

**TECHNISCHE
UNIVERSITÄT
WIEN**

Vienna University of Technology

MASTERARBEIT

Methoden zur Schadenreservierung in der Sachversicherung

Ausgeführt am
Institut für Wirtschaftsmathematik
der Technischen Universität Wien

unter der Anleitung von
Ao.Univ.Prof. DI Dr. Friedrich Hubalek

durch

Barbara Lehner
Schmiedberg 1
4631 Krenglbach

Datum

Unterschrift

Danksagung

Ich danke Ao. Univ.Prof. DI Dr. Friedrich Hubalek für die Betreuung dieser Diplomarbeit. Besonderer Dank gilt auch Dr. Dominique Wagner und Dr. Johann Kronthaler von KPMG Austria AG für die intensive Unterstützung und Hilfeleistung. Des weiteren möchte ich meiner Familie und meinen Freunden für die Unterstützung während des Studiums danken.

Inhaltsverzeichnis

1	Einleitung	5
1.1	Motivation	5
1.2	Zusammenhang mit Solvency II	6
1.3	Aufbau der Arbeit	6
2	Grundlegende Definitionen	8
2.1	Definitionen	8
2.2	Numerisches Beispiel	9
3	Verteilungsfreies Chain-Ladder-Verfahren	14
3.1	Verfahren und Reserve	14
3.2	Mittlere quadratische Abweichung der Prognose	18
3.2.1	Verschärfung der Modellannahmen	18
3.2.2	Optimalität	22
3.2.3	Mittlere quadratische Abweichung der Prognose \hat{R}_i^{CL} für ein Abwicklungs- jahr i	25
3.2.4	Mittlere quadratische Abweichung der Prognose \hat{R}^{CL}	28
3.3	Numerisches Beispiel	30
4	Bornhuetter-Ferguson-Verfahren	35
4.1	Verfahren und Reserve	35
4.2	Vergleich Chain-Ladder- und Bornhuetter-Ferguson-Prognose	36
4.3	Numerisches Beispiel	38
5	Poisson-Modell	40
6	Verallgemeinerte lineare Modelle	45
6.1	Maximum-Likelihood-Schätzer	45
6.2	Modellrahmen	46
6.3	Exponential Dispersion Family	47
6.4	Parameter-Schätzung für die EDF	49
6.4.1	Fisher-Scoring-Algorithmus	51
6.4.2	Mittlere quadratische Abweichung der Prognose	55
6.5	Mittlere quadratische Abweichung der Bornhuetter-Ferguson-Prognose	58
6.5.1	Mittlere quadratische Abweichung der Prognose $\hat{R}_i^{\text{BF}'}$ für ein Abwick- lungsjahr i	59
6.5.2	Mittlere quadratische Abweichung der Prognose $\hat{R}^{\text{BF}'}$	64
6.5.3	Numerisches Beispiel	65

7 Benktander-Hovinen-Verfahren	67
7.1 Verfahren und Reserve	67
7.2 Numerisches Beispiel	69
8 Cape-Cod-Verfahren	71
8.1 Verfahren und Reserve	71
8.2 Numerisches Beispiel	72
9 Credibility-Schadenreservierungsmethoden	75
10 Additives Verfahren	78
10.1 Verfahren und Reserve	78
10.2 Mittlere quadratische Abweichung der Prognose	82
10.2.1 Mittlere quadratische Abweichung der Prognose \hat{R}_i^{AD} für ein Abwicklungs-	
jahr i	82
10.2.2 Mittlere quadratische Abweichung der Prognose \hat{R}^{AD}	83
10.3 Numerisches Beispiel	84
11 Munich-Chain-Ladder-Verfahren	87
11.1 Munich-Chain-Ladder-Modell	87
11.2 Credibility-Ansatz	90
11.3 Parameter-Schätzung	94
11.4 Numerisches Beispiel	99
12 Statistische Datenanalyse	103
12.1 Datenanalyse und -aufbereitung	103
12.2 Tail-Schätzung	103
12.3 Numerisches Beispiel	105
13 Implementierung in Excel	106
A Mittlere quadratische Abweichung einer Prognose	109
B Numerisches Beispiel - Vergleich der Methoden	112

Kapitel 1

Einleitung

1.1 Motivation

Dieses Kapitel lehnt sich an [1] und [2] an.

Am Ende eines Geschäftsjahres eines Versicherungsunternehmens sind im Normalfall nicht alle Schäden, die in diesem Jahr entstanden sind, abgewickelt. Diese Schäden bezeichnet man als **Spätschäden** und können aus zwei Gründen entstehen:

- Ein Schaden ist entstanden, aber noch nicht gemeldet. (Engl.: **IBNR = incurred but not reported**)
- Ein Schaden ist gemeldet, aber die Höhe des Schadens ist noch ungewiss. Es könnte sein, dass die für diesen Schaden gebildete Einzelfallreserve nicht ausreicht. (Engl.: **IBNER = incurred but not enough reserved**)

Die Einzelfallreserve wird für jeden gemeldeten Schaden einzeln vom zuständigen Schadensachbearbeiter aufgrund seiner Erfahrung und der verfügbaren Informationen über die Umstände des Einzelfalles festgesetzt.

Spätschäden gibt es grundsätzlich in allen Versicherungszweigen. In einigen Versicherungszweigen wie etwa der Haftpflichtversicherung kann es aber vorkommen, dass Schäden erst lange nach ihrer Verursachung bemerkt werden, wie z.B. ein Konstruktionsfehler bei einer Brücke, der erst nach deren Einsturz bemerkt wird. Ein weiteres Beispiel sind Personenschäden, da hier die Höhe des Schadens nicht sofort feststellbar ist, weil diese z.B. vom Verlauf der Heilung abhängt.

Das Versicherungsunternehmen muss für die im vergangenen Geschäftsjahr entstandenen Spätschäden eine Reserve bilden, um die Einhaltung seines Leistungsversprechens an den Versicherungsnehmer gewährleisten zu können. Diese Reserve ist im Grunde eine Prämie für die Selbstversicherung gegen Spätschäden. Wie jede andere Prämie auch sollte sie einerseits im Interesse der Versicherten ausreichend sein und andererseits im Interesse der Allgemeinheit, der die Steuern aus Gewinnen zugute kommt, auch nicht zu hoch bemessen sein. Neben der Bestimmung der Reserve für das vergangene Geschäftsjahr ist auch die Aktualisierung der Reserven für weiter zurückliegende Geschäftsjahre von Bedeutung.

Für die Bildung der Einzelfallreserve werden praktisch keine mathematischen Verfahren eingesetzt. Eine Ausnahme bilden Kleinschäden, bei denen vom Sachbearbeiter nur eine Intervallschätzung der Art „voraussichtlicher Schadenaufwand unter 1.000 Euro“ erstellt wird, so dass am Ende des Jahres eine Prognose für die Anzahl der Schäden unter 1.000 Euro vor-

liegt. Die durchschnittliche Schadenhöhe muss dann noch aufgrund entsprechender Statistiken früherer Jahre prognostiziert werden. Da der Schätzfehler hierbei aber nur eine unbedeutende Größenordnung aufweist, werden wir uns mit dieser Problematik nicht weiter befassen.

Dagegen ist die Schätzung der IBNR- und IBNER-Reserve in Versicherungszweigen mit langer Abwicklungsdauer von entscheidender Bedeutung. Diese Arbeit befasst sich mit verschiedenen Methoden zur Schadenreservierung in der Sachversicherung. Allen Verfahren gemeinsam ist, dass sie versuchen, die Erfahrungen früherer Anfallsjahre auf spätere Anfallsjahre zu übertragen.

1.2 Zusammenhang mit Solvency II

Dieses Kapitel orientiert sich an [3].

In den nächsten Jahren wird das zukünftige europäische Aufsichtsregime „Solvency II“ für Versicherungsunternehmen in Kraft treten (ein genauer Zeitpunkt ist aufgrund der zahlreichen Verschiebungen derzeit unklar), welches eine grundlegende Reform des Aufsichtsgesetzes darstellt.

Dabei soll unter anderem das bisher statische System zur Ermittlung des Eigenmittel-Erfordernisses durch einen risikobasierten Ansatz ersetzt werden. Weiters müssen unter Solvency II alle Bilanzpositionen marktkonsistent bewertet werden. Dies bedeutet für die versicherungstechnischen Rückstellungen insbesondere, dass eine Cash-Flow Betrachtung aller erforderlichen Zahlungsströme (Prämien, Kosten, Schäden etc.) notwendig ist, d.h. neben der Höhe der einzelnen Zahlungen muss auch der Zeitpunkt der Zahlungen prognostiziert werden.

Damit kommt den in dieser Arbeit beschriebenen Methoden unter Solvency II eine noch höhere Bedeutung als derzeit zu, denn diese stellen die aktuellsten Methoden zur Prognose der ausstehenden Schadenzahlungen pro Anfalls- und Abwicklungsjahr an den Versicherungsnehmer dar.

1.3 Aufbau der Arbeit

In Kapitel 2 werden das Schadendreieck, auf dem alle Methoden aufbauen, und damit verbundene Begriffe erklärt.

Kapitel 3 handelt vom wohl bekanntesten und weit verbreitetsten Verfahren, dem Chain-Ladder-Verfahren. Zuerst werden aus einem stochastischen Modell Prognosen für die Reserve abgeleitet. Weiters wird die mittlere quadratische Abweichung dieser Prognose hergeleitet und geschätzt.

In Kapitel 4 wird das Bornhuetter-Ferguson-Verfahren erläutert. Bei dieser Methode fließen zusätzlich externe Schätzer für den erwarteten Endschaden und das Abwicklungsmuster, die nicht aus dem Schadendreieck berechnet werden, ein. Am Ende des Kapitels werden das Chain-Ladder- und das Bornhuetter-Ferguson-Verfahren miteinander verglichen.

In Kapitel 5 wird ein anderes stochastisches Modell, das Poisson-Modell, zugrunde gelegt und gezeigt, dass dies ebenfalls auf die Chain-Ladder-Reserve führt.

Kapitel 6 handelt von verallgemeinerten linearen Modellen, die eher theoretischer Natur sind. Es werden sowohl die Reserve als auch die mittlere quadratische Abweichung der Prognose für dieses Modell ermittelt. Diese Ergebnisse werden verwendet, um in Kapitel 6.5 einen Schätzer für die mittlere quadratische Abweichung der Reserve-Prognose im Bornhuetter-Ferguson-Modell zu finden.

Nachdem für die Bornhuetter-Ferguson-Reserve die aktuellste Beobachtung der Schadenzahlungen bzw. -aufwendungen komplett ignoriert wird und beim Chain-Ladder-Verfahren davon ausgegangen wird, dass die aktuelle Beobachtung der Schadenzahlungen bzw. -aufwendungen zuverlässig ist, um zukünftige Aufwendungen zu prognostizieren und eine Prognose des Gesamtaufwandes des Anfallsjahres, die nicht auf Beobachtungen beruht, komplett ignoriert wird, versucht man beim Benktander-Hovinen-Verfahren in Kapitel 7 beide Werte miteinzubeziehen. Dabei ist die Reserve-Prognose eine Konvexkombination der Chain-Ladder- und Bornhuetter-Ferguson-Reserve.

Das Cape-Cod-Verfahren, welches in der Praxis häufig verwendet wird, in Kapitel 8 ist ein Spezialfall des Bornhuetter-Ferguson-Verfahrens, wobei der A-priori-Schätzer des erwarteten Gesamtaufwandes für ein Anfallsjahr als Produkt des Prämienvolumens des entsprechenden Anfallsjahr und eines Schätzers für die erwartete Schadenquote, der mithilfe des Abwicklungsdreiecks geschätzt wird, ermittelt wird.

In Kapitel 9 stellen wir uns die Frage nach der optimalen Konvexkombination aus Chain-Ladder- und Bornhuetter-Ferguson-Reserve, wobei sich die Optimalität hier auf die minimale mittlere quadratische Abweichung bezieht.

In Kapitel 10 wird mit dem Additiven Modell – im Gegensatz zu den bisherigen Kapiteln – ein Modell für die Inkrement der Schadenaufwendungen bzw. -zahlungen zugrunde gelegt. Neben der Reserve wird auch die mittlere quadratische Abweichung der Prognose für die Reserve hergeleitet.

Das Munich-Chain-Ladder-Verfahren in Kapitel 11 versucht die Information aus Schadenaufwendungen und -zahlungen in einem Modell zu verwenden. Die Prognosen des gesamten Schadenaufwandes sollen sich den Prognosen der gesamten Schadenzahlungen annähern.

In Kapitel 12 wird einerseits die für die Praxis sehr relevante Datenaufbereitung behandelt. Insbesondere werden Regressions-Methoden zur Glättung der Chain-Ladder-Faktoren beschrieben. Andererseits werden mögliche Tail-Schätzungen erläutert, die insbesondere dann relevant sind, wenn das Abwicklungsdreieck noch keine ausreichende Größe hat.

Im Anschluss an die Erklärung jeder einzelnen Methode, wird jeweils das Ergebnis dieser Methode basierend auf denselben Input-Daten eines numerischen Beispiels berechnet. Die Ergebnisse der einzelnen Methoden werden im Anhang B gegenübergestellt.

Kapitel 2

Grundlegende Definitionen

2.1 Definitionen

Dieses Kapitel orientiert sich an [4], [5], [2] und [6].

In dieser Arbeit werden alle Zufallsvariablen als quadratisch integrierbare Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) vorausgesetzt und sie sind deshalb Elemente des Hilbertraumes $L^2(\Omega, \mathcal{F}, P)$ mit dem Skalarprodukt $\langle U, V \rangle = \mathbb{E}[UV]$. Man betrachte eine Familie von Zufallsvariablen $S_{i,k}$ wobei $i \in \{0, 1, \dots, n\}, k \in \{0, 1, \dots, n\}$ mit $n \in \mathbb{N}$. Die Zufallsvariable $S_{i,k}$ ist die **kumulierte Schadenhöhe** der Schäden, die im Jahr i anfallen und im Jahr $i+k$ abgewickelt werden. Dabei bezeichnet i das **Anfallsjahr**, also jenes Jahr, in dem der Schaden entsteht und k das **Abwicklungsjahr**, also jenes Jahr, in dem der Schaden abgewickelt wird. Der Einfachheit halber werden für i keine Jahreszahlen, sondern $\{0, 1, \dots, n\}$ betrachtet. Wir nehmen vorerst an, dass alle Schäden bis zum n -ten Abwicklungsjahr vollständig beglichen sind. $S_{i,n}$ ist also der Gesamtschaden für das Anfallsjahr i . Es wird angenommen, dass $S_{i,k}$ für $i+k \leq n$ beobachtbar und für $i+k > n$ nicht beobachtbar ist, wobei $S_{i,k}$ sowohl die Zufallsvariable als auch die Realisation bezeichnet. Aus dem Zusammenhang sollte aber klar sein, worum es sich handelt.

Die beobachtbaren Schadenstände können in einem sogenannten **Abwicklungsdreieck** dargestellt werden:

		Abwicklungsjahr								
		0	1	...	k	...	$n-i$...	$n-1$	n
Anfallsjahr	0	$S_{0,0}$	$S_{0,1}$...	$S_{0,k}$...	$S_{0,n-i}$...	$S_{0,n-1}$	$S_{0,n}$
	1	$S_{1,0}$	$S_{1,1}$...	$S_{1,k}$...	$S_{1,n-i}$...	$S_{1,n-1}$	
	⋮	⋮	⋮		⋮		⋮			
	i	$S_{i,0}$	$S_{i,1}$...	$S_{i,k}$...	$S_{i,n-i}$			
	⋮	⋮	⋮		⋮					
	$n-k$	$S_{n-k,0}$	$S_{n-k,1}$...	$S_{n-k,k}$					
	⋮	⋮	⋮							
	$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
	n	$S_{n,0}$								

Es wird angenommen, dass die Beträge $S_{i,k}$ des Abwicklungsdreiecks um eine monetäre Inflation

bereinigt sind.

Das Abwicklungsdreieck kann aus zweierlei Daten gebildet werden: $S_{i,k}$ kann der bis einschließlich Abwicklungsjahr k bezahlte Schadenbetrag, d.h. ohne Einzelfallreserven, oder der bis einschließlich Abwicklungsjahr k angefallene Schadenbetrag, d.h. alle Zahlungen plus die zu dem Zeitpunkt bestehenden Einzelfallreserven, sein. $S_{i,k}^P$ bezeichne den bezahlten (paid) und $S_{i,k}^I$ den angefallenen (incurred) Schadenbetrag. Falls kein Index P oder I angeführt ist, gelten die Aussagen sowohl für die bezahlten als auch für die angefallenen Schadenbeträge.

Der Vorteil bei Verwendung eines Abwicklungsdreiecks auf Basis der bezahlten Schäden liegt darin, dass die Daten keine Schätzungen beinhalten und deshalb zuverlässiger erscheinen. Andererseits spricht für ein Abwicklungsdreieck auf Basis der angefallenen Schäden, dass die Einbeziehung der Einzelfallreserven die Größenordnung des Endschadens pro Anfallsjahr früher erkennen lässt.

Die **individuellen Abwicklungsfaktoren** für $i \in \{0, \dots, n\}$ und $k \in \{1, \dots, n\}$ sind folgendermaßen definiert:

$$F_{i,k} = \frac{S_{i,k}}{S_{i,k-1}}$$

Durch Schadenreservierungsmethoden soll $S_{i,n}$ prognostiziert bzw. der Erwartungswert $\mathbb{E}[S_{i,n}]$ und die Varianz $\text{Var}(S_{i,n})$ der beobachtbaren kumulierten Schadenhöhen geschätzt werden.

Weiters werden die Inkremente $Z_{i,k}$ für $i, k \in \{0, 1, \dots, n\}$ betrachtet:

$$Z_{i,k} := \begin{cases} S_{i,0} & \text{falls } k = 0 \\ S_{i,k} - S_{i,k-1} & \text{falls } k > 0 \end{cases}$$

Es gilt also

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

Die Familie $(Z_{i,k})_{i,k}$ enthält die selbe Information wie die Familie $(S_{i,k})_{i,k}$.

Falls der bezahlte Schadenbetrag $S_{i,k}^P$ verwendet wird, sind normalerweise alle zugehörigen $Z_{i,k}^P$ nichtnegativ. Anders ist dies jedoch, falls der angefallene Schaden $S_{i,k}^I$ verwendet wird, denn hier können sich sehr wohl negative Inkremente ergeben, z.B. falls die Einzelfallreserve reduziert wird.

2.2 Numerisches Beispiel

In den folgenden Kapitel werden die verschiedenen Verfahren für die Daten aus Tabelle 2.1 (Paid) bzw. 2.2 (Incurred) gerechnet. In den Tabellen 2.3 (Paid) und 2.4 sind die Inkremente dargestellt. Die Tabellen 2.5 (Paid) und 2.6 (Incurred) zeigen die individuellen Abwicklungsfaktoren. Dabei werden alle Methoden für das Paid-Abwicklungsdreieck gerechnet, das Chain-Ladder-Verfahren wird für beide Abwicklungsdreiecke gerechnet und für das Munich-Chain-Ladder-Verfahren werden (definitionsgemäß) beide Dreiecke verwendet. Im Anhang B findet man eine Übersicht der Ergebnisse.

Tabelle 2.2: Kumulierte Schadenstände (Incurred) $S_{i,k}^I$

Anfallsjahr i	Abwicklungsjahr k								
	0	1	2	3	4	5	6	7	8
0	192.318	191.407	184.968	184.879	183.544	184.809	183.741	182.736	182.709
1	234.989	222.817	213.515	210.199	208.995	207.497	204.364	203.868	
2	231.372	244.200	238.523	241.015	239.370	236.747	237.338		
3	265.731	274.320	268.655	258.780	262.889	260.887			
4	287.615	276.767	269.171	265.728	259.396				
5	277.600	272.543	273.254	266.976					
6	280.993	271.075	265.775						
7	276.388	265.032							
8	270.814								

Tabelle 2.3: Zuwächse der Schadenstände (Paid) $Z_{i,k}^P$

i/k	0	1	2	3	4	5	6	7	8
0	115.160	50.863	10.889	1.778	1.381	626	1.614	269	129
1	122.376	62.527	14.584	167	1.515	316	580	876	
2	133.328	71.197	14.682	1.830	3.611	1.884	2.762		
3	155.997	80.811	11.145	3.948	211	656			
4	169.121	61.516	17.814	469	1.337				
5	166.265	64.031	18.615	434					
6	161.851	65.616	15.344						
7	152.112	73.728							
8	150.683								

Tabelle 2.4: Zuwächse der Schadenstände (Incurred) $Z_{i,k}^I$

i/k	0	1	2	3	4	5	6	7	8
0	192.318	-911	-6.439	-89	-1.335	1.265	-1.068	-1.005	-27
1	234.989	-12.172	-9.302	-3.316	-1.204	-1.498	-3.133	-496	
2	231.372	12.828	-5.677	2.492	-1.645	-2.623	591		
3	265.731	8.589	-5.665	-9.875	4.109	-2.002			
4	287.615	-10.848	-7.596	-3.443	-6.332				
5	277.600	-5.057	711	-6.278					
6	280.993	-9.918	-5.300						
7	276.388	-11.356							
8	270.814								

Tabelle 2.5: Individuelle Abwicklungsfaktoren (Paid) $F_{i,k}^P$

i/k	1	2	3	4	5	6	7	8
0	1,4417	1,0656	1,0101	1,0077	1,0035	1,0089	1,0015	1,0007
1	1,5109	1,0789	1,0008	1,0076	1,0016	1,0029	1,0043	
2	1,5340	1,0718	1,0083	1,0163	1,0084	1,0122		
3	1,5180	1,0471	1,0159	1,0008	1,0026			
4	1,3637	1,0772	1,0019	1,0054				
5	1,3851	1,0808	1,0017					
6	1,4054	1,0675						
7	1,4847							

Tabelle 2.6: Individuelle Abwicklungsfaktoren (Incurred) $F_{i,k}^I$

i/k	1	2	3	4	5	6	7	8
0	0,9953	0,9664	0,9995	0,9928	1,0069	0,9942	0,9945	0,9999
1	0,9482	0,9583	0,9845	0,9943	0,9928	0,9849	0,9976	
2	1,0554	0,9768	1,0104	0,9932	0,9890	1,0025		
3	1,0323	0,9793	0,9632	1,0159	0,9924			
4	0,9623	0,9726	0,9872	0,9762				
5	0,9818	1,0026	0,9770					
6	0,9647	0,9804						
7	0,9589							

Kapitel 3

Verteilungsfreies Chain-Ladder-Verfahren

3.1 Verfahren und Reserve

Dieses Kapitel lehnt sich an [6] an.

Das Chain-Ladder-Verfahren ist aufgrund seiner Einfachheit eine der beliebtesten Schadenreservierungsmethoden. Dabei werden keine Annahmen an die Verteilung von $S_{i,k}$ getroffen.

Modellannahmen 3.1:

- Die kumulativen Schadenhöhen $S_{i,k}$ für verschiedene Anfallsjahre i sind unabhängig.
- Es gibt deterministische Abwicklungsfaktoren $f_1, \dots, f_n > 0$, sodass für alle $0 \leq i \leq n$ und für alle $1 \leq k \leq n$ gilt:

$$\mathbb{E}[S_{i,k} | S_{i,0}, \dots, S_{i,k-1}] = \mathbb{E}[S_{i,k} | S_{i,k-1}] = f_k S_{i,k-1} \quad (3.1)$$

Die Daten sollen homogen sein, d.h. es soll kein Trend in den individuellen Abwicklungsfaktoren $F_{i,k}$ sein. Sonst müssen die Daten transformiert werden oder das Portfolio wird in Subportfolios unterteilt, sodass die Subportfolios die Homogenitätsbedingung erfüllen.

Man beachte, dass mit (3.1) nur Annahmen an die ersten Momente gemacht werden. Das ist bereits ausreichend, um die zukünftigen Schadenhöhen zu prognostizieren. Um die Unsicherheit der Prognose zu messen, sind weitere Annahmen an höhere Momente notwendig. Dies erfolgt im Kapitel 3.2.

Aus (3.1) folgt für $1 \leq k \leq n$

$$\mathbb{E}[S_{i,k}] = \mathbb{E}[\mathbb{E}[S_{i,k} | S_{i,0}, \dots, S_{i,k-1}]] = f_k \mathbb{E}[S_{i,k-1}] \quad (3.2)$$

und deshalb

$$f_k = \frac{\mathbb{E}[S_{i,k}]}{\mathbb{E}[S_{i,k-1}]} \quad (3.3)$$

Es sollen nun Schätzer für die Chain-Ladder-Faktoren f_k , $k \in \{1, \dots, n\}$, gefunden werden. $\mathcal{D}_n = \sigma\{S_{i,k}; 0 \leq i \leq n, 0 \leq k \leq n, i+k \leq n\}$ bezeichne die σ -Algebra, die durch die Beobachtungen zur Zeit n erzeugt wird.

Lemma 3.2: Unter den Modellannahmen 3.1 gilt für $k > n - i$

$$\mathbb{E}[S_{i,k} | \mathcal{D}_n] = \mathbb{E}[S_{i,k} | S_{i,n-i}] = S_{i,n-i} f_{n-i+1} \cdots f_k \quad (3.4)$$

Beweis: Aufgrund der Annahme, dass die kumulativen Schadenhöhen für verschiedene Anfallsjahre unabhängig sind, der Turmeigenschaft für die σ -Algebren

$$\sigma(S_{i,0}, \dots, S_{i,n-i}) \subset \sigma(S_{i,0}, \dots, S_{i,k-1})$$

und Annahme (3.1) gilt:

$$\begin{aligned} \mathbb{E}[S_{i,k} | \mathcal{D}_n] &= \mathbb{E}[S_{i,k} | S_{i,0}, \dots, S_{i,n-i}] = \mathbb{E}[\mathbb{E}[S_{i,k} | S_{i,0}, \dots, S_{i,k-1}] | S_{i,0}, \dots, S_{i,n-i}] \\ &= \mathbb{E}[f_k S_{i,k-1} | S_{i,0}, \dots, S_{i,n-i}] = f_k \mathbb{E}[S_{i,k-1} | \mathcal{D}_n] \end{aligned}$$

Durch wiederholte Anwendung kommt man zum Ergebnis. Analog für $\mathbb{E}[S_{i,k} | S_{i,n-i}]$. \square

Für bekannte Chain-Ladder-Faktoren f_k wird die Rückstellung für noch nicht abgewickelte Schäden des Anfallsjahres i basierend auf \mathcal{D}_n also folgendermaßen prognostiziert:

$$\mathbb{E}[S_{i,n} | \mathcal{D}_n] - S_{i,n-i} = S_{i,n-i} (f_{n-i+1} \cdots f_n - 1)$$

Die Chain-Ladder-Faktoren sind jedoch normalerweise nicht bekannt und müssen geschätzt werden. Aufgrund von (3.3) bieten sich folgende Schätzer an:

$$\hat{f}_k = \frac{\sum_{i=0}^{n-k} S_{i,k}}{\sum_{i=0}^{n-k} S_{i,k-1}} = \sum_{i=0}^{n-k} \frac{S_{i,k-1}}{\sum_{j=0}^{n-k} S_{j,k-1}} \frac{S_{i,k}}{S_{i,k-1}} = \sum_{i=0}^{n-k} \frac{S_{i,k-1}}{\sum_{j=0}^{n-k} S_{j,k-1}} F_{i,k}, \quad k \in \{1, \dots, n\}. \quad (3.5)$$

Die Chain-Ladder-Faktoren f_k werden also durch gewichtete Mittel der individuellen Abwicklungsfaktoren $F_{i,k}$ geschätzt.

Der Chain-Ladder-Schätzer für $\mathbb{E}[S_{i,k} | \mathcal{D}_n]$ für $i + k > n$ ist also

$$\hat{S}_{i,k}^{\text{CL}} = \hat{\mathbb{E}}[S_{i,k} | \mathcal{D}_n] = S_{i,n-i} \hat{f}_{n-i+1} \cdots \hat{f}_k. \quad (3.6)$$

Für die Reserve für das Anfallsjahr i gilt

$$R_i = S_{i,n} - S_{i,n-i}. \quad (3.7)$$

Daraus ergibt sich folgende prognostizierte Reserve \hat{R}_i^{CL} für das Anfallsjahr i :

$$\hat{R}_i^{\text{CL}} = S_{i,n-i} \left(\hat{f}_{n-i+1} \cdots \hat{f}_n - 1 \right).$$

Eine Prognose für die Gesamtreserve

$$R = \sum_{i=1}^n R_i$$

ist die Chain-Ladder-Gesamtreserve

$$\hat{R}^{\text{CL}} = \sum_{i=1}^n \hat{R}_i^{\text{CL}}.$$

Sei $\mathcal{B}_k = \sigma\{S_{i,j} : 0 \leq i + j \leq n, 0 \leq j \leq k\} \subseteq \mathcal{D}_n$.

Lemma 3.3: Unter den Modellannahmen 3.1 gilt für $k, i, j \in \{1, \dots, n\}$:

- a) Gegeben \mathcal{B}_{k-1} ist \widehat{f}_k ein unverzerrter Schätzer für f_k , d.h. $\mathbb{E} \left[\widehat{f}_k \mid \mathcal{B}_{k-1} \right] = f_k$.
- b) \widehat{f}_k ist ein unverzerrter Schätzer für f_k , d.h. $\mathbb{E} \left[\widehat{f}_k \right] = f_k$.
- c) $\widehat{f}_1, \dots, \widehat{f}_n$ sind paarweise unkorreliert, d.h. $\mathbb{E} \left[\widehat{f}_i \widehat{f}_j \right] = \mathbb{E} \left[\widehat{f}_i \right] \mathbb{E} \left[\widehat{f}_j \right]$ für $i \neq j$. Außerdem gilt $\mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_k \right] = \mathbb{E} \left[\widehat{f}_1 \right] \cdots \mathbb{E} \left[\widehat{f}_k \right]$.
- d) Gegeben $S_{i,n-i}$ ist $\widehat{S}_{i,n}^{\text{CL}}$ ein unverzerrter Schätzer für $\mathbb{E} [S_{i,n} \mid \mathcal{D}_n] = \mathbb{E} [S_{i,n} \mid S_{i,n-i}]$, d.h. $\mathbb{E} \left[\widehat{S}_{i,n}^{\text{CL}} \mid S_{i,n-i} \right] = \mathbb{E} [S_{i,n} \mid \mathcal{D}_n]$.
- e) $\widehat{S}_{i,n}^{\text{CL}}$ ist ein unverzerrter Schätzer für $\mathbb{E} [S_{i,n}]$, d.h. $\mathbb{E} \left[\widehat{S}_{i,n}^{\text{CL}} \right] = \mathbb{E} [S_{i,n}]$.

Beweis:

- a) Es gilt:

$$\mathbb{E} \left[\widehat{f}_k \mid \mathcal{B}_{k-1} \right] = \mathbb{E} \left[\frac{\sum_{i=0}^{n-k} S_{i,k}}{\sum_{i=0}^{n-k} S_{i,k-1}} \mid \mathcal{B}_{k-1} \right].$$

Nachdem $\sum_{i=0}^{n-k} S_{i,k-1}$ messbar bezüglich \mathcal{B}_{k-1} ist, gilt

$$\mathbb{E} \left[\frac{\sum_{i=0}^{n-k} S_{i,k}}{\sum_{i=0}^{n-k} S_{i,k-1}} \mid \mathcal{B}_{k-1} \right] = \frac{\sum_{i=0}^{n-k} \mathbb{E} [S_{i,k} \mid \mathcal{B}_{k-1}]}{\sum_{i=0}^{n-k} S_{i,k-1}}.$$

Aufgrund der Annahme der Unabhängigkeit der Anfallsjahre und Annahme (3.1) gilt

$$\frac{\sum_{i=0}^{n-k} \mathbb{E} [S_{i,k} \mid \mathcal{B}_{k-1}]}{\sum_{i=0}^{n-k} S_{i,k-1}} = \frac{\sum_{i=0}^{n-k} \mathbb{E} [S_{i,k} \mid S_{i,0}, \dots, S_{i,k-1}]}{\sum_{i=0}^{n-k} S_{i,k-1}} = \frac{\sum_{i=0}^{n-k} f_k S_{i,k-1}}{\sum_{i=0}^{n-k} S_{i,k-1}} = f_k.$$

Damit ist die bedingte Unverzerrtheit gezeigt.

- b) Aus der Definition der bedingten Erwartung und a) folgt

$$\mathbb{E} \left[\widehat{f}_k \right] = \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k \mid \mathcal{B}_k \right] \right] = \mathbb{E} [f_k].$$

- c) O.B.d.A. sei $i < j$. Dann gilt

$$\mathbb{E} \left[\widehat{f}_i \widehat{f}_j \right] = \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_i \widehat{f}_j \mid \mathcal{B}_{j-1} \right] \right].$$

Da \widehat{f}_j messbar bezüglich \mathcal{B}_{j-1} ist, gilt

$$\mathbb{E} \left[\mathbb{E} \left[\widehat{f}_i \widehat{f}_j \mid \mathcal{B}_{j-1} \right] \right] = \mathbb{E} \left[\widehat{f}_i \mathbb{E} \left[\widehat{f}_j \mid \mathcal{B}_{j-1} \right] \right].$$

Aus a) und b) folgt

$$\mathbb{E} \left[\widehat{f}_i \mathbb{E} \left[\widehat{f}_j \mid \mathcal{B}_{j-1} \right] \right] = \mathbb{E} \left[\widehat{f}_i \right] f_j = \mathbb{E} \left[\widehat{f}_i \right] \mathbb{E} \left[\widehat{f}_j \right].$$

Die zweite Aussage folgt analog, denn es gilt

$$\mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_k \right] = \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_k \mid \mathcal{B}_{k-1} \right] \right].$$

Da $\widehat{f}_1, \dots, \widehat{f}_{k-1}$ messbar bezüglich \mathcal{B}_{k-1} sind, gilt

$$\mathbb{E} \left[\mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_k \mid \mathcal{B}_{k-1} \right] \right] = \mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_{k-1} \mathbb{E} \left[\widehat{f}_k \mid \mathcal{B}_{k-1} \right] \right].$$

Aus a) und b) folgt

$$\mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_{k-1} \mathbb{E} \left[\widehat{f}_k \mid \mathcal{B}_{k-1} \right] \right] = \mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_{k-1} \right] f_k = \mathbb{E} \left[\widehat{f}_1 \cdots \widehat{f}_{k-1} \right] \mathbb{E} \left[\widehat{f}_k \right].$$

Durch Wiederholung der Schritte kommt man zum Ergebnis.

d) Es gilt

$$\mathbb{E} \left[\widehat{S}_{i,n}^{\text{CL}} \mid S_{i,n-i} \right] = \mathbb{E} \left[S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_n \mid S_{i,n-i} \right].$$

Aufgrund der Turmeigenschaft gilt

$$\mathbb{E} \left[S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_n \mid S_{i,n-i} \right] = \mathbb{E} \left[\mathbb{E} \left[S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_n \mid \mathcal{B}_{n-1} \right] \mid S_{i,n-i} \right].$$

Nachdem $\widehat{f}_{n-i}, \dots, \widehat{f}_{n-1}$ und $S_{i,n-i}$ messbar bezüglich \mathcal{B}_{n-1} sind, gilt

$$\begin{aligned} & \mathbb{E} \left[\mathbb{E} \left[S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_n \mid \mathcal{B}_{n-1} \right] \mid S_{i,n-i} \right] = \\ & = \mathbb{E} \left[S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_{n-1} \mathbb{E} \left[\widehat{f}_n \mid \mathcal{B}_{n-1} \right] \mid S_{i,n-i} \right]. \end{aligned}$$

Aus a) folgt

$$\mathbb{E} \left[S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_{n-1} \mathbb{E} \left[\widehat{f}_n \mid \mathcal{B}_{n-1} \right] \mid S_{i,n-i} \right] = f_n \mathbb{E} \left[S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_{n-1} \mid S_{i,n-i} \right].$$

Iterativ kommt man auf

$$\mathbb{E} \left[\widehat{S}_{i,n}^{\text{CL}} \mid S_{i,n-i} \right] = S_{i,n-i} f_{n-i+1} \cdots f_n.$$

Aufgrund von Lemma (3.2) gilt

$$\mathbb{E} \left[\widehat{S}_{i,n}^{\text{CL}} \mid S_{i,n-i} \right] = \mathbb{E} [S_{i,n} \mid \mathcal{D}_n].$$

e) Aus der Definition der bedingten Erwartung und d) folgt

$$\mathbb{E} \left[\widehat{S}_{i,n}^{\text{CL}} \right] = \mathbb{E} \left[\mathbb{E} \left[\widehat{S}_{i,n}^{\text{CL}} \mid S_{i,n-i} \right] \right] = \mathbb{E} [\mathbb{E} [S_{i,n} \mid \mathcal{D}_n]] = \mathbb{E} [S_{i,n}].$$

□

Bemerkungen:

- Auf den ersten Blick erscheint es überraschend, dass die Chain-Ladder-Faktoren \widehat{f}_k unkorreliert sind, nachdem zwei Faktoren \widehat{f}_k und \widehat{f}_{k+1} teilweise von den selben Daten abhängen, nämlich einmal im Nenner und einmal im Zähler. Man beachte aber, dass die Abwicklungsfaktoren \widehat{f}_k zwar unkorreliert, aber keineswegs unabhängig sind. In Lemma 3.5 wird gezeigt, dass die quadrierten Schätzer \widehat{f}_k und \widehat{f}_{k+1} negativ korreliert sind.

- Für die Schätzung von Abwicklungsfaktoren f_k mit großem k sind nur wenige Daten vorhanden. Dadurch wird der Schätzer \hat{f}_k in der Praxis nicht sehr zuverlässig sein.
- Die ersten Beobachtungen eines Anfallsjahres sind manchmal nicht sehr repräsentativ für die Entwicklung des Schadenstandes, was Probleme für jüngere Jahre, für die erst wenige Beobachtungen vorhanden sind, verursacht. Im Allgemeinen kann es passieren, dass die letzte Beobachtung ein Ausreißer ist, der sich durch das Chain-Ladder-Verfahren direkt auf den prognostizierten Endschaden $S_{i,n}$ auswirkt. Deshalb wird für jüngere Jahre oft die Bornhuetter-Ferguson-Methode bevorzugt. (siehe Kapitel 4)

3.2 Mittlere quadratische Abweichung der Prognose

Dieses Kapitel lehnt sich an [6] und [7] an.

