-1} \right)^2 + q_{i,j} - \gamma'_{i,j} \Delta_{i,j}^{-1} \gamma_{i,j}$$

This completes the proof.

A.2 Proof of Corollary 2.7

Proof. Equation (2.6) can be verified using Model Assumptions 2.1 (2):

$$\widehat{C}_{i,I-i+h+1}^{I} = \operatorname{Pro}_{I}(C_{i,I-i+h+1}) = g_{I-i+h} \operatorname{Pro}_{I}(C_{i,I-i+h})$$

$$\vdots$$

$$\widehat{C}_{i,I-i+h+1}^{I} = \operatorname{Pro}_{I}(C_{i,I-i+h+1}) = (g_{I-i+h}g_{I-i+h-1} \dots g_{I-i+1}) \operatorname{Pro}_{I}(C_{i,I-i+1})$$

for h = 1, ..., J - I + i - 1.

To verify equation (2.7) we use Model Assumptions 2.1 (2) and get the relation

$$C_{i,I-i+h+1} - \operatorname{Pro}_{I}(C_{i,I-i+h+1}) = g_{I-i+h}(C_{i,I-i+h} - \operatorname{Pro}_{I}(C_{i,I-i+h})) + v_{i,I-i+h}.$$
 (A.16)

Equation (A.16) is equivalent to:

$$C_{i,I-i+h+1} - \widehat{C}_{i,I-i+h+1}^{I} = g_{I-i+h}(C_{i,I-i+h} - \widehat{C}_{i,I-i+h}^{I}) + v_{i,I-i+h}$$

This leads to:

$$E\left[\left(C_{i,I-i+h+1} - \widehat{C}_{i,I-i+h+1}^{I}\right)^{2}\right] = E\left[\left(g_{I-i+h}\left(C_{i,I-i+h} - \widehat{C}_{i,I-i+h}^{I}\right) + \upsilon_{i,I-i+h}\right)^{2}\right]$$
$$= g_{I-i+h}^{2} E\left[\left(C_{i,I-i+h} - \widehat{C}_{i,I-i+h}^{I}\right)^{2}\right]$$
$$+ 2g_{I-i+h} E\left[\left(C_{i,I-i+h} - \widehat{C}_{i,I-i+h}^{I}\right)\upsilon_{i,I-i+h}\right]$$
$$+ E\left[\upsilon_{i,I-i+h}^{2}\right]$$
(A.17)

Using (2.3) and $E[C_{i,j}v_{i,j}] = 0$ equation (A.17) can be shortened to:

$$\left(\hat{\sigma}_{i,I-i+h+1}^{I}\right)^{2} = g_{I-i+h}^{2} \left(\hat{\sigma}_{i,I-i+h}^{I}\right)^{2} + q_{i,I-i+h}$$

for h = 1, ..., J - I + i - 1. This completes the proof.

References

- Atherino, R., Pizzinga, A. & Fernandes, C. (2010): A Row-Wise Stacking of the Runoff Triangle: State Space Alternatives for IBNR Reserve Prediction. *ASTIN Bulletin*, 40: 917– 946.
- [2] Brockwell, P. J. & Davis, R. A. (2002): Introduction to Time Series and Forecasting. Springer.
- [3] Dahms, R. (2008): A loss reserving method for incomplete claim data. *Bulletin of Swiss* Association of Actuaries, 127–148.
- [4] Dahms, R., Merz, M. & Wüthrich, M. V. (2009): Claims development result for combined claims incurred and claims paid data. *Bulletin Francais d'Actuariat*, 9(18): 5–39.
- [5] De Jong, P. & Zehnwirth, B. (1983): Claims reserving, state-space models and the Kalman filter. *Journal of the Institute of Actuaries*, 110: 157–181.
- [6] Halliwell, L. J. (1997): Conjoint Prediction of Paid and Incurred Losses. *CAS E-Forum*, 1: 241–380.

- [7] Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press.
- [8] Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2): 213–225.
- [9] Merz, M. & Wüthrich, M. V. (2010): Paid-incurred chain claims reserving method. *Insurance: Mathematics and Economics*, 46(3): 568–579.
- [10] Quarg, G. & Mack, T. (2004): Munich Chain Ladder. *Blätter der DGVFM*, 26(4): 597–630.
- [11] Verrall, R. J. (1989): A state space representation of the chain ladder linear model. *Journal of the Institute of Actuaries*, 116: 589–609.
- [12] Wright, T. S. (1990): A stochastic method for claims reserving in general insurance. *Journal of the Institute of Actuaries*, 117: 677–731.
- [13] Wüthrich, M. V. (2012): A row-wise stacking of the run-off triangle: revisited.

Paper E

Bootstrapping the ultimate claims of cumulative claims payments and incurred losses data using Kalman-filter theory

Jochen Heberle

Bootstrapping the ultimate claims of cumulative claims payments and incurred losses data using Kalman-filter theory

Jochen Heberle

Abstract

The prediction of outstanding loss liabilities for non-life run-off portfolios and the quantification of the prediction error is one of the most important actuarial tasks in non-life insurance. In this paper we consider this prediction problem in a bivariate context. More precisely, we derive the predictive bootstrap distribution of the aggregated claims reserve based on cumulative claims payments and incurred losses data using Kalman-filter theory. With the predictive bootstrap distribution we are able to calculate all risk measures such as Value-at-Risk, Expected-Shortfall or Conditional-Tail-Expectation. An example is given at the end with a calculation of the bootstrap distribution and various risk measures are presented.

1 Introduction and motivation

Claims reserving is one of the most important actuarial tasks in non-life insurance since

- 1. claims reserves are often the largest position on the liability side of the balance sheet;
- 2. the reserve risk is one of the main risk drivers in non-life insurance and
- 3. the total amount of the claims reserves has a major impact on income statements (cf. Heberle et al. [15]).

Therefore, based on the available information, the prediction of adequate claims reserves for a non-life run-off portfolio as well as the quantification of the prediction uncertainty in these reserves is a major task in actuarial practice and science (cf. Rodermund [22], Teugels & Sundt [25] and Wüthrich & Merz [26]).

There exists a lot of claims reserving methods which bases on only one source of information, for example cumulative claims payments. Our aim is to present a statistical method for combining

cumulative claims payments *and* incurred losses data. Halliwell [12, 13] was one of the first who combined these two sources of information. Here, we use Kalman-filter and bootstrapping theory for combining these two sources of information.

1.1 Predictive distribution of the claims reserves

In this paper we want to estimate the predictive distribution of the aggregated claims reserve for cumulative claims payments and incurred losses data using Kalman-filter and bootstrapping theory. Based on the context of Kalman-filter theory, only the first two moments of the aggregated claims payments can be predicted. Since we are interested in the whole predictive distribution – this is often the main task in stochastic claims reserving (cf. England & Verrall [9]) – we use bootstrapping techniques to achieve the predictive distribution.

In practical applications and solvency considerations the regulator often requires (cf. Federal Office of Private Insurance [10]) predictions for second moments such as the mean square error of prediction (MSEP), the claims development result (CDR) or risk measures such as Value-at-Risk, Expected-Shortfall or Conditional-Tail-Expectation. However, the predictive distribution based on a portfolio of paid and incurred data contains more information than only best estimates of the expected claims reserve and estimates of the MSEP. This additional information is valuable for solvency discussions, capital modelling and risk management considerations. More precisely, any risk measure, such as

- 1. best estimates of the expected loss liabilities;
- 2. estimates of the corresponding MSEP;
- 3. estimates for predicting intervals, quantiles (such as Value-at-Risk) and more sophisticated risk measures such as the Expected-Shortfall or the Conditional-Tail-Expectation

can be evaluated.

Unfortunately, in most stochastic claims reserving models which have been proposed up to now, one is not able to calculate the predictive distribution analytically since the distribution of the sum of random variables is required (cf. Wüthrich & Merz [26], Section 4.2.3 and 4.3.1 for two exceptions). Therefore, in many cases one is forced to adopt simulation techniques such as bootstrapping or Markov Chain Monte Carlo methods to generate a simulated predictive distribution for the claims payments.