3.2.1 Verschärfung der Modellannahmen

Da nun die bedingte mittlere quadratische Abweichung geschätzt werden soll, müssen Annahmen an die zweiten Momente gemacht werden. Wir verschärfen also die Modellannahmen 3.1:

Modellannahmen 3.4:

- Die kumulativen Schadenhöhen $S_{i,k}$ für verschiedene Anfallsjahre i sind unabhängig.
- $(S_{i,k})_{k \geq 0}$ bilden eine Markov-Kette, d.h. für $x, x_1, \dots, x_k \in \mathbb{R}$, für die folgender bedingte Erwartungswert sinnvoll definiert ist, gilt

$$\mathbb{P}[S_{i,k} = x | S_{i,0} = x_0, \dots, S_{i,k-1} = x_{k-1}] = \mathbb{P}[S_{i,k} = x | S_{i,k-1} = x_{k-1}].$$

Es gibt Abwicklungsfaktoren $f_1, \dots, f_n > 0$ und Varianz-Parameter $\sigma_1^2, \dots, \sigma_n^2 > 0$, sodass für alle $0 \leq i \leq n$ und $1 \leq k \leq n$ gilt

$$\begin{aligned} \mathbb{E}[S_{i,k} | S_{i,k-1}] &= f_k S_{i,k-1} \\ \text{Var}(S_{i,k} | S_{i,k-1}) &= \sigma_k^2 S_{i,k-1} \end{aligned}$$

Man beachte, dass keine Annahmen an die Verteilung von $S_{i,k}$ gemacht wird, sondern nur an die ersten zwei Momente (und dass die kumulativen Schäden $S_{i,k-1}$ positiv sind, sonst macht die Annahme an die Varianz keinen Sinn).

Die Resultate aus Kapitel 3 bleiben natürlich gültig, nachdem die Modellannahmen 3.4 ein Spezialfall der Modellannahmen 3.1 sind. Es gilt

$$\hat{f}_k = \frac{\sum_{i=0}^{n-k} S_{i,k}}{\sum_{i=0}^{n-k} S_{i,k-1}} = \sum_{i=0}^{n-k} \frac{S_{i,k-1}}{\sum_{j=0}^{n-k} S_{j,k-1}} \frac{S_{i,k}}{S_{i,k-1}} = \sum_{i=0}^{n-k} \frac{S_{i,k-1}}{\sum_{j=0}^{n-k} S_{j,k-1}} F_{i,k}, \quad k \in \{1, \dots, n\}.$$

Lemma 3.5: Unter den Modellannahmen 3.4 sind für $k \in \{1, \dots, n-1\}$ die Quadrate zweier aufeinanderfolgender Chain-Ladder-Faktoren \hat{f}_k und \hat{f}_{k+1} bedingt auf \mathcal{B}_{k-1} negativ korreliert, d.h.

$$\text{Cov} \left(\hat{f}_k^2, \hat{f}_{k+1}^2 \mid \mathcal{B}_{k-1} \right) < 0.$$

Beweis: Unter Verwendung der Turmeigenschaft für $\mathcal{B}_{k-1} \subset \mathcal{B}_k$ gilt allgemein

$$\begin{aligned}
& \text{Cov} \left(\widehat{f}_k^2, \widehat{f}_{k+1}^2 \mid \mathcal{B}_{k-1} \right) = \mathbb{E} \left[\widehat{f}_k^2 \widehat{f}_{k+1}^2 \mid \mathcal{B}_{k-1} \right] - \mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_{k-1} \right] \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_{k-1} \right] = \\
& = \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k^2 \widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] - \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] = \\
& = \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k^2 \widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] - \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right] \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] + \\
& \quad + \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right] \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] - \\
& \quad - \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] = \\
& = \mathbb{E} \left[\mathbb{E} \left[\widehat{f}_k^2 \widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] - \mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right] \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right] + \\
& \quad + \text{Cov} \left(\mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right], \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right) = \\
& = \mathbb{E} \left[\text{Cov} \left(\widehat{f}_k^2, \widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right) \mid \mathcal{B}_{k-1} \right] + \text{Cov} \left(\mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right], \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right).
\end{aligned}$$

Weil \widehat{f}_k^2 messbar bezüglich \mathcal{B}_k ist, gilt

$$\begin{aligned}
\text{Cov} \left(\widehat{f}_k^2, \widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right) &= \mathbb{E} \left[\left(\widehat{f}_k^2 - \mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right] \right) \left(\widehat{f}_{k+1}^2 - \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \right) \mid \mathcal{B}_k \right] = \\
&= \mathbb{E} \left[\left(\widehat{f}_k^2 - \widehat{f}_k^2 \right) \left(\widehat{f}_{k+1}^2 - \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \right) \mid \mathcal{B}_k \right] = 0.
\end{aligned}$$

Außerdem folgt aus den Modellannahmen 3.4

$$\begin{aligned}
\mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] &= \mathbb{E} \left[\left(\frac{\sum_{i=0}^{n-k-1} S_{i,k+1}}{\sum_{i=0}^{n-k-1} S_{i,k}} \right)^2 \mid \mathcal{B}_k \right] = \\
&= \frac{1}{\left(\sum_{i=0}^{n-k-1} S_{i,k} \right)^2} \mathbb{E} \left[\left(\sum_{i=0}^{n-k-1} S_{i,k+1} \right)^2 \mid \mathcal{B}_k \right] = \\
&= \frac{1}{\left(\sum_{i=0}^{n-k-1} S_{i,k} \right)^2} \left(\sum_{i=0}^{n-k-1} \sigma_{k+1}^2 S_{i,k} + \left(\sum_{i=0}^{n-k-1} f_{k+1} S_{i,k} \right)^2 \right) = \\
&= \frac{\sigma_{k+1}^2}{\sum_{i=0}^{n-k-1} S_{i,k}} + f_{k+1}^2.
\end{aligned}$$

Es folgt also

$$\begin{aligned}
\text{Cov} \left(\widehat{f}_k^2, \widehat{f}_{k+1}^2 \mid \mathcal{B}_{k-1} \right) &= \text{Cov} \left(\mathbb{E} \left[\widehat{f}_k^2 \mid \mathcal{B}_k \right], \mathbb{E} \left[\widehat{f}_{k+1}^2 \mid \mathcal{B}_k \right] \mid \mathcal{B}_{k-1} \right) = \\
&= \text{Cov} \left(\widehat{f}_k^2, \frac{\sigma_{k+1}^2}{\sum_{i=0}^{n-k-1} S_{i,k}} + f_{k+1}^2 \mid \mathcal{B}_{k-1} \right) = \\
&= \sigma_{k+1}^2 \text{Cov} \left(\widehat{f}_k^2, \frac{1}{\sum_{i=0}^{n-k-1} S_{i,k}} \mid \mathcal{B}_{k-1} \right) = \\
&= \frac{\sigma_{k+1}^2}{\left(\sum_{i=0}^{n-k} S_{i,k-1} \right)^2} \text{Cov} \left(\left(\sum_{i=0}^{n-k} S_{i,k} \right)^2, \frac{1}{\sum_{i=0}^{n-k-1} S_{i,k}} \mid \mathcal{B}_{k-1} \right)
\end{aligned}$$

Es gilt

$$\left(\sum_{i=0}^{n-k} S_{i,k}\right)^2 = \left(\sum_{i=0}^{n-k-1} S_{i,k} + S_{n-k,k}\right)^2 = \left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^2 + 2 S_{n-k,k} \sum_{i=0}^{n-k-1} S_{i,k} + S_{n-k,k}^2.$$

Nutzt man nun die Unabhängigkeit der Anfallsjahre aus, so erhält man

$$\begin{aligned} & \text{Cov}\left(\left(\sum_{i=0}^{n-k} S_{i,k}\right)^2, \frac{1}{\sum_{i=0}^{n-k-1} S_{i,k}} \middle| \mathcal{B}_{k-1}\right) = \\ &= \text{Cov}\left(\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^2, \left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right) + \\ &+ 2 \text{Cov}\left(S_{n-k,k} \sum_{i=0}^{n-k-1} S_{i,k}, \left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right) = \\ &= \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] - \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^2 \middle| \mathcal{B}_{k-1}\right] \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right] + \\ &+ 2\left(\mathbb{E}[S_{n-k,k} \middle| \mathcal{B}_{k-1}] - \mathbb{E}\left[S_{n-k,k} \sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right]\right) = \\ &= \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] - \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^2 \middle| \mathcal{B}_{k-1}\right] \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right] + \\ &+ 2 f_k S_{n-k,k-1} \left(1 - \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right]\right). \end{aligned}$$

Nachdem $S_{i,k}$ und somit f_k als positiv vorausgesetzt werden, folgt mithilfe der Jensenschen Ungleichung für die konvexen Funktionen $f(x) = x^2$ bzw. $f(x) = x^{-1}$ (für $x \geq 0$)

$$\begin{aligned} & \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] - \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^2 \middle| \mathcal{B}_{k-1}\right] \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right] + \\ &+ 2 f_k S_{n-k,k-1} \left(1 - \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] \mathbb{E}\left[\left(\sum_{i=0}^{n-k-1} S_{i,k}\right)^{-1} \middle| \mathcal{B}_{k-1}\right]\right) < \\ &< \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] - \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right]^2 \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right]^{-1} + \\ &+ 2 f_k S_{n-k,k-1} \left(1 - \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right] \mathbb{E}\left[\sum_{i=0}^{n-k-1} S_{i,k} \middle| \mathcal{B}_{k-1}\right]^{-1}\right) = 0 \end{aligned}$$

Die Jensensche Ungleichung ist strikt, nachdem angenommen wurde, dass die Varianz-Parameter σ_k^2 strikt positiv sind, d.h. $\sum_{i=0}^{n-k-1} S_{i,k}$ ist zur Zeit $k-1$ nicht deterministisch. Damit ist die Aussage bewiesen. \square

Die **Varianz** wird folgendermaßen geschätzt:

$$\hat{\sigma}_k^2 = \frac{1}{n-k} \sum_{i=0}^{n-k} S_{i,k-1} \left(\frac{S_{i,k}}{S_{i,k-1}} - \hat{f}_k \right)^2, \quad k \in \{1, \dots, n-1\}$$

Als Motivation für diese Wahl dient Lemma 3.6. Für $k = n$ kann mit obiger Formel keine Varianz geschätzt werden, nachdem für den Schätzer \hat{f}_n zu wenige Daten vorhanden sind. Eine Möglichkeit für die Schätzung von $\hat{\sigma}_n$ ist, die normalerweise exponentiell fallende Folge $\hat{\sigma}_1, \dots, \hat{\sigma}_{n-1}$ zu extrapolieren. Das kann z.B. mithilfe einer log-linearen Regression oder einfach nur durch die Forderung, dass

$$\frac{\hat{\sigma}_{n-2}}{\hat{\sigma}_{n-1}} = \frac{\hat{\sigma}_{n-1}}{\hat{\sigma}_n}$$

zumindest für $\hat{\sigma}_{n-2} > \hat{\sigma}_{n-1}$ gilt, erfolgen. Diese letzte Möglichkeit führt auf

$$\hat{\sigma}_n^2 := \min \left(\frac{\hat{\sigma}_{n-1}^4}{\hat{\sigma}_{n-2}^2}, \hat{\sigma}_{n-1}^2, \hat{\sigma}_{n-2}^2 \right). \quad (3.8)$$

Dann gilt folgendes Lemma:

Lemma 3.6: Unter den Modellannahmen 3.4 gilt für $k \in \{1, \dots, n-1\}$:

- a) Bedingt auf \mathcal{B}_{k-1} ist $\hat{\sigma}_k^2$ ein unverzerrter Schätzer für σ_k^2 , d.h. $\mathbb{E}[\hat{\sigma}_k^2 | \mathcal{B}_{k-1}] = \sigma_k^2$.
- b) $\hat{\sigma}_k^2$ ist ein (unbedingt) unverzerrter Schätzer für σ_k^2 , d.h. $\mathbb{E}[\hat{\sigma}_k^2] = \sigma_k^2$.

Beweis:

- a) Wir betrachten zuerst

$$\begin{aligned} & \mathbb{E} \left[\left(\frac{S_{i,k}}{S_{i,k-1}} - \hat{f}_k \right)^2 \middle| \mathcal{B}_{k-1} \right] = \mathbb{E} \left[\left(\left(\frac{S_{i,k}}{S_{i,k-1}} - f_k \right) - (\hat{f}_k - f_k) \right)^2 \middle| \mathcal{B}_{k-1} \right] = \\ & = \mathbb{E} \left[\left(\frac{S_{i,k}}{S_{i,k-1}} - f_k \right)^2 \middle| \mathcal{B}_{k-1} \right] - 2 \mathbb{E} \left[\left(\frac{S_{i,k}}{S_{i,k-1}} - f_k \right) (\hat{f}_k - f_k) \middle| \mathcal{B}_{k-1} \right] \\ & \quad + \mathbb{E} \left[(\hat{f}_k - f_k)^2 \middle| \mathcal{B}_{k-1} \right]. \end{aligned}$$

Wir wollen nun jeden Term der rechten Seite einzeln berechnen. Aus den Modellannahmen 3.4 und (3.10) folgt

$$\mathbb{E} \left[\left(\frac{S_{i,k}}{S_{i,k-1}} - f_k \right)^2 \middle| \mathcal{B}_{k-1} \right] = \text{Var} \left(\frac{S_{i,k}}{S_{i,k-1}} \middle| \mathcal{B}_{k-1} \right) = \text{Var} (F_{i,k} | \mathcal{B}_{k-1}) = \frac{\sigma_k^2}{S_{i,k-1}}.$$

Aus den Modellannahmen 3.4, Lemma 3.3 und (3.10) folgt

$$\begin{aligned}
& \mathbb{E} \left[\left(\frac{S_{i,k}}{S_{i,k-1}} - f_k \right) \left(\hat{f}_k - f_k \right) \middle| \mathcal{B}_{k-1} \right] = \text{Cov} \left(\frac{S_{i,k}}{S_{i,k-1}}, \hat{f}_k \middle| \mathcal{B}_{k-1} \right) = \\
& = \text{Cov} \left(\frac{S_{i,k}}{S_{i,k-1}}, \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}} \middle| \mathcal{B}_{k-1} \right) = \\
& = \frac{1}{S_{i,k-1}} \frac{1}{\sum_{j=0}^{n-k} S_{j,k-1}} \sum_{j=0}^{n-k} \text{Cov} (S_{i,k}, S_{j,k} | \mathcal{B}_{k-1}) = \\
& = \frac{1}{S_{i,k-1}} \frac{1}{\sum_{j=0}^{n-k} S_{j,k-1}} \text{Var} (S_{i,k} | \mathcal{B}_{k-1}) = \frac{S_{i,k-1}}{\sum_{j=0}^{n-k} S_{j,k-1}} \text{Var} (F_{i,k} | \mathcal{B}_{k-1}) = \\
& = \frac{\sigma_k^2}{\sum_{j=0}^{n-k} S_{j,k-1}}.
\end{aligned}$$

Wegen Lemma 3.3 und 3.8 gilt

$$\mathbb{E} \left[\left(\hat{f}_k - f_k \right)^2 \middle| \mathcal{B}_{k-1} \right] = \text{Var} \left(\hat{f}_k \middle| \mathcal{B}_{k-1} \right) = \frac{\sigma_k^2}{\sum_{j=0}^{n-k} S_{j,k-1}}.$$

Insgesamt ergibt sich also

$$\begin{aligned}
\mathbb{E} [\hat{\sigma}_k^2 | \mathcal{B}_{k-1}] &= \mathbb{E} \left[\frac{1}{n-k} \sum_{i=0}^{n-k} S_{i,k-1} \left(\frac{S_{i,k}}{S_{i,k-1}} - \hat{f}_k \right)^2 \middle| \mathcal{B}_{k-1} \right] = \\
&= \frac{1}{n-k} \sum_{i=0}^{n-k} S_{i,k-1} \mathbb{E} \left[\left(\frac{S_{i,k}}{S_{i,k-1}} - \hat{f}_k \right)^2 \middle| \mathcal{B}_{k-1} \right] = \\
&= \frac{1}{n-k} \sum_{i=0}^{n-k} S_{i,k-1} \left(\frac{\sigma_k^2}{S_{i,k-1}} - 2 \frac{\sigma_k^2}{\sum_{j=0}^{n-k} S_{j,k-1}} + \frac{\sigma_k^2}{\sum_{j=0}^{n-k} S_{j,k-1}} \right) = \\
&= \frac{n-k+1}{n-k} \sigma_k^2 - \frac{\sigma_k^2}{n-k} = \sigma_k^2.
\end{aligned}$$

b) Aus Punkt a) folgt unmittelbar

$$\mathbb{E} [\hat{\sigma}_k^2] = \mathbb{E} [\mathbb{E} [\hat{\sigma}_k^2 | \mathcal{B}_{k-1}]] = \mathbb{E} [\sigma_k^2] = \sigma_k^2.$$

□

3.2.2 Optimalität

In diesem Kapitel wollen wir zeigen, dass die Schätzer \hat{f}_k in einem gewissen Sinne optimal sind.

Man beachte, dass $F_{i,k} = S_{i,k}/S_{i,k-1}$ ein (auf \mathcal{B}_{k-1}) bedingt unverzerrter Schätzer für f_k ist, denn aufgrund der Modellannahmen 3.4 gilt

$$\begin{aligned}
\mathbb{E} [F_{i,k} | \mathcal{B}_{k-1}] &= \mathbb{E} \left[\frac{S_{i,k}}{S_{i,k-1}} \middle| \mathcal{B}_{k-1} \right] = \mathbb{E} \left[\frac{S_{i,k}}{S_{i,k-1}} \middle| S_{i,1}, \dots, S_{i,k-1} \right] = \\
&= \frac{1}{S_{i,k-1}} \mathbb{E} [S_{i,k} | S_{i,1}, \dots, S_{i,k-1}] = \frac{1}{S_{i,k-1}} \mathbb{E} [S_{i,k} | S_{i,k-1}] = \\
&= \frac{1}{S_{i,k-1}} f_k S_{i,k-1} = f_k.
\end{aligned} \tag{3.9}$$

Um die Wahl des Schätzers \hat{f}_k zu motivieren, brauchen wir folgendes Hilfslemma:

Lemma 3.7: P_1, \dots, P_H seien stochastisch unabhängige unverzerrte Schätzer für μ mit Varianzen $\sigma_1^2, \dots, \sigma_H^2 > 0$. Dann ist jene unverzerrte Linearkombination von P_1, \dots, P_H mit minimaler Varianz gegeben durch

$$P = \frac{\sum_{h=1}^H \frac{P_h}{\sigma_h^2}}{\sum_{h=1}^H \frac{1}{\sigma_h^2}}$$

mit

$$\text{Var}(P) = \left(\sum_{h=1}^H \frac{1}{\sigma_h^2} \right)^{-1}$$

Beweis: Nachdem P eine Linearkombination von P_1, \dots, P_H sein soll, gilt $P = \sum_{h=1}^H \alpha_h P_h$ für $\alpha_1, \dots, \alpha_H \in \mathbb{R}$. Wir definieren die zwei H -dimensionalen Vektoren

$$\begin{aligned} \mathbf{P} &= (P_1, \dots, P_H)', \\ \boldsymbol{\alpha} &= (\alpha_1, \dots, \alpha_H)'. \end{aligned}$$

Außerdem bezeichne V die Kovarianzmatrix von \mathbf{P} , die in diesem Fall eine Diagonalmatrix mit Einträgen σ_h^2 ist. Dann gilt für $\mathbf{1} = (1, \dots, 1)' \in \mathbb{R}^H$

$$\begin{aligned} \mathbb{E}[\boldsymbol{\alpha}'\mathbf{P}] &= \mu \boldsymbol{\alpha}'\mathbf{1} \\ \text{Var}(\boldsymbol{\alpha}'\mathbf{P}) &= \boldsymbol{\alpha}'V\boldsymbol{\alpha} = \sum_{h=1}^H \alpha_h^2 \sigma_h^2. \end{aligned}$$

Nachdem der Schätzer unverzerrt sein soll, muss als Nebenbedingung

$$\mathbb{E}[\boldsymbol{\alpha}'\mathbf{P}] = \mu$$

bzw.

$$\boldsymbol{\alpha}'\mathbf{1} = 1$$

gelten. Daher ist die Lagrange-Funktion für den unverzerrten Minimum-Varianz-Schätzer gegeben durch

$$\mathcal{L}(\boldsymbol{\alpha}, \lambda) = \frac{1}{2} \boldsymbol{\alpha}'V\boldsymbol{\alpha} - \lambda (\boldsymbol{\alpha}'\mathbf{1} - 1)$$

Um das Optimum zu erhalten, muss folgendes System an Gleichungen gelöst werden:

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\alpha}} \mathcal{L}(\boldsymbol{\alpha}, \lambda) &= V\boldsymbol{\alpha} - \lambda \mathbf{1} = \mathbf{0} \\ \frac{\partial}{\partial \lambda} \mathcal{L}(\boldsymbol{\alpha}, \lambda) &= -\boldsymbol{\alpha}'\mathbf{1} + 1 = 0 \end{aligned}$$

Daraus folgt

$$\begin{aligned} \boldsymbol{\alpha} &= \lambda V^{-1}\mathbf{1}, \\ \boldsymbol{\alpha}'\mathbf{1} &= \mathbf{1}'\boldsymbol{\alpha} = 1 \end{aligned}$$

und

$$\mathbf{1} = \mathbf{1}'\boldsymbol{\alpha} = \lambda\mathbf{1}'V^{-1}\mathbf{1},$$

d.h.

$$\lambda = (\mathbf{1}'V^{-1}\mathbf{1})^{-1}$$

und

$$\boldsymbol{\alpha} = (\mathbf{1}'V^{-1}\mathbf{1})^{-1} V^{-1}\mathbf{1}.$$

Setzt man nun die spezielle Form von V ein, so erhält man

$$\boldsymbol{\alpha} = \left(\sum_{h=1}^H \sigma_h^{-2} \right)^{-1} \cdot (\sigma_1^{-2}, \dots, \sigma_H^{-2})'$$

Die Aussage folgt nun sofort aus $P = \boldsymbol{\alpha}'\mathbf{P}$. □

Folgendes Lemma motiviert die Wahl des Schätzers \widehat{f}_k für die Abwicklungsfaktoren:

Lemma 3.8: Unter den Modellannahmen 3.4 ist \widehat{f}_k ein \mathcal{B}_{k-1} -messbarer unverzerrter Schätzer für f_k mit minimaler bedingte Varianz unter allen unverzerrten Linearkombinationen der unverzerrten Schätzer $(F_{i,k})_{0 \leq i \leq n-k}$ für f_k , bedingt auf \mathcal{B}_{k-1} , d.h.

$$\text{Var} \left(\widehat{f}_k \mid \mathcal{B}_{k-1} \right) = \min_{\substack{\alpha_i \in \mathbb{R}, \\ \sum \alpha_i = 1}} \text{Var} \left(\sum_{i=0}^{n-k} \alpha_i F_{i,k} \mid \mathcal{B}_{k-1} \right)$$

Die bedingte Varianz von \widehat{f}_k ist gegeben durch

$$\text{Var} \left(\widehat{f}_k \mid \mathcal{B}_{k-1} \right) = \frac{\sigma_k^2}{\sum_{i=0}^{n-k} S_{i,k-1}}$$

Beweis: Bedingt auf \mathcal{B}_{k-1} sind die individuellen Abwicklungsfaktoren $(F_{i,k})_{0 \leq i \leq n-k}$ wegen (3.9) unverzerrte und unabhängige Schätzer für f_k mit

$$\begin{aligned} \text{Var} (F_{i,k} \mid \mathcal{B}_{k-1}) &= \mathbb{E} [F_{i,k}^2 \mid \mathcal{B}_{k-1}] - \mathbb{E} [F_{i,k} \mid \mathcal{B}_{k-1}]^2 = \\ &= \mathbb{E} \left[\frac{S_{i,k}^2}{S_{i,k-1}^2} \mid S_{i,1}, \dots, S_{i,k-1} \right] - \mathbb{E} \left[\frac{S_{i,k}}{S_{i,k-1}} \mid S_{i,1}, \dots, S_{i,k-1} \right]^2 = \\ &= \frac{1}{S_{i,k-1}^2} \mathbb{E} [S_{i,k}^2 \mid S_{i,1}, \dots, S_{i,k-1}] - \left(\frac{1}{S_{i,k-1}} \mathbb{E} [S_{i,k} \mid S_{i,1}, \dots, S_{i,k-1}] \right)^2 = \\ &= \frac{1}{S_{i,k-1}^2} \mathbb{E} [S_{i,k}^2 \mid S_{i,k-1}] - \left(\frac{1}{S_{i,k-1}} \mathbb{E} [S_{i,k} \mid S_{i,k-1}] \right)^2 = \\ &= \mathbb{E} \left[\left(\frac{S_{i,k}}{S_{i,k-1}} \right)^2 \mid S_{i,k-1} \right] - \mathbb{E} \left[\frac{S_{i,k}}{S_{i,k-1}} \mid S_{i,k-1} \right]^2 = \text{Var} \left(\frac{S_{i,k}}{S_{i,k-1}} \mid S_{i,k-1} \right) = \\ &= \frac{1}{S_{i,k-1}^2} \text{Var} (S_{i,k} \mid S_{i,k-1}) = \frac{1}{S_{i,k-1}^2} \sigma_k^2 S_{i,k-1} = \frac{\sigma_k^2}{S_{i,k-1}}. \end{aligned} \tag{3.10}$$

Die zu beweisenden Aussagen folgen nun unmittelbar aus Lemma 3.7. □

3.2.3 Mittlere quadratische Abweichung der Prognose $\widehat{R}_i^{\text{CL}}$ für ein Abwicklungsjahr i

Wir wollen nun für das verteilungsfreie Chain-Ladder-Modell für jedes Anfallsjahr $i \in \{1, \dots, n\}$ einen Schätzer für die bedingte mittlere quadratische Abweichung von $\widehat{R}_i^{\text{CL}}$ finden und verwenden dafür die von Mack in [7] vorgeschlagenen Schätzer. Die mittlere quadratische Abweichung wird im Anhang A genauer erklärt.

Lemma 3.9: Unter den Modellannahmen 3.4 kann die mittlere quadratische Abweichung der Chain-Ladder-Reserve $\widehat{R}_i^{\text{CL}}$ für das Abwicklungsjahr i , wobei $i \in \{1, \dots, n\}$ folgendermaßen geschätzt werden:

$$\widehat{\text{MSE}}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{CL}}) = (\widehat{S}_{i,n}^{\text{CL}})^2 \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_{k+1}^2}{\widehat{f}_{k+1}^2} \left(\frac{1}{\widehat{S}_{i,k}^{\text{CL}}} + \frac{1}{\sum_{i=0}^{n-k-1} S_{i,k}} \right)$$

Beweis: Wegen (A.1) gilt

$$\begin{aligned} \text{MSE}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{CL}}) &= \mathbb{E} \left[\left(\widehat{R}_i^{\text{CL}} - R_i \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \mathbb{E} \left[\left(\left(\widehat{R}_i^{\text{CL}} + S_{i,n-i} \right) - \left(R_i + S_{i,n-i} \right) \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \mathbb{E} \left[\left(\widehat{S}_{i,n}^{\text{CL}} - S_{i,n} \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \text{Var}(S_{i,n} | \mathcal{D}_n) + \left(\widehat{S}_{i,n}^{\text{CL}} - \mathbb{E}[S_{i,n} | \mathcal{D}_n] \right)^2. \end{aligned} \quad (3.11)$$

Wir wollen nun beide Terme einzeln berechnen und in weiterer Folge unter den Modellannahmen 3.4 Schätzer dafür finden. Für den bedingten stochastischen Fehler gilt für $i > 0$

$$\begin{aligned} \text{Var}(S_{i,n} | \mathcal{D}_n) &= \text{Var}(S_{i,n} | S_{i,n-i}) = \\ &= \mathbb{E} \left[(S_{i,n} - \mathbb{E}[S_{i,n} | S_{i,n-i}])^2 \middle| S_{i,n-i} \right] = \\ &= \mathbb{E} \left[(S_{i,n} - \mathbb{E}[S_{i,n} | S_{i,n-1}] + \mathbb{E}[S_{i,n} | S_{i,n-1}] - \mathbb{E}[S_{i,n} | S_{i,n-i}])^2 \middle| S_{i,n-i} \right] = \\ &= \mathbb{E} \left[(S_{i,n} - \mathbb{E}[S_{i,n} | S_{i,n-1}])^2 \middle| S_{n-i} \right] + \\ &\quad + \mathbb{E} \left[(\mathbb{E}[S_{i,n} | S_{i,n-1}] - \mathbb{E}[S_{i,n} | S_{i,n-i}])^2 \middle| S_{n-i} \right] + \\ &\quad + 2 \mathbb{E} \left[(S_{i,n} - \mathbb{E}[S_{i,n} | S_{i,n-1}]) (\mathbb{E}[S_{i,n} | S_{i,n-1}] - \mathbb{E}[S_{i,n} | S_{i,n-i}]) \middle| S_{i,n-i} \right]. \end{aligned} \quad (3.12)$$

Für den letzten Term gilt

$$\begin{aligned} &\mathbb{E} \left[(S_{i,n} - \mathbb{E}[S_{i,n} | S_{i,n-1}]) (\mathbb{E}[S_{i,n} | S_{i,n-1}] - \mathbb{E}[S_{i,n} | S_{i,n-i}]) \middle| S_{i,n-i} \right] = \\ &= \mathbb{E} \left[S_{i,n} \mathbb{E}[S_{i,n} | S_{i,n-1}] \middle| S_{i,n-i} \right] - \mathbb{E} \left[\mathbb{E}[S_{i,n} | S_{i,n-1}]^2 \middle| S_{i,n-i} \right] - \\ &\quad - \mathbb{E} \left[S_{i,n} \mathbb{E}[S_{i,n} | S_{i,n-i}] \middle| S_{i,n-i} \right] + \mathbb{E} \left[\mathbb{E}[S_{i,n} | S_{i,n-1}] \mathbb{E}[S_{i,n} | S_{i,n-i}] \middle| S_{i,n-i} \right] = \\ &= \mathbb{E} \left[\mathbb{E}[S_{i,n} \mathbb{E}[S_{i,n} | S_{i,n-1}] \middle| S_{i,n-1}] \middle| S_{i,n-i} \right] - \mathbb{E} \left[\mathbb{E}[S_{i,n} | S_{i,n-1}]^2 \middle| S_{i,n-i} \right] - \\ &\quad - \mathbb{E} \left[S_{i,n} \mathbb{E}[S_{i,n} | S_{i,n-i}]^2 \right] + \mathbb{E} \left[S_{i,n} \mathbb{E}[S_{i,n} | S_{i,n-i}] \mathbb{E}[\mathbb{E}[S_{i,n} | S_{i,n-1}] \middle| S_{i,n-i}] \right] = \\ &= \mathbb{E} \left[\mathbb{E}[S_{i,n} | S_{i,n-1}]^2 \middle| S_{i,n-i} \right] - \mathbb{E} \left[\mathbb{E}[S_{i,n} | S_{i,n-1}]^2 \middle| S_{i,n-i} \right] - \\ &\quad - \mathbb{E} \left[S_{i,n} \mathbb{E}[S_{i,n} | S_{i,n-i}]^2 \right] + \mathbb{E} \left[S_{i,n} \mathbb{E}[S_{i,n} | S_{i,n-i}]^2 \right] = 0. \end{aligned}$$

Also bleibt in (3.12)

$$\begin{aligned}
\text{Var}(S_{i,n} | \mathcal{D}_n) &= \mathbb{E} \left[(S_{i,n} - \mathbb{E}[S_{i,n} | S_{i,n-1}])^2 \middle| S_{i,n-i} \right] + \\
&\quad + \mathbb{E} \left[(\mathbb{E}[S_{i,n} | S_{i,n-1}] - \mathbb{E}[S_{i,n} | S_{i,n-i}])^2 \middle| S_{i,n-i} \right] = \\
&= \mathbb{E} \left[\mathbb{E} \left[(S_{i,n} - \mathbb{E}[S_{i,n} | S_{i,n-1}])^2 \middle| S_{i,n-1} \right] \middle| S_{i,n-i} \right] + \\
&\quad + \mathbb{E} \left[(\mathbb{E}[S_{i,n} | S_{i,n-1}] - \mathbb{E}[\mathbb{E}[S_{i,n} | S_{i,n-1}] | S_{i,n-i}])^2 \middle| S_{i,n-i} \right] = \\
&= \mathbb{E}[\text{Var}(S_{i,n} | S_{i,n-1}) | S_{i,n-i}] + \text{Var}(\mathbb{E}[S_{i,n} | S_{i,n-1}] | S_{i,n-i}) = \\
&= \sigma_n^2 \mathbb{E}[S_{i,n-1} | S_{i,n-i}] + f_n^2 \text{Var}(S_{i,n-1} | S_{i,n-i}) = \\
&= \sigma_n^2 S_{i,n-i} \prod_{m=n-i+1}^{n-1} f_m + f_n^2 \text{Var}(S_{i,n-1} | S_{i,n-i}).
\end{aligned}$$

Iterativ kommt man auf

$$\begin{aligned}
\text{Var}(S_{i,n} | S_{i,n-i}) &= S_{i,n-i} \sum_{j=n-i+1}^n \left(\prod_{k=j+1}^n f_k^2 \right) \sigma_j^2 \prod_{m=n-i+1}^{j-1} f_m = \\
&= \sum_{j=n-i+1}^n \left(\prod_{k=j+1}^n f_k^2 \right) \sigma_j^2 \mathbb{E}[S_{i,j-1} | S_{i,n-i}] = \\
&= \sum_{j=n-i+1}^n \frac{\sigma_j^2}{f_j^2} \left(\prod_{k=j}^n f_k^2 \right) \frac{\mathbb{E}[S_{i,j-1} | S_{i,n-i}]^2}{\mathbb{E}[S_{i,j-1} | S_{i,n-i}]} = \\
&= \mathbb{E}[S_{i,n} | S_{i,n-i}]^2 \sum_{j=n-i+1}^n \frac{\sigma_j^2}{f_j^2 \mathbb{E}[S_{i,j-1} | S_{i,n-i}]}.
\end{aligned}$$

Das motiviert den Schätzer für den bedingten stochastischen Fehler für ein Anfallsjahr i :

$$\widehat{\text{Var}}(R_i | \mathcal{D}_n) = \widehat{\text{Var}}(S_{i,n} | \mathcal{D}_n) = \left(\widehat{S}_{i,n}^{\text{CL}} \right)^2 \sum_{j=n-i}^{n-1} \frac{\widehat{\sigma}_{j+1}^2}{\widehat{f}_{j+1}^2 \widehat{S}_{i,j}^{\text{CL}}}$$

Für den zweiten Term in (3.11), den bedingten Schätzfehler, gilt

$$\begin{aligned}
\left(\widehat{S}_{i,n}^{\text{CL}} - \mathbb{E}[S_{i,n} | \mathcal{D}_n] \right)^2 &= \left(S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_n - S_{i,n-i} f_{n-i+1} \cdots f_n \right)^2 = \\
&= S_{i,n-i}^2 \left(\widehat{f}_{n-i+1} \cdots \widehat{f}_n - f_{n-i+1} \cdots f_n \right)^2 \tag{3.13}
\end{aligned}$$

Hier kann nicht einfach f_k durch den Schätzer \widehat{f}_k ersetzt werden, denn das würde auf einen Schätzfehler von Null führen. Wir verwenden deshalb einen anderen Ansatz. Sei

$$T_k^i = \widehat{f}_{n-i+1} \cdots \widehat{f}_k \cdot (f_{k+1} - \widehat{f}_{k+1}) \cdot f_{k+2} \cdots f_n.$$

Dann gilt

$$\begin{aligned}
\sum_{k=n-i}^{n-1} T_k^i &= \sum_{k=n-i}^{n-1} \left(\widehat{f}_{n-i+1} \cdots \widehat{f}_k f_{k+1} \cdots f_n - \widehat{f}_{n-i+1} \cdots \widehat{f}_{k+1} f_{k+2} \cdots f_n \right) = \\
&= f_{n-i+1} \cdots f_n - \widehat{f}_{n-i+1} \cdots \widehat{f}_n
\end{aligned}$$

und somit

$$\left(\widehat{S}_{i,n}^{\text{CL}} - \mathbb{E}[S_{i,n} | \mathcal{D}_n]\right)^2 = S_{i,n-i}^2 \left(\sum_{k=n-i}^{n-1} T_k^i\right)^2 = S_{i,n-i}^2 \left(\sum_{k=n-i}^{n-1} (T_k^i)^2 + 2 \sum_{n-i \leq j < k \leq n-1} T_j^i T_k^i\right).$$

Um zu einem Schätzer zu gelangen, wird nun $(T_k^i)^2$ durch $\mathbb{E}\left[(T_k^i)^2 | \mathcal{B}_k\right]$ und $T_j^i T_k^i$ für $j < k$ durch $\mathbb{E}\left[T_j^i T_k^i | \mathcal{B}_k\right]$ ersetzt. Es gilt unter Verwendung von Lemma 3.3 und 3.8

$$\begin{aligned} & \mathbb{E}\left[T_j^i T_k^i | \mathcal{B}_k\right] = \\ &= \mathbb{E}\left[\widehat{f}_{n-i+1} \cdots \widehat{f}_j \cdot (f_{j+1} - \widehat{f}_{j+1}) \cdot f_{j+2} \cdots f_n \cdot \right. \\ & \quad \left. \widehat{f}_{n-i+1} \cdots \widehat{f}_k \cdot (f_{k+1} - \widehat{f}_{k+1}) \cdot f_{k+2} \cdots f_n | \mathcal{B}_k\right] = \\ &= \widehat{f}_{n-i+1} \cdots \widehat{f}_j \cdot (f_{j+1} - \widehat{f}_{j+1}) \cdot f_{j+2} \cdots f_n \cdot \widehat{f}_{n-i+1} \cdots \widehat{f}_k \cdot f_{k+2} \cdots f_n \cdot \\ & \quad \mathbb{E}\left[f_{k+1} - \widehat{f}_{k+1} | \mathcal{B}_k\right] = 0 \end{aligned}$$

und

$$\begin{aligned} \mathbb{E}\left[(T_k^i)^2 | \mathcal{B}_k\right] &= \mathbb{E}\left[\widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot (f_{k+1} - \widehat{f}_{k+1})^2 \cdot f_{k+2}^2 \cdots f_n^2 | \mathcal{B}_k\right] = \\ &= \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot f_{k+2}^2 \cdots f_n^2 \cdot \mathbb{E}\left[(f_{k+1} - \widehat{f}_{k+1})^2 | \mathcal{B}_k\right] = \\ &= \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot f_{k+2}^2 \cdots f_n^2 \cdot \text{Var}\left(\widehat{f}_{k+1} | \mathcal{B}_k\right) = \\ &= \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot f_{k+2}^2 \cdots f_n^2 \cdot \frac{\sigma_{k+1}^2}{\sum_{i=0}^{n-k-1} S_{i,k}}. \end{aligned}$$

Der Schätzer für den bedingten Schätzfehler (3.13) ist also gegeben durch

$$S_{i,n-i}^2 \sum_{k=n-i}^{n-1} \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot f_{k+2}^2 \cdots f_n^2 \cdot \frac{\sigma_{k+1}^2}{\sum_{i=0}^{n-k-1} S_{i,k}}.$$

Ersetzt man nun die unbekannt Parameter f_k und σ_k^2 durch ihre Schätzer \widehat{f}_k und $\widehat{\sigma}_k^2$, so erhält man als Schätzer für die bedingte mittlere quadratische Abweichung (3.11) von $\widehat{S}_{i,n}^{\text{CL}}$ für ein Abwicklungsjahr $i \in \{1, \dots, n\}$

$$\begin{aligned} \widehat{\text{MSE}}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{CL}}\right) &= \left(\widehat{S}_{i,n}^{\text{CL}}\right)^2 \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_{k+1}^2}{\widehat{f}_{k+1}^2 \widehat{S}_{i,k}^{\text{CL}}} \\ & \quad + S_{i,n-i}^2 \sum_{k=n-i}^{n-1} \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot \widehat{f}_{k+2}^2 \cdots \widehat{f}_n^2 \cdot \frac{\widehat{\sigma}_{k+1}^2}{\sum_{i=0}^{n-k-1} S_{i,k}} = \\ &= \left(\widehat{S}_{i,n}^{\text{CL}}\right)^2 \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_{k+1}^2}{\widehat{f}_{k+1}^2 \widehat{S}_{i,k}^{\text{CL}}} + \left(\widehat{S}_{i,n}^{\text{CL}}\right)^2 \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_{k+1}^2}{\widehat{f}_{k+1}^2 \sum_{i=0}^{n-k-1} S_{i,k}} = \\ &= \left(\widehat{S}_{i,n}^{\text{CL}}\right)^2 \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_{k+1}^2}{\widehat{f}_{k+1}^2} \left(\frac{1}{\widehat{S}_{i,k}^{\text{CL}}} + \frac{1}{\sum_{i=0}^{n-k-1} S_{i,k}}\right). \end{aligned}$$

□

3.2.4 Mittlere quadratische Abweichung der Prognose \widehat{R}^{CL}

Oft ist aber nicht nur die mittlere quadratische Abweichung für ein Anfallsjahr i von Interesse, sondern auch die mittlere quadratische Abweichung der gesamten Reserve $\widehat{R} = \sum_{i=1}^n \widehat{R}_i^{\text{CL}}$. Auch hier verwenden wir wieder die von Mack in [7] vorgeschlagenen Schätzer. Dabei gilt folgendes Lemma:

Lemma 3.10: Unter den Modellannahmen 3.4 kann die mittlere quadratische Abweichung der Gesamtreserve $\widehat{R} = \sum_{i=1}^n \widehat{R}_i^{\text{CL}}$ folgendermaßen geschätzt werden:

$$\begin{aligned} \widehat{\text{MSE}}_{\sum_i R_i | \mathcal{D}_n} \left(\sum_{i=1}^n \widehat{R}_i^{\text{CL}} \right) &= \sum_{i=1}^n \widehat{\text{MSE}}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{CL}} \right) \\ &\quad + 2 \sum_{1 \leq i < j \leq n} \widehat{S}_{i,n}^{\text{CL}} \widehat{S}_{j,n}^{\text{CL}} \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_{k+1}^2}{\widehat{f}_{k+1}^2 \sum_{l=0}^{n-k-1} S_{l,k}} \end{aligned}$$

Beweis: Wegen (3.7) und (A.1) gilt

$$\begin{aligned} \text{MSE}_{\sum_i R_i | \mathcal{D}_n} \left(\sum_{i=1}^n \widehat{R}_i^{\text{CL}} \right) &= \mathbb{E} \left[\left(\sum_{i=1}^n \widehat{R}_i^{\text{CL}} - \sum_{i=1}^n R_i \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \mathbb{E} \left[\left(\sum_{i=1}^n (\widehat{R}_i^{\text{CL}} + S_{i,n-i}) - \sum_{i=1}^n (R_i + S_{i,n-i}) \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \mathbb{E} \left[\left(\sum_{i=1}^n \widehat{S}_{i,n}^{\text{CL}} - \sum_{i=1}^n S_{i,n} \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \text{Var} \left(\sum_{i=1}^n S_{i,n} \middle| \mathcal{D}_n \right) + \left(\mathbb{E} \left[\sum_{i=1}^n S_{i,n} \middle| \mathcal{D}_n \right] - \sum_{i=1}^n \widehat{S}_{i,n}^{\text{CL}} \right)^2 \end{aligned} \quad (3.14)$$

Aufgrund der Unabhängigkeit der unterschiedlichen Anfallsjahre gilt

$$\text{Var} \left(\sum_{i=1}^n S_{i,n} \middle| \mathcal{D}_n \right) = \sum_{i=1}^n \text{Var} (S_{i,n} | \mathcal{D}_n). \quad (3.15)$$

Außerdem gilt

$$\begin{aligned} \left(\mathbb{E} \left[\sum_{i=1}^n S_{i,n} \middle| \mathcal{D}_n \right] - \sum_{i=1}^n \widehat{S}_{i,n}^{\text{CL}} \right)^2 &= \left(\sum_{i=1}^n \left(\mathbb{E} [S_{i,n} | \mathcal{D}_n] - \widehat{S}_{i,n}^{\text{CL}} \right) \right)^2 = \\ &= \sum_{i=1}^n \left(\mathbb{E} [S_{i,n} | \mathcal{D}_n] - \widehat{S}_{i,n}^{\text{CL}} \right)^2 + 2 \sum_{1 \leq i < j \leq n} \left(\mathbb{E} [S_{i,n} | \mathcal{D}_n] - \widehat{S}_{i,n}^{\text{CL}} \right) \left(\mathbb{E} [S_{j,n} | \mathcal{D}_n] - \widehat{S}_{j,n}^{\text{CL}} \right). \end{aligned} \quad (3.16)$$

Für die mittlere quadratische Abweichung erhält man aus (3.14), (3.15) und (3.16) also

$$\begin{aligned}
& \text{MSE}_{\sum_i R_i | \mathcal{D}_n} \left(\sum_{i=1}^n \widehat{R}_i^{\text{CL}} \right) = \\
&= \sum_{i=1}^n \text{Var} (S_{i,n} | \mathcal{D}_n) + \sum_{i=1}^n \left(\mathbb{E} [S_{i,n} | \mathcal{D}_n] - \widehat{S}_{i,n}^{\text{CL}} \right)^2 + \\
&+ 2 \sum_{1 \leq i < j \leq n} \left(\mathbb{E} [S_{i,n} | \mathcal{D}_n] - \widehat{S}_{i,n}^{\text{CL}} \right) \left(\mathbb{E} [S_{j,n} | \mathcal{D}_n] - \widehat{S}_{j,n}^{\text{CL}} \right) = \\
&= \sum_{i=1}^n \text{MSE}_{S_{i,n} | \mathcal{D}_n} \left(\widehat{S}_{i,n}^{\text{CL}} \right) + 2 \sum_{1 \leq i < j \leq n} \left(\mathbb{E} [S_{i,n} | \mathcal{D}_n] - \widehat{S}_{i,n}^{\text{CL}} \right) \left(\mathbb{E} [S_{j,n} | \mathcal{D}_n] - \widehat{S}_{j,n}^{\text{CL}} \right) = \\
&= \sum_{i=1}^n \text{MSE}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{CL}} \right) + 2 \sum_{1 \leq i < j \leq n} \left(S_{i,n-i} f_{n-i+1} \cdots f_n - S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_n \right) \cdot \\
&\quad \left(S_{j,n-j} f_{n-j+1} \cdots f_n - S_{j,n-j} \widehat{f}_{n-j+1} \cdots \widehat{f}_n \right) = \\
&= \sum_{i=1}^n \text{MSE}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{CL}} \right) + 2 \sum_{1 \leq i < j \leq n} S_{i,n-i} S_{j,n-j} \left(f_{n-i+1} \cdots f_n - \widehat{f}_{n-i+1} \cdots \widehat{f}_n \right) \cdot \\
&\quad \left(f_{n-j+1} \cdots f_n - \widehat{f}_{n-j+1} \cdots \widehat{f}_n \right) = \\
&= \sum_{i=1}^n \text{MSE}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{CL}} \right) + 2 \sum_{1 \leq i < j \leq n} S_{i,n-i} S_{j,n-j} \left(\sum_{k=n-i}^{n-1} T_k^i \right) \left(\sum_{l=n-j}^{n-1} T_l^j \right) = \\
&= \sum_{i=1}^n \text{MSE}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{CL}} \right) + 2 \sum_{1 \leq i < j \leq n} S_{i,n-i} S_{j,n-j} \left(\sum_{k=n-i}^{n-1} T_k^i T_k^j + \sum_{k=n-i}^{n-1} \sum_{\substack{l=n-j \\ l \neq k}}^{n-1} T_k^i T_l^j \right)
\end{aligned}$$

Wir gehen nun analog wie oben vor und ersetzen $T_k^i T_l^j$ für $k \neq l$ durch $\mathbb{E} \left[T_k^i T_l^j \mid \mathcal{B}_{\max(k,l)} \right]$ und $T_k^i T_k^j$ durch $\mathbb{E} \left[T_k^i T_k^j \mid \mathcal{B}_k \right]$. Man erhält für $k \neq l$

$$\begin{aligned}
& \mathbb{E} \left[T_k^i T_l^j \mid \mathcal{B}_{\max(k,l)} \right] = \\
&= \mathbb{E} \left[\widehat{f}_{n-i+1} \cdots \widehat{f}_k \cdot \left(f_{k+1} - \widehat{f}_{k+1} \right) \cdot f_{k+2} \cdots f_n \cdot \right. \\
&\quad \left. \widehat{f}_{n-j+1} \cdots \widehat{f}_l \cdot \left(f_{l+1} - \widehat{f}_{l+1} \right) \cdot f_{l+2} \cdots f_n \mid \mathcal{B}_{\max(k,l)} \right] = \\
&= \widehat{f}_{n-i+1} \cdots \widehat{f}_k \cdot \left(f_{\min(k,l)+1} - \widehat{f}_{\min(k,l)+1} \right) \cdot f_{k+2} \cdots f_n \cdot \widehat{f}_{n-j+1} \cdots \widehat{f}_l \cdot f_{l+2} \cdots f_n \cdot \\
&\quad \mathbb{E} \left[\left(f_{\max(k,l)+1} - \widehat{f}_{\max(k,l)+1} \right) \mid \mathcal{B}_{\max(k,l)} \right] = 0
\end{aligned}$$

und

$$\begin{aligned}
& \mathbb{E} \left[T_k^i T_k^j \mid \mathcal{B}_k \right] = \\
&= \mathbb{E} \left[\widehat{f}_{n-i+1} \cdots \widehat{f}_k \cdot \left(f_{k+1} - \widehat{f}_{k+1} \right) \cdot f_{k+2} \cdots f_n \cdot \right. \\
&\quad \left. \widehat{f}_{n-j+1} \cdots \widehat{f}_k \cdot \left(f_{k+1} - \widehat{f}_{k+1} \right) \cdot f_{k+2} \cdots f_n \mid \mathcal{B}_k \right] = \\
&= \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot f_{k+2}^2 \cdots f_n^2 \cdot \widehat{f}_{n-j+1} \cdots \widehat{f}_{n-i} \cdot \mathbb{E} \left[\left(f_{k+1} - \widehat{f}_{k+1} \right)^2 \mid \mathcal{B}_k \right] =
\end{aligned}$$

$$= \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot f_{k+2}^2 \cdots f_n^2 \cdot \widehat{f}_{n-j+1} \cdots \widehat{f}_{n-i} \cdot \frac{\sigma_{k+1}^2}{\sum_{i=0}^{n-k-1} S_{i,k}}.$$

Ersetzt man nun wieder f_k durch \widehat{f}_k , σ_k^2 durch $\widehat{\sigma}_k^2$ und $\text{MSE}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{CL}})$ durch $\widehat{\text{MSE}}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{CL}})$, so erhält man als Schätzer

$$\begin{aligned} & \sum_{i=1}^n \text{MSE}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{CL}}) \\ & + 2 \sum_{1 \leq i < j \leq n} S_{i,n-i} S_{j,n-j} \sum_{k=n-i}^{n-1} \widehat{f}_{n-i+1}^2 \cdots \widehat{f}_k^2 \cdot f_{k+2}^2 \cdots f_n^2 \cdot \widehat{f}_{n-j+1} \cdots \widehat{f}_{n-i} \cdot \frac{\widehat{\sigma}_{k+1}^2}{\sum_{i=0}^{n-k-1} S_{i,k}} = \\ & = \sum_{i=1}^n \text{MSE}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{CL}}) + 2 \sum_{1 \leq i < j \leq n} \widehat{S}_{i,n}^{\text{CL}} \widehat{S}_{j,n}^{\text{CL}} \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_{k+1}^2}{f_{k+1}^2 \sum_{l=0}^{n-k-1} S_{l,k}}. \end{aligned}$$

□

3.3 Numerisches Beispiel

Es wird nun für die Paid-Daten aus Tabelle 2.1 das Chain-Ladder-Verfahren gerechnet und die zugehörige mittlere quadratische Abweichung der Prognosen berechnet. $\widehat{\sigma}_n^2$ wird mit (3.8) geschätzt. Die Ergebnisse sind in den Tabellen 3.1 und 3.2 dargestellt.

Zum Vergleich wird das Chain-Ladder-Verfahren auch noch für die Incurred-Daten aus Tabelle 2.2 gerechnet. Die Ergebnisse sind in Tabelle 3.3 dargestellt. In Tabelle 3.4 sind die gesamten Reserven (d.h. inklusive Einzelfallreserve) dargestellt, d.h. die Differenz zwischen dem prognostizierten Incurred-Endstand und der aktuellen Beobachtung der Zahlungen:

$$\begin{aligned} \widetilde{R}_0^{\text{CL}} &= S_{0,n}^I - S_{0,n}^P \\ \widetilde{R}_i^{\text{CL}} &= \widehat{S}_{i,n}^{\text{CL}} - S_{i,n-i}^P \quad \text{für } i \geq 1 \\ \widetilde{R}^{\text{CL}} &= \sum_{i=0}^n \widetilde{R}_i^{\text{CL}} \end{aligned}$$

Vergleicht man die Gesamtreserve in Tabelle 3.2 mit jener in Tabelle 3.4, so erkennt man, dass die Reserve um 39.955 höher ist, wenn mit Incurred-Daten gerechnet wird. In Kapitel 11 wird ein Verfahren vorgestellt, dass diese Differenz verringern soll.

Tabelle 3.2: Reserven und mittlere quadratische Abweichung der Reserven des Chain-Ladder-Verfahrens für Paid-Daten

i	\hat{R}_i^{CL}	$\widehat{\text{Var}}(R_i \mathcal{D}_n)$	$\widehat{\text{MSE}}_{R_i \mathcal{D}_n}(\hat{R}_i^{\text{CL}})$	$\sqrt{\widehat{\text{MSE}}_{R_i \mathcal{D}_n}(\hat{R}_i^{\text{CL}})}$
1	143	26.494	55.943	237
2	846	209.997	355.175	596
3	2.998	1.431.597	2.108.466	1.452
4	3.996	1.939.743	2.759.310	1.661
5	5.836	3.847.821	5.104.216	2.259
6	7.281	5.922.544	7.518.912	2.742
7	22.965	13.214.478	15.911.321	3.989
8	90.166	132.651.620	150.533.836	12.269
	\hat{R}^{CL}	$\widehat{\text{Var}}(R \mathcal{D}_n)$	$\widehat{\text{MSE}}_{R \mathcal{D}_n}(\hat{R}^{\text{CL}})$	$\sqrt{\widehat{\text{MSE}}_{R \mathcal{D}_n}(\hat{R}^{\text{CL}})}$
	134.230	159.244.294	218.531.610	14.783

Tabelle 3.4: Gesamte Reserven für das Chain-Ladder-Verfahren für Incurred-Daten

i	\tilde{R}_i^{CL}
0	0
1	897
2	7.091
3	5.580
4	5.220
5	12.146
6	13.817
7	24.339
8	101.351
Gesamt	171.555

Kapitel 4

Bornhuetter-Ferguson-Verfahren

Dieses Kapitel lehnt sich an [6] und [1] an.

4.1 Verfahren und Reserve

Dieses Verfahren geht zurück auf Bornhuetter und Ferguson (1972). Die Bornhuetter-Ferguson-Methode ist grundsätzlich ein mechanischer Algorithmus. Es gibt verschiedene Möglichkeit, ein zugrundeliegendes stochastisches Modell zu bilden, das das Bornhuetter-Ferguson-Verfahren motiviert. Im Folgenden wird ein Modell vorgestellt. Die Bornhuetter-Ferguson-Methode ist im Allgemeinen sehr robust, nachdem sie Ausreißer in den Beobachtungen nicht miteinbezieht.

Modellannahmen 4.1:

- Die kumulativen Schadenhöhen $S_{i,k}$ für verschiedene Anfallsjahre i sind unabhängig.
- Es gibt Parameter $\mu_0, \dots, \mu_n > 0$ und ein Abwicklungsmuster $\alpha_0, \dots, \alpha_n > 0$ mit $\alpha_n = 1$, sodass für alle $0 \leq i \leq n$, $0 \leq k \leq n - 1$ und $1 \leq j \leq n - k$ gilt

$$\begin{aligned}\mathbb{E}[S_{i,0}] &= \alpha_0 \mu_i \\ \mathbb{E}[S_{i,k+j} | S_{i,0}, \dots, S_{i,k}] &= S_{i,k} + (\alpha_{k+j} - \alpha_k) \mu_i.\end{aligned}$$

Unter den Modellannahmen 4.1 gilt

$$\mathbb{E}[S_{i,k}] = \mathbb{E}[\mathbb{E}[S_{i,k} | S_{i,0}]] = \mathbb{E}[S_{i,0} + (\alpha_k - \alpha_0) \mu_i] = \alpha_0 \mu_i + \alpha_k \mu_i - \alpha_0 \mu_i = \alpha_k \mu_i$$

und

$$\mathbb{E}[S_{i,n}] = \mu_i.$$

Die Familie $(\alpha_k)_{k=0, \dots, n}$ bezeichnet das **Schadenabwicklungsmuster**. α_k gibt an, welcher Anteil am Gesamtschaden nach k Jahren bereits abgewickelt ist.

Aus den Modellannahmen 4.1 folgen die

Modellannahmen 4.2:

- Die kumulativen Schadenhöhen $S_{i,k}$ für verschiedene Anfallsjahre i sind unabhängig.

- Es gibt Parameter $\mu_0, \dots, \mu_n > 0$ und ein Abwicklungsmuster $\alpha_0, \dots, \alpha_n > 0$ mit $\alpha_n = 1$, sodass für alle $0 \leq i \leq n$ und $0 \leq k \leq n$ gilt

$$\mathbb{E}[S_{i,k}] = \alpha_k \mu_i.$$

Unter den Modellannahmen 4.2 gilt

$$\mathbb{E}[S_{i,n}] = \mu_i$$

und daher

$$\alpha_k = \frac{\mathbb{E}[S_{i,k}]}{\mathbb{E}[S_{i,n}]}.$$
 (4.1)

Man beachte, dass die Modellannahmen 4.1 stärker als die Modellannahmen 4.2 sind. Oft wird das Bornhuetter-Ferguson-Verfahren nur mithilfe der Modellannahmen 4.2 definiert. Dabei ist es jedoch schwierig, die Bornhuetter-Ferguson-Prognose zu begründen. Es gilt

$$\mathbb{E}[S_{i,n} | \mathcal{D}_n] = \mathbb{E}[S_{i,n} | S_{i,0}, \dots, S_{i,n-i}] = S_{i,n-i} + \mathbb{E}[S_{i,n} - S_{i,n-i} | S_{i,0}, \dots, S_{i,n-i}].$$

Ohne weitere Annahmen, so wie es bei den Modellannahmen 4.2 der Fall ist, ist unklar, was mit dem letzten Term geschehen soll. Unter den Modellannahmen 4.1 gilt jedoch

$$S_{i,n-i} + \mathbb{E}[S_{i,n} - S_{i,n-i} | S_{i,0}, \dots, S_{i,n-i}] = S_{i,n-i} + (\alpha_n - \alpha_{n-i})\mu_i = S_{i,n-i} + (1 - \alpha_{n-i})\mu_i.$$

Im Bornhuetter-Ferguson-Modell ist der Schätzer für $\mathbb{E}[S_{i,n} | \mathcal{D}_n]$ für $1 \leq i \leq n$ gegeben durch

$$\widehat{S}_{i,n}^{\text{BF}} = \widehat{\mathbb{E}}[S_{i,n} | \mathcal{D}_n] = S_{i,n-i} + (1 - \widehat{\alpha}_{n-i})\widehat{\mu}_i,$$
 (4.2)

wobei $\widehat{\alpha}_{n-i}$ ein geeigneter Schätzer für α_{n-i} und $\widehat{\mu}_i$ ein A-priori-Schätzer für den erwarteten Endschaten $\mathbb{E}[S_{i,n}]$ ist. $\widehat{S}_{i,n}^{\text{BF}}$ kann als A-posteriori-Schätzer für $\mu_i = \mathbb{E}[S_{i,n}]$ gesehen werden. Er stellt also eine Korrektur des A-priori-Schätzers $\widehat{\mu}_i$ dar, wobei die Korrektur durch Verwendung der aktuellen Schadenstände und der Schätzer des Abwicklungsmusters erfolgt.

Für die Bornhuetter-Ferguson-Reserven für das Anfallsjahr i ergibt sich

$$\widehat{R}_i^{\text{BF}} = \widehat{S}_{i,n}^{\text{BF}} - S_{i,n-i} = (1 - \widehat{\alpha}_{n-i})\widehat{\mu}_i.$$
 (4.3)

Die Bornhuetter-Ferguson-Gesamtreserve ist

$$\widehat{R}^{\text{BF}} = \sum_{i=1}^n \widehat{R}_i^{\text{BF}}.$$

4.2 Vergleich Chain-Ladder- und Bornhuetter-Ferguson-Prognose

Es stellt sich nun die Frage, wie α_k und μ_i geschätzt werden sollen. Dazu vergleichen wir die Chain-Ladder- und die Bornhuetter-Ferguson-Prognose. Aus (3.2) folgt

$$\mathbb{E}[S_{i,k}] = \mathbb{E}[S_{i,0}] \prod_{j=1}^k f_j,$$

$$\mathbb{E}[S_{i,n}] = \mathbb{E}[S_{i,0}] \prod_{j=1}^n f_j.$$

Daraus folgt

$$\mathbb{E}[S_{i,k}] = \prod_{j=k+1}^n f_j^{-1} \mathbb{E}[S_{i,n}].$$

Vergleicht man das mit den Modellannahmen 4.2, so sieht man, dass $\prod_{j=k+1}^n f_j^{-1}$ an der Stelle von α_k steht. Deshalb werden $\prod_{j=k+1}^n f_j^{-1}$ und α_k oft so verwendet, als gelte $\prod_{j=k+1}^n f_j^{-1} = \alpha_k$. Das kann aber nur mit den Modellannahmen 4.2 erfolgen, denn die Modellannahmen 4.1 folgen nicht aus den Chain-Ladder-Modellannahmen und umgekehrt. Sprich, wenn man die Chain-Ladder-Faktoren f_k kennt, kann man das Abwicklungsmuster $(\alpha_k)_k$ daraus schätzen und umgekehrt. Aufgrund von (3.3) und (4.1) gilt

$$\alpha_k = \prod_{l=k+1}^n \frac{1}{f_l}.$$

Deshalb kann die Bornhuetter-Ferguson-Prognose (4.2) folgendermaßen dargestellt werden:

$$\widehat{S}_{i,n}^{\text{BF}} = S_{i,n-i} + \left(1 - \widehat{\prod_{l=n-i+1}^n \frac{1}{f_l}}\right) \widehat{\mu}_i$$

Weiters betrachten wir die Chain-Ladder-Prognose:

$$\begin{aligned} \widehat{S}_{i,n}^{\text{CL}} &= S_{i,n-i} \prod_{l=n-i+1}^n \widehat{f}_l = S_{i,n-i} + S_{i,n-i} \left(\prod_{l=n-i+1}^n \widehat{f}_l - 1 \right) \\ &= S_{i,n-i} + \frac{\widehat{S}_{i,n}^{\text{CL}}}{\prod_{l=n-i+1}^n \widehat{f}_l} \left(\prod_{l=n-i+1}^n \widehat{f}_l - 1 \right) \\ &= S_{i,n-i} + \left(1 - \prod_{l=n-i+1}^n \frac{1}{\widehat{f}_l}\right) \widehat{S}_{i,n}^{\text{CL}} \end{aligned}$$

Wenn man also das Abwicklungsmuster beider Verfahren als gleich annimmt, besteht der Unterschied darin, dass man beim Bornhuetter-Ferguson-Verfahren an den geschätzten Endstand $\widehat{\mu}_i$ glaubt. Dieser wird jedoch beim Chain-Ladder-Verfahren durch $\widehat{S}_{i,n}^{\text{CL}}$, welcher nur von Beobachtungen abhängt, ersetzt.

Für μ_i kann

$$\widehat{\mu}_i = \pi_i \widehat{\kappa}_i$$

als Schätzer verwendet werden, wobei π_i das Prämienvolumen im Anfallsjahr i und $\widehat{\kappa}_i$ ein Schätzer der Endschaadenquote $\kappa_i = \mathbb{E}[S_{i,n}/\pi_i]$ ist. Dieser Wert sollte geschätzt werden, ohne Beobachtungen miteinzubeziehen. Wenn strikt die Bornhuetter-Ferguson-Methode angewandt wird, sollte auch der Wert α_k unabhängig von den Beobachtungen geschätzt werden. Aber in vielen Fällen wird trotzdem folgender Schätzer verwendet:

$$\widehat{\alpha}_k^{\text{CL}} = \widehat{\alpha}_k = \left(\widehat{\prod_{l=k+1}^n \frac{1}{f_l}} \right) = \prod_{l=k+1}^n \frac{1}{\widehat{f}_l} \quad (4.4)$$

In diesem Fall unterscheiden sich die Bornhuetter-Ferguson- und die Chain-Ladder-Methode nur in der Wahl der Prognose für den Endschaden $S_{i,n}$. Beim Bornhuetter-Ferguson-Verfahren wird $\hat{\mu}_i$, beim Chain-Ladder-Verfahren $\hat{S}_{i,n}^{\text{CL}}$ verwendet. Es gilt

$$\hat{S}_{i,n}^{\text{BF}} = S_{i,n-i} + (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{\mu}_i \quad (4.5)$$

$$\hat{S}_{i,n}^{\text{CL}} = S_{i,n-i} + (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{S}_{i,n}^{\text{CL}}. \quad (4.6)$$

4.3 Numerisches Beispiel

Es wird nun für die Daten aus Tabelle 2.1 das Bornhuetter-Ferguson-Verfahren gerechnet. Dabei werden die Schätzer $\hat{\mu}_i$ für die erwarteten Endschäden $\mathbb{E}[S_{i,n}]$ aus Tabelle 4.1 verwendet. Nachdem kein Abwicklungsmuster $(\hat{\alpha}_k)_k$ vorgegeben wird, wird dieses mithilfe von (4.4) berechnet. Die Ergebnisse sind in Tabelle 4.2 dargestellt.

Tabelle 4.1: Schätzungen für Endschäden

i	$\hat{\mu}_i$
0	181.630
1	220.162
2	230.990
3	278.794
4	272.417
5	285.277
6	295.331
7	277.801
8	248.329

Kapitel 5

Poisson-Modell

Dieses Kapitel orientiert sich an [8] und [6].

Das Poisson-Modell wird hauptsächlich angewendet, um die Schadenanzahl zu modellieren. Im Gegensatz zum verteilungsfreien Chain-Ladder-Modell werden hier Annahmen an die Verteilung der Zufallsvariablen gemacht. Wir werden sehen, dass der Maximum-Likelihood-Schätzer in diesem Modell zur selben Reserve wie im Chain-Ladder-Modell führt.

Modellannahmen 5.1:

- Die Familie $(Z_{i,k})_{i,k}$ der Zuwächse ist unabhängig.
- Es gibt Parameter $\mu_0, \dots, \mu_n \in (0, \infty)$ und $\theta_0, \dots, \theta_n \in (0, 1)$ mit $\sum_{i=0}^n \theta_i = 1$, sodass für alle $i, k \in \{0, \dots, n\}$ $Z_{i,k}$ Poisson-verteilt mit Parameter $\mu_i \theta_k$ ist, d.h.

$$\mathbb{P}[Z_{i,k} = x] = \frac{(\mu_i \theta_k)^x}{x!} \exp(-\mu_i \theta_k) \quad \forall x \in \mathbb{N}.$$

Man beachte, dass aus der Poisson-Annahme unmittelbar folgt, dass die Zuwächse nichtnegativ sind. Aus den Annahmen ergibt sich

$$\mathbb{E}[Z_{i,k}] = \mu_i \theta_k \tag{5.1}$$

und durch Summation über $k = 0, \dots, n$ erhält man

$$\mathbb{E}[S_{i,n}] = \mu_i. \tag{5.2}$$

Nachdem die Summe unabhängiger Poisson-verteilter Zufallsvariablen wieder Poisson-verteilt ist, ist $S_{i,n}$, der gesamte Schadenaufwand des Abwicklungsjahres i , Poisson-verteilt mit Parameter μ_i . μ_i steht also für den erwarteten Gesamtschaden des Abwicklungsjahres i . Wegen (5.1) und (5.2) gilt

$$\mathbb{E}[Z_{i,k}] = \theta_k \mathbb{E}[S_{i,n}].$$

$(\theta_k)_k$ definiert also das erwartete Abwicklungsmuster über die verschiedenen Abwicklungsjahre. Außerdem ist folgender Ausdruck unabhängig von i :

$$\frac{\mathbb{E}[Z_{i,k}]}{\mathbb{E}[Z_{i,0}]} = \frac{\theta_k}{\theta_0}$$

Lemma 5.2: Das Poisson-Modell erfüllt die Modellannahmen 4.1.

Beweis: Aus der Unabhängigkeit der $Z_{i,k}$ folgt, dass die kumulativen Schadenhöhen $S_{i,k}$ für verschiedene Anfallsjahre unabhängig sind.

Weiters gilt

$$\mathbb{E}[S_{i,0}] = \mathbb{E}[Z_{i,0}] = \mu_i \alpha_0 \quad \text{mit} \quad \alpha_0 = \theta_0$$

und

$$\begin{aligned} \mathbb{E}[S_{i,k+j} | S_{i,0}, \dots, S_{i,k}] &= \mathbb{E}[S_{i,k} + Z_{i,k+1} + \dots + Z_{i,k+j} | S_{i,0}, \dots, S_{i,k}] = \\ &= S_{i,k} + \sum_{l=1}^j \mathbb{E}[Z_{i,k+l} | Z_{i,0}, \dots, Z_{i,k}] = S_{i,k} + \sum_{l=1}^j \mathbb{E}[Z_{i,k+l}] = \\ &= S_{i,k} + \mu_i \sum_{l=1}^j \theta_{k+l} = S_{i,k} + \mu_i (\beta_{k+j} - \beta_k) \quad \text{mit} \quad \beta_k = \sum_{l=0}^k \theta_l. \end{aligned}$$

□

Das bedeutet, dass das Poisson-Modell die Annahmen an das Bornhuetter-Ferguson-Modell erfüllt. Man könnte also die Bornhuetter-Ferguson-Prognose verwenden, um die Reserve im Poisson-Modell zu prognostizieren. Wir werden hier aber anders vorgehen und zuerst die Parameter $(\mu_i)_i$ und $(\theta_k)_k$ schätzen. Dafür verwenden wir die Maximum-Likelihood-Schätzer. Die verfügbare Information wird durch $\mathcal{D}_n = \sigma \{Z_{i,k}; 0 \leq i, k \leq n, i+k \leq n\}$ dargestellt. Aufgrund der Unabhängigkeit der Zuwächse ist die Likelihood-Funktion L gegeben durch

$$L(\mu_0, \dots, \mu_n, \theta_0, \dots, \theta_n | \mathcal{D}_n) = \prod_{i=0}^n \prod_{k=0}^{n-i} \left(\frac{(\mu_i \theta_k)^{Z_{i,k}}}{Z_{i,k}!} \exp(-\mu_i \theta_k) \right).$$

Durch Logarithmieren erhält man

$$\begin{aligned} \ln L(\mu_0, \dots, \mu_n, \theta_0, \dots, \theta_n | \mathcal{D}_n) &= \sum_{i=0}^n \sum_{k=0}^{n-i} (Z_{i,k} \ln(\mu_i \theta_k) - \ln(Z_{i,k}!) - \mu_i \theta_k) = \\ &= \sum_{k=0}^n \sum_{i=0}^{n-k} (Z_{i,k} \ln(\mu_i \theta_k) - \ln(Z_{i,k}!) - \mu_i \theta_k). \end{aligned}$$

Durch Nullsetzen der partiellen Ableitungen kommt man auf

$$\begin{aligned} \sum_{k=0}^{n-i} \left(Z_{i,k} \frac{\hat{\theta}_k}{\hat{\mu}_i \hat{\theta}_k} - \hat{\theta}_k \right) &= 0 \quad \forall i \in \{0, \dots, n\}, \\ \sum_{i=0}^{n-k} \left(Z_{i,k} \frac{\hat{\mu}_i}{\hat{\mu}_i \hat{\theta}_k} - \hat{\mu}_i \right) &= 0 \quad \forall k \in \{0, \dots, n\}, \end{aligned}$$

unter der Nebenbedingung $\sum_{k=0}^n \hat{\theta}_k = 1$, was äquivalent ist zu

$$\begin{aligned} \sum_{k=0}^{n-i} \hat{\mu}_i \hat{\theta}_k &= \sum_{k=0}^{n-i} Z_{i,k} = S_{i,n-i} \quad \forall i \in \{0, \dots, n\}, \\ \sum_{i=0}^{n-k} \hat{\mu}_i \hat{\theta}_k &= \sum_{i=0}^{n-k} Z_{i,k} \quad \forall k \in \{0, \dots, n\} \end{aligned} \tag{5.3}$$

unter der Nebenbedingung $\sum_{k=0}^n \hat{\theta}_k = 1$.

Falls das System (5.3) eine eindeutige Lösung hat, erhält man daraus die Maximum-Likelihood-Schätzer $\hat{\mu}_i$ und $\hat{\theta}_k$ für μ_i bzw. θ_k .

Im Poisson-Modell werden $\mathbb{E}[Z_{i,k}]$ bzw. $\mathbb{E}[S_{i,n}|\mathcal{D}_n]$ folgendermaßen geschätzt:

$$\begin{aligned}\hat{Z}_{i,k}^{\text{Poi}} &= \hat{\mathbb{E}}[Z_{i,k}] = \hat{\mu}_i \hat{\theta}_k && \forall i+k > n \\ \hat{S}_{i,n}^{\text{Poi}} &= \hat{\mathbb{E}}[S_{i,n}|\mathcal{D}_n] = S_{i,n-i} + \sum_{k=n-i+1}^n \hat{Z}_{i,k}^{\text{Poi}} && \forall i > 0\end{aligned}\quad (5.4)$$

Es gilt

$$\hat{S}_{i,n}^{\text{Poi}} = S_{i,n-i} + \sum_{k=n-i+1}^n \hat{\mu}_i \hat{\theta}_k = S_{i,n-i} + \left(1 - \sum_{k=0}^{n-i} \hat{\theta}_k\right) \hat{\mu}_i, \quad (5.5)$$

d.h. die Poisson-Prognose hat die selbe Form wie die Bornhuetter-Ferguson-Prognose (4.2), mit dem Unterschied, dass hier Maximum-Likelihood-Schätzer für μ_i und θ_k verwendet werden, die von den Daten \mathcal{D}_n abhängen.

Es wird nun gezeigt, dass das Poisson-Modell auf die selbe Reserve wie das verteilungsfreie Chain-Ladder-Modell führt. Dazu nehmen wir an, dass das System (5.3) eine positive Lösung hat. Dazu wird zuerst folgendes Hilfslemma bewiesen:

Lemma 5.3: Unter den Modellannahmen 5.1 gilt auf \mathcal{D}_n

$$\sum_{i=0}^{n-k} S_{i,k} = \sum_{i=0}^{n-k} \hat{\mu}_i \sum_{j=0}^k \hat{\theta}_j$$

Beweis: Der Beweis erfolgt durch Induktion.

- Induktionsanfang: $k = n$
Aus (5.3) erhält man für $i = 0$

$$S_{i,n} = \sum_{k=0}^n Z_{0,k} = \hat{\mu}_0 \sum_{k=0}^n \hat{\theta}_k.$$

- Induktionsschritt: $l \rightarrow l-1$

Induktionsvoraussetzung: Für $k = l$ gilt

$$\sum_{i=0}^{n-l} S_{i,l} = \sum_{i=0}^{n-l} \hat{\mu}_i \sum_{j=0}^l \hat{\theta}_j.$$

Es gilt

$$\begin{aligned}\sum_{i=0}^{n-(l-1)} S_{i,l-1} &= \sum_{i=0}^{n-l} S_{i,l-1} + S_{n-l+1,l-1} = \sum_{i=0}^{n-l} (S_{i,l} - Z_{i,l}) + S_{n-l+1,l-1} = \\ &= \sum_{i=0}^{n-l} S_{i,l} - \sum_{i=0}^{n-l} Z_{i,l} + \sum_{j=0}^{l-1} Z_{n-l+1,j}.\end{aligned}$$

Aus der Induktionsvoraussetzung und (5.3) folgt

$$\begin{aligned} & \sum_{i=0}^{n-l} S_{i,l} - \sum_{i=0}^{n-l} Z_{i,l} + \sum_{j=0}^{l-1} Z_{n-l+1,j} = \sum_{i=0}^{n-l} \hat{\mu}_i \sum_{j=0}^l \hat{\theta}_j - \sum_{i=0}^{n-l} \hat{\mu}_i \hat{\theta}_l + \sum_{j=0}^{l-1} \hat{\mu}_{n-l+1} \hat{\theta}_j = \\ & = \sum_{i=0}^{n-l} \hat{\mu}_i \sum_{j=0}^{l-1} \hat{\theta}_j + \hat{\mu}_{n-l+1} \sum_{j=0}^{l-1} \hat{\theta}_j = \sum_{i=0}^{n-(l-1)} \hat{\mu}_i \sum_{j=0}^{l-1} \hat{\theta}_j. \end{aligned}$$

□

Somit können wir nun folgendes Lemma beweisen:

Lemma 5.4: Die Chain-Ladder-Prognose (3.6) und die Prognose (5.4) mit den Maximum-Likelihood-Schätzern im Poisson-Modell stimmen überein, d.h. $\hat{S}_{i,n}^{\text{CL}} = \hat{S}_{i,n}^{\text{Poi}}$.