1.2 Claims reserving based on cumulative claims payments and incurred losses data

The prediction of ultimate claims for different accident years leads to a prediction problem based on different sources of information: cumulative claims payments and incurred losses data. For example the Munich Chain-Ladder method (cf. Quarg & Mack [21]) is one method that describes a technique that deals with this problem. The weakness of the Munich Chain-Ladder method lies in different predictions for the ultimate claims. That means the Munich Chain-Ladder reduces the gap between two independent Chain-Ladder predictions based on paid data and on incurred losses data, respectively. Another method, which is not distribution-free, is the PIC model (cf. Merz & Wüthrich [20]). This model combines cumulative claims payments and incurred losses data as well. The main idea behind the PIC model is the combination of Hertig's model (cf. Hertig [16]) and Gogol's model (cf. Gogol [11]) which leads to a log-normal paid-incurred chain (PIC) model.

In this paper we present a distribution-free and simulation-based technique to achieve the whole predictive distribution of the aggregated reserve using both sources of information. In Liu & Verrall [18] a similar approach is used, i.e. the two sources of information, namely the cumulative claims payments and the incurred losses data, are used to predict reserves. Liu & Verrall also bootstrap paid and incurred data, but the main difference between their approach and the approach used in this paper is that their underlying model is the Munich Chain-Ladder model (cf. Quarg & Mack [21]) while here a Kalman-filter framework is used.

1.3 Bootstrapping in claims reserving

The bootstrap technique, which was introduced by Efron [5] (see also Efron & Tibshirani [7] or Efron [6]), is a powerful statistical technique. It can be used for example to obtain information about an aggregate distribution only using a single dataset. In the claims reserving literature bootstrap methods are widespread (cf. for example Brehm [1], Taylor [24], Kirschner et al. [17], Heberle et al. [15] or England & Verrall [8]) and have become very popular in actuarial practice, since bootstrapping

- 1. is a well established statistical technique;
- 2. is very easy to apply since computer resources are very cheap nowadays;
- 3. can often be derived in a few lines of programming code.

In the following section we assume a model structure which allows us to apply bootstrapping techniques. We use the given set of observations to re-sample (pseudo) observations and to fit the given model to the (pseudo) observations. This means, that the bootstrapping approach can be seen as "simulating from an estimated model" by the resampling approach (cf. Wüthrich & Merz [26]). In the following we use the Kalman-filter for the prediction of ultimate claims. Afterwards we use bootstrapping techniques to derive the predictive distribution of the aggregated claims payments.

2 Kalman-filter theory applied to development triangles

In the following we consider the Kalman-filter approach provided by Heberle [14]. This approach is suitable for two run-off portfolios, namely

- 1. cumulative claims payments and
- 2. incurred losses data.

2.1 Notation

We denote by $C_{i,j}^{\text{Pa}}$ cumulative claims payments for accident year $i \in \{0, \ldots, I\}$ until development year $j \in \{0, \ldots, J\}$ and $C_{i,j}^{\text{In}}$ denotes incurred losses data for accident year $i \in \{0, \ldots, I\}$ until development year $j \in \{0, \ldots, J\}$ (cf. Figure 1). For reasons of simplicity, we assume that I = J. Of course, all formulas similarly hold true for I > J (development trapezoids). We assume that



Figure 1: Development triangle for cumulative claims payments data and incurred losses data.

we are at time t = I, i.e. the upper left part in each run-off triangle is observable, while the lower right part has to be predicted (cf. Figure 1). At time t = I we have observations

$$\mathcal{D}_{I}^{\mathrm{Pa}} = \left\{ C_{i,j}^{\mathrm{Pa}} \mid i+j \leq I \right\}$$

and

$$\mathcal{D}_{I}^{\mathrm{In}} = \left\{ C_{i,j}^{\mathrm{In}} \mid i+j \leq I \right\}.$$

Altogether, at time t = I, we have observations

$$\mathcal{D}_I = \mathcal{D}_I^{\operatorname{Pa}} \cup \mathcal{D}_I^{\operatorname{In}}.$$

Since we use Kalman-filter theory in our framework we assume an unobservable (state-) triangle (cf. Heberle [14]) with values $C_{i,j}$ (i = 0, ..., I and j = 0, ..., J). Predictions of these unobservable values are denoted by $\widehat{C}_{i,j}^{I}$ (i = 0, ..., I and j = 0, ..., J), while the superscript "I" denotes the business period, i.e. the current business period the prediction is based on. The estimates of the variance-terms is denoted in the same way.

We make the following Model Assumptions (cf. Heberle [14]):

Model Assumptions 2.1:

(1) There exist vectors $h_j = (h_j^{\text{Pa}}, h_j^{\text{In}})' \in \mathbb{R}^2$ and positive semidefinite 2×2 -matrices $T_{i,j} \in \mathbb{R}^{2 \times 2}$, such that

$$C_{ij}^{\text{Pa,In}} = \begin{pmatrix} C_{ij}^{\text{Pa}} \\ C_{ij}^{\text{In}} \end{pmatrix} = \begin{pmatrix} h_j^{\text{Pa}} \\ h_j^{\text{In}} \end{pmatrix} C_{ij} + \begin{pmatrix} w_{ij}^{\text{Pa}} \\ w_{ij}^{\text{In}} \end{pmatrix} = \boldsymbol{h}_j C_{ij} + \boldsymbol{w}_{ij},$$

for i = 0, ..., I and j = 0, ..., J where $(\mathbf{w}_{i,j})_{j=0,...,J}^{i=0,...,I}$ constitutes a two-dimensional whitenoise-process with

$$E[\boldsymbol{w}_{i,j}] = \mathbf{0} \quad \text{and} \quad E[\boldsymbol{w}_{i,j}\boldsymbol{w}'_{k,l}] = \begin{cases} T_{i,j} & \text{if } i = k \text{ and } j = l \\ O & \text{otherwise} \end{cases}$$

The 2 \times 2-matrix *O* contains zeros only.

(2) There exist parameters $g_j \in \mathbb{R}$ and $q_{i,j} \in \mathbb{R}$, such that

$$C_{i,j+1} = g_j C_{i,j} + v_{i,j},$$

for i = 0, ..., I and j = 0, ..., J - 1, where $(v_{i,j})_{j=0,...,J-1}^{i=0,...,I}$ constitutes a one-dimensional white-noise-process with

$$\mathbf{E}\left[v_{i,j}\right] = 0 \quad \text{and} \quad \mathbf{E}\left[v_{i,j}v_{k,l}\right] = \begin{cases} q_{i,j} & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise} \end{cases}.$$

(3) The white-noise-processes $(\mathbf{w}_{i,j})_{j=0,\dots,J}^{i=0,\dots,I}$ and $(v_{i,j})_{j=0,\dots,J-1}^{i=0,\dots,I}$ are uncorrelated, i.e.

$$\mathrm{E}\left[\upsilon_{i,j}\boldsymbol{w}_{k,l}\right] = \mathbf{0}$$

for i, k = 0, ..., I, j = 0, ..., J - 1 and l = 0, ..., J.

(4) Cumulative claims $C_{i,j}$ of different accident years i = 0, ..., I are independent.

For achieving predicted values for the unobservable (state-) triangle using Kalman-filter theory we make the following definition.