Beweis: Für die Poisson-Prognose gilt unter mehrmaliger Berücksichtigung von (5.3)

$$\begin{aligned} \hat{S}_{i,n}^{\text{Poi}} &= \widehat{\mathbb{E}}[S_{i,n} | \mathcal{D}_n] = S_{i,n-i} + \sum_{k=n-i+1}^n \hat{\mu}_i \hat{\theta}_k = \sum_{k=0}^{n-i} \hat{\mu}_i \hat{\theta}_k + \sum_{k=n-i+1}^n \hat{\mu}_i \hat{\theta}_k = \hat{\mu}_i \sum_{k=0}^n \hat{\theta}_k = \\ &= \frac{S_{i,n-i}}{\sum_{k=0}^{n-i} \hat{\theta}_k} \sum_{k=0}^n \hat{\theta}_k = S_{i,n-i} \frac{\sum_{k=0}^{n-i+1} \hat{\theta}_k}{\sum_{k=0}^{n-i} \hat{\theta}_k} \dots \frac{\sum_{k=0}^n \hat{\theta}_k}{\sum_{k=0}^{n-1} \hat{\theta}_k}. \end{aligned} \quad (5.6)$$

Aufgrund von Lemma 5.3 gilt auf \mathcal{D}_n

$$\sum_{i=0}^{n-k} S_{i,k} = \sum_{i=0}^{n-k} \hat{\mu}_i \sum_{j=0}^k \hat{\theta}_j. \quad (5.7)$$

Daraus folgt unter Berücksichtigung von (5.3)

$$\sum_{i=0}^{n-k} S_{i,k-1} = \sum_{i=0}^{n-k} (S_{i,k} - Z_{i,k}) = \sum_{i=0}^{n-k} \hat{\mu}_i \sum_{j=0}^k \hat{\theta}_j - \sum_{i=0}^{n-k} \hat{\mu}_i \hat{\theta}_k = \sum_{i=0}^{n-k} \hat{\mu}_i \sum_{j=0}^{k-1} \hat{\theta}_j. \quad (5.8)$$

(5.7) und (5.8) implizieren

$$\frac{\sum_{j=0}^k \hat{\theta}_j}{\sum_{j=0}^{k-1} \hat{\theta}_j} = \frac{\sum_{i=0}^{n-k} S_{i,k}}{\sum_{i=0}^{n-k} S_{i,k-1}} = \hat{f}_k$$

Aus (5.6) folgt also

$$\hat{S}_{i,n}^{\text{Poi}} = S_{i,n-i} \hat{f}_{n-i+1} \cdots \hat{f}_n = \hat{S}_{i,n}^{\text{CL}}. \quad (5.9)$$

□

Lemma 5.5: Unter den Modellannahmen 5.1 gilt

$$\sum_{k=0}^j \hat{\theta}_k = \hat{\alpha}_j^{\text{CL}} = \prod_{k=j+1}^n \frac{1}{\hat{f}_k} \quad \forall 0 \leq j \leq n,$$

wobei ein leeres Produkt definitionsgemäß gleich 1 ist.

Beweis: Aus (5.6) und (5.9) folgt

$$\widehat{S}_{i,n}^{\text{Poi}} = S_{i,n-i} \frac{\sum_{k=0}^n \widehat{\theta}_k}{\sum_{k=0}^{n-i} \widehat{\theta}_k} = \widehat{S}_{i,n}^{\text{CL}} = S_{i,n-i} \widehat{f}_{n-i+1} \cdots \widehat{f}_n.$$

Nachdem $\sum_{k=0}^n \widehat{\theta}_k = 1$, erhält man

$$\sum_{k=0}^{n-i} \widehat{\theta}_k = \prod_{k=n-i+1}^n \frac{1}{\widehat{f}_k} = \widehat{\alpha}_{n-i}^{\text{CL}}.$$

□

Bemerkungen:

- Lemma 5.5 besagt, dass die verteilungsfreie Chain-Ladder-Methode und das Poisson-Modell auf das selbe Abwicklungsmuster $(\widehat{\alpha}_k^{\text{CL}})_k$ führen und deshalb auf die selbe Bornhuetter-Ferguson-Reserve, falls dieses Abwicklungsmuster als Schätzung für $(\alpha_k)_k$ herangezogen wird. Falls also das Abwicklungsmuster $(\widehat{\alpha}_k^{\text{CL}})_k$ für das Bornhuetter-Ferguson-Verfahren verwendet wird, unterscheiden sich das Poisson-Modell und das Bornhuetter-Ferguson-Verfahren nur in der Wahl des erwarteten Gesamtschadens μ_i , denn es folgt aus (5.5)

$$\widehat{S}_{i,n}^{\text{Poi}} = S_{i,n-i} + (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i,$$

wobei $\widehat{\mu}_i$ der Maximum-Likelihood-Schätzer aus (5.3) ist.

- Die Lösung von (5.3) erhält man in diesem Fall sehr einfach mithilfe der Chain-Ladder-Faktoren \widehat{f}_k , und zwar

$$\widehat{\theta}_k = \widehat{\alpha}_k^{\text{CL}} - \widehat{\alpha}_{k-1}^{\text{CL}} = \prod_{j=k+1}^n \frac{1}{\widehat{f}_j} - \prod_{j=k}^n \frac{1}{\widehat{f}_j} = \prod_{j=k+1}^n \frac{1}{\widehat{f}_j} \left(1 - \frac{1}{\widehat{f}_k}\right)$$

und

$$\widehat{\mu}_i = \frac{S_{i,n-i}}{\sum_{k=0}^{n-i} \widehat{\theta}_k} = \frac{S_{i,n-i}}{\widehat{\alpha}_{n-i}^{\text{CL}}} = \widehat{S}_{i,n}^{\text{CL}},$$

d.h. es ist keine numerische Approximation zur Lösung des Gleichungssystems notwendig.

- Man beachte, dass das verteilungsfreie Chain-Ladder-Modell auf allgemeinere Datensätze angewendet werden kann. Das Poisson-Modell kann nur gelten, wenn die Inkremente $Z_{i,k}$ positiv sind, nachdem die Poisson-Verteilung nur natürliche Zahlen als Werte annimmt. Im verteilungsfreien Chain-Ladder-Modell dürfen die Inkremente auch negativ sein, solange die kumulativen Schadenstände positiv sind.

Kapitel 6

Verallgemeinerte lineare Modelle

Dieses Kapitel lehnt sich an [6] an.

6.1 Maximum-Likelihood-Schätzer

Lemma 6.1: Seien X_1, \dots, X_n i.i.d. Zufallsvariablen mit Dichte $f_\theta(\cdot)$ und $\hat{\theta}_n^{\text{MLE}}$ der Maximum-Likelihood-Schätzer für den unbekannt Parameter der Verteilung θ . Unter milden Voraussetzungen gilt folgende Konvergenz in der Verteilung:

$$\sqrt{n} \left(\hat{\theta}_n^{\text{MLE}} - \theta \right) \xrightarrow{(d)} \mathcal{N} \left(0, \frac{1}{H(\theta)} \right), \quad \text{für } n \rightarrow \infty$$

mit

$$H(\theta) = -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} \ln f_\theta(X) \right] = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_\theta(X) \right)^2 \right] \quad (6.1)$$

Beweis: siehe Theorem 2.3. in Kapitel 6 in [9], wobei sich die zweite Gleichheit in (6.1) folgendermaßen ergibt:

$$\begin{aligned} -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} \ln f_\theta(X) \right] &= -\mathbb{E} \left[\frac{\partial}{\partial \theta} \left(\frac{1}{f_\theta(X)} \frac{\partial}{\partial \theta} f_\theta(X) \right) \right] = \\ &= -\mathbb{E} \left[-\frac{1}{f_\theta(X)^2} \left(\frac{\partial}{\partial \theta} f_\theta(X) \right)^2 + \frac{1}{f_\theta(X)} \frac{\partial^2}{\partial \theta^2} f_\theta(X) \right] = \\ &= \mathbb{E} \left[\left(\frac{1}{f_\theta(X)} \frac{\partial}{\partial \theta} f_\theta(X) \right)^2 \right] - \int_x \frac{1}{f_\theta(x)} \left(\frac{\partial^2}{\partial \theta^2} f_\theta(x) \right) f_\theta(x) dx = \\ &= \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_\theta(X) \right)^2 \right] - \int_x \frac{\partial^2}{\partial \theta^2} f_\theta(x) dx = \\ &= \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_\theta(X) \right)^2 \right] - \frac{\partial^2}{\partial \theta^2} \int_x f_\theta(x) dx = \\ &= \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_\theta(X) \right)^2 \right] - \frac{\partial^2}{\partial \theta^2} 1 = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_\theta(X) \right)^2 \right] \end{aligned}$$

□

Bemerkungen:

- $H(\theta)$ wird **Fisher-Information** genannt.
- Dieses Lemma besagt, dass Maximum-Likelihood-Schätzer asymptotisch unverzerrt mit einer asymptotischen Varianz, die durch die Inverse der Fisher-Information gegeben ist, sind. Wir werden also die Fisher-Information benutzen, um den Schätzfehler zu approximieren. Man sei sich aber bewusst, dass dieses Ergebnis nur asymptotisch gilt. Die Frage, die man sich also zu stellen hat, ist, ob man bereits im asymptotischen Bereich ist.
- Die milden Voraussetzungen sind in [9] in Kapitel 6.2. beschrieben. Die Verteilungen, die im Folgenden verwendet werden, erfüllen diese Bedingungen.
- Es gilt auch das multivariate Analogon. Seien X_1, \dots, X_n i.i.d. Zufallsvariablen mit Dichte $f_\theta(\cdot)$ und Parameter $\theta = (\theta_1, \dots, \theta_p)$ und die Fisher-Informationsmatrix folgendermaßen definiert:

$$H(\theta)_{r,s} = -\mathbb{E} \left[\frac{\partial^2}{\partial \theta_r \partial \theta_s} \ln f_\theta(X) \right] = \mathbb{E} \left[\frac{\partial}{\partial \theta_r} \ln f_\theta(X) \frac{\partial}{\partial \theta_s} \ln f_\theta(X) \right]$$

Dann ist $\sqrt{n} \left(\hat{\theta}_n^{\text{MLE}} - \theta \right)$ asymptotisch multivariat normalverteilt mit Mittelwertvektor Null und Kovarianzmatrix $H(\theta)^{-1}$.

6.2 Modellrahmen

In diesem Kapitel soll der Modellrahmen für die verallgemeinerten linearen Modelle festgelegt werden:

- 1. Zufallskomponente:** Für die Zufallsvariablen $Z_{i,k}$, $i \in \{0, \dots, n\}$ und $k \in \{0, \dots, n\}$, gilt

$$\begin{aligned} \mathbb{E}[Z_{i,k}] &= z_{i,k} \\ \text{Var}(Z_{i,k}) &= \frac{\phi_{i,k}}{w_{i,k}} V(z_{i,k}), \end{aligned} \tag{6.2}$$

wobei $\phi_{i,k} > 0$ ein Streuungsparameter, $V(\cdot)$ eine passende Varianzfunktion und $w_{i,k}$ ein bekanntes Gewicht oder Volumsmaß ist.

Zur Zeit hat das Modell noch $(n+1)^2$ Parameter. Es muss also noch die Dimension der unbekannt Parameter reduziert werden, nachdem zu wenige Daten vorhanden sind, um diese zu schätzen.

- 2. Systematische Komponente:** Es wird angenommen, dass $z_{i,k}$ durch eine kleinere Anzahl an unbekannt Parameter $\mathbf{b} = (b_1, \dots, b_p)'$ und einen passenden deterministischen p -dimensionalen Zeilenvektor $\Gamma_{i,k}$, die $\boldsymbol{\eta} = (\eta_{i,k})_{i,k}$ ergeben, bestimmt werden kann:

$$\eta_{i,k} = \Gamma_{i,k} \mathbf{b} \tag{6.3}$$

- 3. Überleitungsfunktion:** Die Verbindung zwischen der Zufallskomponente und der systematischen Komponente ist durch die Überleitungsfunktion $h(\cdot)$ und ihre Inverse $g = h^{-1}$ gegeben:

$$z_{i,k} = h(\eta_{i,k}) \tag{6.4}$$

$$g(z_{i,k}) = \eta_{i,k} = \Gamma_{i,k} \mathbf{b} \tag{6.5}$$

6.3 Exponential Dispersion Family

Wir nehmen an, dass die Verteilung der Inkremente $Z_{i,k}$ zur Exponential Dispersion Family, kurz EDF, gehört, d.h. die Dichte (6.6) erfüllt.

Modellannahmen 6.2:

- Die Inkremente $Z_{i,k}$ verschiedener Anfallsjahre i und/oder verschiedener Abwicklungsjahre k sind unabhängig.
- Die Inkremente $Z_{i,k}$ haben eine Dichte (bezüglich des Zähl- oder Lebesgue-Maßes) der folgenden Form:

$$f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) = a\left(x, \frac{\phi_{i,k}}{w_{i,k}}\right) \exp\left\{\frac{x\gamma_{i,k} - d(\gamma_{i,k})}{\phi_{i,k}/w_{i,k}}\right\}, \quad (6.6)$$

wobei $d(\cdot)$ eine reellwertige zwei mal differenzierbare Funktion des Parameters $\gamma_{i,k}$ ist, so dass $(d')^{-1}(\cdot)$ existiert, $\phi_{i,k} > 0$ ein Streuungsparameter, $w_{i,k} > 0$ ein bekanntes Gewicht und $a(\cdot, \cdot)$ eine passende reellwertige Normalisierungsfunktion ist.

Bemerkungen:

- Jede Verteilung, die zur EDF gehört, ist durch die spezielle Form der Funktionen $a(\cdot, \cdot)$ und $d(\cdot)$ charakterisiert. Beispiele für solche Verteilungen sind die Poisson-, Overdispersed Poisson- (siehe Kapitel 6.5), Bernoulli-, Binomial-, Negativ-Binomial-, Gamma-, Normal- und Inverse Normal-Verteilung.

Beispiele:

- $Z_{i,k} \sim \text{Poi}(\mu_i \theta_k)$, d.h.

$$f(x; \mu_i, \theta_k) = \mathbb{P}[Z_{i,k} = x] = \frac{(\mu_i \theta_k)^x}{x!} e^{-\mu_i \theta_k} = \frac{1}{x!} \exp\{x \ln(\mu_i \theta_k) - (\mu_i \theta_k)\}.$$

Es gilt also

$$\begin{aligned} \frac{\phi_{i,k}}{w_{i,k}} &= 1, \\ \gamma_{i,k} &= \ln(\mu_i \theta_k), \\ d(\cdot) &= \exp(\cdot), \\ a(x) &= \frac{1}{x!}. \end{aligned}$$

- $Z_{i,k} \sim \text{N}(\mu_{i,k}, \sigma_{i,k})$, d.h.

$$\begin{aligned} f(x; \mu_{i,k}, \sigma_{i,k}) &= \frac{1}{\sqrt{2\pi\sigma_{i,k}^2}} \exp\left\{-\frac{(x - \mu_{i,k})^2}{2\sigma_{i,k}^2}\right\} = \\ &= \frac{1}{\sqrt{2\pi\sigma_{i,k}^2}} \exp\left\{-\frac{x^2}{2\sigma_{i,k}^2}\right\} \exp\left\{\frac{x\mu_{i,k}}{\sigma_{i,k}^2} - \frac{\mu_{i,k}^2}{2\sigma_{i,k}^2}\right\}. \end{aligned}$$

Es gilt also

$$\begin{aligned}\frac{\phi_{i,k}}{w_{i,k}} &= \sigma_{i,k}^2, \\ \gamma_{i,k} &= \mu_{i,k}, \\ d(x) &= \frac{x^2}{2}, \\ a(x, \sigma_{i,k}^2) &= \frac{1}{\sqrt{2\pi\sigma_{i,k}^2}} \exp\left\{-\frac{x^2}{2\sigma_{i,k}^2}\right\}.\end{aligned}$$

- Ein verallgemeinertes lineares Modell wird durch folgende drei Komponenten charakterisiert:
 - Typ der EDF-Verteilung für die Zufallsvariable $Z_{i,k}$
 - Überleitungsfunktion h oder die Inverse g
 - Zeilenvektor $\Gamma_{i,k}$
- Die Wahl einer Überleitungsfunktion h hängt von der speziellen EDF ab, wobei für jede EDF kanonische Überleitungsfunktionen existieren:

$$\eta_{i,k} = g(z_{i,k}) = \gamma_{i,k}.$$

Nachdem $z_{i,k} = d'(\gamma_{i,k})$ (siehe (6.7) weiter unten), folgt daraus $g(d'(\gamma_{i,k})) = \gamma_{i,k}$ und folglich ist die kanonische Überleitungsfunktion gegeben durch $g = (d')^{-1}$. Es gilt dann $\gamma_{i,k} = \Gamma_{i,k}\mathbf{b}$.

Lemma 6.3: Unter den Modellannahmen 6.2 gilt

$$\mathbb{E}[Z_{i,k}] = z_{i,k} = d'(\gamma_{i,k}) \quad (6.7)$$

$$\text{Var}(Z_{i,k}) = \frac{\phi_{i,k}}{w_{i,k}} d''(\gamma_{i,k}), \quad (6.8)$$

wobei $d'(\gamma_{i,k})$ und $d''(\gamma_{i,k})$ die erste bzw. zweite Ableitung von d nach $\gamma_{i,k}$ sind, die aufgrund der Annahmen existieren.

Beweis: Sei

$$l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) := \ln f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}).$$

O.B.d.A. nehmen wir an, dass $f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k})$ eine Dichte bezüglich des Lebesgue-Maßes ist. Dann gilt

$$\begin{aligned}0 &= \frac{\partial}{\partial \gamma_{i,k}} 1 = \frac{\partial}{\partial \gamma_{i,k}} \int_x f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) dx = \int_x \frac{\partial}{\partial \gamma_{i,k}} f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) dx = \\ &= \int_x \frac{\partial l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k})}{\partial \gamma_{i,k}} f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) dx = \mathbb{E} \left[\frac{\partial l(Z_{i,k}; \gamma_{i,k}, \phi_{i,k}, w_{i,k})}{\partial \gamma_{i,k}} \right] = \\ &= \mathbb{E} \left[\frac{\partial}{\partial \gamma_{i,k}} \left(\ln a \left(Z_{i,k}, \frac{\phi_{i,k}}{w_{i,k}} \right) + \frac{Z_{i,k} \gamma_{i,k} - d(\gamma_{i,k})}{\phi_{i,k}/w_{i,k}} \right) \right] = \frac{w_{i,k}}{\phi_{i,k}} (\mathbb{E}[Z_{i,k}] - d'(\gamma_{i,k})),\end{aligned}$$

woraus sofort die Aussage für den Erwartungswert folgt.

Analog gilt unter Verwendung von $\mathbb{E}[Z_{i,k}] = d'(\gamma_{i,k})$

$$\begin{aligned}
0 &= \frac{\partial^2}{\partial \gamma_{i,k}^2} 1 = \frac{\partial^2}{\partial \gamma_{i,k}^2} \int_x f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) dx = \int_x \frac{\partial^2}{\partial \gamma_{i,k}^2} f(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) dx = \\
&= \mathbb{E} \left[\frac{\partial^2 l(Z_{i,k}; \gamma_{i,k}, \phi_{i,k}, w_{i,k})}{\partial \gamma_{i,k}^2} \right] + \mathbb{E} \left[\left(\frac{\partial l(Z_{i,k}; \gamma_{i,k}, \phi_{i,k}, w_{i,k})}{\partial \gamma_{i,k}} \right)^2 \right] = \\
&= -\frac{w_{i,k}}{\phi_{i,k}} d''(\gamma_{i,k}) + \left(\frac{w_{i,k}}{\phi_{i,k}} \right)^2 \mathbb{E} \left[(Z_{i,k} - d'(\gamma_{i,k}))^2 \right] = \\
&= \frac{w_{i,k}}{\phi_{i,k}} \left(-d''(\gamma_{i,k}) + \frac{w_{i,k}}{\phi_{i,k}} \text{Var}(Z_{i,k}) \right),
\end{aligned}$$

woraus die Aussage über die Varianz folgt.

Bei beiden Rechnungen wurde verwendet, dass die Differentiation und die Integration vertauscht werden dürfen, falls die Verteilung zur EDF gehört. Dazu sei auf [10] oder [11] verwiesen. \square

Wegen (6.7) gilt

$$\gamma_{i,k} = (d')^{-1}(z_{i,k}) \quad (6.9)$$

und wegen (6.2) und (6.8) definiert man

$$V(z_{i,k}) = d'' \left((d')^{-1}(z_{i,k}) \right). \quad (6.10)$$

6.4 Parameter-Schätzung für die EDF

In diesem Kapitel wird eine multiplikative Struktur angenommen:

$$z_{i,k} = \mu_i \theta_k, \quad (6.11)$$

wobei μ_i für das Exposure des Anfallsjahres i (z.B. erwartete Schadenlast, erwartete Anzahl der Schäden, gesamte Anzahl an Polizzen) und θ_k für ein erwartetes Schaden-/Reporting-/Cashflow-Muster über die unterschiedlichen Anfallsjahre steht. Somit hat sich die Anzahl der Parameter auf $2n + 2$ reduziert. Außerdem wählen wir

$$g(\cdot) = h^{-1}(\cdot) = \ln(\cdot).$$

Es ergibt sich also

$$\eta_{i,k} = \ln(z_{i,k}) = \ln(\mu_i) + \ln(\theta_k).$$

Gleichung (6.11) zeigt, dass μ_i und θ_k nur bis auf einen konstanten Faktor eindeutig sind, denn $\tilde{\mu}_i = c\mu_i$ und $\tilde{\theta}_k = \theta_k/c$ würden für jedes $c > 0$ den selben Schätzer für $z_{i,k}$ ergeben. In Kapitel 5 wird angenommen, dass $(\theta_k)_k$ ein Schadenabwicklungsmuster ist und die Summe deshalb 1 ist. In diesem Kapitel ist es einfacher, $\mu_0 = 1$ und deshalb $\ln(\mu_0) = 0$ zu setzen. Das führt auf

$$\mathbf{b} = (\ln(\mu_1), \dots, \ln(\mu_n), \ln(\theta_0), \dots, \ln(\theta_n))'$$

und für $k \in \{0, \dots, n\}$

$$\begin{aligned}
\Gamma_{0,k} &= (0, \dots, 0, 0, 0, \dots, 0, e_{n+k+1}, 0, \dots, 0) \\
\Gamma_{i,k} &= (0, \dots, 0, e_i, 0, \dots, 0, e_{n+k+1}, 0, \dots, 0), \quad i \in \{1, \dots, n\},
\end{aligned}$$

wobei die Einträge $e_i = 1$ und $e_{n+k+1} = 1$ an der i -ten und an der $(n+k+1)$ -sten Position sind und es gilt

$$\eta_{i,k} = \Gamma_{i,k} \mathbf{b} \quad (6.12)$$

Wir haben also die Anzahl der Parameter auf $2n+1$ reduziert.

Oft ist es nützlich, die Log-Likelihood-Funktion mithilfe des Mittelwertes auszudrücken. Wegen (6.9) gilt

$$l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) = l(x; (d^l)^{-1}(z_{i,k}), \phi_{i,k}, w_{i,k}) = \tilde{l}(x; z_{i,k}, \phi_{i,k}, w_{i,k}).$$

Je nachdem, ob $\gamma_{i,k}$ oder $z_{i,k}$ verwendet wird, sollte klar sein, um welche der beiden Funktionen es sich handelt und wir wollen die Tilde weglassen und identifizieren

$$l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) \Leftrightarrow l(x; z_{i,k}, \phi_{i,k}, w_{i,k}).$$

Wir verwenden nun die Maximum-Likelihood-Schätzung auf der Menge der Beobachtungen \mathcal{D}_n , um den unbekannt Parametervektor \mathbf{b} zu schätzen, d.h. wir maximieren folgende Funktion:

$$\begin{aligned} \ln L(\mathbf{b} | \mathcal{D}_n) &= \ln \prod_{i+k \leq n} f(Z_{i,k}; (d^l)^{-1}(z_{i,k}), \phi_{i,k}, w_{i,k}) = \sum_{i+k \leq n} l(Z_{i,k}; z_{i,k}, \phi_{i,k}, w_{i,k}) = \\ &= \sum_{i+k \leq n} l(Z_{i,k}; \mu_i \theta_k, \phi_{i,k}, w_{i,k}) \end{aligned} \quad (6.13)$$

Wir maximieren diese Funktion, indem die $2n+1$ partiellen Ableitungen bezüglich der unbekannt Parameter μ_i und θ_k gleich Null gesetzt werden. Dadurch erhalten wir Schätzungen $\widehat{\mu}_i$ und $\widehat{\theta}_k$ und deshalb

$$\widehat{\mathbf{b}} = \left(\widehat{\ln(\mu_1)}, \dots, \widehat{\ln(\mu_n)}, \widehat{\ln(\theta_0)}, \dots, \widehat{\ln(\theta_n)} \right)$$

mit $\widehat{\ln(\mu_i)} = \ln(\widehat{\mu}_i)$ und $\widehat{\ln(\theta_k)} = \ln(\widehat{\theta}_k)$. Dadurch erhält man einen Schätzer für den Erwartungswert $z_{i,k}$:

$$\widehat{z}_{i,k} = \exp\{\widehat{\eta}_{i,k}\} = \exp\left\{\widehat{\ln(\mu_i)} + \widehat{\ln(\theta_k)}\right\} = \widehat{\mu}_i \widehat{\theta}_k$$

Dadurch ergeben sich folgende Prognosen in diesem Modell:

$$\widehat{Z}_{i,k}^{\text{EDF}} = \widehat{z}_{i,k} = \widehat{\mu}_i \widehat{\theta}_k, \quad i+k > n \quad (6.14)$$

$$\widehat{S}_{i,n}^{\text{EDF}} = S_{i,n-i} + \sum_{k=n-i+1}^n \widehat{Z}_{i,k}^{\text{EDF}} \quad i \in \{1, \dots, n\} \quad (6.15)$$

Bemerkungen:

- Wir erhalten ein Gleichungssystem für $\widehat{\mu}_i$ und $\widehat{\theta}_k$, das wir lösen müssen. Für das Poisson-Modell ist es gegeben durch (5.3).

6.4.1 Fisher-Scoring-Algorithmus

Es wird nun ein Algorithmus erklärt, der für die Berechnung der Maximum-Likelihood-Schätzer für den Parametervektor \mathbf{b} verwendet werden kann.

Lemma 6.4: Unter den Modellannahmen 6.2 erfüllt die Log-Likelihood-Funktion

$$l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) = \ln a \left(x, \frac{\phi_{i,k}}{w_{i,k}} \right) + \frac{x\gamma_{i,k} - d(\gamma_{i,k})}{\phi_{i,k}/w_{i,k}}$$

für $j = 1, \dots, 2n + 1$ folgende Gleichung

$$\frac{\partial}{\partial b_j} l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) = W(z_{i,k}) (x - z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)},$$

wobei $\Gamma_{i,k}^{(j)}$ die j -te Koordinate des Zeilenvektors $\Gamma_{i,k}$ ist und

$$W(z_{i,k}) = \frac{w_{i,k}}{\phi_{i,k}} \frac{1}{V(z_{i,k})} \left(\frac{\partial \eta_{i,k}}{\partial z_{i,k}} \right)^{-2},$$

wobei $V(z_{i,k})$ die in (6.10) definierte Funktion ist.

Beweis: Aufgrund der Kettenregel gilt

$$\frac{\partial}{\partial b_j} l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) = \frac{\partial l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k})}{\partial \gamma_{i,k}} \frac{\partial \gamma_{i,k}}{\partial z_{i,k}} \frac{\partial z_{i,k}}{\partial \eta_{i,k}} \frac{\partial \eta_{i,k}}{\partial b_j}.$$

Wir berechnen nun jeden Term. Mithilfe von (6.7) folgt

$$\begin{aligned} \frac{\partial l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k})}{\partial \gamma_{i,k}} &= \frac{w_{i,k}}{\phi_{i,k}} (x - d'(\gamma_{i,k})) = \frac{w_{i,k}}{\phi_{i,k}} (x - z_{i,k}), \\ \frac{\partial z_{i,k}}{\partial \gamma_{i,k}} &= d''(\gamma_{i,k}) = V(z_{i,k}), \\ \frac{\partial \eta_{i,k}}{\partial b_j} &= \Gamma_{i,k}^{(j)}. \end{aligned}$$

Es folgt also

$$\frac{\partial}{\partial b_j} l(x; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) = \frac{w_{i,k}}{\phi_{i,k}} (x - z_{i,k}) \frac{1}{V(z_{i,k})} \frac{\partial z_{i,k}}{\partial \eta_{i,k}} \Gamma_{i,k}^{(j)} = W(z_{i,k}) (x - z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)}.$$

□

Nachdem der Maximum-Likelihood-Schätzer $\hat{\mathbf{b}}$ die Log-Likelihood-Funktion (6.13) maximiert, verschwindet die Ableitung nach b_j für $j = 1, \dots, 2n + 1$. Aus Lemma 6.4 folgt für $j = 1, \dots, 2n + 1$

$$0 = \frac{\partial}{\partial b_j} \sum_{i+k \leq n} l(Z_{i,k}; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) \Bigg|_{\hat{\mathbf{b}}} = \sum_{i+k \leq n} W(z_{i,k}) (Z_{i,k} - z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} \Bigg|_{\hat{\mathbf{b}}}. \quad (6.16)$$

Deshalb haben wir den Maximum-Likelihood-Schätzer $\hat{\mathbf{b}}$ für \mathbf{b} gefunden, wenn die sogenannte Score-Funktion

$$\begin{aligned} \mathbf{u} = \mathbf{u}(\mathbf{b}) &= \left(\frac{\partial}{\partial b_j} \sum_{i+k \leq n} l(Z_{i,k}; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) \right)_{j=1, \dots, 2n+1} = \\ &= \left(\sum_{i+k \leq n} W(z_{i,k}) (Z_{i,k} - z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} \right)_{j=1, \dots, 2n+1} \end{aligned} \quad (6.17)$$

gleich dem Nullvektor ist und die Hessematrix von $l(\cdot; \gamma_{i,k}, \phi_{i,k}, w_{i,k})$ auf \mathcal{D}_n

$$\mathbf{W} = \mathbf{W}(\mathbf{b}) = \left(\sum_{i+k \leq n} \frac{\partial^2}{\partial b_j \partial b_l} l(Z_{i,k}; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) \right)_{j,l=1, \dots, 2n+1}$$

negativ definit ist. Im Folgenden betrachten wir die Hessematrix \mathbf{W} als eine Funktion der Zufallsvariablen $Z_{i,k}$ und definieren die Fisher-Informationsmatrix folgendermaßen:

$$H = H(\mathbf{b}) = -\mathbb{E}[\mathbf{W}(\mathbf{b})] = \left(-\mathbb{E} \left[\sum_{i+k \leq n} \frac{\partial^2}{\partial b_j \partial b_l} l(Z_{i,k}; \gamma_{i,k}, \phi_{i,k}, w_{i,k}) \right] \right)_{j,l=1, \dots, 2n+1}$$

Bemerkungen:

- Man beachte, dass durch die Summation über $i+k \leq n$ die Anzahl der Beobachtungen in H enthalten ist.
- Die Matrizen $-\mathbf{W}$ und $H = -\mathbb{E}[\mathbf{W}]$ werden oft beobachtete und erwartete Fisher-Informationsmatrix genannt.

Lemma 6.5: Für die Fisher-Informationsmatrix $H = (H_{j,l})_{j,l=1, \dots, 2n+1}$ gilt unter den Modellannahmen 6.2

$$H_{j,l} = H_{j,l}(\mathbf{b}) = \sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \Gamma_{i,k}^{(l)}$$

Beweis: Es gilt

$$H_{j,l}(\mathbf{b}) = -\mathbb{E} \left[\frac{\partial}{\partial b_l} u_j(\mathbf{b}) \right]$$

und

$$\begin{aligned} \frac{\partial}{\partial b_l} u_j(\mathbf{b}) &= \frac{\partial}{\partial b_l} \sum_{i+k \leq n} W(z_{i,k}) (Z_{i,k} - z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} = \\ &= \sum_{i+k \leq n} \left(\frac{\partial}{\partial b_l} (Z_{i,k} - z_{i,k}) \right) W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} \\ &\quad + \sum_{i+k \leq n} (Z_{i,k} - z_{i,k}) \frac{\partial}{\partial b_l} \left(W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} \right). \end{aligned}$$

Für den ersten Term gilt aufgrund von (6.12)

$$\begin{aligned} & \sum_{i+k \leq n} \left(\frac{\partial}{\partial b_l} (Z_{i,k} - z_{i,k}) \right) W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} = - \sum_{i+k \leq n} \left(\frac{\partial z_{i,k}}{\partial b_l} \right) W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} = \\ & = - \sum_{i+k \leq n} W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial b_l} \Gamma_{i,k}^{(j)} = - \sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(l)} \Gamma_{i,k}^{(j)}. \end{aligned}$$

Der zweite Term verschwindet im Erwartungswert:

$$\begin{aligned} & \mathbb{E} \left[\sum_{i+k \leq n} (Z_{i,k} - z_{i,k}) \frac{\partial}{\partial b_l} \left(W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} \right) \right] = \\ & = \sum_{i+k \leq n} (\mathbb{E}[Z_{i,k}] - z_{i,k}) \frac{\partial}{\partial b_l} \left(W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} \right) = \\ & = \sum_{i+k \leq n} (z_{i,k} - z_{i,k}) \frac{\partial}{\partial b_l} \left(W(z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} \right) = 0 \end{aligned}$$

□

Man betrachte nun einen Vektor $\mathbf{b}^* = (b_1^*, \dots, b_{2n+1}^*)'$. Angenommen die Inverse der Fisher-Informationsmatrix $H = H(\mathbf{b}^*)$ existiert. Wir haben also wegen (6.3) und (6.4) $\eta_{i,k} = \eta_{i,k}(\mathbf{b}^*) = \Gamma_{i,k} \mathbf{b}^*$ und $z_{i,k} = z_{i,k}(\mathbf{b}^*)$ und aus Lemma 6.5 folgt für $j = 1, \dots, 2n+1$

$$(H\mathbf{b}^*)_j = \sum_{l=1}^{2n+1} H_{j,l} b_l^* = \sum_{l=1}^{2n+1} \sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \Gamma_{i,k}^{(l)} b_l^* = \sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \eta_{i,k}(\mathbf{b}^*). \quad (6.18)$$

Sei

$$\delta \mathbf{b}^* = H^{-1} \mathbf{u}(\mathbf{b}^*), \quad (6.19)$$

dann gilt wegen (6.19), (6.18), (6.17) und (6.5) für $j = 1, \dots, 2n+1$

$$\begin{aligned} (H(\mathbf{b}^* + \delta \mathbf{b}^*))_j &= (H\mathbf{b}^*)_j + (H\delta \mathbf{b}^*)_j = (H\mathbf{b}^*)_j + u_j(\mathbf{b}^*) = \\ &= \sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \eta_{i,k} + \sum_{i+k \leq n} W(z_{i,k}) (Z_{i,k} - z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \Gamma_{i,k}^{(j)} = \\ &= \sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \left(\eta_{i,k} + (Z_{i,k} - z_{i,k}) \frac{\partial \eta_{i,k}}{\partial z_{i,k}} \right) = \\ &= \sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} (g(z_{i,k}) + g'(z_{i,k})(Z_{i,k} - z_{i,k})) \end{aligned} \quad (6.20)$$

Dabei gilt für den letzten Term

$$g(Z_{i,k}) \approx g(z_{i,k}) + g'(z_{i,k})(Z_{i,k} - z_{i,k}).$$

Man beachte, dass $u_j(\mathbf{b}^*)$ für den Maximum-Likelihood-Schätzer $\mathbf{b}^* = \hat{\mathbf{b}}$ verschwindet, d.h. wir suchen eine Nullstelle von $u_j(\mathbf{b}^*)$, für die $H(\mathbf{b}^* + \delta \mathbf{b}^*) = H\mathbf{b}^*$ gilt.

Einerseits kann das mithilfe des Newton-Algorithmus erfolgen, wobei $\hat{\mathbf{b}}^{(m+1)}$ jener Schätzer ist, der durch

$$\mathbf{u}(\hat{\mathbf{b}}^{(m)}) + \mathbf{W}(\hat{\mathbf{b}}^{(m)}) (\hat{\mathbf{b}}^{(m+1)} - \hat{\mathbf{b}}^{(m)}) = \mathbf{0}$$

definiert ist, nachdem $\mathbf{W}(\hat{\mathbf{b}}^{(m)})$ die Jacobi-Matrix von $\mathbf{u}(\hat{\mathbf{b}}^{(m)})$, also die Matrix der partiellen Ableitungen von $\mathbf{u}(\hat{\mathbf{b}}^{(m)})$, ist.

Der Fisher-Scoring-Algorithmus ist eine Variante davon, wobei die (möglicherweise indefinite) Hessematrix $\mathbf{W}(\mathbf{b})$ durch minus die Fisher-Informationsmatrix $H(\mathbf{b})$ ersetzt wird:

$$\mathbf{u}(\hat{\mathbf{b}}^{(m)}) - H(\hat{\mathbf{b}}^{(m)}) (\hat{\mathbf{b}}^{(m+1)} - \hat{\mathbf{b}}^{(m)}) = \mathbf{0}$$

Das hat den Vorteil, dass diese Matrix leichter auszuwerten ist und immer positiv semidefinit ist. Verwendet man einen Startwert $\hat{\mathbf{b}}^{(0)}$, so gilt wegen (6.19) folgende Iteration für $m > 0$:

$$\hat{\mathbf{b}}^{(m+1)} = \hat{\mathbf{b}}^{(m)} + H(\hat{\mathbf{b}}^{(m)})^{-1} \mathbf{u}(\hat{\mathbf{b}}^{(m)}) = \hat{\mathbf{b}}^{(m)} + \delta \hat{\mathbf{b}}^{(m)} \quad (6.21)$$

Die Iteration wird abgebrochen, sobald ein Abbruchkriterium erfüllt ist. Dies kann z.B.

$$\frac{\|\hat{\mathbf{b}}^{(m+1)} - \hat{\mathbf{b}}^{(m)}\|}{\|\hat{\mathbf{b}}^{(m)}\|} \leq \varepsilon$$

für ein kleines $\varepsilon > 0$ sein. Aus (6.21) und (6.20) folgt für $m > 0$

$$\hat{\mathbf{b}}^{(m+1)} = H(\hat{\mathbf{b}}^{(m)})^{-1} \left(\sum_{i+k \leq n} W(z_{i,k}^{(m)}) \Gamma_{i,k}^{(j)} x^{(m)} \right)_{j=1, \dots, 2n+1},$$

wobei

$$\begin{aligned} x^{(m)} &= x^{(m)}(Z_{i,k}) = g(z_{i,k}^{(m)}) + g'(z_{i,k}^{(m)})(Z_{i,k} - z_{i,k}^{(m)}), \\ z_{i,k}^{(m)} &= h(\hat{\eta}_{i,k}^{(m)}) = h(\Gamma_{i,k} \hat{\mathbf{b}}^{(m)}) \end{aligned}$$

Für den Maximum-Likelihood-Schätzer $\hat{\mathbf{b}}$ gilt also wegen (6.5) und (6.16)

$$\hat{\mathbf{b}} = H(\hat{\mathbf{b}})^{-1} \left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} (g(z_{i,k}) + g'(z_{i,k})(Z_{i,k} - z_{i,k})) \right) \Big|_{\hat{\mathbf{b}}} = \quad (6.22)$$

$$= H(\hat{\mathbf{b}})^{-1} \left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \eta_{i,k} \right) \Big|_{\hat{\mathbf{b}}} \quad (6.23)$$

$$+ H(\hat{\mathbf{b}})^{-1} \left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \frac{\partial \eta_{i,k}}{\partial z_{i,k}} (Z_{i,k} - z_{i,k}) \right) \Big|_{\hat{\mathbf{b}}} \quad (6.24)$$

$$= H(\hat{\mathbf{b}})^{-1} \left(\sum_{i+k \leq n} W(\hat{z}_{i,k}) \Gamma_{i,k}^{(j)} \hat{\eta}_{i,k} \right)_{j=1, \dots, 2n+1}. \quad (6.25)$$

6.4.2 Mittlere quadratische Abweichung der Prognose

In diesem Kapitel soll ein Schätzer für die mittlere quadratische Abweichung der Prognose gefunden werden. Für die bedingte mittlere quadratische Abweichung gilt wegen (A.1)

$$\begin{aligned} \text{MSE}_{\sum_i S_{i,n} | \mathcal{D}_n} \left(\sum_{i=1}^n \widehat{S}_{i,n}^{\text{EDF}} \right) &= \mathbb{E} \left[\left(\sum_{i=1}^n \widehat{S}_{i,n}^{\text{EDF}} - \sum_{i=1}^n S_{i,n} \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \mathbb{E} \left[\left(\sum_{i+k>n} \widehat{Z}_{i,k}^{\text{EDF}} - \sum_{i+k>n} Z_{i,k} \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \text{Var} \left(\sum_{i+k>n} Z_{i,k} \middle| \mathcal{D}_n \right) + \left(\sum_{i+k>n} \left(\widehat{Z}_{i,k}^{\text{EDF}} - \mathbb{E}[Z_{i,k} | \mathcal{D}_n] \right) \right)^2, \end{aligned}$$

woraus aus den Modellannahmen 6.2 und (6.2)

$$\begin{aligned} &= \text{Var} \left(\sum_{i+k>n} Z_{i,k} \right) + \left(\sum_{i+k>n} \left(\widehat{Z}_{i,k}^{\text{EDF}} - \mathbb{E}[Z_{i,k}] \right) \right)^2 = \\ &= \sum_{i+k>n} \text{Var}(Z_{i,k}) + \left(\sum_{i+k>n} \left(\widehat{Z}_{i,k}^{\text{EDF}} - \mathbb{E}[Z_{i,k}] \right) \right)^2 = \\ &= \sum_{i+k>n} \frac{\phi_{i,k}}{w_{i,k}} V(z_{i,k}) + \left(\sum_{i+k>n} \left(\widehat{Z}_{i,k}^{\text{EDF}} - \mathbb{E}[Z_{i,k}] \right) \right)^2 \end{aligned}$$

folgt. Der erste Term kann für gegebene Varianzfunktion $V(z_{i,k})$ und Parameter $\phi_{i,k}$ und $w_{i,k}$ einfach geschätzt werden. Die Schwierigkeit ist wieder, die bedingten Schätzfehler zu schätzen.

Die unbedingte mittlere quadratische Abweichung hat folgende Form:

$$\begin{aligned} \text{MSE}_{\sum_i S_{i,n}} \left(\sum_{i=1}^n \widehat{S}_{i,n}^{\text{EDF}} \right) &= \mathbb{E} \left[\text{MSE}_{\sum_i S_{i,n} | \mathcal{D}_n} \left(\sum_{i=1}^n \widehat{S}_{i,n}^{\text{EDF}} \right) \right] = \\ &= \sum_{i+k>n} \frac{\phi_{i,k}}{w_{i,k}} V(z_{i,k}) + \mathbb{E} \left[\left(\sum_{i+k>n} \left(\widehat{Z}_{i,k}^{\text{EDF}} - \mathbb{E}[Z_{i,k}] \right) \right)^2 \right] \end{aligned}$$

Die Schwierigkeit besteht wiederum darin, den letzten Term, also den Schätzfehler, zu schätzen. Es gilt

$$\mathbb{E} \left[\left(\sum_{i+k>n} \left(\widehat{Z}_{i,k}^{\text{EDF}} - \mathbb{E}[Z_{i,k}] \right) \right)^2 \right] = \sum_{\substack{i+k>n, \\ l+m>n}} \mathbb{E} \left[\left(\widehat{Z}_{i,k}^{\text{EDF}} - \mathbb{E}[Z_{i,k}] \right) \left(\widehat{Z}_{l,m}^{\text{EDF}} - \mathbb{E}[Z_{l,m}] \right) \right] \quad (6.26)$$

Man beachte, dass $\widehat{Z}_{i,k}^{\text{EDF}}$ im Allgemeinen kein unverzerrter Schätzer für $\mathbb{E}[Z_{i,k}]$ ist, d.h. die Terme in (6.26) können einen zusätzlichen systematischen Fehler haben, der aber oft vernachlässigt werden kann, sodass die quadratischen Terme folgendermaßen approximiert werden:

$$\begin{aligned} \text{Var} \left(\widehat{Z}_{i,k}^{\text{EDF}} \right) &= \text{Var} \left(\exp\{\widehat{\eta}_{i,k}\} \right) = \exp\{2\eta_{i,k}\} \text{Var} \left(\exp\{\widehat{\eta}_{i,k} - \eta_{i,k}\} \right) \approx \\ &\approx \exp\{2\eta_{i,k}\} \text{Var} \left(1 + \widehat{\eta}_{i,k} - \eta_{i,k} \right) = \exp\{2\eta_{i,k}\} \text{Var} \left(\widehat{\eta}_{i,k} \right) = \\ &= z_{i,k}^2 \text{Var} \left(\Gamma_{i,k} \widehat{\mathbf{b}} \right) = z_{i,k}^2 \Gamma_{i,k} \text{Cov} \left(\widehat{\mathbf{b}}, \widehat{\mathbf{b}} \right) \Gamma_{i,k}' \end{aligned}$$

Dabei wurde benutzt, dass $\exp(x) \approx 1 + x$ für $x \approx 0$. Analog werden die gemischten Terme approximiert.

$$\begin{aligned} \text{Cov} \left(\widehat{Z}_{i,k}^{\text{EDF}}, \widehat{Z}_{l,m}^{\text{EDF}} \right) &= \text{Cov} \left(\exp\{\widehat{\eta}_{i,k}\}, \exp\{\widehat{\eta}_{l,m}\} \right) \approx \exp\{\eta_{i,k}\} \exp\{\eta_{l,m}\} \text{Cov} \left(\widehat{\eta}_{i,k}, \widehat{\eta}_{l,m} \right) = \\ &= z_{i,k} z_{l,m} \Gamma_{i,k} \text{Cov} \left(\widehat{\mathbf{b}}, \widehat{\mathbf{b}} \right) \Gamma'_{l,m} \end{aligned}$$

Wir müssen also die Kovarianzmatrix von $\widehat{\mathbf{b}}$ berechnen. Wir nehmen an, dass $Z_{i,k}$ zur EDF gehört mit gegebener Varianzfunktion $V(z_{i,k})$ und Parameter \mathbf{b} und definieren die Zufallsvariable

$$\mathbf{B} = H(\mathbf{b})^{-1} \left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} X_{i,k} \right)_{j=1, \dots, 2n+1},$$

wobei

$$X_{i,k} = X_{i,k}(\mathbf{b}) = g(z_{i,k}) + g'(z_{i,k})(Z_{i,k} - z_{i,k})$$

und $z_{i,k} = z_{i,k}(\mathbf{b}) = h(\Gamma_{i,k} \mathbf{b})$. Dann gilt

Lemma 6.6: Unter den Modellannahmen 6.2 gilt

$$\text{Cov}(\mathbf{B}, \mathbf{B}) = H(\mathbf{b})^{-1}.$$

Beweis: Nachdem die Inkremente $Z_{i,k}$ unabhängig sind, ist die Kovarianz zwischen $X_{i,k}$ und $X_{l,m}$ gleich Null für $(i, k) \neq (l, m)$. Wir betrachten also nur

$$\text{Var}(X_{i,k}) = \text{Var}(g'(z_{i,k})Z_{i,k}) = \left(\frac{\partial \eta_{i,k}}{\partial z_{i,k}} \right)^2 \text{Var}(Z_{i,k}) = W(z_{i,k})^{-1}.$$

Mithilfe von Lemma 6.5 folgt also

$$\begin{aligned} &\text{Cov} \left(\left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} X_{i,k} \right)_j, \left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(l)} X_{i,k} \right)_l \right) = \\ &= \left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \text{Var}(X_{i,k}) \Gamma_{i,k}^{(l)} W(z_{i,k}) \right)_{j,l} = \left(\sum_{i+k \leq n} W(z_{i,k}) \Gamma_{i,k}^{(j)} \Gamma_{i,k}^{(l)} \right)_{j,l} = \\ &= (H_{j,l})_{j,l} = H(\mathbf{b}) \end{aligned}$$

und deshalb

$$\text{Cov}(\mathbf{B}, \mathbf{B}) = H(\mathbf{b})^{-1} H(\mathbf{b}) H(\mathbf{b})^{-1} = H(\mathbf{b})^{-1}.$$

□

Für den Maximum-Likelihood-Schätzer gilt wegen (6.22) und (6.16)

$$\begin{aligned}
\widehat{\mathbf{b}} &= H(\widehat{\mathbf{b}})^{-1} \left(\sum_{i+k \leq n} W(\widehat{z}_{i,k}) \Gamma_{i,k}^{(j)} \widehat{\eta}_{i,k} \right)_{j=1, \dots, 2n+1} = \\
&= H(\widehat{\mathbf{b}})^{-1} \left(\sum_{i+k \leq n} W(\widehat{z}_{i,k}) \Gamma_{i,k}^{(j)} g(\widehat{z}_{i,k}) \right)_{j=1, \dots, 2n+1} = \\
&= H(\widehat{\mathbf{b}})^{-1} \left(\sum_{i+k \leq n} W(\widehat{z}_{i,k}) \Gamma_{i,k}^{(j)} X_{i,k}(\widehat{\mathbf{b}}) \right)_{j=1, \dots, 2n+1}. \tag{6.27}
\end{aligned}$$

Ersetzt man in (6.27) auf der rechten Seite den geschätzten Parameter $\widehat{\mathbf{b}}$ durch den wahren \mathbf{b} , so erhält man die Zufallsvariable \mathbf{B} . Deshalb schätzen wir die Kovarianzmatrix des Maximum-Likelihood-Schätzers $\widehat{\mathbf{b}}$ durch

$$\widehat{\text{Cov}}(\widehat{\mathbf{b}}, \widehat{\mathbf{b}}) = H(\widehat{\mathbf{b}})^{-1} = \left(\left(\sum_{i+k \leq n} W(\widehat{z}_{i,k}) \Gamma_{i,k}^{(j)} \Gamma_{i,k}^{(l)} \right)_{j,l=1, \dots, 2n+1} \right)^{-1}. \tag{6.28}$$

Insgesamt ergibt sich also folgender Schätzer für die unbedingte mittlere quadratische Abweichung:

$$\widehat{\text{MSE}}_{\sum_i S_{i,n}} \left(\sum_{i=1}^n \widehat{S}_{i,n}^{\text{EDF}} \right) = \sum_{i+k > n} \frac{\phi_{i,k}}{w_{i,k}} V(\widehat{z}_{i,k}) + \sum_{\substack{i+k > n, \\ l+m > n}} \widehat{z}_{i,k} \widehat{z}_{l,m} \Gamma_{i,k} H(\widehat{\mathbf{b}})^{-1} \Gamma'_{l,m},$$

wobei $H(\widehat{\mathbf{b}})^{-1}$ durch (6.28) gegeben ist.

Bemerkungen:

- Falls die Inverse der Überleitungsfunktion der Logarithmus ist, d.h. $\eta_{i,k} = \ln(z_{i,k})$, dann gilt

$$\frac{\partial \eta_{i,k}}{\partial z_{i,k}} = z_{i,k}^{-1}$$

und es folgt

$$W(z_{i,k}) = \frac{w_{i,k}}{\phi_{i,k}} \frac{z_{i,k}^2}{V(z_{i,k})}.$$

- Für die Schätzung von $\phi_{i,k}/w_{i,k}$ nehmen wir an, dass der Term konstant ist, d.h. $\phi_{i,k}/w_{i,k} = \phi$ für alle i, k , und verwenden die Pearson-Residuen

$$R_{i,k}^{\text{Pe}}(z_{i,k}) = \frac{Z_{i,k} - z_{i,k}}{V(z_{i,k})^{1/2}},$$

für die

$$\mathbb{E} \left[(R_{i,k}^{\text{Pe}}(z_{i,k}))^2 \right] = \frac{\phi_{i,k}}{w_{i,k}} = \phi$$

gilt. Wir setzen

$$\widehat{R}_{i,k}^{\text{Pe}} = \frac{Z_{i,k} - \widehat{z}_{i,k}}{V(\widehat{z}_{i,k})^{1/2}}$$

und schätzen ϕ folgendermaßen:

$$\widehat{\phi}^{\text{Pe}} = \frac{\sum_{i+k \leq n} \left(\widehat{R}_{i,k}^{\text{Pe}} \right)^2}{N - p}, \quad (6.29)$$

wobei N die Anzahl der Beobachtungen $Z_{i,k}$ in \mathcal{D}_n und p die Anzahl der geschätzten Parameter b_j , d.h. $p = 2n + 1$, ist.

6.5 Mittlere quadratische Abweichung der Bornhuetter-Ferguson-Prognose

In diesem Abschnitt soll ein Schätzer für die mittlere quadratische Abweichung der Bornhuetter-Ferguson-Prognose gefunden werden.

Das Abwicklungsmuster α_k wird häufig mithilfe der Chain-Ladder-Faktoren wie in (4.4) geschätzt, obwohl der Schätzer durch kein gemeinsames stochastisches Modell begründet ist. Der Hauptkritikpunkt dieses Ansatzes ist, dass die Verwendung der \widehat{f}_k der Grundidee des Bornhuetter-Ferguson-Verfahrens, dass $S_{i,n} - S_{i,n-i}$ unabhängig von $S_{i,n-i}$ prognostiziert wird, widerspricht. Wir wollen nun ein stochastisches Modell zugrunde legen, das (4.4) motiviert.

Modellannahmen 6.7:

- Die Inkremente $Z_{i,k}$ sind unabhängig mit einer **Overdispersed Poisson-Verteilung**, d.h. es gibt positive Parameter $\theta_0, \dots, \theta_n, \mu_0, \dots, \mu_n$ mit $\sum_{k=0}^n \theta_k = 1$ und $\phi > 0$, sodass

$$\frac{Z_{i,k}}{\phi} \sim \text{Poi} \left(\frac{\theta_k \mu_i}{\phi} \right),$$

d.h.

$$\begin{aligned} \mathbb{E}[Z_{i,k}] &= z_{i,k} = \theta_k \mu_i, \\ \text{Var}(Z_{i,k}) &= \phi z_{i,k} = \phi \theta_k \mu_i. \end{aligned}$$

- $\widehat{\mu}_i, i \in \{0, \dots, n\}$, sind unabhängige und unverzerrte A-priori-Schätzer für $\mu_i = \mathbb{E}[S_{i,n}]$.
- $Z_{j,k}$ und $\widehat{\mu}_i$ sind unabhängig für alle i, j, k .

Bemerkungen:

- $Z_{i,k}$ gehört zur EDF (siehe Modellannahmen 6.2) mit Varianzfunktion $V(z) = z$ und $d(\cdot) = \exp(\cdot)$, denn es gilt

$$\begin{aligned} \mathbb{E}[Z_{i,k} = x] &= \mathbb{E} \left[\frac{Z_{i,k}}{\phi} = \frac{x}{\phi} \right] = \frac{1}{\left(\frac{x}{\phi}\right)!} \left(\frac{\theta_k \mu_i}{\phi} \right)^{\frac{x}{\phi}} \exp \left(-\frac{\theta_k \mu_i}{\phi} \right) = \\ &= \frac{1}{\left(\frac{x}{\phi}\right)!} \exp \left(-\frac{\theta_k \mu_i}{\phi} + \frac{x}{\phi} \ln \left(\frac{\theta_k \mu_i}{\phi} \right) \right) = \\ &= \frac{1}{\left(\frac{x}{\phi}\right)!} \exp \left(-\frac{\theta_k \mu_i}{\phi} + \frac{x}{\phi} \ln(\theta_k \mu_i) - \frac{x}{\phi} \ln \phi \right), \end{aligned}$$

d.h. mit den Bezeichnungen aus den Modellannahmen 6.2 gilt

$$\begin{aligned}\frac{\phi_{i,k}}{w_{i,k}} &= \phi, \\ \gamma_{i,k} &= \ln(\mu_i \theta_k), \\ d(\cdot) &= \exp(\cdot), \\ a(x, \phi) &= \frac{1}{\left(\frac{x}{\phi}\right)!} \exp\left(-\frac{x}{\phi} \ln \phi\right).\end{aligned}$$

- Unter den Modellannahmen 6.7 sind auch die Annahmen 4.1 erfüllt, und zwar mit $\alpha_k = \sum_{j=0}^k \theta_j$, d.h. dieses Overdispersed-Poisson-Modell kann verwendet werden, um die Bornhuetter-Ferguson-Methode zu motivieren.
- Führt man für das Overdispersed-Poisson-Modell eine Maximum-Likelihood-Schätzung durch, so erhält man $\hat{\mu}_i^{\text{MLE}}$ und $\hat{\theta}_k^{\text{MLE}}$ als Lösungen der Gleichungen (5.3), nachdem sich die ϕ weggelassen.
- $\hat{\mu}_i$ ist ein A-priori-Schätzer für den Endschaden μ_i , d.h. die Schätzung erfolgt nur auf Basis externer Daten und Expertenmeinung und wir nehmen deshalb an, dass sie unabhängig von $Z_{i,k}$ ist.

Verwenden wir nun die Maximum-Likelihood-Schätzer $\hat{\theta}_k^{\text{MLE}}$ für die Schätzung des Abwicklungsmusters, so erhalten wir folgenden Bornhuetter-Ferguson-Schätzer:

$$\hat{S}_{i,n}^{\text{BF}'} = S_{i,n-i} + \left(1 - \sum_{k=0}^{n-i} \hat{\theta}_k^{\text{MLE}}\right) \hat{\mu}_i$$

Wegen Lemma 5.5 gilt

$$\hat{S}_{i,n}^{\text{BF}'} = S_{i,n-i} + (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{\mu}_i,$$

was komplett durch die Modellannahmen 6.7 und Verwendung des Maximum-Likelihood-Schätzers für θ_k motiviert ist.

Bemerkung:

- Ersetzt man den A-priori-Schätzer $\hat{\mu}_i$ durch den Maximum-Likelihood-Schätzer $\hat{\mu}_i^{\text{MLE}}$, so erhält man

$$\hat{S}_{i,n}^{\text{EDF}} = S_{i,n-i} + \left(1 - \sum_{k=0}^{n-i} \hat{\theta}_k^{\text{MLE}}\right) \hat{\mu}_i^{\text{MLE}}$$

wie in (6.14) und (6.15).

6.5.1 Mittlere quadratische Abweichung der Prognose $\hat{R}_i^{\text{BF}'}$ für ein Abwicklungsjahr i

Unser Ziel ist es, die bedingte mittlere quadratische Abweichung der Prognose

$$\hat{R}_i^{\text{BF}'} = \hat{S}_{i,n}^{\text{BF}'} - S_{i,n-i}$$

zu schätzen:

$$\begin{aligned}
\text{MSE}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{BF}'}) &= \mathbb{E} \left[\left(R_i - \widehat{R}_i^{\text{BF}'} \right)^2 \middle| \mathcal{D}_n \right] = \\
&= \mathbb{E} \left[\left(\sum_{k=n-i+1}^n Z_{i,k} - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i \right)^2 \middle| \mathcal{D}_n \right] = \\
&= \mathbb{E} \left[\left(\sum_{k=n-i+1}^n (Z_{i,k} - \mathbb{E}[Z_{i,k}] + \mathbb{E}[Z_{i,k}]) - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i \right)^2 \middle| \mathcal{D}_n \right]
\end{aligned}$$

Nachdem die $Z_{i,k}$ unabhängig sind, gilt

$$\begin{aligned}
&\mathbb{E} \left[\left(\sum_{k=n-i+1}^n (Z_{i,k} - \mathbb{E}[Z_{i,k}] + \mathbb{E}[Z_{i,k}]) - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i \right)^2 \middle| \mathcal{D}_n \right] = \\
&= \sum_{k=n-i+1}^n \text{Var}(Z_{i,k}) + \mathbb{E} \left[\left(\sum_{k=n-i+1}^n \mathbb{E}[Z_{i,k}] - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i \right)^2 \middle| \mathcal{D}_n \right] \\
&\quad + 2 \mathbb{E} \left[\left(\sum_{k=n-i+1}^n (Z_{i,k} - \mathbb{E}[Z_{i,k}]) \right) \left(\sum_{l=n-i+1}^n \mathbb{E}[Z_{i,l}] - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i \right) \middle| \mathcal{D}_n \right].
\end{aligned}$$

Nachdem $\widehat{\mu}_i$ unabhängig von $Z_{j,k}$ für alle i, j, k , $\widehat{\alpha}_{n-i}^{\text{CL}}$ messbar bezüglich \mathcal{D}_n und $\mathbb{E}[\widehat{\mu}_i] = \mu_i$ ist, gilt

$$\begin{aligned}
&\mathbb{E} \left[\left(\sum_{k=n-i+1}^n \mathbb{E}[Z_{i,k}] - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i \right)^2 \middle| \mathcal{D}_n \right] = \\
&= \left(\sum_{k=n-i+1}^n \theta_k \mu_i \right)^2 - 2 \left(\sum_{k=n-i+1}^n \theta_k \mu_i \right) (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \mu_i + (1 - \widehat{\alpha}_{n-i}^{\text{CL}})^2 \mathbb{E}[\widehat{\mu}_i^2] = \\
&= \mu_i^2 \left(\left(\sum_{k=n-i+1}^n \theta_k \right)^2 - 2 \left(\sum_{k=n-i+1}^n \theta_k \right) (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) + (1 - \widehat{\alpha}_{n-i}^{\text{CL}})^2 \right) \\
&\quad + (1 - \widehat{\alpha}_{n-i}^{\text{CL}})^2 \text{Var}(\widehat{\mu}_i) = \\
&= \mu_i^2 \left(\sum_{k=n-i+1}^n \theta_k - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \right)^2 + (1 - \widehat{\alpha}_{n-i}^{\text{CL}})^2 \text{Var}(\widehat{\mu}_i) = \\
&= \mu_i^2 \left(\sum_{k=n-i+1}^n \theta_k - \left(1 - \sum_{k=0}^{n-i} \widehat{\theta}_k^{\text{MLE}} \right) \right)^2 + (1 - \widehat{\alpha}_{n-i}^{\text{CL}})^2 \text{Var}(\widehat{\mu}_i) = \\
&= \mu_i^2 \left(\sum_{k=n-i+1}^n \theta_k - \sum_{k=n-i+1}^n \widehat{\theta}_k^{\text{MLE}} \right)^2 + (1 - \widehat{\alpha}_{n-i}^{\text{CL}})^2 \text{Var}(\widehat{\mu}_i)
\end{aligned}$$

und

$$\begin{aligned}
& \mathbb{E} \left[\left(\sum_{k=n-i+1}^n (Z_{i,k} - \mathbb{E}[Z_{i,k}]) \right) \left(\sum_{l=n-i+1}^n \mathbb{E}[Z_{i,l}] - (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{\mu}_i \right) \middle| \mathcal{D}_n \right] = \\
&= \sum_{k=n-i+1}^n \mathbb{E} \left[(Z_{i,k} - \mathbb{E}[Z_{i,k}]) \left(\sum_{l=n-i+1}^n \mathbb{E}[Z_{i,l}] - (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{\mu}_i \right) \middle| \mathcal{D}_n \right] = \\
&= \sum_{k=n-i+1}^n \mathbb{E}[Z_{i,k} - \mathbb{E}[Z_{i,k}] | \mathcal{D}_n] \mathbb{E} \left[\left(\sum_{l=n-i+1}^n \mathbb{E}[Z_{i,l}] - (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{\mu}_i \right) \middle| \mathcal{D}_n \right] = \\
&= \sum_{k=n-i+1}^n \mathbb{E}[Z_{i,k} - \mathbb{E}[Z_{i,k}]] \mathbb{E} \left[\left(\sum_{l=n-i+1}^n \mathbb{E}[Z_{i,l}] - (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{\mu}_i \right) \middle| \mathcal{D}_n \right] = 0.
\end{aligned}$$

Damit ergibt sich

$$\begin{aligned}
\text{MSE}_{R_i | \mathcal{D}_n}(\hat{R}_i^{\text{BF}'}) &= \sum_{k=n-i+1}^n \text{Var}(Z_{i,k}) + (1 - \hat{\alpha}_{n-i}^{\text{CL}})^2 \text{Var}(\hat{\mu}_i) \\
&\quad + \mu_i^2 \left(\sum_{k=n-i+1}^n \theta_k - \sum_{k=n-i+1}^n \hat{\theta}_k^{\text{MLE}} \right)^2.
\end{aligned}$$

Wir müssen also die drei Terme der rechten Seite schätzen. Ein Schätzer für $\text{Var}(Z_{i,k})$ ist

$$\widehat{\text{Var}}(Z_{i,k}) = \hat{\phi}^{\text{Pe}} \hat{\theta}_k^{\text{MLE}} \hat{\mu}_i,$$

wobei $\hat{\phi}^{\text{Pe}}$ der Schätzer aus (6.29) ist. Der zweite Term enthält die Unsicherheit des A-priori-Schätzers $\hat{\mu}_i$ und kann deshalb nur mithilfe externer Daten, Markterfahrung und Expertenmeinung bestimmt werden. Eine Möglichkeit ist, einen Variationskoeffizient

$$\text{Vco}(\hat{\mu}_i) = \frac{\sqrt{\widehat{\text{Var}}(\hat{\mu}_i)}}{\mathbb{E}[\hat{\mu}_i]} = \frac{\sqrt{\widehat{\text{Var}}(\hat{\mu}_i)}}{\mu_i}$$

zu schätzen, wobei dieser in der Praxis oft zwischen 5% und 10% liegt. (vgl. Swiss Solvency Test [12]) Dann ergibt sich für den zweiten Term folgende Schätzung

$$(1 - \hat{\alpha}_{n-i}^{\text{CL}})^2 \widehat{\text{Var}}(\hat{\mu}_i) = (1 - \hat{\alpha}_{n-i}^{\text{CL}})^2 \hat{\mu}_i^2 \widehat{\text{Vco}}(\hat{\mu}_i)^2.$$

Um den dritten Term zu schätzen, müssen wir die Maximum-Likelihood-Schätzer $\hat{\theta}_k^{\text{MLE}}$ abwandeln. Wir haben folgende Normalisierung angenommen:

$$\sum_{k=0}^n \hat{\theta}_k^{\text{MLE}} = 1$$

Bei den verallgemeinerten linearen Modelle haben wir $\mu_0 = 1$ angenommen, womit wir folgende Maximum-Likelihood-Schätzer erhalten

$$\hat{\mathbf{b}} = \left(\widehat{\ln(\mu_1)}^{\text{GLM}}, \dots, \widehat{\ln(\mu_n)}^{\text{GLM}}, \widehat{\ln(\theta_0)}^{\text{GLM}}, \dots, \widehat{\ln(\theta_n)}^{\text{GLM}} \right)',$$

womit wir ein zweites Abwicklungsmuster $\hat{\theta}_0^{\text{GLM}}, \dots, \hat{\theta}_n^{\text{GLM}}$ erhalten, wobei $\widehat{\ln(\theta_k)}^{\text{GLM}} = \ln(\hat{\theta}_k^{\text{GLM}})$ für alle $k = 0, \dots, n$. Wir wissen, dass $\hat{\mathbf{b}}$ asymptotisch normalverteilt ist und die Varianz durch

die Inverse der Fisher-Information $H^{-1}(\mathbf{b})$ approximiert wird. (siehe Lemma 6.1) Außerdem gilt aufgrund der unterschiedlichen Normierungen

$$\widehat{\theta}_k^{\text{MLE}} = \frac{\widehat{\theta}_k^{\text{GLM}}}{\sum_{l=0}^n \widehat{\theta}_l^{\text{GLM}}}.$$

Indem wir eine mögliche Verzerrung der Schätzer vernachlässigen, schätzen wir

$$\left(\sum_{k=n-i+1}^n \theta_k - \sum_{k=n-i+1}^n \widehat{\theta}_k^{\text{MLE}} \right)^2$$

durch

$$\begin{aligned} \text{Var} \left(\sum_{k=n-i+1}^n \widehat{\theta}_k^{\text{MLE}} \right) &= \sum_{k=n-i+1}^n \sum_{l=n-i+1}^n \text{Cov} \left(\widehat{\theta}_k^{\text{MLE}}, \widehat{\theta}_l^{\text{MLE}} \right) = \\ &= \sum_{k=n-i+1}^n \sum_{l=n-i+1}^n \text{Cov} \left(\frac{\widehat{\theta}_k^{\text{GLM}}}{\sum_{j=0}^n \widehat{\theta}_j^{\text{GLM}}}, \frac{\widehat{\theta}_l^{\text{GLM}}}{\sum_{j=0}^n \widehat{\theta}_j^{\text{GLM}}} \right) = \\ &= \sum_{k=n-i+1}^n \sum_{l=n-i+1}^n \text{Cov} \left(\frac{1}{1 + \sum_{j \neq k} \frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}}}, \frac{1}{1 + \sum_{j \neq l} \frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_l^{\text{GLM}}}} \right) = \\ &= \sum_{k=n-i+1}^n \sum_{l=n-i+1}^n \text{Cov} \left(\frac{1}{1 + \Delta_k}, \frac{1}{1 + \Delta_l} \right) \end{aligned}$$

für

$$\Delta_k = \sum_{j \neq k} \frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}} \quad \text{und} \quad \delta_k = \mathbb{E}[\Delta_k].$$

Wir wollen nun eine Taylor-Approximation mit Entwicklungspunkt δ_k machen. Sei

$$f(x) = \frac{1}{1+x} \quad \text{und} \quad f'(x) = -\frac{1}{(1+x)^2}.$$

Dann gilt

$$f(x) \approx f(\delta_k) + f'(\delta_k)(x - \delta_k) = \frac{1}{1 + \delta_k} - \frac{1}{(1 + \delta_k)^2}(x - \delta_k)$$

und

$$\begin{aligned} \text{Cov} \left(\frac{1}{1 + \Delta_k}, \frac{1}{1 + \Delta_l} \right) &\approx \frac{1}{(1 + \delta_k)^2} \frac{1}{(1 + \delta_l)^2} \text{Cov}(\Delta_k, \Delta_l) = \\ &= \frac{1}{(1 + \delta_k)^2} \frac{1}{(1 + \delta_l)^2} \sum_{j \neq k} \sum_{m \neq l} \text{Cov} \left(\frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}}, \frac{\widehat{\theta}_m^{\text{GLM}}}{\widehat{\theta}_l^{\text{GLM}}} \right) \end{aligned}$$

Falls eine mögliche Verzerrung vernachlässigt wird, gilt wegen $\exp(x) \approx 1 + x$ für $x \approx 0$

$$\begin{aligned} &\text{Cov} \left(\frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}}, \frac{\widehat{\theta}_m^{\text{GLM}}}{\widehat{\theta}_l^{\text{GLM}}} \right) = \\ &= \frac{\theta_j}{\theta_k} \frac{\theta_m}{\theta_l} \text{Cov} \left(\exp \left\{ \ln \left(\frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}} \right) - \ln \left(\frac{\theta_j}{\theta_k} \right) \right\}, \exp \left\{ \ln \left(\frac{\widehat{\theta}_m^{\text{GLM}}}{\widehat{\theta}_l^{\text{GLM}}} \right) - \ln \left(\frac{\theta_m}{\theta_l} \right) \right\} \right) \approx \\ &\approx \frac{\theta_j}{\theta_k} \frac{\theta_m}{\theta_l} \text{Cov} \left(\ln \left(\frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}} \right), \ln \left(\frac{\widehat{\theta}_m^{\text{GLM}}}{\widehat{\theta}_l^{\text{GLM}}} \right) \right). \end{aligned}$$

Wir definieren für $k = 0, \dots, n$ die Matrix

$$\tilde{\Gamma}_k = \Gamma_{0,k} = (0, \dots, 0, e_{n+k+1}, 0, \dots, 0),$$

sodass

$$\ln \left(\hat{\theta}_k^{\text{GLM}} \right) = \tilde{\Gamma}_k \hat{\mathbf{b}}$$

gilt. Also folgt

$$\begin{aligned} \text{Cov} \left(\ln \left(\frac{\hat{\theta}_j^{\text{GLM}}}{\hat{\theta}_k^{\text{GLM}}} \right), \ln \left(\frac{\hat{\theta}_m^{\text{GLM}}}{\hat{\theta}_l^{\text{GLM}}} \right) \right) &= \text{Cov} \left(\tilde{\Gamma}_j \hat{\mathbf{b}} - \tilde{\Gamma}_k \hat{\mathbf{b}}, \tilde{\Gamma}_m \hat{\mathbf{b}} - \tilde{\Gamma}_l \hat{\mathbf{b}} \right) = \\ &= \left(\tilde{\Gamma}_j - \tilde{\Gamma}_k \right) \text{Cov} \left(\hat{\mathbf{b}}, \hat{\mathbf{b}} \right) \left(\tilde{\Gamma}_m - \tilde{\Gamma}_l \right)' \end{aligned}$$

und

$$\begin{aligned} \text{Var} \left(\sum_{k=n-i+1}^n \hat{\theta}_k^{\text{MLE}} \right) &\approx \sum_{k=n-i+1}^n \sum_{l=n-i+1}^n \frac{1}{(1 + \delta_k)^2} \frac{1}{(1 + \delta_l)^2} \\ &\times \sum_{j \neq k} \sum_{m \neq l} \frac{\theta_j}{\theta_k} \frac{\theta_m}{\theta_l} \left(\tilde{\Gamma}_j - \tilde{\Gamma}_k \right) H \left(\hat{\mathbf{b}} \right)^{-1} \left(\tilde{\Gamma}_m - \tilde{\Gamma}_l \right)'. \end{aligned}$$

Wir definieren also folgenden Schätzer

$$\begin{aligned} \widehat{\text{Var}} \left(\sum_{k=n-i+1}^n \hat{\theta}_k^{\text{MLE}} \right) &= \sum_{k=n-i+1}^n \sum_{l=n-i+1}^n \frac{1}{(1 + \hat{\delta}_k)^2} \frac{1}{(1 + \hat{\delta}_l)^2} \\ &\times \sum_{j \neq k} \sum_{m \neq l} \frac{\hat{\theta}_j^{\text{GLM}}}{\hat{\theta}_k^{\text{GLM}}} \frac{\hat{\theta}_m^{\text{GLM}}}{\hat{\theta}_l^{\text{GLM}}} \left(\tilde{\Gamma}_j - \tilde{\Gamma}_k \right) H \left(\hat{\mathbf{b}} \right)^{-1} \left(\tilde{\Gamma}_m - \tilde{\Gamma}_l \right)', \end{aligned}$$

wobei δ_k folgendermaßen geschätzt wird:

$$\hat{\delta}_k = \sum_{j \neq k} \frac{\hat{\theta}_j^{\text{GLM}}}{\hat{\theta}_k^{\text{GLM}}}$$

Unter den Modellannahmen 6.7 ist ein Schätzer für die bedingte mittlere quadratische Abweichung für ein Anfallsjahr $i \in \{1, \dots, n\}$ gegeben durch

$$\begin{aligned} \widehat{\text{MSE}}_{R_i | \mathcal{D}_n} \left(\hat{R}_i^{\text{BF}'} \right) &= \sum_{k=n-i+1}^n \hat{\phi}^{\text{Pe}} \hat{\theta}_k^{\text{MLE}} \hat{\mu}_i + (1 - \hat{\alpha}_{n-i}^{\text{CL}})^2 \hat{\mu}_i^2 \widehat{\text{Vco}} \left(\hat{\mu}_i \right)^2 \\ &+ \hat{\mu}_i^2 \widehat{\text{Var}} \left(\sum_{k=n-i+1}^n \hat{\theta}_k^{\text{MLE}} \right) = \\ &= \hat{\phi}^{\text{Pe}} (1 - \hat{\alpha}_{n-i}^{\text{CL}}) \hat{\mu}_i + (1 - \hat{\alpha}_{n-i}^{\text{CL}})^2 \hat{\mu}_i^2 \widehat{\text{Vco}} \left(\hat{\mu}_i \right)^2 \\ &+ \hat{\mu}_i^2 \widehat{\text{Var}} \left(\sum_{k=n-i+1}^n \hat{\theta}_k^{\text{MLE}} \right). \end{aligned}$$

6.5.2 Mittlere quadratische Abweichung der Prognose $\widehat{R}^{\text{BF}'}$

Ziel dieses Kapitels ist es, einen Schätzer für die bedingte mittlere quadratische Abweichung von

$$\widehat{R}^{\text{BF}'} = \sum_{i=1}^n \widehat{R}_i^{\text{BF}'}$$

zu finden. Wir starten mit zwei Anfallsjahren $i < j$.

$$\begin{aligned} \text{MSE}_{R_i+R_j|\mathcal{D}_n} \left(\widehat{R}_i^{\text{BF}'} + \widehat{R}_j^{\text{BF}'} \right) &= \mathbb{E} \left[\left(R_i + R_j - \widehat{R}_i^{\text{BF}'} - \widehat{R}_j^{\text{BF}'} \right)^2 \middle| \mathcal{D}_n \right] = \\ &= \text{MSE}_{R_i|\mathcal{D}_n} \left(\widehat{R}_i^{\text{BF}'} \right) + \text{MSE}_{R_j|\mathcal{D}_n} \left(\widehat{R}_j^{\text{BF}'} \right) + 2 \mathbb{E} \left[\left(R_i - \widehat{R}_i^{\text{BF}'} \right) \left(R_j - \widehat{R}_j^{\text{BF}'} \right) \middle| \mathcal{D}_n \right] \end{aligned}$$