Definition 2.2 (Best affine-linear predictors): In the Kalman-filter framework in business period I the predictors $\widehat{C}_{i,j}^{I}$ for the unobservable claims amounts $C_{i,j}$ are defined by

$$\widehat{C}_{i,j}^{I} = \operatorname{Pro}(C_{i,j} \mid \mathcal{A}^{I})$$

(i = 0, ..., I and j = 0, ..., J) where "Pro" is the orthogonal projection operator on the subspace \mathcal{A}^{I} (cf. Brockwell & Davis [2, p. 272]). The set \mathcal{A}^{I} is given by

$$\mathcal{A}^{I} = \left\{ A \mid A = a + \sum_{l=0}^{I-i} a'_{l} C_{i,l}^{\mathrm{Pa,In}} \text{ with } a \in \mathbb{R}, a_{0}, \ldots, a_{I-i} \in \mathbb{R}^{2} \text{ and } i \in \{0, \ldots, I\} \right\}.$$

This is the subspace of all affine-linear predictors.

Remarks 2.3: Predictions of these unobservable values are called

- Kalman-smoothed values if i + j < I,
- Kalman-filtered values if i + j = I and
- Kalman-predicted values if i + j > I.

The Kalman-smoothed values are computed through an iterative process which is described together with the Kalman-filtered and Kalman-predicted values in Theorem 2.4.

The Kalman-estimators for the error-variances are denoted with $\left(\hat{\sigma}_{i,j}^{I}\right)^{2}$ and defined by

$$\left(\hat{\sigma}_{i,j}^{I}\right)^{2} = \mathbf{E}\left[\left(C_{i,j} - \widehat{C}_{i,j}^{I}\right)^{2}\right]$$

for i = 0, ..., I and j = 0, ..., J. Additionally a covariance term is needed for the calculation of the Kalman-smoothed values. This term is defined by

$$\left(\check{\sigma}_{i,j}^{i+j+k}\right)^{2} = \mathbb{E}\left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1}\right)\left(C_{i,j+k} - \widehat{C}_{i,j+k}^{i+j+k-1}\right)\right]$$

for $k \ge 1$, i = 0, ..., I and j = 0, ..., J.

Using Definition 2.2 the predictors for $C_{i,j}$ ($i + j \le I$) are given in Theorem 2.4.

Theorem 2.4 (Kalman formulas): Given Model Assumptions 2.1 and the initial values $\widehat{C}_{i,0}^{I}$ and $\left(\widehat{\sigma}_{i,0}^{I}\right)^{2}$ the smoothed, filtered and (one-step) Kalman-predicted values are given by

$$\widehat{C}_{i,j}^{i+j+k} = \widehat{C}_{i,j}^{i+j+k-1} + \left(\check{\sigma}_{i,j}^{i+j+k}\right)^2 \boldsymbol{h}_{j+k}^{\prime} \Delta_{i,j+k}^{-1} \left(C_{i,j+k}^{\text{Pa,In}} - \boldsymbol{h}_{j+k} \widehat{C}_{i,j+k}^{i+j+k-1} \right)$$
(2.1)

$$\widehat{C}_{i,I-i}^{I} = \widehat{C}_{i,I-i}^{I-1} + \left(\widehat{\sigma}_{i,I-i}^{I-1}\right)^{2} \mathbf{h}_{I-i}^{\prime} \Delta_{i,I-i}^{-1} \left(C_{i,I-i}^{\text{Pa,In}} - \mathbf{h}_{I-i}\widehat{C}_{i,I-i}^{I-1}\right)$$

$$\widehat{C}_{i,I-i}^{I} = \widehat{C}_{i,I-i}^{I-1} + \left(\widehat{\sigma}_{i,I-i}^{I-1}\right) - \widehat{C}_{i,I-i}^{I-1} + \widehat{C}_{i,I-i}^{I-1} - \widehat{C}_{i,I-i}^{I-1} + \widehat{C}_{i,I-i}^{I-1} - \widehat{C}_{i,I-i}^{I-1}$$

$$\widehat{C}_{i,j+1}^{i+j} = g_j \widehat{C}_{i,j}^{i+j-1} + \gamma'_{i,j} \Delta_{i,j}^{-1} \left(C_{i,j}^{\text{Pa,In}} - h_j \widehat{C}_{i,j}^{i+j-1} \right)$$
(2.3)

with

$$\left(\check{\sigma}_{i,j}^{i+j+k}\right)^{2} = \left(\check{\sigma}_{i,j}^{i+j+k-1}\right)^{2} \left(g_{j+k-1} - \gamma_{i,j+k-1}^{\prime} \Delta_{i,j+k-1}^{-1} h_{j+k-1}\right)$$
(2.4)

$$\left(\hat{\sigma}_{i,j}^{i+j+k}\right)^{2} = \left(\hat{\sigma}_{i,j}^{i+j+k-1}\right)^{2} - \left(\check{\sigma}_{i,j}^{i+j+k}\right)^{4} \mathbf{h}_{j+k}^{\prime} \mathbf{\Delta}_{i,j+k}^{-1} \mathbf{h}_{j+k}$$
(2.5)

$$\left(\hat{\sigma}_{i,I-i}^{I}\right)^{2} = \left(\hat{\sigma}_{i,I-i}^{I-1}\right)^{2} - \left(\hat{\sigma}_{i,I-i}^{I-1}\right)^{4} \boldsymbol{h}_{I-i}^{\prime} \boldsymbol{\Delta}_{i,I-i}^{-1} \boldsymbol{h}_{I-i}$$

$$(2.6)$$

$$\left(\hat{\sigma}_{i,j+1}^{i+j}\right)^2 = g_j^2 \left(\hat{\sigma}_{i,j}^{i+j-1}\right)^2 + q_{i,j} - \boldsymbol{\gamma}_{i,j}' \boldsymbol{\Delta}_{i,j}^{-1} \boldsymbol{\gamma}_{i,j}$$

for $k \ge 1$ and

$$\boldsymbol{\gamma}_{i,j} = g_j \left(\hat{\sigma}_{i,j}^{i+j-1} \right)^2 \boldsymbol{h}_j$$
$$\boldsymbol{\Delta}_{i,j} = \boldsymbol{h}_j \left(\hat{\sigma}_{i,j}^{i+j-1} \right)^2 \boldsymbol{h}'_j + \boldsymbol{T}_{i,j}$$

for
$$i = 0, ..., I$$
 and $j = 0, ..., J$.

Proof. See Heberle [14] for the Kalman-prediction formulas. The proof for Kalman-filtering and Kalman-smoothing is given in Appendix A.

Remarks 2.5:

- Theorem 2.4 is an extended version of the theorem given in Heberle [14]. Here the smoothed and filtered Kalman values are also given since in Section 3 they are used for the calculation of bootstrapping residuals.
- The initial values in Theorem 2.4 are denoted as $\widehat{C}_{i,0}^{I}$ and $(\widehat{\sigma}_{i,0}^{I})^{2}$ since these initial values are needed for smoothing, filtering and prediction.
- The error-variance-terms for the predicted values are defined as in Heberle [14] and are calculated with the given information \mathcal{D}_I .

• The error-variance-terms for the filtered values do not need additional information for calculation since these terms are based on the error-variance-terms of the predicted values.

3 Derivation of the predictive (bootstrap) distribution for the claims reserve of aggregated accident years

In this section we describe the main steps shown in Figure 2, which leads to the predictive distribution of the ultimate claims for aggregated accident years for the unobservable run-off triangle, given Model Assumptions 2.1. Therefore, we proceed with the following steps:

- 1. In the first step, we generate (standardized) residuals for bootstrapping purposes after computing the unobservable triangle.
- 2. In the second step, we bootstrap the residuals and generate new "observations", i.e. we generate new development triangle pairs (paid/incurred pairs).
- 3. In the third step, we apply the Kalman-filter recursions to each new triangle pair. This step considers the estimation error in our model.
- 4. The fourth step is to consider the process variance, which describes the pure random part of our claims reserving problem (since we deal with stochastic processes).
- 5. In the last step, we calculate a kernel density estimation of the aggregated claims reserves and compute some risk measures.

The first step to apply the bootstrap is to generate bootstrap residuals. In the bootstrapping framework it is essential that the observations are realizations of identically and independently distributed random variables. Obviously, this is not the case for run-off triangles. Therefore, in the following we do not bootstrap the claims data itself, but residuals, based on the Kalman-model. For each of the two given run-off triangles we define the $K = \frac{1}{2}J(2I - J - 1) + I + 1$ residuals by

$$\boldsymbol{r}_{i,j}^{\mathrm{Pa,In}} = \begin{pmatrix} r_{i,j}^{\mathrm{Pa}} \\ r_{i,j}^{\mathrm{In}} \end{pmatrix} = \begin{pmatrix} C_{i,j}^{\mathrm{Pa}} - h_{j}^{\mathrm{Pa}} \widehat{C}_{i,j}^{l} \\ C_{i,j}^{\mathrm{In}} - h_{j}^{\mathrm{In}} \widehat{C}_{i,j}^{l} \end{pmatrix} = \boldsymbol{C}_{i,j}^{\mathrm{Pa,In}} - \boldsymbol{h}_{j} \widehat{C}_{i,j}^{l}$$
(3.1)

for $i = 0, \ldots, I$ and $j = 0, \ldots, J$ with $i + j \leq I$.

These residuals have to be standardized to guarantee that all model residual have, at least, the same first two moments (cf. Stoffer & Wall [23]).



Figure 2: Obtaining the predictive distribution using bootstrapping.

Remarks 3.1:

- The estimated Kalman-value $\widehat{C}_{i,j}^{I}$ is either a Kalman-smoothed value or a Kalman-filtered value. The smoothed value is used if i + j < I holds true and, respectively, the filtered value is used if i + j = I holds true.
- The residuals $r_{i,j}^{\text{Pa,In}}$ are observable, whereas the random variables $w_{i,j}$ (cf. Model Assumptions 2.1 (1)) are not observable for unknown $C_{i,j}$.
- The definition of the residuals given in equation (3.1) is motivated by Model Assumptions 2.1 (1) and by the innovation given in equation (A.2).

The calculation of the standardized residuals is given in Stoffer & Wall [23]. In our framework, the standardized residuals $\bar{r}_{i,j}^{\text{Pa,In}}$ (i = 0, ..., I and j = 0, ..., J with $i + j \leq I$) are given by:

$$\bar{\boldsymbol{r}}_{ij}^{\text{Pa,In}} = \boldsymbol{\Delta}_{ij}^{-1/2} \boldsymbol{r}_{ij}^{\text{Pa,In}}$$
(3.2)

To verify the equality of the first two moments we show this in the "prediction"-case. The "smoothed" and "filtered"-case is almost similar. We use the notation given in Appendix A.1.

The first moment is given by:

$$\mathbf{E}\left[\bar{\boldsymbol{r}}_{i,j}^{\mathrm{Pa,In}}\right] = \mathbf{E}\left[\Delta_{i,j}^{-1/2}\boldsymbol{r}_{i,j}^{\mathrm{Pa,In}}\right] = \Delta_{i,j}^{-1/2}\mathbf{E}\left[\boldsymbol{L}_{i,j}\right] = \mathbf{0}$$

The second moment is given by:

$$\begin{aligned} \operatorname{Var}\left[\bar{\boldsymbol{r}}_{i,j}^{\operatorname{Pa,In}}\right] &= \operatorname{Var}\left[\Delta_{i,j}^{-1/2}\boldsymbol{r}_{i,j}^{\operatorname{Pa,In}}\right] = \Delta_{i,j}^{-1/2}\operatorname{Var}\left[\boldsymbol{L}_{i,j}\right]\left(\Delta_{i,j}^{-1/2}\right)' \\ &= \Delta_{i,j}^{-1/2}\operatorname{E}\left[\boldsymbol{L}_{i,j}\boldsymbol{L}_{i,j}'\right]\left(\Delta_{i,j}^{-1/2}\right)' = \Delta_{i,j}^{-1/2}\Delta_{i,j}\left(\Delta_{i,j}^{-1/2}\right)' = \mathbb{1}_{2\times 2}\end{aligned}$$

Thereby, $\mathbb{1}_{2\times 2}$ denotes the identity-matrix with dimension 2×2 .

The set

$$\bar{\boldsymbol{r}}^{\mathrm{Pa,In}} = \left\{ \bar{\boldsymbol{r}}^{\mathrm{Pa,In}}_{i,j} \mid i = 0, \dots, I \text{ and } j = 0, \dots, J \text{ with } i+j \leq I \right\}$$

of all standardized residuals constitute the bootstrap distribution $G^{\text{Pa,In}}$ for both run-off triangles.

3.1 Step 1: Bootstrapping the residuals for the derivation of the estimation error

Given the set $\bar{r}^{\text{Pa,In}}$ of the standardized residuals $\bar{r}^{\text{Pa}}_{i,j}$ and $\bar{r}^{\text{In}}_{i,j}$ (i = 0, ..., I and j = 0 ..., J with $i + j \leq I$) for both run-off triangles, we re-sample these residuals from $\bar{r}^{\text{Pa,In}}$ for both run-off triangles.

From equation (3.1) it is straightforward to generate pseudo/bootstrap observations $C_{i,j}^{\text{Pa}}$ and $C_{i,j}^{\text{In}}$ (i = 0, ..., I, j = 0, ..., J and $i + j \leq I$) for both run-off triangles. We will also denote the re-sampled (pseudo) observations by $C_{i,j}^{\text{Pa}}$ and $C_{i,j}^{\text{In}}$ to avoid notational overload.

Remarks 3.2:

- The described bootstrap approach means that we build two new run-off triangles of pseudo cumulative claims payments and pseudo incurred losses.
- With the described approach we are able to generate a large number of run-off triangle pairs containing both, cumulative claims payments and incurred losses data. On both run-off triangles we simultaneously adopt our model for cumulative claims payments and incurred losses data and obtain the upper left part of the unobservable run-off triangle.
3.2 Step 2: Sampling approach for including the process variance

Since we are not only interested in the estimation error, but also in the process variance, we proceed with Step 2 (cf. Figure 2). This means that we combine the first step where we regarded the estimation error with an additional step to include the process variance.

In Step 1 we used Model Assumptions 2.1 (1) to generate the bootstrap-residuals. Now, in Step 2 we use Model Assumptions 2.1 (2) to include the process variance in our bootstrapping model. Therefore, we use the Kalman-filtered values $\widehat{C}_{i,I-i}^{I}$ (i = 1, ..., I) on the diagonal to obtain one-step forecasts

$$\widehat{\widehat{C}}_{i,I-i+1} = g_{I-i}\widehat{C}_{i,I-i}^I + \widetilde{\upsilon}_{i,I-i}$$

for i = 1, ..., I as well as two- and more-step forecasts

$$\widetilde{\widehat{C}}_{i,j+1} = g_j \widetilde{\widehat{C}}_{i,j} + \widetilde{v}_{i,j}$$

for i = 2, ..., I and j = I - i + 1, ..., J - 1.

As in our Model Assumptions 2.1 we assume for the random variables $\tilde{v}_{i,j}$

$$\mathbb{E}\left[\tilde{\upsilon}_{i,j}\right] = 0 \qquad \text{and} \qquad \mathbb{E}\left[\tilde{\upsilon}_{i,j}\tilde{\upsilon}_{k,l}\right] = \begin{cases} q_{i,j} & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise} \end{cases}$$

In the following we use a Normal distribution for $\tilde{v}_{i,j}$, i.e. $\tilde{v}_{i,j} \sim \mathcal{N}(0, q_{i,j})$.