Es gilt

$$\begin{aligned} &\mathbb{E} \left[\left(R_i - \widehat{R}_i^{\text{BF}'} \right) \left(R_j - \widehat{R}_j^{\text{BF}'} \right) \middle| \mathcal{D}_n \right] = \\ &= \mathbb{E} \left[\left(\sum_{k=n-i+1}^n Z_{i,k} - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \widehat{\mu}_i \right) \left(\sum_{l=n-j+1}^n Z_{j,l} - (1 - \widehat{\alpha}_{n-j}^{\text{CL}}) \widehat{\mu}_j \right) \middle| \mathcal{D}_n \right] = \\ &= \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \mathbb{E} [Z_{i,k}] \mathbb{E} [Z_{j,l}] - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \mu_i \sum_{l=n-j+1}^n \mathbb{E} [Z_{j,l}] \\ &\quad - \sum_{k=n-i+1}^n \mathbb{E} [Z_{i,k}] (1 - \widehat{\alpha}_{n-j}^{\text{CL}}) \mu_j + (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \mu_i (1 - \widehat{\alpha}_{n-j}^{\text{CL}}) \mu_j = \\ &= \mu_i \mu_j \sum_{k=n-i+1}^n \theta_k \sum_{l=n-j+1}^n \theta_l - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \mu_i \mu_j \sum_{l=n-j+1}^n \theta_l \\ &\quad - (1 - \widehat{\alpha}_{n-j}^{\text{CL}}) \mu_i \mu_j \sum_{k=n-i+1}^n \theta_k + (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \mu_i (1 - \widehat{\alpha}_{n-j}^{\text{CL}}) \mu_j = \\ &= \mu_i \mu_j \left(\sum_{k=n-i+1}^n \theta_k - (1 - \widehat{\alpha}_{n-i}^{\text{CL}}) \right) \left(\sum_{l=n-j+1}^n \theta_l - (1 - \widehat{\alpha}_{n-j}^{\text{CL}}) \right) = \\ &= \mu_i \mu_j \left(\sum_{k=n-i+1}^n \theta_k - \sum_{k=n-i+1}^n \widehat{\theta}_k^{\text{MLE}} \right) \left(\sum_{l=n-j+1}^n \theta_l - \sum_{l=n-j+1}^n \widehat{\theta}_l^{\text{MLE}} \right). \end{aligned}$$

Also erhält man

$$\begin{aligned} \text{MSE}_{R_i+R_j|\mathcal{D}_n} \left(\widehat{R}_i^{\text{BF}'} + \widehat{R}_j^{\text{BF}'} \right) &= \sum_{l=i,j} \text{MSE}_{R_l|\mathcal{D}_n} \left(\widehat{R}_l^{\text{BF}'} \right) \\ &\quad + 2\mu_i \mu_j \left(\sum_{k=n-i+1}^n \theta_k - \sum_{k=n-i+1}^n \widehat{\theta}_k^{\text{MLE}} \right) \left(\sum_{l=n-j+1}^n \theta_l - \sum_{l=n-j+1}^n \widehat{\theta}_l^{\text{MLE}} \right). \end{aligned}$$

Indem eine mögliche Verzerrung vernachlässigt wird, wird der letzte Term wieder folgendermaßen geschätzt:

$$\Upsilon_{i,j} = \text{Cov} \left(\sum_{k=n-i+1}^n \widehat{\theta}_k^{\text{MLE}}, \sum_{l=n-j+1}^n \widehat{\theta}_l^{\text{MLE}} \right) = \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \text{Cov} \left(\widehat{\theta}_k^{\text{MLE}}, \widehat{\theta}_l^{\text{MLE}} \right)$$

Durch die selben Schritte wie in Kapitel 6.5.1 erhält man

$$\begin{aligned}
\Upsilon_{i,j} &= \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \text{Cov} \left(\widehat{\theta}_k^{\text{MLE}}, \widehat{\theta}_l^{\text{MLE}} \right) = \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \text{Cov} \left(\frac{1}{1 + \Delta_k}, \frac{1}{1 + \Delta_l} \right) \approx \\
&\approx \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \frac{1}{(1 + \delta_k)^2} \frac{1}{(1 + \delta_l)^2} \sum_{j \neq k} \sum_{m \neq l} \text{Cov} \left(\frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}}, \frac{\widehat{\theta}_m^{\text{GLM}}}{\widehat{\theta}_l^{\text{GLM}}} \right) = \\
&= \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \frac{1}{(1 + \delta_k)^2} \frac{1}{(1 + \delta_l)^2} \sum_{j \neq k} \sum_{m \neq l} \frac{\theta_j}{\theta_k} \frac{\theta_m}{\theta_l} \left(\widetilde{\Gamma}_j - \widetilde{\Gamma}_k \right) H \left(\widehat{\mathbf{b}} \right)^{-1} \left(\widetilde{\Gamma}_m - \widetilde{\Gamma}_l \right)'.
\end{aligned}$$

Das motiviert folgenden Schätzer:

$$\begin{aligned}
\widehat{\Upsilon}_{i,j} &= \widehat{\text{Cov}} \left(\sum_{k=n-i+1}^n \widehat{\theta}_k^{\text{MLE}}, \sum_{l=n-j+1}^n \widehat{\theta}_l^{\text{MLE}} \right) = \\
&= \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \frac{1}{(1 + \widehat{\delta}_k)^2} \frac{1}{(1 + \widehat{\delta}_l)^2} \\
&\quad \times \sum_{j \neq k} \sum_{m \neq l} \frac{\widehat{\theta}_j^{\text{GLM}}}{\widehat{\theta}_k^{\text{GLM}}} \frac{\widehat{\theta}_m^{\text{GLM}}}{\widehat{\theta}_l^{\text{GLM}}} \left(\widetilde{\Gamma}_j - \widetilde{\Gamma}_k \right) H \left(\widehat{\mathbf{b}} \right)^{-1} \left(\widetilde{\Gamma}_m - \widetilde{\Gamma}_l \right)'
\end{aligned}$$

Insgesamt ergibt sich unter den Modellannahmen 6.7 folgender Schätzer für die bedingte mittlere quadratische Abweichung für die Prognose der Gesamtreserve:

$$\widehat{\text{MSE}}_{\sum_{i=1}^n R_i | \mathcal{D}_n} \left(\sum_{i=1}^n \widehat{R}_i^{\text{BF}'} \right) = \sum_{i=1}^n \widehat{\text{MSE}}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{BF}'} \right) + 2 \sum_{1 \leq i < j \leq n} \widehat{\mu}_i \widehat{\mu}_j \widehat{\Upsilon}_{i,j}$$

6.5.3 Numerisches Beispiel

Es werden nun für die Daten aus Tabelle 2.1 Schätzer für die mittlere quadratische Abweichung der Reserven aus Tabelle 4.2 berechnet. Der Variationskoeffizient der Schätzer für die erwarteten Endschäden $\text{Vco}(\widehat{\mu}_i)$ wird mit 5% angenommen. Die Ergebnisse sind in Tabelle 6.1 dargestellt.

Tabelle 6.1: Mittlere quadratische Abweichung der Bornhuetter-Ferguson-Reserven

i	\hat{R}_i^{BF}	$\widehat{\text{MSE}}_{R_i \mathcal{D}_n}(\hat{R}_i^{\text{BF}})$	$\sqrt{\widehat{\text{MSE}}_{R_i \mathcal{D}_n}(\hat{R}_i^{\text{BF}})}$
1	155	280.779	530
2	849	1.200.250	1.096
3	3.268	4.250.248	2.062
4	4.281	5.295.367	2.301
5	6.524	7.754.943	2.785
6	8.598	9.987.875	3.160
7	25.642	27.075.680	5.203
8	92.966	104.680.628	10.231
	\hat{R}^{BF}	$\widehat{\text{MSE}}_{R \mathcal{D}_n}(\hat{R}^{\text{BF}})$	$\sqrt{\widehat{\text{MSE}}_{R \mathcal{D}_n}(\hat{R}^{\text{BF}})}$
	142.282	220.929.971	14.864

Kapitel 7

Benktander-Hovinen-Verfahren

7.1 Verfahren und Reserve

Dieses Kapitel lehnt sich an [6] und [13] an.

Die Benktander-Hovinen-Methode geht zurück auf Benktander (1976) und Hovinen (1981). Sie haben unabhängig von einander ein Verfahren entwickelt, das auf die selbe Prognose für den Gesamtschaden eines Jahres führt.

Bei dieser Methode werden wie beim Bornhuetter-Ferguson-Verfahren die Modellannahmen 4.2 unterstellt. Dabei ist μ_i ein A-priori-Schätzer für den erwarteten Gesamtaufwand $\mathbb{E}[S_{i,n}]$ des Anfallsjahres i und weiters geht man davon aus, dass das Abwicklungsmuster $(\alpha_k)_{0 \leq k \leq n}$ bekannt ist, wobei $\mathbb{E}[S_{i,k}] = \mu_i \alpha_k$ gilt.

Nachdem für die Bornhuetter-Ferguson-Reserve die aktuellste Beobachtung $S_{i,n-i}$ komplett ignoriert wird und beim Chain-Ladder-Verfahren davon ausgegangen wird, dass die aktuelle Beobachtung $S_{i,n-i}$ zuverlässig ist, um zukünftige Aufwendungen zu schätzen und $\hat{\mu}_i$, eine Schätzung des erwarteten Gesamtaufwandes des Abwicklungsjahres i , die nicht auf Beobachtungen beruht, komplett ignoriert wird, versucht man beim Benktander-Hovinen-Verfahren beide Werte miteinzubeziehen. Deshalb definiert man für $c \in [0, 1]$ und $1 \leq i \leq n$ die Credibility-Mischung

$$u_i(c) = c \hat{S}_{i,n}^{\text{CL}} + (1 - c) \mu_i.$$

Der Credibility-Faktor c sollte wachsen, je mehr Beobachtungen vorhanden sind. Benktander schlägt deshalb $c = \alpha_{n-i}$ und folgende Prognose für die Reserve R_i vor:

$$\hat{R}_i^{\text{BH}} = (1 - \alpha_{n-i}) \left(\alpha_{n-i} \hat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i}) \mu_i \right)$$

Daraus ergibt sich die Benktander-Hovinen-Prognose für den Gesamtaufwand des Abwicklungsjahres i , wobei $i \in \{1, \dots, n\}$:

$$\hat{S}_{i,n}^{\text{BH}} = S_{i,n-i} + (1 - \alpha_{n-i}) \left(\alpha_{n-i} \hat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i}) \mu_i \right)$$

Wenn wir die Modellannahmen 4.2 verwenden, kann das Abwicklungsmuster $(\alpha_k)_{0 \leq k \leq n}$ mit den Chain-Ladder-Faktoren $(f_k)_{1 \leq k \leq n}$ identifiziert werden. Für das restliche Kapitel nehmen wir

$$\alpha_k = \prod_{l=k+1}^n \frac{1}{f_l} \tag{7.1}$$

an. Nachdem die α_k bekannt sind, sind es die f_k für $k \in \{1, \dots, n\}$ ebenfalls. Es folgt unter Berücksichtigung von (4.5) und (4.6)

$$\begin{aligned}\widehat{S}_{i,n}^{\text{BH}} &= (1 - \alpha_{n-i} + \alpha_{n-i}) S_{i,n-i} + (1 - \alpha_{n-i}) \left(\alpha_{n-i} \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i}) \mu_i \right) = \\ &= (1 - \alpha_{n-i}) (S_{i,n-i} + (1 - \alpha_{n-i}) \mu_i) + \alpha_{n-i} \left(S_{i,n-i} + (1 - \alpha_{n-i}) \widehat{S}_{i,n}^{\text{CL}} \right) = \\ &= (1 - \alpha_{n-i}) \widehat{S}_{i,n}^{\text{BF}} + \alpha_{n-i} \widehat{S}_{i,n}^{\text{CL}} = (1 - \alpha_{n-i}) \widehat{S}_{i,n}^{\text{BF}} + S_{i,n-i}\end{aligned}\quad (7.2)$$

und

$$\begin{aligned}\widehat{R}_i^{\text{BH}} &= (1 - \alpha_{n-i}) \widehat{S}_{i,n}^{\text{BF}} = (1 - \alpha_{n-i}) \left(S_{i,n-i} + \widehat{R}_i^{\text{BF}} \right) = \\ &= S_{i,n-i} - \alpha_{n-i} S_{i,n-i} + (1 - \alpha_{n-i}) \widehat{R}_i^{\text{BF}} = \\ &= \alpha_{n-i} \widehat{S}_{i,n}^{\text{CL}} - \alpha_{n-i} S_{i,n-i} + (1 - \alpha_{n-i}) \widehat{R}_i^{\text{BF}} = \alpha_{n-i} \widehat{R}_i^{\text{CL}} + (1 - \alpha_{n-i}) \widehat{R}_i^{\text{BF}}.\end{aligned}$$

Gleichung (7.2) zeigt, dass die Benktander-Hovinen-Prognose als eine iterierte Bornhuetter-Ferguson-Prognose betrachtet werden kann.

Lemma 7.1: Unter den obigen Annahmen gilt

$$\widehat{S}_{i,n}^{\text{BH}} = u_i \left(1 - (1 - \alpha_{n-i})^2 \right).$$

Beweis: Es gilt

$$\begin{aligned}\widehat{S}_{i,n}^{\text{BH}} &= S_{i,n-i} + (1 - \alpha_{n-i}) \left(\alpha_{n-i} \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i}) \mu_i \right) = \\ &= \alpha_{n-i} \widehat{S}_{i,n}^{\text{CL}} + (\alpha_{n-i} - \alpha_{n-i}^2) \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i})^2 \mu_i = \\ &= \left(1 - (1 - \alpha_{n-i})^2 \right) \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i})^2 \mu_i = u_i \left(1 - (1 - \alpha_{n-i})^2 \right).\end{aligned}$$

□

Das folgende Lemma zeigt, dass man die Chain-Ladder-Reserve erhält, wenn man weiter iteriert.

Lemma 7.2: Unter den Annahmen, dass das Abwicklungsmuster $(\alpha_k)_{0 \leq k \leq n}$ bekannt ist und unter (7.1) gilt für $\alpha_{n-i} > 0$

$$\lim_{m \rightarrow \infty} \widehat{S}^{(m)} = \widehat{S}_{i,n}^{\text{CL}},$$

wobei

$$\begin{aligned}\widehat{S}^{(0)} &= \mu_i \quad \text{und} \\ \widehat{S}^{(m+1)} &= S_{i,n-i} + (1 - \alpha_{n-i}) \widehat{S}^{(m)} \quad \text{für } m \geq 0.\end{aligned}$$

Beweis: Durch Induktion wird gezeigt, dass für $m \geq 1$ folgende Gleichung gilt:

$$\widehat{S}^{(m)} = (1 - (1 - \alpha_{n-i})^m) \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i})^m \mu_i \quad (7.3)$$

- Induktionsanfang: $m = 1$
Es gilt

$$\widehat{S}^{(1)} = S_{i,n-i} + (1 - \alpha_{n-i}) \mu_i = \alpha_{n-i} \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i}) \mu_i.$$

- Induktionsschritt: $m \rightarrow m + 1$

Induktionsvoraussetzung: Für $m \geq 1$ gilt

$$\widehat{S}^{(m)} = (1 - (1 - \alpha_{n-i})^m) \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i})^m \mu_i.$$

Dann gilt

$$\begin{aligned} \widehat{S}^{(m+1)} &= S_{i,n-i} + (1 - \alpha_{n-i}) \widehat{S}^{(m)} = \\ &= S_{i,n-i} + (1 - \alpha_{n-i}) \left((1 - (1 - \alpha_{n-i})^m) \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i})^m \mu_i \right) = \\ &= \alpha_{n-i} \widehat{S}_{i,n}^{\text{CL}} + \left((1 - \alpha_{n-i}) - (1 - \alpha_{n-i})^{m+1} \right) \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i})^{m+1} \mu_i = \\ &= \left(1 - (1 - \alpha_{n-i})^{m+1} \right) \widehat{S}_{i,n}^{\text{CL}} + (1 - \alpha_{n-i})^{m+1} \mu_i. \end{aligned}$$

Somit ist (7.3) gezeigt. Daraus folgt die Aussage sofort für $m \rightarrow \infty$. □

7.2 Numerisches Beispiel

Es wird nun für die Daten aus Tabelle 2.1 das Benktander-Hovinen-Verfahren gerechnet. Dabei werden die Schätzer $\widehat{\mu}_i$ für die erwarteten Endschäden $\mathbb{E}[S_{i,n}]$ aus Tabelle 4.1 verwendet. Nachdem kein Abwicklungsmuster $(\widehat{\alpha}_k)_k$ vorgegeben wird, wird dieses mithilfe von (4.4) berechnet. Die Ergebnisse sind in Tabelle 7.1 dargestellt.

Kapitel 8

Cape-Cod-Verfahren

8.1 Verfahren und Reserve

Dieses Kapitel orientiert sich an [6] und [14].

Eine Schwachstelle des Chain-Ladder-Verfahrens ist, dass die Prognose des Gesamtschaden eines Anfallsjahres vollkommen von der letzten Beobachtung abhängt. Falls diese ein Ausreißer ist, überträgt sich das direkt auf den prognostizierten Gesamtschaden. Eine Möglichkeit, Ausreißer zu glätten, ist die Benktander-Hovinen- oder eben die Cape-Cod-Methode. Hierbei werden folgende Annahmen getroffen:

Modellannahmen 8.1:

- Die kumulativen Schadenhöhen $S_{i,k}$ für verschiedene Anfallsjahre i sind unabhängig.
- Es gibt Parameter $\pi_0, \dots, \pi_n > 0$, $\kappa > 0$ und ein Abwicklungsmuster $(\alpha_k)_{0 \leq k \leq n}$ mit $\alpha_n = 1$, sodass für $i \in \{0, \dots, n\}$

$$\mathbb{E}[S_{i,k}] = \kappa \pi_i \alpha_k.$$

Das Abwicklungsmuster $(\alpha_k)_{0 \leq k \leq n}$ wird für dieses Verfahren wieder als bekannt vorausgesetzt. Man beachte, dass die Modellannahmen 8.1 des Cape-Cod-Verfahrens mit jenen des Bornhuetter-Ferguson-Verfahrens 4.2 übereinstimmen, wenn man $\mu_i = \kappa \pi_i$ setzt. π_i kann also als Prämienvolumen des Anfallsjahres i und κ als durchschnittliche Endschadenquote interpretiert werden. Es wird angenommen, dass die Schadenquote unabhängig vom Prämienvolumen ist und deshalb für alle Jahre gleich ist und dass die Parameter π_0, \dots, π_n bekannt sind und der Parameter κ unbekannt ist.

Aus den Annahmen folgt sofort für $i \in \{0, \dots, n\}$

$$\kappa = \frac{\mathbb{E}[S_{i,n}]}{\pi_i}.$$

In diesem Kapitel wird wieder (7.1) unterstellt. Wir schätzen für jedes Anfallsjahr $i \in \{0, \dots, n\}$ die Schadenquote folgendermaßen:

$$\hat{\kappa}_i = \frac{\hat{S}_{i,n}^{\text{CL}}}{\pi_i} = \frac{S_{i,n-i} \prod_{k=n-i+1}^n f_k}{\pi_i} = \frac{S_{i,n-i}}{\alpha_{n-i} \pi_i} \quad (8.1)$$

κ_i ist ein unverzerrter Schätzer für κ , denn aus den Modellannahmen 8.1 folgt

$$\mathbb{E}[\widehat{\kappa}_i] = \mathbb{E}\left[\frac{S_{i,n-i}}{\alpha_{n-i}\pi_i}\right] = \frac{\mathbb{E}[S_{i,n-i}]}{\alpha_{n-i}\pi_i} = \kappa.$$

Aber nachdem der Schätzer auf dem aktuellen Schadenstand beruht, ist er nicht robust gegen Ausreißer. Um diese Empfindlichkeit gegenüber Ausreißern zu verringern, geht man zu einem gewichteten Mittel der Schätzer $\widehat{\kappa}_i$ über:

$$\widehat{\kappa}^{\text{CC}} = \sum_{i=0}^n \frac{\alpha_{n-i}\pi_i}{\sum_{j=0}^n \alpha_{n-j}\pi_j} \widehat{\kappa}_i = \frac{\sum_{i=0}^n S_{i,n-i}}{\sum_{i=0}^n \alpha_{n-i}\pi_i}$$

Man beachte, dass $\widehat{\kappa}^{\text{CC}}$ also ein unverzerrter Schätzer für κ ist. Wegen (8.1) gilt für $i \in \{0, \dots, n\}$ folgende Beziehung

$$S_{i,n-i} = \alpha_{n-i}\pi_i\widehat{\kappa}_i,$$

welche zusammen mit der Robustifizierung der $\widehat{\kappa}_i$ durch die Cape-Cod-Endschadenquote $\widehat{\kappa}^{\text{CC}}$ folgende Prognose für den Schadenstand $S_{i,n-i}$

$$\widehat{S}_{i,n-i}^{\text{CC}} = \alpha_{n-i}\pi_i\widehat{\kappa}^{\text{CC}}$$

und folgende Prognose für den Endschaden $S_{i,n}$ für $i \in \{1, \dots, n\}$

$$\widehat{S}_{i,n}^{\text{CC}} = S_{i,n-i} + \prod_{k=n-i+1}^n f_k \widehat{S}_{i,n-i}^{\text{CC}} - \widehat{S}_{i,n-i}^{\text{CC}}$$

nahelegt. Es gilt

$$\begin{aligned} \widehat{S}_{i,n}^{\text{CC}} &= S_{i,n-i} + \prod_{k=n-i+1}^n f_k \widehat{S}_{i,n-i}^{\text{CC}} - \widehat{S}_{i,n-i}^{\text{CC}} = S_{i,n-i} + \left(\prod_{k=n-i+1}^n f_k - 1 \right) \alpha_{n-i}\pi_i\widehat{\kappa}^{\text{CC}} = \\ &= S_{i,n-i} + (1 - \alpha_{n-i})\pi_i\widehat{\kappa}^{\text{CC}}. \end{aligned}$$

Die Cape-Cod-Prognose ist also vom selben Typ wie die Bornhuetter-Ferguson-Prognose, und zwar mit dem A-priori-Schätzer $\widehat{\kappa}^{\text{CC}}\pi_i$ für den erwarteten Endschaden. Die Reserve hat dann folgende Gestalt

$$\widehat{R}_i^{\text{CC}} = (1 - \alpha_{n-i})\pi_i\widehat{\kappa}^{\text{CC}}.$$

Man erhält folgendes Resultat:

Lemma 8.2: Unter den Modellannahmen 8.1 und der Identifikation (7.1) ist $\widehat{R}_i^{\text{CC}}$ ein unverzerrter Schätzer für $\mathbb{E}[R_i] = \mathbb{E}[S_{i,n} - S_{i,n-i}] = \kappa\pi_i(1 - \alpha_{n-i})$.

Beweis: Es gilt

$$\mathbb{E}[\widehat{R}_i^{\text{CC}}] = \mathbb{E}[(1 - \alpha_{n-i})\pi_i\widehat{\kappa}^{\text{CC}}] = (1 - \alpha_{n-i})\pi_i\mathbb{E}[\widehat{\kappa}^{\text{CC}}] = (1 - \alpha_{n-i})\pi_i\kappa.$$

□

8.2 Numerisches Beispiel

Es wird nun für die Daten aus Tabelle 2.1 das Cape-Cod-Verfahren gerechnet. Dabei werden die Prämien π_i der einzelnen Anfallsjahre aus Tabelle 8.1 verwendet. Nachdem kein Abwicklungsmuster $(\widehat{\alpha}_k)_k$ vorgegeben wird, wird dieses mithilfe von (4.4) berechnet. Die Ergebnisse sind in Tabelle 8.2 dargestellt.

Tabelle 8.1: Prämien nach Anfallsjahren

i	π_i
0	188.589
1	245.181
2	281.656
3	326.994
4	343.242
5	347.530
6	338.702
7	319.224
8	326.366

Kapitel 9

Credibility- Schadenreservierungsmethoden

Dieses Kapitel orientiert sich an [6] und [13].

In diesem Kapitel stellen wir uns folgende Frage: Was ist die optimale Mischung zwischen Chain-Ladder- und Bornhuetter-Ferguson-Prognose? Sei

$$\widehat{R}_i^c = c\widehat{R}_i^{\text{CL}} + (1 - c)\widehat{R}_i^{\text{BF}}$$

für $c \in [0, 1]$ eine Prognose für R_i , die Reserve des Anfallsjahres i , wobei \widehat{R}_i^c eine Konvexkombination der Chain-Ladder- und Bornhuetter-Ferguson-Reserve ist. Wir nehmen an, dass das Abwicklungsmuster $(\alpha_k)_k$ bekannt ist, (7.1) gilt und dass μ_i ein A-priori-Schätzer für den erwarteten Gesamtaufwand $\mathbb{E}[S_{i,n}]$ ist. Für $c = 1$ erhält man die Chain-Ladder-, für $c = 0$ die Bornhuetter-Ferguson- und für $c = \alpha_{n-i}$ die Benktander-Hovinen-Reserve. Unter Berücksichtigung von (4.5), (4.6) und (4.3) gilt

$$\begin{aligned}\widehat{R}_i^c &= c\widehat{R}_i^{\text{CL}} + (1 - c)\widehat{R}_i^{\text{BF}} = c\left(\widehat{S}_{i,n}^{\text{CL}} - S_{i,n-i}\right) + (1 - c)\left(\widehat{S}_{i,n}^{\text{BF}} - S_{i,n-i}\right) = \\ &= c\left(\widehat{S}_{i,n}^{\text{CL}} - \widehat{S}_{i,n}^{\text{BF}}\right) - S_{i,n-i} + \widehat{S}_{i,n}^{\text{BF}} = c(1 - \alpha_{n-i})\left(\widehat{S}_{i,n}^{\text{CL}} - \mu_i\right) + (1 - \alpha_{n-i})\mu_i = \\ &= (1 - \alpha_{n-i})\left(c\widehat{S}_{i,n}^{\text{CL}} + (1 - c)\mu_i\right) = (1 - \alpha_{n-i})u_i(c)\end{aligned}$$

Es stellt sich nun die Frage nach dem optimalen c , wobei Optimalität hier mithilfe der unbedingten mittleren quadratischen Abweichung definiert ist. Es soll also folgende Funktion minimiert werden:

$$\text{MSE}_{R_i}\left(\widehat{R}_i^c\right) = \mathbb{E}\left[\left(R_i - \widehat{R}_i^c\right)^2\right] \quad (9.1)$$

Für diese Minimierung wollen wir zuerst einige Annahmen treffen:

Modellannahmen 9.1:

- Die kumulativen Schadenhöhen $S_{i,k}$ für verschiedene Anfallsjahre i sind unabhängig.
- Es gibt eine Folge $(\alpha_k)_{0 \leq k \leq n}$ mit $\alpha_n = 1$, sodass für alle $k \in \{0, \dots, n\}$ gilt

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

- Es gibt Zufallsvariablen U_1, \dots, U_n , welche unverzerrte Schätzer für $\mathbb{E}[S_{i,n}]$ sind, wobei $\mathbb{E}[U_i] = \mathbb{E}[S_{i,n}]$ und U_i unabhängig von $S_{i,n-i}$ und $S_{i,n}$ ist.

Bemerkungen:

- Die Modellannahmen 9.1 sind mit den Modellannahmen 4.2 ident, falls $U_i = \mu_i > 0$ als deterministisch angenommen wird. Hier wird im Allgemeinen also nicht mehr angenommen, dass man einen bekannten deterministischen Wert für den a priori erwarteten Endschaden hat, sondern einen Schätzer, der unverzerrt ist. Es gibt also auch Unsicherheit bezüglich der Wahl des „richtigen“ Mittelwertes $\mathbb{E}[S_{i,n}]$.
- Man beachte, dass die Chain-Ladder-Modellannahmen 3.1 zwar diese Modellannahmen 9.1 erfüllen, aber nicht umgekehrt. Wir nehmen an, dass f_k durch α_k mithilfe von (7.1) definiert ist. Die Chain-Ladder- und Bornhuetter-Ferguson-Prognosen sind gegeben durch

$$\widehat{S}_{i,n}^{\text{CL}} = \frac{S_{i,n-i}}{\alpha_{n-i}} \quad \text{und} \quad \widehat{S}_{i,n}^{\text{BF}} = S_{i,n-i} + (1 - \alpha_{n-i})U_i \quad (9.2)$$

und die Credibility-gewichtete Reserve hat folgende Gestalt

$$\widehat{R}_i^c = (1 - \alpha_{n-i}) \left(c \widehat{S}_{i,n}^{\text{CL}} + (1 - c) U_i \right).$$

Lemma 9.2: Unter den Modellannahmen 9.1 und (9.2) ist der optimale Credibility-Faktor c_i^* , der die unbedingte mittlere quadratische Abweichung (9.1) minimiert, gegeben durch

$$c_i^* = \frac{\alpha_{n-i}}{1 - \alpha_{n-i}} \frac{\text{Cov}(S_{i,n-i}, R_i) + \alpha_{n-i} (1 - \alpha_{n-i}) \text{Var}(U_i)}{\text{Var}(S_{i,n-i}) + \alpha_{n-i}^2 \text{Var}(U_i)}$$

Beweis: Es gilt

$$\begin{aligned} \text{MSE}_{R_i}(\widehat{R}_i^{c_i}) &= \mathbb{E} \left[\left(\widehat{R}_i^{c_i} - R_i \right)^2 \right] = \mathbb{E} \left[\left(c_i \widehat{R}_i^{\text{CL}} + (1 - c_i) \widehat{R}_i^{\text{BF}} - R_i \right)^2 \right] = \\ &= \mathbb{E} \left[\left(c_i \left(\widehat{R}_i^{\text{CL}} - \widehat{R}_i^{\text{BF}} \right) + \widehat{R}_i^{\text{BF}} - R_i \right)^2 \right] = \\ &= c_i^2 \mathbb{E} \left[\left(\widehat{R}_i^{\text{CL}} - \widehat{R}_i^{\text{BF}} \right)^2 \right] - 2c_i \mathbb{E} \left[\left(\widehat{R}_i^{\text{CL}} - \widehat{R}_i^{\text{BF}} \right) \left(R_i - \widehat{R}_i^{\text{BF}} \right) \right] \\ &\quad + \mathbb{E} \left[\left(R_i - \widehat{R}_i^{\text{BF}} \right)^2 \right] \end{aligned}$$

und

$$\frac{\partial}{\partial c_i} \text{MSE}_{R_i}(\widehat{R}_i^{c_i}) = 2c_i \mathbb{E} \left[\left(\widehat{R}_i^{\text{CL}} - \widehat{R}_i^{\text{BF}} \right)^2 \right] - 2 \mathbb{E} \left[\left(\widehat{R}_i^{\text{CL}} - \widehat{R}_i^{\text{BF}} \right) \left(R_i - \widehat{R}_i^{\text{BF}} \right) \right] \stackrel{!}{=} 0.$$

Die zweite Ableitung ist positiv, d.h. durch Nullsetzen erhält man ein Minimum. Für c_i^* ergibt

sich unter Verwendung von (9.2)

$$\begin{aligned}
c_i^* &= \frac{\mathbb{E} \left[\left(\widehat{R}_i^{\text{CL}} - \widehat{R}_i^{\text{BF}} \right) \left(R_i - \widehat{R}_i^{\text{BF}} \right) \right]}{\mathbb{E} \left[\left(\widehat{R}_i^{\text{CL}} - \widehat{R}_i^{\text{BF}} \right)^2 \right]} = \\
&= \frac{\mathbb{E} \left[\left(\frac{S_{i,n-i}}{\alpha_{n-i}} - S_{i,n-i} - (1 - \alpha_{n-i}) U_i \right) \left(R_i - (1 - \alpha_{n-i}) U_i \right) \right]}{\mathbb{E} \left[\left(\frac{S_{i,n-i}}{\alpha_{n-i}} - S_{i,n-i} - (1 - \alpha_{n-i}) U_i \right)^2 \right]} = \\
&= \frac{\alpha_{n-i}}{1 - \alpha_{n-i}} \frac{\mathbb{E} \left[\left(S_{i,n-i} - \alpha_{n-i} U_i \right) \left(R_i - (1 - \alpha_{n-i}) U_i \right) \right]}{\mathbb{E} \left[\left(S_{i,n-i} - \alpha_{n-i} U_i \right)^2 \right]}.
\end{aligned}$$

Nachdem

$$\begin{aligned}
\mathbb{E} [S_{i,n-i} - \alpha_{n-i} U_i] &= \alpha_{n-i} \mathbb{E} [S_{i,n}] - \alpha_{n-i} \mathbb{E} [U_i] = \alpha_{n-i} \mathbb{E} [S_{i,n}] - \alpha_{n-i} \mathbb{E} [S_{i,n}] = 0, \\
\mathbb{E} [R_i - (1 - \alpha_{n-i}) U_i] &= \mathbb{E} [S_{i,n} - S_{i,n-i}] - (1 - \alpha_{n-i}) \mathbb{E} [U_i] = \\
&= \mathbb{E} [U_i] - \alpha_{n-i} \mathbb{E} [U_i] - (1 - \alpha_{n-i}) \mathbb{E} [U_i] = 0
\end{aligned}$$

und weil U_i unabhängig von $S_{i,n-i}$ und $S_{i,n}$ und deshalb auch von R_i ist, erhält man

$$\begin{aligned}
c_i^* &= \frac{\alpha_{n-i}}{1 - \alpha_{n-i}} \frac{\text{Cov} (S_{i,n-i} - \alpha_{n-i} U_i, R_i - (1 - \alpha_{n-i}) U_i)}{\text{Var} (S_{i,n-i} - \alpha_{n-i} U_i)} = \\
&= \frac{\alpha_{n-i}}{1 - \alpha_{n-i}} \frac{\text{Cov} (S_{i,n-i}, R_i) + \alpha_{n-i} (1 - \alpha_{n-i}) \text{Var} (U_i)}{\text{Var} (S_{i,n-i}) + \alpha_{n-i}^2 \text{Var} (U_i)}.
\end{aligned}$$

□

Bemerkungen:

- Man beachte, dass wir hier nicht den Schätzfehler des Abwicklungsmusters α_k und f_k miteinbezogen haben. In diesem Sinn liefert Lemma 9.2 das optimale Credibility-Gewicht unter Einbeziehung des stochastischen Fehlers und der Unsicherheit des A-priori-Schätzers U_i .
- Um c_i^* in Lemma 9.2 explizit zu berechnen, wird ein Modell für $\text{Cov} (S_{i,n-i}, R_i)$, $\text{Var} (U_i)$ und $\text{Var} (S_{i,n-i})$ benötigt.

Kapitel 10

Additives Verfahren

10.1 Verfahren und Reserve

Dieses Kapitel lehnt sich an [14], [2] und [1] an.

Bei diesem Verfahren werden die Zuwächse $Z_{i,k}$, $i, k \in \{0, \dots, n\}$, betrachtet und folgende Annahmen getroffen:

Modellannahmen 10.1:

- Die Zuwächse $Z_{i,k}$ sind für $i, k \in \{0, \dots, n\}$ unabhängig.
- Es gibt Parameter π_0, \dots, π_n , sodass für $i, k \in \{0, \dots, n\}$

$$\mathbb{E} \left[\frac{Z_{i,k}}{\pi_i} \right] = \zeta_k$$

gilt.

- Es gibt Varianzparameter $\sigma_0, \dots, \sigma_n$, sodass für $i, k \in \{0, \dots, n\}$

$$\text{Var} \left(\frac{Z_{i,k}}{\pi_i} \right) = \frac{\sigma_k^2}{\pi_i}$$

gilt.

Wir setzen voraus, dass die Parameter π_0, \dots, π_n bekannt und die Parameter ζ_0, \dots, ζ_n unbekannt sind. Hierbei stellen π_0, \dots, π_n Volumenmaße dar, π_i kann z.B. als Prämie des Anfallsjahres i interpretiert werden.

Für jedes Abwicklungsjahr wird der Parameter ζ_k folgendermaßen geschätzt:

$$\hat{\zeta}_k^{\text{AD}} = \frac{\sum_{i=0}^{n-k} Z_{i,k}}{\sum_{i=0}^{n-k} \pi_i} = \sum_{i=0}^{n-k} \frac{\pi_i}{\sum_{j=0}^{n-k} \pi_j} \frac{Z_{i,k}}{\pi_i}$$

Der Schätzer ist also ein gewichtetes Mittel der normierten beobachtbaren Zuwächse $Z_{i,k}/\pi_i$ des Abwicklungsjahres k .

Lemma 10.2: Unter den Modellannahmen 10.1 gilt

- a) $\widehat{\zeta}_k^{\text{AD}}$ ist ein erwartungstreuer Schätzer für ζ_k , d.h. es gilt $\mathbb{E}[\widehat{\zeta}_k^{\text{AD}}] = \zeta_k$.
- b) Für die Varianz von $\widehat{\zeta}_k^{\text{AD}}$ gilt

$$\text{Var}(\widehat{\zeta}_k^{\text{AD}}) = \frac{\sigma_k^2}{\sum_{i=0}^{n-k} \pi_i}.$$

- c) $\widehat{\zeta}_k^{\text{AD}}$ und $\widehat{\zeta}_j^{\text{AD}}$ sind für $k \neq j$ unabhängig.

Beweis:

- a) Es gilt

$$\mathbb{E}[\widehat{\zeta}_k^{\text{AD}}] = \mathbb{E}\left[\frac{\sum_{i=0}^{n-k} Z_{i,k}}{\sum_{i=0}^{n-k} \pi_i}\right] = \frac{1}{\sum_{i=0}^{n-k} \pi_i} \sum_{i=0}^{n-k} \mathbb{E}[Z_{i,k}] = \frac{1}{\sum_{i=0}^{n-k} \pi_i} \sum_{i=0}^{n-k} \zeta_k \pi_i = \zeta_k.$$

- b) Aufgrund der Unabhängigkeit der $Z_{i,k}$ gilt

$$\begin{aligned} \text{Var}(\widehat{\zeta}_k^{\text{AD}}) &= \text{Var}\left(\frac{\sum_{i=0}^{n-k} Z_{i,k}}{\sum_{i=0}^{n-k} \pi_i}\right) = \frac{1}{\left(\sum_{i=0}^{n-k} \pi_i\right)^2} \sum_{i=0}^{n-k} \text{Var}(Z_{i,k}) = \\ &= \frac{1}{\left(\sum_{i=0}^{n-k} \pi_i\right)^2} \sum_{i=0}^{n-k} \sigma_k^2 \pi_i = \frac{\sigma_k^2}{\sum_{i=0}^{n-k} \pi_i}. \end{aligned}$$

- c) Folgt unmittelbar aus der Definition von $\widehat{\zeta}_k^{\text{AD}}$ und der Unabhängigkeit der Inkremente. □

Der Varianzparameter σ_k^2 wird für $k \in \{0, \dots, n-1\}$ folgendermaßen geschätzt:

$$\widehat{\sigma}_k^2 = \frac{1}{n-k} \sum_{i=0}^{n-k} \pi_i \left(\frac{Z_{i,k}}{\pi_i} - \widehat{\zeta}_k^{\text{AD}}\right)^2$$

Als Motivation dafür gilt folgendes Lemma:

Lemma 10.3: Unter den Modellannahmen 10.1 ist $\widehat{\sigma}_k^2$ ein erwartungstreuer Schätzer für σ_k^2 , d.h. es gilt $\mathbb{E}[\widehat{\sigma}_k^2] = \sigma_k^2$.