Step 1, as described above, together with Step 2 leads to as many filled development triangles as bootstrap iterations are done. For each filled bootstrap triangle the claims reserves

$$\widetilde{\widehat{R}}_i = \widetilde{\widehat{C}}_{i,J} - C^{\text{Pa}}_{i,I-i}$$

for i = 1, ..., I can be computed as well as the aggregated claims reserve

$$\widetilde{\widehat{R}} = \sum_{i=1}^{I} \widetilde{\widehat{R}}_i = \sum_{i=1}^{I} \left(\widetilde{\widehat{C}}_{i,J} - C^{\mathrm{Pa}}_{i,I-i} \right).$$

3.3 Achieving the (bootstrap) distribution for the aggregated claims reserve

The estimated bootstrap claims reserves $\widetilde{\widehat{R}}_i$ $(i \in \{1, \ldots, I\})$ as well as the aggregated bootstrap reserve $\widetilde{\widehat{R}}$, depends on the estimation of the ultimate claims $\widetilde{\widehat{C}}_{i,J}$ $(i = 1, \ldots, I)$ and these ultimate claims are mostly different for each generated bootstrap triangle.

In the following we focus on the bootstrap distribution of the whole claims reserve, i.e. we focus on

$$\widetilde{\widehat{R}} = \sum_{i=1}^{I} \widetilde{\widehat{R}}_i = \sum_{i=1}^{I} \left(\widetilde{\widehat{C}}_{i,J} - C_{i,I-i}^{\operatorname{Pa}} \right).$$

This reserve is, in most cases, different for each bootstrap triangle and is used to generate the bootstrap distribution. That means the more bootstrap triangles we generate, the "better" our (bootstrap) distribution gets. Since in most applications bootstrapping methods achieve acceptable results for less than 200 replications, the use of bootstrapping techniques in risk management requires a larger number of replications. The reason for that is that we are not interested in the mean value of the simulated bootstrap, but in its higher percentiles. A larger number of bootstrap replications is needed for the tails of the distribution in order to have the chance to get filled in.

The following steps must be performed to achieve the bootstrap distribution of the aggregated claims reserve:

- 1. Apply the Kalman model (see Heberle [14]) to both triangles, i.e. to the cumulative claims payments and incurred losses data. The result is an estimation of the upper left part of the unobservable (state-) triangle, which in our framework is the "original" triangle.
- 2. Compute the residuals from equation (3.1) and the scaled residuals from equation (3.2) for both run-off development-triangles.
- 3. Build (pseudo) observations $\widetilde{C}_{i,j}^{\text{Pa}}$ and $\widetilde{C}_{i,j}^{\text{In}}$ (i = 0, ..., I and j = 0, ..., J with $i + j \leq I$) by resampling the scaled residuals. This has to be done a large number of times (in our example we use 10,000 bootstrap iterations).
- 4. Compute for each "new" triangle pair the upper left part of the unobservable development triangle (cf. Heberle [14]) using the given model-parameters.
- 5. Include the process variance as described in Section 3.2.
- 6. Compute for each bootstrapped and filled claims development triangle the aggregated claims reserve \tilde{R} and generate an empirical bootstrap distribution (with kernel density estimation) from all computed claims reserves.

The described procedure completes the approach for deriving the bootstrap distribution of the aggregated claims reserve.

4 Example

For our example we use two run-off triangles from Dahms [3] (see Table 1 and Table 2).

	0	1	2	3	4	5	6	7	8	9
0	1,216,632	1,347,072	1,786,877	2,281,606	2,656,224	2,909,307	3,283,388	3,587,549	3,754,403	3,921,258
1	798,924	1,051,912	1,215,785	1,349,939	1,655,312	1,926,210	2,132,833	2,287,311	2,567,056	
2	1,115,636	1,387,387	1,930,867	2,177,002	2,513,171	2,931,930	3,047,368	3,182,511		
3	1,052,161	1,321,206	1,700,132	1,971,303	2,298,349	2,645,113	3,003,425			
4	808,864	1,029,523	1,229,626	1,590,338	1,842,662	2,150,351				
5	1,016,862	1,251,420	1,698,052	2,105,143	2,385,339					
6	948,312	1,108,791	1,315,524	1,487,577						
7	917,530	1,082,426	1,484,405							
8	1,001,238	1,376,124								
9	841,930									

Table 1: Observed cumulative claims payments $C_{i,j}^{\text{Pa}}$ (triangle A).

	0	1	2	3	4	5	6	7	8	9
0	3,362,115	5,217,243	4,754,900	4,381,677	4,136,883	4,094,140	4,018,736	3,971,591	3,941,391	3,921,258
1	2,640,443	4,643,860	3,869,954	3,248,558	3,102,002	3,019,980	2,976,064	2,946,941	2,919,955	
2	2,879,697	4,785,531	4,045,448	3,467,822	3,377,540	3,341,934	3,283,928	3,257,827		
3	2,933,345	5,299,146	4,451,963	3,700,809	3,553,391	3,469,505	3,413,921			
4	2,768,181	4,658,933	3,936,455	3,512,735	3,385,129	3,298,998				
5	3,228,439	5,271,304	4,484,946	3,798,384	3,702,427					
6	2,927,033	5,067,768	4,066,526	3,704,113						
7	3,083,429	4,790,944	4,408,097							
8	2,761,163	4,132,757								
9	3,045,376									

Table 2: Observed incurred losses data $C_{i,j}^{\text{In}}$ (triangle B).

For predicting the unobservable (state-) triangle and generating new bootstrap triangles we have to make assumptions about the initial values $\widehat{C}_{i,0}^{I}$ and the initial error-variance-terms $(\widehat{\sigma}_{i,0}^{I})^{2}$ (i = 0, ..., I). For the Kalman-parameters \mathbf{h}_{j} , $T_{i,j}$, g_{j} and $q_{i,j}$ (i = 0, ..., I and j = 0, ..., J - 1) we also have to make assumptions. Since we use the same two run-off triangles as in Heberle [14] and since we deal with the same model assumptions, we make the same assumptions about parameters and initial values as in Heberle [14]. Therefore, we set

$$\boldsymbol{h}_{j} = \begin{pmatrix} h_{j}^{\mathrm{Pa}} \\ h_{j}^{\mathrm{In}} \end{pmatrix} = \begin{pmatrix} \frac{j}{2J} + \frac{1}{2} \\ -\frac{j}{2J} + \frac{3}{2} \end{pmatrix}$$

to weight the incurred losses data at the beginning stronger than the paid data. For j = J both weights are equal. Additionally, the values g_j (j = 0, ..., J - 1) are set to:

$$g_j = \sqrt{\hat{f}_j^{\text{CL,Pa}} \hat{f}_j^{\text{CL,In}}}$$

The initial values $\widehat{C}_{i,0}^{I}$ as well as $(\widehat{\sigma}_{i,0}^{I})^{2}$ (i = 0, ..., I) are set to $(\alpha = 0.5)$:

$$\begin{split} \widehat{C}_{i,0}^{I} &= \alpha C_{i,0}^{\text{Pa}} + (1-\alpha) C_{i,0}^{\text{In}} \\ \widehat{\sigma}_{i,0}^{I} &= \alpha \widehat{\sigma}_{0}^{\text{CL,Pa}} + (1-\alpha) \widehat{\sigma}_{0}^{\text{CL,In}} \end{split}$$

Thereby, $\hat{f}_j^{\text{CL,Pa}}$ (respectively $\hat{f}_j^{\text{CL,In}}$) as well as $\hat{\sigma}_0^{\text{CL,Pa}}$ (respectively $\hat{\sigma}_0^{\text{CL,In}}$) are the well known Chain-Ladder estimators (cf. Mack [19]).

With these parameters and initial values we are able to estimate the unobservable developmenttriangle, i.e. we are able to estimate $\widehat{C}_{i,j}^{I}$ for i = 0, ..., I and j = 0, ..., J. The estimated values of the upper left triangle, which are needed to generate residuals (cf. Section 3), are given in Table 3.