Beweis: Es gilt

$$\mathbb{E}[\widehat{\sigma}_k^2] = \mathbb{E}\left[\frac{1}{n-k} \sum_{i=0}^{n-k} \pi_i \left(\frac{Z_{i,k}}{\pi_i} - \widehat{\zeta}_k^{\text{AD}}\right)^2\right] = \frac{1}{n-k} \sum_{i=0}^{n-k} \pi_i \mathbb{E}\left[\left(\frac{Z_{i,k}}{\pi_i} - \widehat{\zeta}_k^{\text{AD}}\right)^2\right] \quad (10.1)$$

und

$$\begin{aligned}
& \mathbb{E} \left[\left(\frac{Z_{i,k}}{\pi_i} - \widehat{\zeta}_k^{\text{AD}} \right)^2 \right] = \\
&= \mathbb{E} \left[\left(\frac{Z_{i,k}}{\pi_i} \right)^2 \right] + \mathbb{E} \left[\left(\frac{\sum_{j=0}^{n-k} Z_{j,k}}{\sum_{j=0}^{n-k} \pi_j} \right)^2 \right] - 2 \mathbb{E} \left[\frac{Z_{i,k}}{\pi_i} \frac{\sum_{j=0}^{n-k} Z_{j,k}}{\sum_{j=0}^{n-k} \pi_j} \right] = \\
&= \text{Var} \left(\frac{Z_{i,k}}{\pi_i} \right) + \mathbb{E} \left[\frac{Z_{i,k}}{\pi_i} \right]^2 + \frac{1}{\left(\sum_{j=0}^{n-k} \pi_j \right)^2} \sum_{j=0}^{n-k} \sum_{l=0}^{n-k} \mathbb{E} [Z_{j,k} Z_{l,k}] - \\
&\quad - \frac{2}{\pi_i \sum_{j=0}^{n-k} \pi_j} \sum_{j=0}^{n-k} \mathbb{E} [Z_{i,k} Z_{j,k}] = \\
&= \frac{\sigma_k^2}{\pi_i} + \zeta_k^2 + \frac{1}{\left(\sum_{j=0}^{n-k} \pi_j \right)^2} \left(\sum_{j=0}^{n-k} \sum_{\substack{l=0, \\ l \neq j}}^{n-k} \mathbb{E} [Z_{j,k}] \mathbb{E} [Z_{l,k}] + \sum_{j=0}^{n-k} \mathbb{E} [Z_{j,k}^2] \right) - \\
&\quad - \frac{2}{\pi_i \sum_{j=0}^{n-k} \pi_j} \left(\sum_{\substack{j=0, \\ j \neq i}}^{n-k} \mathbb{E} [Z_{i,k}] \mathbb{E} [Z_{j,k}] + \mathbb{E} [Z_{i,k}^2] \right) = \\
&= \frac{\sigma_k^2}{\pi_i} + \zeta_k^2 + \frac{1}{\left(\sum_{j=0}^{n-k} \pi_j \right)^2} \left(\sum_{j=0}^{n-k} \sum_{\substack{l=0, \\ l \neq j}}^{n-k} \pi_j \zeta_k \pi_l \zeta_k + \sum_{j=0}^{n-k} \left(\text{Var} (Z_{j,k}) + \mathbb{E} [Z_{j,k}^2] \right) \right) \\
&\quad - \frac{2}{\pi_i \sum_{j=0}^{n-k} \pi_j} \left(\sum_{\substack{j=0, \\ j \neq i}}^{n-k} \pi_i \zeta_k \pi_j \zeta_k + \text{Var} (Z_{i,k}) + \mathbb{E} [Z_{i,k}^2] \right) = \\
&= \frac{\sigma_k^2}{\pi_i} + \zeta_k^2 + \frac{1}{\left(\sum_{j=0}^{n-k} \pi_j \right)^2} \left(\sum_{j=0}^{n-k} \sum_{\substack{l=0, \\ l \neq j}}^{n-k} \pi_j \zeta_k \pi_l \zeta_k + \sum_{j=0}^{n-k} \left(\pi_j \sigma_k^2 + \pi_j^2 \zeta_k^2 \right) \right) - \\
&\quad - \frac{2}{\pi_i \sum_{j=0}^{n-k} \pi_j} \left(\sum_{\substack{j=0, \\ j \neq i}}^{n-k} \pi_i \zeta_k \pi_j \zeta_k + \pi_i \sigma_k^2 + \pi_i^2 \zeta_k^2 \right) = \\
&= \frac{\sigma_k^2}{\pi_i} + \zeta_k^2 + \frac{1}{\left(\sum_{j=0}^{n-k} \pi_j \right)^2} \left(\zeta_k^2 \left(\sum_{j=0}^{n-k} \pi_j \right)^2 + \sigma_k^2 \sum_{j=0}^{n-k} \pi_j \right) - \\
&\quad - \frac{2}{\pi_i \sum_{j=0}^{n-k} \pi_j} \left(\zeta_k^2 \pi_i \sum_{j=0}^{n-k} \pi_j + \pi_i \sigma_k^2 \right) = \\
&= \frac{\sigma_k^2}{\pi_i} + \zeta_k^2 + \zeta_k^2 + \frac{\sigma_k^2}{\sum_{j=0}^{n-k} \pi_j} - 2\zeta_k^2 - \frac{2\sigma_k^2}{\sum_{j=0}^{n-k} \pi_j} = \\
&= \frac{\sigma_k^2}{\pi_i} - \frac{\sigma_k^2}{\sum_{j=0}^{n-k} \pi_j}.
\end{aligned}$$

Setzt man das in (10.1) ein, so folgt

$$\mathbb{E}[\hat{\sigma}_k^2] = \frac{1}{n-k} \sum_{i=0}^{n-k} \pi_i \left(\frac{\sigma_k^2}{\pi_i} - \frac{\sigma_k^2}{\sum_{j=0}^{n-k} \pi_j} \right) = \frac{n-k+1}{n-k} \sigma_k^2 - \frac{1}{n-k} \sigma_k^2 = \sigma_k^2.$$

□

$\hat{\sigma}_n^2$ muss wieder wie in Kapitel 3.2.1 mit anderen Methoden geschätzt werden. Für jedes Anfallsjahr $i \in \{1, \dots, n\}$ und jedes Abwicklungsjahr $k \in \{n-i+1, \dots, n\}$ wird der zukünftige Schadenstand $S_{i,k}$ folgendermaßen prognostiziert:

$$\hat{S}_{i,k}^{\text{AD}} = S_{i,n-i} + \pi_i \sum_{j=n-i+1}^k \hat{\zeta}_j^{\text{AD}}$$

Lemma 10.4: Unter den Modellannahmen 10.1 ist $\hat{S}_{i,k}^{\text{AD}}$ für $i \in \{1, \dots, n\}$ und $k \in \{n-i+1, \dots, n\}$ ein erwartungstreuer Schätzer für $\mathbb{E}[S_{i,k}]$, d.h. es gilt $\mathbb{E}[\hat{S}_{i,k}^{\text{AD}}] = \mathbb{E}[S_{i,k}]$.

Beweis: Mithilfe von Lemma 10.2 a) erhält man

$$\begin{aligned} \mathbb{E}[\hat{S}_{i,k}^{\text{AD}}] &= \mathbb{E} \left[S_{i,n-i} + \pi_i \sum_{j=n-i+1}^k \hat{\zeta}_j^{\text{AD}} \right] = \mathbb{E}[S_{i,n-i}] + \pi_i \sum_{j=n-i+1}^k \mathbb{E}[\hat{\zeta}_j^{\text{AD}}] = \\ &= \mathbb{E}[S_{i,n-i}] + \pi_i \sum_{j=n-i+1}^k \zeta_j = \mathbb{E}[S_{i,n-i}] + \sum_{j=n-i+1}^k \mathbb{E}[Z_{i,j}] = \\ &= \mathbb{E} \left[S_{i,n-i} + \sum_{j=n-i+1}^k Z_{i,j} \right] = \mathbb{E}[S_{i,k}]. \end{aligned}$$

□

Somit ergeben sich im Additiven Verfahren folgende Prognosen für die Reserven:

$$\hat{R}_i^{\text{AD}} = \hat{S}_{i,n}^{\text{AD}} - S_{i,n-i} = \pi_i \sum_{j=n-i+1}^n \hat{\zeta}_j^{\text{AD}} \quad (10.2)$$

$$\hat{R}^{\text{AD}} = \sum_{i=1}^n \hat{R}_i^{\text{AD}} \quad (10.3)$$

Aus Lemma 10.4 folgt sofort, dass auch \hat{R}_i^{AD} ein erwartungstreuer Schätzer für $\mathbb{E}[R_i]$ ist.

Im Allgemeinen gilt für eine Reserve R_i wegen (A.1)

$$\text{MSE}_{R_i|\mathcal{D}_n}(\hat{R}_i) = \mathbb{E} \left[\left(\hat{R}_i - R_i \right)^2 \middle| \mathcal{D}_n \right] = \text{Var}(R_i | \mathcal{D}_n) + \left(\hat{R}_i - \mathbb{E}[R_i | \mathcal{D}_n] \right)^2$$

Die mittlere quadratische Abweichung ist also minimal für $\hat{R}_i = \mathbb{E}[R_i | \mathcal{D}_n]$. Nachdem die Inkremente unabhängig sind, gilt in unserem Fall

$$\mathbb{E}[R_i | \mathcal{D}_n] = \mathbb{E}[Z_{i,n-i+1} + \dots + Z_{i,n} | \mathcal{D}_n] = \mathbb{E}[Z_{i,n-i+1} + \dots + Z_{i,n}] = \pi_i \sum_{j=n-i+1}^n \zeta_j$$

Ersetzt man nun ζ_j durch $\widehat{\zeta}_j^{\text{AD}}$, so erhält man die Reserve $\widehat{R}_i^{\text{AD}}$ aus (10.2) und erkennt, dass diese in einem gewissen Sinne optimal ist.

Definiert man

$$\widehat{\alpha}_k^{\text{AD}} := \frac{\sum_{j=0}^k \widehat{\zeta}_j^{\text{AD}}}{\sum_{j=0}^n \widehat{\zeta}_j^{\text{AD}}},$$

$$\widehat{\mu}_i^{\text{AD}} := \pi_i \sum_{j=0}^n \widehat{\zeta}_j^{\text{AD}}$$

so erhält man

$$\widehat{R}_i^{\text{AD}} = \pi_i \sum_{j=n-i+1}^n \widehat{\zeta}_j^{\text{AD}} = (1 - \widehat{\alpha}_{n-i}^{\text{AD}}) \widehat{\mu}_i^{\text{AD}},$$

d.h. die Reserve des Additiven Verfahrens hat die selbe Form wie die des Bornhuetter-Ferguson-Verfahrens.

Definiert man weiters

$$\widehat{\kappa}^{\text{AD}} := \sum_{j=0}^n \widehat{\zeta}_j^{\text{AD}},$$

so erhält man

$$\widehat{R}_i^{\text{AD}} = \pi_i \sum_{j=n-i+1}^n \widehat{\zeta}_j^{\text{AD}} = (1 - \widehat{\alpha}_{n-i}^{\text{AD}}) \pi_i \widehat{\kappa}^{\text{AD}},$$

d.h. die Reserve des Additiven Verfahrens hat die selbe Form wie die des Cape-Cod-Verfahrens.

10.2 Mittlere quadratische Abweichung der Prognose

Dieses Kapitel orientiert sich an [15]. Es sollen Schätzer für die mittlere quadratische Abweichung von $\widehat{R}_i^{\text{AD}}$ und \widehat{R}^{AD} gefunden werden.

10.2.1 Mittlere quadratische Abweichung der Prognose $\widehat{R}_i^{\text{AD}}$ für ein Abwicklungsjahr i

Wegen (A.1) gilt

$$\text{MSE}_{R_i|\mathcal{D}_n}(\widehat{R}_i^{\text{AD}}) = \text{Var}(R_i|\mathcal{D}_n) + \left(\widehat{R}_i^{\text{AD}} - \mathbb{E}[R_i|\mathcal{D}_n]\right)^2.$$

Es müssen also Schätzer für den stochastischen Fehler und den Schätzfehler gefunden werden. Für den stochastischen Fehler gilt aufgrund der Unabhängigkeit der Inkremente

$$\begin{aligned} \text{Var}(R_i|\mathcal{D}_n) &= \text{Var}\left(\sum_{k=n-i+1}^n Z_{i,k} \middle| \mathcal{D}_n\right) = \text{Var}\left(\sum_{k=n-i+1}^n Z_{i,k}\right) = \sum_{k=n-i+1}^n \text{Var}(Z_{i,k}) = \\ &= \pi_i \sum_{k=n-i+1}^n \sigma_k^2. \end{aligned}$$

Um den stochastischen Fehler zu schätzen, ersetzen wir nun alle σ_k^2 durch die Schätzer $\hat{\sigma}_k^2$ und erhalten

$$\widehat{\text{Var}}(R_i | \mathcal{D}_n) = \pi_i \sum_{k=n-i+1}^n \hat{\sigma}_k^2.$$

Um auf einen Schätzer für den Schätzfehler $\left(\hat{R}_i^{\text{AD}} - \mathbb{E}[R_i | \mathcal{D}_n]\right)^2$ zu gelangen, ersetzen wir diesen durch $\mathbb{E}\left[\left(\hat{R}_i^{\text{AD}} - \mathbb{E}[R_i | \mathcal{D}_n]\right)^2\right]$. Dann folgt mithilfe von Lemma 10.2 c) und b)

$$\begin{aligned} \mathbb{E}\left[\left(\hat{R}_i^{\text{AD}} - \mathbb{E}[R_i | \mathcal{D}_n]\right)^2\right] &= \mathbb{E}\left[\left(\hat{R}_i^{\text{AD}} - \mathbb{E}[R_i]\right)^2\right] = \text{Var}\left(\hat{R}_i^{\text{AD}}\right) = \\ &= \text{Var}\left(\pi_i \sum_{k=n-i+1}^n \hat{\zeta}_k^{\text{AD}}\right) = \pi_i^2 \sum_{k=n-i+1}^n \text{Var}\left(\hat{\zeta}_k^{\text{AD}}\right) = \\ &= \pi_i^2 \sum_{k=n-i+1}^n \frac{\sigma_k^2}{\sum_{l=0}^{n-k} \pi_l}. \end{aligned}$$

Ersetzt man auch hier die σ_k^2 wieder durch $\hat{\sigma}_k^2$, so erhält man folgenden Schätzer für die mittlere quadratische Abweichung von \hat{R}_i^{AD}

$$\widehat{\text{MSE}}_{R_i | \mathcal{D}_n}(\hat{R}_i^{\text{AD}}) = \pi_i \sum_{k=n-i+1}^n \hat{\sigma}_k^2 + \pi_i^2 \sum_{k=n-i+1}^n \frac{\hat{\sigma}_k^2}{\sum_{l=0}^{n-k} \pi_l}.$$

10.2.2 Mittlere quadratische Abweichung der Prognose \hat{R}^{AD}

Ziel dieses Kapitels ist es, einen Schätzer für die bedingte mittlere quadratische Abweichung von

$$\hat{R}^{\text{AD}} = \sum_{i=1}^n \hat{R}_i^{\text{AD}}$$

zu finden. Wir starten mit zwei Anfallsjahren $i < j$. Dann gilt aufgrund der Unabhängigkeit der Inkremente

$$\begin{aligned} \text{MSE}_{R_i+R_j | \mathcal{D}_n}(\hat{R}_i^{\text{AD}} + \hat{R}_j^{\text{AD}}) &= \mathbb{E}\left[\left(\hat{R}_i^{\text{AD}} + \hat{R}_j^{\text{AD}} - R_i - R_j\right)^2 \middle| \mathcal{D}_n\right] = \\ &= \mathbb{E}\left[\left(\hat{R}_i^{\text{AD}} - R_i\right)^2 \middle| \mathcal{D}_n\right] + \mathbb{E}\left[\left(\hat{R}_j^{\text{AD}} - R_j\right)^2 \middle| \mathcal{D}_n\right] + 2 \mathbb{E}\left[\left(\hat{R}_i^{\text{AD}} - R_i\right)\left(\hat{R}_j^{\text{AD}} - R_j\right) \middle| \mathcal{D}_n\right] = \\ &= \text{MSE}_{R_i | \mathcal{D}_n}(\hat{R}_i^{\text{AD}}) + \text{MSE}_{R_j | \mathcal{D}_n}(\hat{R}_j^{\text{AD}}) + \\ &\quad + 2\left(\hat{R}_i^{\text{AD}} \hat{R}_j^{\text{AD}} - \hat{R}_i^{\text{AD}} \mathbb{E}[R_j] - \hat{R}_j^{\text{AD}} \mathbb{E}[R_i] + \mathbb{E}[R_i] \mathbb{E}[R_j]\right) = \\ &= \text{MSE}_{R_i | \mathcal{D}_n}(\hat{R}_i^{\text{AD}}) + \text{MSE}_{R_j | \mathcal{D}_n}(\hat{R}_j^{\text{AD}}) + 2\left(\hat{R}_i^{\text{AD}} - \mathbb{E}[R_i]\right)\left(\hat{R}_j^{\text{AD}} - \mathbb{E}[R_j]\right). \end{aligned}$$

Der letzte Term wird mithilfe von $\mathbb{E} \left[\left(\widehat{R}_i^{\text{AD}} - \mathbb{E}[R_i] \right) \left(\widehat{R}_j^{\text{AD}} - \mathbb{E}[R_j] \right) \right]$ geschätzt. Dann gilt unter Verwendung von Lemma 10.2 b)

$$\begin{aligned}
& \mathbb{E} \left[\left(\widehat{R}_i^{\text{AD}} - \mathbb{E}[R_i] \right) \left(\widehat{R}_j^{\text{AD}} - \mathbb{E}[R_j] \right) \right] = \\
&= \mathbb{E} \left[\pi_i \left(\sum_{k=n-i+1}^n \widehat{\zeta}_k^{\text{AD}} - \sum_{k=n-i+1}^n \zeta_k \right) \pi_j \left(\sum_{l=n-j+1}^n \widehat{\zeta}_l^{\text{AD}} - \sum_{l=n-j+1}^n \zeta_l \right) \right] = \\
&= \pi_i \pi_j \mathbb{E} \left[\sum_{k=n-i+1}^n \left(\widehat{\zeta}_k^{\text{AD}} - \zeta_k \right) \sum_{l=n-j+1}^n \left(\widehat{\zeta}_l^{\text{AD}} - \zeta_l \right) \right] = \\
&= \pi_i \pi_j \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \mathbb{E} \left[\left(\widehat{\zeta}_k^{\text{AD}} - \zeta_k \right) \left(\widehat{\zeta}_l^{\text{AD}} - \zeta_l \right) \right] = \\
&= \pi_i \pi_j \sum_{k=n-i+1}^n \sum_{l=n-j+1}^n \text{Cov} \left(\widehat{\zeta}_k^{\text{AD}}, \widehat{\zeta}_l^{\text{AD}} \right) = \\
&= \pi_i \pi_j \sum_{k=n-i+1}^n \text{Var} \left(\widehat{\zeta}_k^{\text{AD}} \right) = \\
&= \pi_i \pi_j \sum_{k=n-i+1}^n \frac{\sigma_k^2}{\sum_{l=0}^{n-k} \pi_l}
\end{aligned}$$

Ersetzt man nun wieder σ_k^2 durch $\widehat{\sigma}_k^2$, so erhält man folgenden Schätzer für die mittlere quadratische Abweichung von \widehat{R}^{AD} :

$$\widehat{\text{MSE}}_{\sum_i R_i | \mathcal{D}_n} \left(\sum_{i=1}^n \widehat{R}_i^{\text{AD}} \right) = \sum_{i=1}^n \widehat{\text{MSE}}_{R_i | \mathcal{D}_n} \left(\widehat{R}_i^{\text{AD}} \right) + 2 \sum_{1 \leq i < j \leq n} \pi_i \pi_j \sum_{k=n-i+1}^n \frac{\widehat{\sigma}_k^2}{\sum_{l=0}^{n-k} \pi_l}$$

10.3 Numerisches Beispiel

Es wird nun für die Daten aus Tabelle 2.1 das Additive Verfahren gerechnet und die zugehörige mittlere quadratische Abweichung der Prognose berechnet. Dabei werden die Prämien π_i der einzelnen Anfallsjahre aus Tabelle 8.1 verwendet. $\widehat{\sigma}_n^2$ wird mit (3.8) geschätzt. Ergebnisse sind in Tabelle 10.1 und 10.2 dargestellt.

Tabelle 10.1: Additives Verfahren (Schadenzuwächse)

	Abwicklungsjahr k									
	0	1	2	3	4	5	6	7	8	$\widehat{R}_i^{\text{AD}}$
$\widehat{\zeta}_k$	0,4883	0,2218	0,0497	0,0050	0,0058	0,0033	0,0069	0,0026	0,0007	
i										
0	115.160	50.863	10.889	1.778	1.381	626	1.614	269	129	168
1	122.376	62.527	14.584	167	1.515	316	580	876	168	936
2	133.328	71.197	14.682	1.830	3.611	1.884	2.762	743	193	3.352
3	155.997	80.811	11.145	3.948	211	656	2.265	863	224	4.665
4	169.121	61.516	17.814	469	1.337	1.147	2.378	906	235	6.744
5	166.265	64.031	18.615	434	2.020	1.161	2.407	917	238	8.258
6	161.851	65.616	15.344	1.686	1.969	1.131	2.346	894	232	23.664
7	152.112	73.728	15.881	1.589	1.856	1.066	2.211	843	218	96.573
8	150.683	72.380	16.236	1.624	1.897	1.090	2.261	861	223	
										\widehat{R}^{AD} 144.360

Tabelle 10.2: Reserven und mittlere quadratische Abweichung der Reserven des Additiven Verfahrens

i	\hat{R}_i^{AD}	$\text{MSE}_{R_i \mathcal{D}_n}(\hat{R}_i^{\text{AD}})$	$\sqrt{\text{MSE}_{R_i \mathcal{D}_n}(\hat{R}_i^{\text{AD}})}$
1	168	34.272	185
2	936	270.834	520
3	3.352	2.227.072	1.492
4	4.665	3.102.609	1.761
5	6.744	5.793.617	2.407
6	8.258	8.416.571	2.901
7	23.664	15.837.453	3.980
8	96.573	153.476.475	12.389
	\hat{R}^{AD}	$\text{MSE}_{R \mathcal{D}_n}(\hat{R}^{\text{AD}})$	$\sqrt{\text{MSE}_{R \mathcal{D}_n}(\hat{R}^{\text{AD}})}$
	144.360	228.568.315	15.118

Kapitel 11

Munich-Chain-Ladder-Verfahren

11.1 Munich-Chain-Ladder-Modell

Dieses Kapitel orientiert sich an [16], [6] und [4].

Das Munich-Chain-Ladder-Modell basiert auf dem verteilungsfreien Chain-Ladder-Modell 3.4 und verallgemeinert es.

Normalerweise wird für die Prognose der Chain-Ladder-Reserve für ein Portfolio sowohl das Abwicklungsdreieck der Schadenzahlungen (Paid) als auch der Schadenaufwendungen (Incurred), d.h. der Summe aus Schadenzahlungen und Einzelfallreserven, herangezogen. Dabei entsteht aber oft das Problem, dass die Paid- und die Incurred-Prognose weit auseinander liegen. Es kann auch von Anfallsjahr zu Anfallsjahr wechseln, welcher Typ die höhere Endschadenprognose liefert.

Mit **(P/I)-Stand** wird im Folgenden der Quotient aus Schadenzahlungen und Schadenaufwendungen des i -ten Anfallsjahres bis zum k -ten Abwicklungsjahr bezeichnet:

$$Q_{i,k} = \frac{S_{i,k}^P}{S_{i,k}^I} \quad \text{und} \quad Q_{i,k}^{-1} = \frac{S_{i,k}^I}{S_{i,k}^P}$$

Daten (siehe z.B. [16]) zeigen, dass auf einen relativ niedrigen (P/I)-Stand entweder relativ hohe Abwicklungsfaktoren bei Paid oder relativ niedrige Abwicklungsfaktoren bei Incurred folgen und umgekehrt analog. Dieser grundlegende Zusammenhang wird beim Chain-Ladder-Verfahren komplett ignoriert. Das Munich-Chain-Ladder-Verfahren kombiniert die Datentypen Paid (P) und Incurred (I), indem es die (P/I)-Stände in die Prognoserechnung miteinbezieht: Je nachdem, ob der aktuelle (P/I)-Stand über- bzw. unterdurchschnittlich ist, sollte ein unter- bzw. überdurchschnittlicher Paid-Abwicklungsfaktor bzw. ein über- bzw. unterdurchschnittlicher Incurred-Abwicklungsfaktor verwendet werden.

Es stehen folgende Daten zur Verfügung:

$$\begin{aligned} \mathcal{D}_n^P &= \sigma\{S_{i,k}^P; 0 \leq i \leq n, 0 \leq k \leq n, 0 \leq i+k \leq n\} \\ \mathcal{D}_n^I &= \sigma\{S_{i,k}^I; 0 \leq i \leq n, 0 \leq k \leq n, 0 \leq i+k \leq n\} \\ \tilde{\mathcal{D}}_n &= \sigma\{S_{i,k}^P, S_{i,k}^I; 0 \leq i \leq n, 0 \leq k \leq n, 0 \leq i+k \leq n\} \end{aligned}$$

Man beachte, dass sich $\tilde{\mathcal{D}}_n$ von \mathcal{D}_n aus den vorherigen Kapiteln unterscheidet. Weiters sei

$$\begin{aligned}\mathcal{B}_k^P &= \sigma\{S_{i,j}^P; 0 \leq j \leq k, 0 \leq i+j \leq n\}, \\ \mathcal{B}_k^I &= \sigma\{S_{i,j}^I; 0 \leq j \leq k, 0 \leq i+j \leq n\}, \\ \tilde{\mathcal{B}}_k &= \sigma\{S_{i,j}^P, S_{i,j}^I; 0 \leq j \leq k, 0 \leq i+j \leq n\}\end{aligned}$$

Für das Munich-Chain-Ladder-Modell werden folgende Annahmen getroffen:

Modellannahmen 11.1:

- Die kumulierten Zahlungen $S_{i,k}^P$ verschiedener Anfallsjahre sind unabhängig. Die kumulierten Schadenaufwendungen $S_{i,k}^I$ verschiedener Anfallsjahre sind unabhängig.
- Es gibt Faktoren $f_1^P, \dots, f_n^P > 0$, $f_1^I, \dots, f_n^I > 0$ und Varianz-Parameter $\sigma_1^P, \dots, \sigma_n^P > 0$, $\sigma_1^I, \dots, \sigma_n^I > 0$, sodass für alle $i \in \{0, \dots, n\}$ und $k \in \{1, \dots, n\}$ gilt

$$\begin{aligned}\mathbb{E}[S_{i,k}^P | \mathcal{B}_{k-1}^P] &= f_k^P S_{i,k-1}^P, & \text{Var}(S_{i,k}^P | \mathcal{B}_{k-1}^P) &= (\sigma_k^P)^2 S_{i,k-1}^P, \\ \mathbb{E}[S_{i,k}^I | \mathcal{B}_{k-1}^I] &= f_k^I S_{i,k-1}^I, & \text{Var}(S_{i,k}^I | \mathcal{B}_{k-1}^I) &= (\sigma_k^I)^2 S_{i,k-1}^I.\end{aligned}$$

- Es gibt Konstanten λ^P, λ^I , sodass für alle $i \in \{0, \dots, n\}$ und $k \in \{1, \dots, n\}$

$$\mathbb{E}\left[\frac{S_{i,k}^P}{S_{i,k-1}^P} \middle| \tilde{\mathcal{B}}_{k-1}\right] = f_k^P + \lambda^P \text{Var}\left(\frac{S_{i,k}^P}{S_{i,k-1}^P} \middle| \mathcal{B}_{k-1}^P\right)^{1/2} \cdot \frac{Q_{i,k-1}^{-1} - \mathbb{E}[Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P]}{\text{Var}(Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P)^{1/2}} \quad (11.1)$$

und

$$\mathbb{E}\left[\frac{S_{i,k}^I}{S_{i,k-1}^I} \middle| \tilde{\mathcal{B}}_{k-1}\right] = f_k^I + \lambda^I \text{Var}\left(\frac{S_{i,k}^I}{S_{i,k-1}^I} \middle| \mathcal{B}_{k-1}^I\right)^{1/2} \cdot \frac{Q_{i,k-1} - \mathbb{E}[Q_{i,k-1} | \mathcal{B}_{k-1}^I]}{\text{Var}(Q_{i,k-1} | \mathcal{B}_{k-1}^I)^{1/2}} \quad (11.2)$$

gilt.

- Die Schadenstände unterschiedlicher Anfallsjahre sind unabhängig, d.h. die Mengen

$$\{S_{0,k}^P, S_{0,k}^I; k = 0, \dots, n\}, \dots, \{S_{n,k}^P, S_{n,k}^I; k = 0, \dots, n\}$$

sind unabhängig.

Bemerkungen:

- Die ersten beiden Punkte in den Modellannahmen 11.1 sind die klassischen Chain-Ladder-Annahmen für kumulierte Schadenzahlungen bzw. kumulierte Schadenaufwendungen. Die zusätzlichen Annahmen im dritten Punkt widersprechen diesen nicht, denn

$$\mathbb{E}\left[\frac{S_{i,k}^P}{S_{i,k-1}^P} \middle| \mathcal{B}_{k-1}^P\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{S_{i,k}^P}{S_{i,k-1}^P} \middle| \tilde{\mathcal{B}}_{k-1}\right] \middle| \mathcal{B}_{k-1}^P\right] = f_k^P$$

und analog für die Schadenaufwendungen.

- Gleichung (11.2) stellt den bedingten Erwartungswert als Summe des gewöhnlichen Abwicklungsfaktors f_k^I und eines in $Q_{i,k-1}$ linearen Korrekturterms dar. Analog für (11.1).

- Die Parameter λ^P und λ^I sind bedingte Korrelationskoeffizienten, denn aus (11.1) folgt für $i \in \{0, \dots, n\}$ und $k \in \{1, \dots, n\}$

$$\mathbb{E} \left[\frac{\frac{S_{i,k}^P}{S_{i,k-1}^P} - \mathbb{E} \left[\frac{S_{i,k}^P}{S_{i,k-1}^P} \middle| \mathcal{B}_{k-1}^P \right]}{\text{Var} \left(\frac{S_{i,k}^P}{S_{i,k-1}^P} \middle| \mathcal{B}_{k-1}^P \right)^{1/2}} \middle| \tilde{\mathcal{B}}_{k-1} \right] = \lambda^P \frac{Q_{i,k-1}^{-1} - \mathbb{E} \left[Q_{i,k-1}^{-1} \middle| \mathcal{B}_{k-1}^P \right]}{\text{Var} \left(Q_{i,k-1}^{-1} \middle| \mathcal{B}_{k-1}^P \right)^{1/2}}. \quad (11.3)$$

Multipliziert man nun beide Seiten mit dem $\tilde{\mathcal{B}}_{k-1}$ -messbaren Ausdruck

$$\frac{Q_{i,k-1}^{-1} - \mathbb{E} \left[Q_{i,k-1}^{-1} \middle| \mathcal{B}_{k-1}^P \right]}{\text{Var} \left(Q_{i,k-1}^{-1} \middle| \mathcal{B}_{k-1}^P \right)^{1/2}}$$

und nimmt auf beiden Seiten die bedingte Erwartung $\mathbb{E} [\cdot | \mathcal{B}_{k-1}^P]$, so erhält man

$$\text{Cor} \left(Q_{i,k-1}^{-1}, \frac{S_{i,k}^P}{S_{i,k-1}^P} \middle| \mathcal{B}_{k-1}^P \right) = \text{Cor} (S_{i,k-1}^I, S_{i,k}^P | \mathcal{B}_{k-1}^P) = \lambda^P,$$

d.h. es wird angenommen, dass die bedingten Korrelationskoeffizienten unabhängig vom Abwicklungsjahr k sind und nicht von den Beobachtungen \mathcal{B}_{k-1}^P abhängen. Analog für λ^I .

Gemäß den Überlegungen zu Beginn des Kapitels wird im Allgemeinen $\lambda^P, \lambda^I \geq 0$ sein. Falls der bedingte Korrelationskoeffizient klein ist, wird der Korrekturterm auch klein sein und die Munich-Chain-Ladder-Prognose wird sich nur wenig von der gewöhnlichen Chain-Ladder-Prognose unterscheiden.

- Gleichung (11.1) ist äquivalent zu

$$\mathbb{E} \left[S_{i,k}^P \middle| \tilde{\mathcal{B}}_{k-1} \right] = f_k^P S_{i,k-1}^P + \lambda^P \text{Var} (S_{i,k}^P | \mathcal{B}_{k-1}^P)^{1/2} \frac{S_{i,k-1}^I - \mathbb{E} \left[S_{i,k-1}^I \middle| \mathcal{B}_{k-1}^P \right]}{\text{Var} \left(S_{i,k-1}^I \middle| \mathcal{B}_{k-1}^P \right)^{1/2}}$$

und (11.2) zu

$$\mathbb{E} \left[S_{i,k}^I \middle| \tilde{\mathcal{B}}_{k-1} \right] = f_k^I S_{i,k-1}^I + \lambda^I \text{Var} (S_{i,k}^I | \mathcal{B}_{k-1}^I)^{1/2} \frac{S_{i,k-1}^P - \mathbb{E} \left[S_{i,k-1}^P \middle| \mathcal{B}_{k-1}^I \right]}{\text{Var} \left(S_{i,k-1}^P \middle| \mathcal{B}_{k-1}^I \right)^{1/2}}.$$

- Der Standardabweichungsfaktor

$$\frac{\text{Var} \left(\frac{S_{i,k}^I}{S_{i,k-1}^I} \middle| \mathcal{B}_{k-1}^I \right)^{1/2}}{\text{Var} \left(Q_{i,k-1} \middle| \mathcal{B}_{k-1}^I \right)^{1/2}},$$

ist der Quotient der bedingten Standardabweichung des Incurred-Abwicklungsfaktors und des momentanen (P/I)-Standes. Je größer die Standardabweichung des Abwicklungsfaktors ist, desto wahrscheinlicher ist eine Abweichung vom Mittelwert und desto größer ist der Korrekturterm. Umgekehrt analog.

Analoges gilt für den Standardabweichungsfaktor in (11.1).

- Durch den linearen Term

$$Q_{i,k-1} - \mathbb{E} [Q_{i,k-1} | \mathcal{B}_{k-1}^I]$$

treiben überdurchschnittliche momentane (P/I)-Stände den Korrekturterm in die Höhe und unterdurchschnittliche nach unten. Falls der (P/I)-Stand durchschnittlich ist, so wird wie beim Chain-Ladder-Verfahren ein durchschnittlicher Abwicklungsfaktor verwendet. Analoges gilt für den linearen Term in (11.1).

- Die letzte Annahme im Munich-Chain-Ladder-Modell 11.1 ist stärker als die erste. Sie wird nur benötigt, um Schätzer für

$$\begin{aligned} & \mathbb{E} [Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P], \text{Var} (Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P) \text{ und} \\ & \mathbb{E} [Q_{i,k-1} | \mathcal{B}_{k-1}^I], \text{Var} (Q_{i,k-1} | \mathcal{B}_{k-1}^I) \end{aligned}$$

zu motivieren.

- Die Annahmen (11.1) und (11.2) sind in [16] durch Analyse von Daten motiviert.

11.2 Credibility-Ansatz

Dieses Kapitel lehnt sich an [4] an.

Es soll mithilfe eines Credibility-Ansatzes das Munich-Chain-Ladder-Modell aus Kapitel 11.1 begründet werden. Wir verzichten auf die Annahmen (11.1) und (11.2) und zeigen, dass diese aus unserem Ansatz folgen. Wir definieren die linearen Unterräume

$$\begin{aligned} \mathcal{L}_{i,k}^P &= \{C \in L^2(\Omega, \mathcal{F}, P); C = A_0 + A_1 \cdot S_{i,k-1}^I, A_0, A_1 \text{ sind } \mathcal{B}_{k-1}^P\text{-messbar}\}, \\ \mathcal{L}_{i,k}^I &= \{C \in L^2(\Omega, \mathcal{F}, P); C = A_0 + A_1 \cdot S_{i,k-1}^P, A_0, A_1 \text{ sind } \mathcal{B}_{k-1}^I\text{-messbar}\} \end{aligned}$$

und deren Abschluss

$$\begin{aligned} \overline{\mathcal{L}_{i,k}^P} &= \left\{ C \in L^2(\Omega, \mathcal{F}, P); \exists (C_n)_{n \in \mathbb{N}} \subseteq \mathcal{L}_{i,k}^P \text{ mit } \lim_{n \rightarrow \infty} \|C - C_n\| = 0 \right\}, \\ \overline{\mathcal{L}_{i,k}^I} &= \left\{ C \in L^2(\Omega, \mathcal{F}, P); \exists (C_n)_{n \in \mathbb{N}} \subseteq \mathcal{L}_{i,k}^I \text{ mit } \lim_{n \rightarrow \infty} \|C - C_n\| = 0 \right\}. \end{aligned}$$

Dann gilt folgendes Lemma:

Lemma 11.2: Die Teilmengen $\overline{\mathcal{L}_{i,k}^P}$ und $\overline{\mathcal{L}_{i,k}^I}$ von $L^2(\Omega, \mathcal{F}, P)$ sind abgeschlossene lineare Unterräume von $L^2(\Omega, \mathcal{F}, P)$.

Beweis: O.B.d.A. beweisen wir die Aussage für $\overline{\mathcal{L}_{i,k}^P}$.

Diese Menge ist per Definition abgeschlossen.

Leicht überprüft man, dass $0 \in \overline{\mathcal{L}_{i,k}^P}$ und deshalb $\overline{\mathcal{L}_{i,k}^P} \neq \emptyset$.