	0	1	2	3	4	5	6	7	8	9
0	1,999,733	2,853,907	2,999,620	3,057,545	3,233,783	3,428,580	3,567,794	3,672,410	3,793,668	3,867,137
1	1,471,078	2,099,445	2,206,638	2,249,252	2,378,896	2,522,192	2,624,599	2,701,555	2,790,752	
2	1,759,518	2,511,095	2,639,304	2,690,275	2,845,337	3,016,727	3,139,214	3,231,259		
3	1,738,793	2,481,527	2,608,230	2,658,607	2,811,844	2,981,213	3,102,252			
4	1,524,015	2,175,012	2,286,065	2,330,220	2,464,528	2,612,974				
5	1,824,088	2,603,265	2,736,183	2,789,029	2,949,778					
6	1,670,545	2,384,139	2,505,869	2,554,265						
7	1,708,083	2,437,716	2,562,179							
8	1,643,611	2,345,701								
9	1,646,188									

Table 3: Predicted cumulative claims payments $\widehat{C}_{i,j}^{I}$ for i = 0, ..., I and j = 0, ..., J with $i + j \le I$.

With the estimated values $\widehat{C}_{i,j}^{I}$ given in Table 3 we are able to calculate the residuals (cf. equation (3.1)) as well as the standardized residuals given in equation (3.2). With these standardized residuals we created 10,000 new bootstrap triangle pairs and for each triangle pair we computed the upper left part of the unobservable (Kalman-) development-triangle (cf. Heberle [14]). Afterwards, for each estimated development-triangle the process variance was included (cf. Section 3.2) and the aggregated claims reserve was computed. A kernel density estimation of the aggregated claims reserves with GAUSS-kernel is plotted in Figure 3 and a summary of the aggregated claims reserves is given in Table 4.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
9,364,694	10,285,333	10,495,412	10,496,510	10,709,390	11,722,257

Table 4: Summary of the 10,000 bootstrapped aggregated claims reserves.

Different risk measures, i.e. the Value-at-Risk (*VaR*), the Expected-Shortfall (*ES*) and the Conditional-Tail-Expectation (*CTE*) for various values $p \in (0, 1)$, are given in Table 5.



Figure 3: Density plot of the bootstrapped aggregated claims reserves.

	_	p	
	0.90	0.95	0.99
VaR _p	10,897,227	11,006,989	11,222,495
ES_p	10,911,764	11,013,545	11,223,646
CTE_p	11,042,597	11,138,113	11,337,539

Table 5: Value-at-Risk, Expected-Shortfall and Conditional-Tail-Expectation for different choices $p \in \{0.90, 0.95, 0.99\}$.

Compared with the results given in Dahms et al. [4] the mean of our reserve (cf. Table 4) lies between the two Chain-Ladder estimates and also between the two Munich-Chain-Ladder estimates.

A Appendix: Proof of Theorem 2.4

In the following the proofs for the Kalman-smoothing (cf. equations (2.1), (2.4) and (2.5)) and for the Kalman-filtering (cf. equation (2.2) and (2.6)) are given. The proof for Kalman-prediction is given in Heberle [14].

A.1 Proof of the Kalman-smoothing formulas

Proof. The proof starts similar as the proof given in Heberle [14].

The following equation describes the best affine-linear predictor of $C_{i,j}$ in terms of all observations \mathcal{D}_I :

$$\operatorname{Pro}_{I}(C_{i,j}) = \operatorname{Pro}(C_{i,j} \mid A^{I}) = \operatorname*{arg\,min}_{A \in \mathcal{A}^{I}} \mathbb{E}\left[(C_{i,j} - A)^{2} \right]$$
(A.1)

with

$$\mathcal{A}^{I} = \left\{ A \in \mathbb{R} \; \middle| \; A = a + \sum_{l=0}^{I-i} a'_{l} C_{i,l}^{\mathrm{Pa,In}} \text{ with } a \in \mathbb{R}, a_{0}, \dots, a_{k} \in \mathbb{R}^{2} \text{ and } i \in \{0, \dots, I\} \right\}.$$

Overall, $Pro(A | \cdot)$ is the best affine-linear predictor of *A* in terms of " \cdot ". Predictor (A.1) is an orthogonal projection.

We start the proof with the innovation $L_{i,j}$ defined by $L_{i,0} = C_{i,0}^{\text{Pa,In}}$ and

$$L_{i,j} = C_{i,j}^{\text{Pa,In}} - \text{Pro}_{i+j-1}(C_{i,j}^{\text{Pa,In}}) = C_{i,j}^{\text{Pa,In}} - h_j \widehat{C}_{i,j}^{i+j-1} = h_j \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) + w_{i,j}$$
(A.2)

for i = 0, ..., I and j = 1, ..., J. Furthermore, we define

$$\boldsymbol{\gamma}_{i,j} = \mathbf{E}\left[C_{i,j+1}\boldsymbol{L}_{i,j}\right]$$

and

$$\Delta_{i,j} = \mathbb{E}\left[L_{i,j}L'_{i,j}\right] \tag{A.3}$$

for i = 0, ..., I and j = 0, ..., J.

Equation (A.6) as well as equation (A.7) holds true, since

$$E [C_{i,j}v_{i,j}] = 0,$$

$$E [C_{i,j}w_{i,j}] = 0$$
(A.4)

and

$$\mathbb{E}\left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1}\right)\widehat{C}_{i,j}^{i+j-1}\right] = 0 \tag{A.5}$$

holds true.

Using Model Assumptions 2.1 (2) and equation (A.2) $\pmb{\gamma}_{i,j}$ can be written as:

$$\boldsymbol{\gamma}_{i,j} = \mathbb{E} \left[(g_j C_{i,j} + v_{i,j}) \left(\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \boldsymbol{h}_j + \boldsymbol{w}_{i,j} \right) \right] \\ = g_j \mathbb{E} \left[C_{i,j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \right] \boldsymbol{h}_j + g_j \mathbb{E} \left[C_{i,j} \boldsymbol{w}_{i,j} \right] \\ + \mathbb{E} \left[v_{i,j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \right] \boldsymbol{h}_j + \mathbb{E} \left[v_{i,j} \boldsymbol{w}_{i,j} \right] \\ = g_j \left(\widehat{\sigma}_{i,j}^{i+j-1} \right)^2 \boldsymbol{h}_j$$
(A.6)

With equation (A.2) we get for $\Delta_{i,j}$:

$$\Delta_{i,j} = \mathbb{E} \left[\left(\boldsymbol{h}_{j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) + \boldsymbol{w}_{i,j} \right) \left(\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \boldsymbol{h}_{j}' + \boldsymbol{w}_{i,j}' \right) \right] \\ = \boldsymbol{h}_{j} \mathbb{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right)^{2} \right] \boldsymbol{h}_{j}' + \boldsymbol{h}_{j} \mathbb{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \boldsymbol{w}_{i,j}' \right] \right] \\ + \mathbb{E} \left[\boldsymbol{w}_{i,j} \left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \right] \boldsymbol{h}_{j}' + \mathbb{E} \left[\boldsymbol{w}_{i,j} \boldsymbol{w}_{i,j}' \right] \\ = \boldsymbol{h}_{j} \mathbb{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right)^{2} \right] \boldsymbol{h}_{j}' + \mathbb{E} \left[\boldsymbol{w}_{i,j} \boldsymbol{w}_{i,j}' \right] \\ = \boldsymbol{h}_{j} \left(\widehat{\sigma}_{i,j}^{i+j-1} \right)^{2} \boldsymbol{h}_{j}' + T_{i,j}$$
(A.7)

In the Kalman-filter framework, every $y \in \mathbb{R}$ of the form

$$y = a_0 C_{i,0} + \ldots + a_k C_{i,k}$$

can be expressed as

$$y = b_0 C_{i,0} + \ldots + b_{k-1} C_{i,k-1} + d'_k L_{i,k}$$

where $b_0, \ldots, b_{k-1} \in \mathbb{R}$ and $d_k \in \mathbb{R}^2$ only depends on $a_0, \ldots, a_k \in \mathbb{R}$. Together with $\mathbb{E}[C_{i,j}L_{i,k}] = \mathbf{0}$ for all j < k it follows that:

$$\operatorname{Pro}_{i+k}(\cdot) = \operatorname{Pro}_{i+k-1}(\cdot) + \operatorname{Pro}(\cdot \mid L_{i,k})$$
(A.8)

The expression $Pro(\cdot | L_{i,k})$ is meant as best affine-linear predictor only depending on $L_{i,k}$ (cf. (A.1)). Therefore, we get

$$Pro_{i+j+k}(C_{i,j}) = Pro_{i+j+k-1}(C_{i,j}) + b'L_{i,j+k},$$
(A.9)

for $k \ge 1$, and $(k \ge 1)$

$$\boldsymbol{b}' = \mathbb{E} \left[C_{i,j} \left(\boldsymbol{h}_{j+k} \left(C_{i,j+k} - \widehat{C}_{i,j+k}^{i+j+k-1} \right) + \boldsymbol{w}_{i,j+k} \right)' \right] \left(\mathbb{E} \left[\boldsymbol{L}_{i,j+k} \boldsymbol{L}'_{i,j+k} \right] \right)^{-1} \\ = \left(\check{\sigma}_{i,j}^{i+j+k} \right)^2 \boldsymbol{h}'_{j+k} \Delta_{i,j+k}^{-1},$$
(A.10)

where $\left(\check{\sigma}_{i,j}^{i+j+k}\right)^2$ is defined as $(k \ge 1)$:

$$\left(\check{\sigma}_{i,j}^{i+j+k}\right)^2 = \mathbb{E}\left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1}\right)\left(C_{i,j+k} - \widehat{C}_{i,j+k}^{i+j+k-1}\right)\right].$$

Equation (A.10) holds true, since: If X and Y are random vectors with finite second moments, then

$$Pro(X \mid Y) = MY$$

where the matrix *M* is given through $M = E[XY'] (E[YY'])^{-1}$ with $(E[YY'])^{-1}$ any generalized inverse of E[YY'].

Equation (A.9) together with (A.10) leads to the required Kalman-smoothing equation (2.1).

To verify equation (2.4) we use Model Assumptions 2.1 (1) and (2) and equation (2.3) together with the fact that $\mathbb{E}\left[v_{i,k}(C_{i,j} - \widehat{C}_{i,j}^{i+j-1})\right] = 0$ as well as $\mathbb{E}\left[w_{i,k}(C_{i,j} - \widehat{C}_{i,j}^{i+j-1})\right] = \mathbf{0}$ holds true. With the definition of $\left(\check{\sigma}_{i,j}^{i+j+k}\right)^2$ we get

$$\left(\check{\sigma}_{i,j}^{i+j+k}\right)^2 = \mathbf{E} \left[\left(C_{i,j} - \widehat{C}_{i,j}^{i+j-1} \right) \left(C_{i,j+k-1} - \widehat{C}_{i,j+k-1}^{i+j+k-2} \right) \right. \\ \left. \times \left(g_{j+k-1} - \boldsymbol{\gamma}_{i,j+k-1}' \Delta_{i,j+k-1} \boldsymbol{h}_{j+k-1} \right) \right] \\ = \left(\check{\sigma}_{i,j}^{i+j+k-1} \right)^2 \left(g_{j+k-1} - \boldsymbol{\gamma}_{i,j+k-1}' \Delta_{i,j+k-1} \boldsymbol{h}_{j+k-1} \right).$$

To establish equation (2.5) we start with (cf. equation (A.8))

$$C_{i,j} - \operatorname{Pro}_{i+j+k}(C_{i,j}) = C_{i,j} - \operatorname{Pro}_{i+j+k-1}(C_{i,j}) - b' L_{i,j+k}.$$
(A.11)

With equations (A.2) and (A.10) and the fact that $E\left[(C_{i,j} - \operatorname{Pro}_{i+j+k-1}(C_{i,j}))\boldsymbol{b}'\boldsymbol{L}_{i,j+k}\right] = 0$ holds true, equation (A.11) can be written as:

$$\begin{pmatrix} C_{i,j} - \operatorname{Pro}_{i+j+k}(C_{i,j}) + \boldsymbol{b}'\boldsymbol{L}_{i,j+k} \end{pmatrix}^{2} = \begin{pmatrix} C_{i,j} - \operatorname{Pro}_{i+j+k-1}(C_{i,j}) \end{pmatrix}^{2}$$

$$\Leftrightarrow \qquad \begin{pmatrix} C_{i,j} - \operatorname{Pro}_{i+j+k}(C_{i,j}) \end{pmatrix}^{2} + \boldsymbol{b}'\boldsymbol{L}_{i,j+k}\boldsymbol{L}'_{i,j+k}\boldsymbol{b} = \begin{pmatrix} C_{i,j} - \operatorname{Pro}_{i+j+k-1}(C_{i,j}) \end{pmatrix}^{2}$$

$$\Leftrightarrow \qquad \begin{pmatrix} C_{i,j} - \operatorname{Pro}_{i+j+k}(C_{i,j}) \end{pmatrix}^{2} = \begin{pmatrix} C_{i,j} - \operatorname{Pro}_{i+j+k-1}(C_{i,j}) \end{pmatrix}^{2}$$

$$- \boldsymbol{b}'\boldsymbol{L}_{i,j+k}\boldsymbol{L}'_{i,j+k}\boldsymbol{b} \qquad (A.12)$$

Taking the expectation on both sides of (A.12) leads to

$$\mathbb{E}\left[\left(C_{i,j} - \operatorname{Pro}_{i+j+k}(C_{i,j})\right)^{2}\right] = \mathbb{E}\left[\left(C_{i,j} - \operatorname{Pro}_{i+j+k-1}(C_{i,j})\right)^{2}\right] - \mathbb{E}\left[\boldsymbol{b}'\boldsymbol{L}_{i,j+k}\boldsymbol{L}'_{i,j+k}\boldsymbol{b}\right].$$
 (A.13)

Equation (A.13) together with the definition of $\left(\hat{\sigma}_{i,j}^{i+j+k}\right)^2$ leads to:

$$\Leftrightarrow \qquad \left(\hat{\sigma}_{i,j}^{i+j+k}\right)^{2} = \left(\hat{\sigma}_{i,j}^{i+j+k-1}\right)^{2} - \mathbf{b}' \operatorname{E} \left[\mathbf{L}_{i,j+k}\mathbf{L}'_{i,j+k}\right] \mathbf{b} \\ \Leftrightarrow \qquad \left(\hat{\sigma}_{i,j}^{i+j+k}\right)^{2} = \left(\hat{\sigma}_{i,j}^{i+j+k-1}\right)^{2} - \left(\check{\sigma}_{i,j}^{i+j+k}\right)^{2} \mathbf{h}'_{j+k} \Delta_{i,j+k}^{-1} \Delta_{i,j+k} \left(\Delta_{i,j+k}^{-1}\right)' \mathbf{h}_{j+k} \left(\check{\sigma}_{i,j}^{i+j+k}\right)^{2} \\ \Leftrightarrow \qquad \left(\hat{\sigma}_{i,j}^{i+j+k}\right)^{2} = \left(\hat{\sigma}_{i,j}^{i+j+k-1}\right)^{2} - \left(\check{\sigma}_{i,j}^{i+j+k}\right)^{4} \mathbf{h}'_{j+k} \Delta_{i,j+k}^{-1} \mathbf{h}_{j+k}$$

This completes the proof.

A.2 Proof of the Kalman-filtering formulas

Proof. In this proof we use the same definitions and the same notation as in the proof in Appendix A.1.

From relation (A.8) follows that (cf. equation (A.9)):

$$Pro_{I}(C_{i,I-i}) = Pro_{I-1}(C_{i,I-i}) + d'L_{i,I-i}$$
(A.14)

The vector $\mathbf{d} = (d_1, d_2)' \in \mathbb{R}^2$ is very similar to equation (A.10) and – using equations (A.2), (A.3), (A.4) and (A.5) – is given by:

$$d' = E [C_{i,I-i}L'_{i,I-i}] \left(E [L_{i,I-i}L'_{i,I-i}] \right)^{-1}$$

= $E \left[C_{i,I-i} \left(h_{I-i} \left(C_{i,I-i} - \widehat{C}_{i,I-i}^{I-1} \right) + w_{i,I-i} \right)' \right] \Delta_{i,I-i}^{-1}$
= $E \left[C_{i,I-i} \left(C_{i,I-i} - \widehat{C}_{i,I-i}^{I-1} \right) \right] h'_{I-i} \Delta_{i,I-i}^{-1}$
= $\left(\widehat{\sigma}_{i,I-i}^{I-1} \right)^2 h'_{I-i} \Delta_{i,I-i}^{-1}$ (A.15)

Equation (2.2) now follows by simply merging equation (A.14) and (A.15) and using equation (A.2), i.e.:

$$\widehat{C}_{i,I-i}^{I} = \widehat{C}_{i,I-i}^{I-1} + \left(\widehat{\sigma}_{i,I-i}^{I-1}\right)^{2} \boldsymbol{h}_{I-i}^{\prime} \Delta_{i,I-i}^{-1} \left(\boldsymbol{C}_{i,I-i}^{\text{Pa,In}} - \boldsymbol{h}_{I-i} \widehat{C}_{i,I-i}^{I-1}\right)$$

To verify equation (2.6) we use equation (A.14) and obtain:

$$C_{i,I-i} - \operatorname{Pro}_{I-1}(C_{i,I-i}) = C_{i,I-i} - \operatorname{Pro}_{I}(C_{i,I-i}) + \operatorname{Pro}_{I}(C_{i,I-i}) - \operatorname{Pro}_{I-1}(C_{i,I-i})$$

= $C_{i,I-i} - \operatorname{Pro}_{I}(C_{i,I-i}) + d' L_{i,I-i}$ (A.16)

Taking the expectation on both sides of (A.16) and using

$$\mathbb{E}\left[(C_{i,I-i} - \operatorname{Pro}_{I}(C_{i,I-i}))\boldsymbol{d}'\boldsymbol{L}_{i,I-i}\right] = 0$$

we derive:

$$(\hat{\sigma}_{i,I-i}^{I-1})^2 = \mathbb{E} \left[(C_{i,I-i} - \operatorname{Pro}_{I-1}(C_{i,I-i}))^2 \right] = \mathbb{E} \left[(C_{i,I-i} - \operatorname{Pro}_I(C_{i,I-i} + d' \mathbf{L}_{i,I-i})^2 \right] = \mathbb{E} \left[(C_{i,I-i} - \operatorname{Pro}_I(C_{i,I-i}))^2 \right] + \mathbb{E} \left[d' \mathbf{L}_{i,I-i} \mathbf{L}'_{i,I-i} d \right] = (\hat{\sigma}_{i,I-i}^I)^2 + (\hat{\sigma}_{i,I-i}^{I-1})^2 \mathbf{h}'_{I-i} \Delta_{i,I-i}^{-1} \mathbb{E} \left[\mathbf{L}_{i,I-i} \mathbf{L}'_{i,I-i} \right] \Delta_{i,I-i}^{-1} \mathbf{h}_{I-i} (\hat{\sigma}_{i,I-i}^I)^2 = (\hat{\sigma}_{i,I-i}^I)^2 + (\hat{\sigma}_{i,I-i}^{I-1})^4 \mathbf{h}'_{I-i} \Delta_{i,I-i}^{-1} \mathbf{h}_{I-i}$$

This completes the proof.

References

- [1] Brehm, P. J. (2002): Correlation and the aggregation of unpaid loss distributions. *CAS E-Forum*, 1–24.
- [2] Brockwell, P. J. & Davis, R. A. (2002): Introduction to Time Series and Forecasting. Springer.
- [3] Dahms, R. (2008): A loss reserving method for incomplete claim data. *Bulletin of Swiss* Association of Actuaries, 127–148.
- [4] Dahms, R., Merz, M. & Wüthrich, M. V. (2009): Claims development result for combined claims incurred and claims paid data. *Bulletin Francais d'Actuariat*, 9(18): 5–39.
- [5] Efron, B. (1979): Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7: 1–26.
- [6] Efron, B. (1994): *The Jackknife, the Bootstrap and Other Resampling Plans.* 6th ed. Regional conference series in applied mathematics. Philadelphia, Pa.: Society for Industrial and Applied Mathematics.
- [7] Efron, B. & Tibshirani, R. J. (1993): *An Introduction to the Bootstrap*. 1st ed. Monographs on statistics and applied probability. Champam & Hall.
- [8] England, P. D. & Verrall, R. J. (1999): Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, 25(3): 281–293.
- [9] England, P. D. & Verrall, R. J. (2007): Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1(2): 221–270.
- [10] Federal Office of Private Insurance (2006): Swiss Solvency Test. Switzerland.
- [11] Gogol, D. (1993): Using expected loss ratios in reserving. *Insurance: Mathematics and Economics*, 12(3): 297–299.
- [12] Halliwell, L. J. (1997): Conjoint Prediction of Paid and Incurred Losses. *CAS E-Forum*, 1: 241–380.

- [13] Halliwell, L. J. (2009): Modeling Paid and Incurred Losses Together. CAS E-Forum,
- [14] Heberle, J. (2012): Prediction of the ultimate claims based on cumulative payments and claims-incurred data using Kalman-filter theory.
- [15] Heberle, J., Huergo, L. & Merz, M. (2010): Bootstrapping the Chain-Ladder-Method of Several Correlated Run-Off Portfolios. *Zeitschrift für die gesamte Versicherungswissenschaft*, 98(5): 541–564.
- [16] Hertig, J. (1985): A Statistical Approach to IBNR-Reserves in Marine Reinsurance. ASTIN Bulletin, 15(2): 171–183.
- [17] Kirschner, G. S., Kerley, C. & Isaacs, B. (2008): Two approaches to calculating correlated reserve indications across multiple lines of business. *Variance*, 2: 15–38.
- [18] Liu, H. & Verrall, R. (2010): Bootstrap Estimation of the Predictive Distributions of Reserves Using Paid and Incurred Claims. *Variance*, 4: 121–135.
- [19] Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2): 213–225.
- [20] Merz, M. & Wüthrich, M. V. (2010): Paid-incurred chain claims reserving method. *Insurance: Mathematics and Economics*, 46(3): 568–579.
- [21] Quarg, G. & Mack, T. (2004): Munich Chain Ladder. *Blätter der DGVFM*, 26(4): 597–630.
- [22] Rodermund, M. (1990): *Foundations of Casualty Actuarial Science*. Casualty Actuarial Society (CAS).
- [23] Stoffer, D. S. & Wall, K. D. (1991): Bootstrapping State-Space Models: Gaussian Maximum Likelihood Estimation and the Kalman Filter. *Journal of the American Statistical Association*, 86(416): 1024–1033.
- [24] Taylor, G. C. (2000): *Loss Reserving: An Actuarial Perspective*. Kluwer Academic Publishers, Boston.
- [25] Teugels, J. L. & Sundt, B. (2004): *Encyclopedia of Actuarial Science*. Vol. 1. Wiley & Sons Ltd.
- [26] Wüthrich, M. V. & Merz, M. (2008): *Stochastic Claims Reserving Methods in Insurance*. Chichester: Wiley Finance.

Eidesstattliche Versicherung

Hiermit erkläre ich, Jochen Heberle, an Eides statt, dass ich die Dissertation mit dem Titel

Verteilungsfreie Verfahren in der Schadenreservierung

selbstständig und ohne fremde Hilfe verfasst habe.

Andere als die von mir angegebenen Quellen und Hilfsmittel habe ich nicht benutzt. Die den herangezogenen Werken wörtlich oder sinngemäß entnommenen Stellen sind als solche gekennzeichnet.

Hamburg, Februar 2014

Ort, Datum

Unterschrift