Seien $C_1, C_2 \in \overline{\mathcal{L}_{i,k}^P}$. Dann existieren $(C_{1,n})_{n \in \mathbb{N}} \subseteq \mathcal{L}_{i,k}^P$ und $(C_{2,n})_{n \in \mathbb{N}} \subseteq \mathcal{L}_{i,k}^P$ mit

$$\begin{aligned} \lim_{n \rightarrow \infty} \|C_1 - C_{1,n}\| &= 0 \\ \lim_{n \rightarrow \infty} \|C_2 - C_{2,n}\| &= 0. \end{aligned}$$

Weil $\mathcal{L}_{i,k}^P$ ein linearer Unterraum von $L^2(\Omega, \mathcal{F}, P)$ ist, folgt $(\alpha C_{1,n} + C_{2,n})_{n \in \mathbb{N}} \subseteq \mathcal{L}_{i,k}^P$ für $\alpha \in \mathbb{R}$ und es gilt

$$\lim_{n \rightarrow \infty} \|\alpha C_1 + C_2 - (\alpha C_{1,n} + C_{2,n})\| \leq \lim_{n \rightarrow \infty} \|\alpha C_1 - \alpha C_{1,n}\| + \lim_{n \rightarrow \infty} \|C_2 - C_{2,n}\| = 0$$

und deshalb ist $\alpha C_1 + C_2 \in \overline{\mathcal{L}_{i,k}^P}$. \square

Man beachte, dass der letzte Punkt der Modellannahmen 11.1 nahelegt, dass für $l \neq i$ $S_{l,k-1}^I$ nicht in $\mathcal{L}_{i,k}^P$ und $S_{l,k-1}^P$ nicht in $\mathcal{L}_{i,k}^I$ enthalten ist.

Wir wollen Prognosen $C_{i,k}^P \in \overline{\mathcal{L}_{i,k}^P}$ und $C_{i,k}^I \in \overline{\mathcal{L}_{i,k}^I}$ für $S_{i,k}^P$ und $S_{i,k}^I$ finden, sodass die mittlere quadratische Abweichung minimiert wird:

$$\begin{aligned} \text{MSE}_{S_{i,k}^P}(C_{i,k}^P) &= \inf_{C \in \overline{\mathcal{L}_{i,k}^P}} \text{MSE}_{S_{i,k}^P}(C), \\ \text{MSE}_{S_{i,k}^I}(C_{i,k}^I) &= \inf_{C \in \overline{\mathcal{L}_{i,k}^I}} \text{MSE}_{S_{i,k}^I}(C) \end{aligned}$$

Nachdem $\overline{\mathcal{L}_{i,k}^P}$ und $\overline{\mathcal{L}_{i,k}^I}$ abgeschlossene lineare Teilräume des Hilbertraumes $L^2(\Omega, \mathcal{F}, P)$ sind und $S_{i,k}^P, S_{i,k}^I \in L^2(\Omega, \mathcal{F}, P)$, folgt, dass $C_{i,k}^P$ und $C_{i,k}^I$ existieren und fast sicher eindeutig sind (siehe Theorem 2.3.1 in [17]).

Außerdem ist $C_{i,k}^P$ durch die Normalgleichungen

$$\mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot C] = 0 \quad \forall C \in \overline{\mathcal{L}_{i,k}^P} \quad (11.4)$$

charakterisiert (siehe Theorem 3.15 in [18]).

Die selben Gleichungen gelten analog für $C_{i,k}^I$.

Lemma 11.3: Falls die ersten beiden Punkte der Modellannahmen 11.1 gelten, sind die besten affin-linearen Prognosen $C_{i,k}^P \in \overline{\mathcal{L}_{i,k}^P}$ und $C_{i,k}^I \in \overline{\mathcal{L}_{i,k}^I}$ für $S_{i,k}^P$ und $S_{i,k}^I$ für $0 \leq i \leq n$ und $1 \leq k \leq n$ gegeben durch

$$\begin{aligned} C_{i,k}^P &= f_k^P S_{i,k-1}^P + \text{Cor}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \text{Var}(S_{i,k}^P | \mathcal{B}_{k-1}^P)^{1/2} \frac{S_{i,k-1}^I - \mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P]}{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)^{1/2}}, \quad (11.5) \\ C_{i,k}^I &= f_k^I S_{i,k-1}^I + \text{Cor}(S_{i,k}^I, S_{i,k-1}^P | \mathcal{B}_{k-1}^I) \text{Var}(S_{i,k}^I | \mathcal{B}_{k-1}^I)^{1/2} \frac{S_{i,k-1}^P - \mathbb{E}[S_{i,k-1}^P | \mathcal{B}_{k-1}^I]}{\text{Var}(S_{i,k-1}^P | \mathcal{B}_{k-1}^I)^{1/2}}, \end{aligned}$$

wobei die beste affin-lineare Prognose jene Prognose mit der minimalen mittleren quadratischen Abweichung $\text{MSE}_{S_{i,k}^P}(\cdot)$ bzw. $\text{MSE}_{S_{i,k}^I}(\cdot)$ ist.

Beweis: O.B.d.A. beweisen wir die Aussage für die kumulierten Zahlungen $S_{i,k}^P$ und definieren $C_{i,k}^P$ wie in (11.5).

Nachdem alle Ausdrücke außer $S_{i,k-1}^I$ messbar bezüglich \mathcal{B}_{k-1}^P sind, erkennt man, dass $C_{i,k}^P \in \overline{\mathcal{L}_{i,k}^P}$ und wir müssen nur noch (11.4) zeigen.

Dafür sei $C \in \overline{\mathcal{L}_{i,k}^P}$ beliebig, aber fix. Wegen der Stetigkeit des Skalarproduktes auf dem Hilbertraum $L^2(\Omega, \mathcal{F}, P)$ gilt

$$\mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot C] = \mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot \lim_{n \rightarrow \infty} C_n] = \lim_{n \rightarrow \infty} \mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot C_n], \quad (11.6)$$

wobei $(C_n)_{n \in \mathbb{N}} \subseteq \mathcal{L}_{i,k}^P$ mit $\lim_{n \rightarrow \infty} \|C - C_n\| = 0$.

Nachdem $C_n = A_{n,0} + A_{n,1} \cdot S_{i,k-1}^I$ und $A_{n,0}$ und $A_{n,1}$ für alle $n \in \mathbb{N}$ messbar bezüglich \mathcal{B}_{k-1}^P sind, gilt

$$\begin{aligned} \mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot C_n] &= \mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot A_{n,0}] + \mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot A_{n,1} \cdot S_{i,k-1}^I] = \\ &= \mathbb{E}[A_{n,0} \cdot \mathbb{E}[S_{i,k}^P - C_{i,k}^P | \mathcal{B}_{k-1}^P]] \\ &\quad + \mathbb{E}[A_{n,1} \cdot \mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot S_{i,k-1}^I | \mathcal{B}_{k-1}^P]]. \end{aligned}$$

Es gilt nun

$$\begin{aligned} \mathbb{E}[S_{i,k}^P - C_{i,k}^P | \mathcal{B}_{k-1}^P] &= \mathbb{E}[S_{i,k}^P | \mathcal{B}_{k-1}^P] - \mathbb{E}[f_k^P S_{i,k-1}^P | \mathcal{B}_{k-1}^P] + \\ &\quad - \mathbb{E} \left[\text{Cor}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \text{Var}(S_{i,k}^P | \mathcal{B}_{k-1}^P)^{1/2} \frac{S_{i,k-1}^I - \mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P]}{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)^{1/2}} \middle| \mathcal{B}_{k-1}^P \right] = \\ &= f_k^P S_{i,k-1}^P - f_k^P S_{i,k-1}^P - \\ &\quad - \text{Cor}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \text{Var}(S_{i,k}^P | \mathcal{B}_{k-1}^P)^{1/2} \frac{\mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P] - \mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P]}{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)^{1/2}} = \\ &= 0 \end{aligned}$$

und

$$\begin{aligned} \mathbb{E}[(S_{i,k}^P - C_{i,k}^P) \cdot S_{i,k-1}^I | \mathcal{B}_{k-1}^P] &= \mathbb{E}[S_{i,k}^P \cdot S_{i,k-1}^I | \mathcal{B}_{k-1}^P] - \mathbb{E}[C_{i,k}^P \cdot S_{i,k-1}^I | \mathcal{B}_{k-1}^P] = \\ &= \mathbb{E}[S_{i,k}^P \cdot S_{i,k-1}^I | \mathcal{B}_{k-1}^P] - f_k^P S_{i,k-1}^P \mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P] - \text{Cor}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \cdot \\ &\quad \cdot \text{Var}(S_{i,k}^P | \mathcal{B}_{k-1}^P)^{1/2} \cdot \frac{\mathbb{E}[(S_{i,k-1}^I - \mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P]) S_{i,k-1}^I | \mathcal{B}_{k-1}^P]}{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)^{1/2}} = \\ &= \mathbb{E}[S_{i,k}^P \cdot S_{i,k-1}^I | \mathcal{B}_{k-1}^P] - \mathbb{E}[S_{i,k}^P | \mathcal{B}_{k-1}^P] \mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P] - \\ &\quad - \text{Cov}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \frac{\mathbb{E}[(S_{i,k-1}^I)^2 | \mathcal{B}_{k-1}^P] - \mathbb{E}[\mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P] S_{i,k-1}^I | \mathcal{B}_{k-1}^P]}{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)} = \\ &= \text{Cov}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) - \\ &\quad - \text{Cov}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \frac{\mathbb{E}[(S_{i,k-1}^I)^2 | \mathcal{B}_{k-1}^P] - \mathbb{E}[S_{i,k-1}^I | \mathcal{B}_{k-1}^P]^2}{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)} = \\ &= \text{Cov}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) - \text{Cov}(S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \frac{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)}{\text{Var}(S_{i,k-1}^I | \mathcal{B}_{k-1}^P)} = \\ &= 0, \end{aligned}$$

woraus

$$\mathbb{E} [(S_{i,k}^P - C_{i,k}^P) \cdot C_n] = 0 \quad \forall n \in \mathbb{N}$$

und wegen (11.6) auch

$$\mathbb{E} [(S_{i,k}^P - C_{i,k}^P) \cdot C] = 0$$

folgt. □

Damit gilt folgendes Lemma:

Lemma 11.4: Unter den Annahmen aus Lemma 11.3 gilt

- a) Gegeben \mathcal{B}_{k-1}^P und \mathcal{B}_{k-1}^I sind $C_{i,k}^P$ und $C_{i,k}^I$ unverzerrte Schätzer für $\mathbb{E} [S_{i,k}^P | S_{i,k-1}^P]$ bzw. $\mathbb{E} [S_{i,k}^I | S_{i,k-1}^I]$.
- b) $C_{i,k}^P$ und $C_{i,k}^I$ sind (unbedingt) unverzerrte Schätzer für $\mathbb{E} [S_{i,k}^P]$ bzw. $\mathbb{E} [S_{i,k}^I]$.

Beweis:

- a) Aus (11.5) folgt

$$\begin{aligned} \mathbb{E} [C_{i,k}^P | \mathcal{B}_{k-1}^P] &= f_k^P S_{i,k-1}^P + \\ &+ \text{Cor} (S_{i,k}^P, S_{i,k-1}^I | \mathcal{B}_{k-1}^P) \text{Var} (S_{i,k}^P | \mathcal{B}_{k-1}^P)^{1/2} \cdot \frac{\mathbb{E} [S_{i,k-1}^I | \mathcal{B}_{k-1}^P] - \mathbb{E} [S_{i,k-1}^I | \mathcal{B}_{k-1}^P]}{\text{Var} (S_{i,k-1}^I | \mathcal{B}_{k-1}^P)^{1/2}} = \\ &= f_k^P S_{i,k-1}^P = \mathbb{E} [S_{i,k}^P | \mathcal{B}_{k-1}^P]. \end{aligned}$$

Die Aussage für $C_{i,k}^I$ folgt analog.

- b) Aus der Definition der bedingten Erwartung und a) folgt

$$\mathbb{E} [C_{i,k}^P] = \mathbb{E} [\mathbb{E} [C_{i,k}^P | \mathcal{B}_{k-1}^P]] = \mathbb{E} [\mathbb{E} [S_{i,k}^P | \mathcal{B}_{k-1}^P]] = \mathbb{E} [S_{i,k}^P].$$

Die Aussage für $C_{i,k}^I$ folgt analog. □

Bemerkungen:

- Der einzige Unterschied im Vergleich zur Annahme (11.1) bzw. (11.2) ist, dass λ^P bzw. λ^I nicht konstant sein müssen.
- Lemma 11.3 ist auch ohne den letzten Punkt der Modellannahmen 11.1 gültig. Jedoch wird diese Annahme benötigt, um passende Schätzer zu finden und die linearen Unterräume $\mathcal{L}_{i,k}^P$ und $\mathcal{L}_{i,k}^I$ zu motivieren.

11.3 Parameter-Schätzung

Dieses Kapitel orientiert sich an [16] und [6].

Die Abwicklungsfaktoren werden wie in Kapitel 3 für $k \in \{1, \dots, n\}$ geschätzt:

$$\begin{aligned}\widehat{f}_k^P &= \frac{\sum_{i=0}^{n-k} S_{i,k}^P}{\sum_{i=0}^{n-k} S_{i,k-1}^P}, \\ \widehat{f}_k^I &= \frac{\sum_{i=0}^{n-k} S_{i,k}^I}{\sum_{i=0}^{n-k} S_{i,k-1}^I}.\end{aligned}$$

Die Varianz-Parameter werden für $k \in \{1, \dots, n-1\}$ ebenfalls wie üblich geschätzt:

$$\begin{aligned}(\widehat{\sigma}_k^P)^2 &= \frac{1}{n-k} \sum_{i=0}^{n-k} S_{i,k-1}^P \left(\frac{S_{i,k}^P}{S_{i,k-1}^P} - \widehat{f}_k^P \right)^2, \\ (\widehat{\sigma}_k^I)^2 &= \frac{1}{n-k} \sum_{i=0}^{n-k} S_{i,k-1}^I \left(\frac{S_{i,k}^I}{S_{i,k-1}^I} - \widehat{f}_k^I \right)^2.\end{aligned}$$

Es werden also noch Schätzer für die bedingten Korrelationskoeffizienten λ^P und λ^I sowie die bedingten Momente $\mathbb{E}[Q_{i,k-1} | \mathcal{B}_{k-1}^I]$, $\text{Var}(Q_{i,k-1} | \mathcal{B}_{k-1}^I)$, $\mathbb{E}[Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P]$ und $\text{Var}(Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P)$ benötigt.

Dazu nehmen wir an, dass die bedingten Momente von $Q_{i,k}^{-1}$ und $Q_{i,k}$ konstant in i sind und nur von $k = 0, \dots, n$ abhängen.

Wir setzen für $k \in \{0, \dots, n\}$

$$\widehat{q}_k = \frac{1}{\sum_{i=0}^{n-k} S_{i,k}^I} \sum_{i=0}^{n-k} S_{i,k}^I Q_{i,k} = \frac{\sum_{i=0}^{n-k} S_{i,k}^P}{\sum_{i=0}^{n-k} S_{i,k}^I},$$

also das gewichtete Mittel der (P/I)-Stände und

$$\widehat{q}_k^{-1} = \frac{1}{\sum_{i=0}^{n-k} S_{i,k}^P} \sum_{i=0}^{n-k} S_{i,k}^P Q_{i,k}^{-1} = \frac{\sum_{i=0}^{n-k} S_{i,k}^I}{\sum_{i=0}^{n-k} S_{i,k}^P} = (\widehat{q}_k)^{-1}.$$

Als Motivation dient folgendes Lemma:

Lemma 11.5: Unter der Annahme, dass die bedingten ersten Momente von $Q_{i,k}$ und $Q_{i,k}^{-1}$ konstant in i sind und nur von k abhängen, gilt für alle $k \in \{0, \dots, n\}$:

- Bedingt auf \mathcal{B}_k^I ist \widehat{q}_k ein unverzerrter Schätzer für $q_k = \mathbb{E}[Q_{i,k} | \mathcal{B}_k^I]$.
- Bedingt auf \mathcal{B}_k^P ist \widehat{q}_k^{-1} ein unverzerrter Schätzer für $q_k^{-1} = \mathbb{E}[Q_{i,k}^{-1} | \mathcal{B}_k^P]$.

Beweis:

a) Nachdem $\mathbb{E} [Q_{i,k} | \mathcal{B}_k^I]$ nicht von i abhängt, gilt

$$\begin{aligned} \mathbb{E} [\widehat{q}_k | \mathcal{B}_k^I] &= \mathbb{E} \left[\frac{\sum_{i=0}^{n-k} S_{i,k}^P}{\sum_{i=0}^{n-k} S_{i,k}^I} \middle| \mathcal{B}_k^I \right] = \frac{1}{\sum_{i=0}^{n-k} S_{i,k}^I} \sum_{i=0}^{n-k} \mathbb{E} [S_{i,k}^P | \mathcal{B}_k^I] = \\ &= \frac{1}{\sum_{i=0}^{n-k} S_{i,k}^I} \sum_{i=0}^{n-k} \mathbb{E} [Q_{i,k} S_{i,k}^I | \mathcal{B}_k^I] = \frac{1}{\sum_{i=0}^{n-k} S_{i,k}^I} \sum_{i=0}^{n-k} S_{i,k}^I \mathbb{E} [Q_{i,k} | \mathcal{B}_k^I] = \\ &= \frac{1}{\sum_{i=0}^{n-k} S_{i,k}^I} \mathbb{E} [Q_{i,k} | \mathcal{B}_k^I] \sum_{i=0}^{n-k} S_{i,k}^I = \mathbb{E} [Q_{i,k} | \mathcal{B}_k^I] = q_k. \end{aligned}$$

b) analog zu a)

□

Die Schätzer für die Varianzen $\text{Var} (Q_{i,k} | \mathcal{B}_k^I)$ und $\text{Var} (Q_{i,k}^{-1} | \mathcal{B}_k^P)$ sind für $0 \leq k < n$ gegeben durch

$$\widehat{\text{Var}} (Q_{i,k} | \mathcal{B}_k^I) = \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I} \right)^{-1} \sum_{j=0}^{n-k} S_{j,k}^I (Q_{j,k} - \widehat{q}_k)^2$$

und

$$\widehat{\text{Var}} (Q_{i,k}^{-1} | \mathcal{B}_k^P) = \left(\sum_{j=0}^{n-k} S_{j,k}^P - \frac{\sum_{j=0}^{n-k} (S_{j,k}^P)^2}{\sum_{j=0}^{n-k} S_{j,k}^P} \right)^{-1} \sum_{j=0}^{n-k} S_{j,k}^P (Q_{j,k}^{-1} - (\widehat{q}_k)^{-1})^2.$$

Man beachte, dass die Schätzer nicht vom Anfallsjahr i , sondern nur vom Abwicklungsjahr k abhängen.

Lemma 11.6: Unter den Annahmen aus Lemma 11.5 und der zusätzlichen Annahme, dass $\text{Var} (Q_{i,k} | \mathcal{B}_k^I)$, $\text{Var} (Q_{i,k}^{-1} | \mathcal{B}_k^P)$, $\text{Cov} (Q_{i,k}, \widehat{q}_k | \mathcal{B}_k^I)$ und $\text{Cov} (Q_{i,k}^{-1}, \widehat{q}_k | \mathcal{B}_k^P)$ nur von k (und nicht von i) abhängen, gilt für $0 \leq k \leq n$

- a) bedingt auf \mathcal{B}_k^I ist $\widehat{\text{Var}} (Q_{i,k} | \mathcal{B}_k^I)$ ein unverzerrter Schätzer für $\text{Var} (Q_{i,k} | \mathcal{B}_k^I)$ und
- b) bedingt auf \mathcal{B}_k^P ist $\widehat{\text{Var}} (Q_{i,k}^{-1} | \mathcal{B}_k^P)$ ein unverzerrter Schätzer für $\text{Var} (Q_{i,k}^{-1} | \mathcal{B}_k^P)$.

Beweis:

a) Aus Lemma 11.5 a) und der Annahme, dass $\text{Var} (Q_{j,k} | \mathcal{B}_k^I)$ nicht von j abhängt, folgt

$$\begin{aligned} \mathbb{E} [(Q_{j,k} - \widehat{q}_k)^2 | \mathcal{B}_k^I] &= \text{Var} (Q_{j,k} - \widehat{q}_k | \mathcal{B}_k^I) + \mathbb{E} [Q_{j,k} - \widehat{q}_k | \mathcal{B}_k^I]^2 = \\ &= \text{Var} (Q_{j,k} - \widehat{q}_k | \mathcal{B}_k^I) = \\ &= \text{Var} (Q_{j,k} | \mathcal{B}_k^I) + \text{Var} (\widehat{q}_k | \mathcal{B}_k^I) - 2 \text{Cov} (Q_{j,k}, \widehat{q}_k | \mathcal{B}_k^I) = \\ &= \text{Var} (Q_{i,k} | \mathcal{B}_k^I) + \text{Var} (\widehat{q}_k | \mathcal{B}_k^I) - 2 \text{Cov} (Q_{j,k}, \widehat{q}_k | \mathcal{B}_k^I). \end{aligned}$$

Weil die kumulierten Schadenzahlungen $S_{l,k}^P$ für verschiedene Anfallsjahre unabhängig sind und weil $\text{Var}(Q_{j,k} | \mathcal{B}_k^I)$ nicht von j abhängt, gilt

$$\begin{aligned}
\text{Var}(\hat{q}_k | \mathcal{B}_k^I) &= \text{Var}\left(\frac{\sum_{l=0}^{n-k} S_{l,k}^P}{\sum_{l=0}^{n-k} S_{l,k}^I} \middle| \mathcal{B}_k^I\right) = \frac{1}{\left(\sum_{l=0}^{n-k} S_{l,k}^I\right)^2} \text{Var}\left(\sum_{l=0}^{n-k} S_{l,k}^P \middle| \mathcal{B}_k^I\right) = \\
&= \frac{1}{\left(\sum_{l=0}^{n-k} S_{l,k}^I\right)^2} \sum_{l=0}^{n-k} \text{Var}(S_{l,k}^P | \mathcal{B}_k^I) = \\
&= \frac{1}{\left(\sum_{l=0}^{n-k} S_{l,k}^I\right)^2} \sum_{l=0}^{n-k} \text{Var}(Q_{l,k} S_{l,k}^I | \mathcal{B}_k^I) = \\
&= \frac{1}{\left(\sum_{l=0}^{n-k} S_{l,k}^I\right)^2} \sum_{l=0}^{n-k} (S_{l,k}^I)^2 \text{Var}(Q_{l,k} | \mathcal{B}_k^I) = \\
&= \frac{\sum_{l=0}^{n-k} (S_{l,k}^I)^2}{\left(\sum_{l=0}^{n-k} S_{l,k}^I\right)^2} \text{Var}(Q_{i,k} | \mathcal{B}_k^I).
\end{aligned}$$

und

$$\begin{aligned}
\text{Cov}(Q_{j,k}, \hat{q}_k | \mathcal{B}_k^I) &= \text{Cov}\left(Q_{j,k}, \frac{\sum_{l=0}^{n-k} S_{l,k}^P}{\sum_{l=0}^{n-k} S_{l,k}^I} \middle| \mathcal{B}_k^I\right) = \\
&= \frac{1}{\sum_{l=0}^{n-k} S_{l,k}^I} \sum_{l=0}^{n-k} \text{Cov}(Q_{j,k}, S_{l,k}^P | \mathcal{B}_k^I) = \\
&= \frac{1}{\sum_{l=0}^{n-k} S_{l,k}^I} \text{Cov}(Q_{j,k}, S_{j,k}^P | \mathcal{B}_k^I) = \\
&= \frac{1}{\sum_{l=0}^{n-k} S_{l,k}^I} \text{Cov}(Q_{j,k}, Q_{j,k} S_{j,k}^I | \mathcal{B}_k^I) = \\
&= \frac{S_{j,k}^I}{\sum_{l=0}^{n-k} S_{l,k}^I} \text{Var}(Q_{j,k} | \mathcal{B}_k^I) = \\
&= \frac{S_{j,k}^I}{\sum_{l=0}^{n-k} S_{l,k}^I} \text{Var}(Q_{i,k} | \mathcal{B}_k^I).
\end{aligned}$$

Es folgt also

$$\begin{aligned}
&\mathbb{E}\left[\widehat{\text{Var}}(Q_{i,k} | \mathcal{B}_k^I) \middle| \mathcal{B}_k^I\right] = \\
&= \mathbb{E}\left[\left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I}\right)^{-1} \sum_{j=0}^{n-k} S_{j,k}^I (Q_{j,k} - \hat{q}_k)^2 \middle| \mathcal{B}_k^I\right] = \\
&= \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I}\right)^{-1} \sum_{j=0}^{n-k} S_{j,k}^I \mathbb{E}\left[(Q_{j,k} - \hat{q}_k)^2 \middle| \mathcal{B}_k^I\right] =
\end{aligned}$$

$$\begin{aligned}
&= \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I} \right)^{-1} \sum_{j=0}^{n-k} S_{j,k}^I \left(\text{Var} (Q_{i,k} | \mathcal{B}_k^I) + \right. \\
&\quad \left. \frac{\sum_{l=0}^{n-k} (S_{l,k}^I)^2}{\left(\sum_{l=0}^{n-k} S_{l,k}^I \right)^2} \text{Var} (Q_{i,k} | \mathcal{B}_k^I) - 2 \frac{S_{j,k}^I}{\sum_{l=0}^{n-k} S_{l,k}^I} \text{Var} (Q_{i,k} | \mathcal{B}_k^I) \right) = \\
&= \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I} \right)^{-1} \cdot \text{Var} (Q_{i,k} | \mathcal{B}_k^I) \cdot \\
&\quad \sum_{j=0}^{n-k} S_{j,k}^I \left(1 + \frac{\sum_{l=0}^{n-k} (S_{l,k}^I)^2}{\left(\sum_{l=0}^{n-k} S_{l,k}^I \right)^2} - 2 \frac{S_{j,k}^I}{\sum_{l=0}^{n-k} S_{l,k}^I} \right) = \\
&= \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I} \right)^{-1} \cdot \text{Var} (Q_{i,k} | \mathcal{B}_k^I) \cdot \\
&\quad \left(\sum_{j=0}^{n-k} S_{j,k}^I + \frac{\sum_{j=0}^{n-k} S_{j,k}^I \sum_{l=0}^{n-k} (S_{l,k}^I)^2}{\left(\sum_{l=0}^{n-k} S_{l,k}^I \right)^2} - 2 \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{l=0}^{n-k} S_{l,k}^I} \right) = \\
&= \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I} \right)^{-1} \cdot \text{Var} (Q_{i,k} | \mathcal{B}_k^I) \cdot \\
&\quad \left(\sum_{j=0}^{n-k} S_{j,k}^I + \frac{\sum_{l=0}^{n-k} (S_{l,k}^I)^2}{\sum_{l=0}^{n-k} S_{l,k}^I} - 2 \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{l=0}^{n-k} S_{l,k}^I} \right) = \\
&= \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I} \right)^{-1} \cdot \text{Var} (Q_{i,k} | \mathcal{B}_k^I) \cdot \\
&\quad \cdot \left(\sum_{j=0}^{n-k} S_{j,k}^I - \frac{\sum_{j=0}^{n-k} (S_{j,k}^I)^2}{\sum_{j=0}^{n-k} S_{j,k}^I} \right) = \\
&= \text{Var} (Q_{i,k} | \mathcal{B}_k^I)
\end{aligned}$$

b) analog zu a)

□

Wir müssen nun noch λ^P und λ^I schätzen. Aus (11.3) und den Modellannahmen 11.1 folgt für $i \in \{0, \dots, n\}$ und $k \in \{1, \dots, n\}$

$$\lambda^P = \mathbb{E} \left[\frac{\frac{S_{i,k}^P}{S_{i,k-1}^P} - f_k^P}{\sqrt{\frac{(\sigma_k^P)^2}{S_{i,k-1}^P}}} \middle/ \frac{Q_{i,k-1}^{-1} - \mathbb{E} [Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P]}{\text{Var} (Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^P)^{1/2}} \middle| \tilde{\mathcal{B}}_{k-1} \right]$$

und

$$\lambda^I = \mathbb{E} \left[\frac{\frac{S_{i,k}^I - f_k^I}{S_{i,k-1}^I}}{\sqrt{\frac{(\sigma_k^I)^2}{S_{i,k-1}^I}}} \middle/ \frac{Q_{i,k-1} - \mathbb{E}[Q_{i,k-1} | \mathcal{B}_{k-1}^I]}{\text{Var}(Q_{i,k-1} | \mathcal{B}_{k-1}^I)^{1/2}} \middle| \tilde{\mathcal{B}}_{k-1} \right].$$

Wir prognostizieren die Residuen von $Q_{i,k-1}$ und $Q_{i,k-1}^{-1}$, also die Nenner in den obigen Erwartungswerten, für $i \in \{0, \dots, n-1\}$ und $k \in \{1, \dots, n-i\}$ folgendermaßen:

$$\tilde{Q}_{i,k-1} = \frac{Q_{i,k-1} - \hat{q}_{k-1}}{\widehat{\text{Var}}(Q_{i,k-1} | \mathcal{B}_{k-1}^I)^{1/2}} \quad \text{und} \quad \tilde{Q}_{i,k-1}^{-1} = \frac{Q_{i,k-1}^{-1} - \hat{q}_{k-1}^{-1}}{\widehat{\text{Var}}(Q_{i,k-1}^{-1} | \mathcal{B}_{k-1}^I)^{1/2}}.$$

Die Residuen der Chain-Ladder-Faktoren f_k^P und f_k^I werden folgendermaßen prognostiziert:

$$\tilde{F}_{i,k}^P = \frac{\frac{S_{i,k}^P - \hat{f}_k^P}{S_{i,k-1}^P}}{\frac{\hat{\sigma}_k^P}{\sqrt{S_{i,k-1}^P}}} \quad \text{und} \quad \tilde{F}_{i,k}^I = \frac{\frac{S_{i,k}^I - \hat{f}_k^I}{S_{i,k-1}^I}}{\frac{\hat{\sigma}_k^I}{\sqrt{S_{i,k-1}^I}}}.$$

Wir können nun Schätzer für die bedingten Korrelationskoeffizienten λ^P und λ^I definieren:

$$\hat{\lambda}^P = \frac{1}{\sum_{i=0}^{n-1} \sum_{k=1}^{n-i} (\tilde{Q}_{i,k-1}^{-1})^2} \sum_{i=0}^{n-1} \sum_{k=1}^{n-i} (\tilde{Q}_{i,k-1}^{-1})^2 \frac{\tilde{F}_{i,k}^P}{\tilde{Q}_{i,k-1}^{-1}} = \frac{\sum_{i=0}^{n-1} \sum_{k=1}^{n-i} \tilde{Q}_{i,k-1}^{-1} \tilde{F}_{i,k}^P}{\sum_{i=0}^{n-1} \sum_{k=1}^{n-i} (\tilde{Q}_{i,k-1}^{-1})^2},$$

$$\hat{\lambda}^I = \frac{1}{\sum_{i=0}^{n-1} \sum_{k=1}^{n-i} \tilde{Q}_{i,k-1}^2} \sum_{i=0}^{n-1} \sum_{k=1}^{n-i} \tilde{Q}_{i,k-1}^2 \frac{\tilde{F}_{i,k}^I}{\tilde{Q}_{i,k-1}} = \frac{\sum_{i=0}^{n-1} \sum_{k=1}^{n-i} \tilde{Q}_{i,k-1} \tilde{F}_{i,k}^I}{\sum_{i=0}^{n-1} \sum_{k=1}^{n-i} \tilde{Q}_{i,k-1}^2}.$$

Dadurch ergeben sich für $i+k > n$ folgende Prognosen:

$$\widehat{S}_{i,k}^{P, \text{MCL}} = \hat{\mathbb{E}}[S_{i,k}^P | \tilde{\mathcal{D}}_n] = \hat{\mathbb{E}}[S_{i,k-1}^P | \tilde{\mathcal{D}}_n] \left(\hat{f}_k^P + \hat{\lambda}^P \frac{\hat{\sigma}_k^P}{\hat{\mathbb{E}}[S_{i,k-1}^P | \tilde{\mathcal{D}}_n]^{1/2}} \tilde{Q}_{i,k-1}^{-1} \right),$$

$$\widehat{S}_{i,k}^{I, \text{MCL}} = \hat{\mathbb{E}}[S_{i,k}^I | \tilde{\mathcal{D}}_n] = \hat{\mathbb{E}}[S_{i,k-1}^I | \tilde{\mathcal{D}}_n] \left(\hat{f}_k^I + \hat{\lambda}^I \frac{\hat{\sigma}_k^I}{\hat{\mathbb{E}}[S_{i,k-1}^I | \tilde{\mathcal{D}}_n]^{1/2}} \tilde{Q}_{i,k-1} \right)$$

und $\hat{\mathbb{E}}[S_{i,n-i}^P | \tilde{\mathcal{D}}_n] = S_{i,n-i}^P$ und $\hat{\mathbb{E}}[S_{i,n-i}^I | \tilde{\mathcal{D}}_n] = S_{i,n-i}^I$, wobei für $k+i > n$ folgende Prognosen verwendet werden:

$$\tilde{Q}_{i,k}^{-1} = \frac{\hat{Q}_{i,k}^{-1} - \hat{q}_k^{-1}}{\widehat{\text{Var}}(Q_{i,k}^{-1} | \mathcal{B}_k^I)^{1/2}}$$

$$\tilde{Q}_{i,k} = \frac{\hat{Q}_{i,k} - \hat{q}_k}{\widehat{\text{Var}}(Q_{i,k} | \mathcal{B}_k^I)^{1/2}}$$

und

$$\hat{Q}_{i,k} = \frac{(\widehat{S}_{i,k}^P)^{\text{MCL}}}{(\widehat{S}_{i,k}^I)^{\text{MCL}}}.$$

Dann ergeben sich für $i \geq 1$ folgende Prognosen für die Reserven:

$$\begin{aligned}\widehat{R}_i^P{}^{\text{MCL}} &= \widehat{S}_{i,n}^P{}^{\text{MCL}} - S_{i,n-i}^P, & \widehat{R}_i^I{}^{\text{MCL}} &= \widehat{S}_{i,n}^I{}^{\text{MCL}} - S_{i,n-i}^I, \\ \widehat{R}^P{}^{\text{MCL}} &= \sum_{i=1}^n \widehat{R}_i^P{}^{\text{MCL}}, & \widehat{R}^I{}^{\text{MCL}} &= \sum_{i=1}^n \widehat{R}_i^I{}^{\text{MCL}}.\end{aligned}$$

und

$$\begin{aligned}\widetilde{R}_0^I{}^{\text{MCL}} &= S_{0,n}^I - S_{i,n}^P, \\ \widetilde{R}_i^I{}^{\text{MCL}} &= \widehat{S}_{i,n}^I{}^{\text{MCL}} - S_{i,n-i}^P, \\ \widetilde{R}^I{}^{\text{MCL}} &= \sum_{i=0}^n \widetilde{R}_i^I{}^{\text{MCL}}.\end{aligned}$$

11.4 Numerisches Beispiel

Es wird nun für die Daten aus den Tabellen 2.1 und 2.2 das Munich-Chain-Ladder-Verfahren gerechnet. Dabei ist

$$\begin{aligned}\widehat{\lambda}^P &= 0,413, \\ \widehat{\lambda}^I &= 0,273.\end{aligned}$$

Die Ergebnisse für $\widetilde{Q}_{i,k}$ bzw. $\widetilde{Q}_{i,k}^{-1}$ sind in Tabelle 11.1 und 11.2, jene für \widehat{q}_k , $\widehat{\text{Var}}(Q_{i,k} | \mathcal{B}_k^I)$ und $\widehat{\text{Var}}(Q_{i,k}^{-1} | \mathcal{B}_k^P)$ in Tabelle 11.3 dargestellt.

Tabelle 11.1: Residuen $\widetilde{Q}_{i,k}$

i/k	0	1	2	3	4	5	6	7
0	1,050	1,639	2,265	0,995	1,417	1,131	0,758	0,746
1	-2,062	-1,182	0,715	0,211	0,160	0,349	0,520	-0,668
2	0,151	-0,604	-0,354	-1,326	-1,473	-1,299	-1,034	-8,860
3	0,581	1,329	-0,079	1,319	-0,081	0,100	0,058	-3,256
4	0,620	-0,920	-0,073	-0,404	0,310	0,584	0,216	-2,438
5	1,055	-0,044	-0,920	-0,535	-0,752	-0,890	-0,270	-4,942
6	0,140	-0,484	-0,733	-0,372	-0,577	-0,644	-0,186	-4,513
7	-0,882	0,493	0,313	0,026	-0,152	-0,055	0,007	-3,518
8	-0,641	-0,858	-0,337	-0,220	-0,413	-0,415	-0,110	-4,117

Die Munich-Chain-Ladder-Prognosen findet man in Tabelle 11.4 (Paid) bzw. 11.5 (Incurred).

Vergleicht man die Gesamtreserve in Tabelle 11.4 mit jener in Tabelle 11.5, so erkennt man, dass die Incurred-Reserve nur noch um 21.470 höher ist als die Paid-Reserve, d.h. im Vergleich zum Chain-Ladder-Verfahren in Kapitel 3.3 ist der Unterschied zwischen Paid- und Incurred-Gesamtreserve gesunken.

Tabelle 11.2: Residuen $\tilde{Q}_{i,k}^{-1}$

i/k	0	1	2	3	4	5	6	7
0	-0,984	-1,604	-2,210	-0,974	-1,384	-1,118	-0,748	-0,744
1	2,222	1,209	-0,715	-0,210	-0,159	-0,348	-0,515	0,670
2	-0,147	0,612	0,359	1,369	1,504	1,312	1,048	9,072
3	-0,556	-1,307	0,080	-1,282	0,081	-0,099	-0,057	3,285
4	-0,591	0,937	0,074	0,408	-0,308	-0,581	-0,215	2,455
5	-0,989	0,044	0,943	0,542	0,760	0,896	0,270	5,008
6	-0,137	0,490	0,749	0,376	0,581	0,647	0,187	4,567
7	0,899	-0,490	-0,315	-0,026	0,152	0,055	-0,007	3,551
8	0,646	0,873	0,342	0,221	0,415	0,415	0,110	4,163

Tabelle 11.3: Parameter für das Munich-Chain-Ladder-Verfahren

k	\hat{q}_k	$\widehat{\text{Var}}(Q_{i,k} \mathcal{B}_k^I)$	$\widehat{\text{Var}}(Q_{i,k}^{-1} \mathcal{B}_k^P)$
0	0,5725	0,00063	0,00609
1	0,8456	0,00018	0,00034
2	0,9241	0,00020	0,00027
3	0,9453	0,00045	0,00057
4	0,9602	0,00022	0,00026
5	0,9680	0,00007	0,00008
6	0,9812	0,00021	0,00023
7	0,9972	0,00001	0,00001
8	1,0000		

Kapitel 12

Statistische Datenanalyse

Dieses Kapitel folgt in der Darstellung [19] und [20].

12.1 Datenanalyse und -aufbereitung

Nachdem die Daten für das Schadendreieck homogen sein sollen, werden in der Praxis Großschäden oft entfernt und extra prognostiziert. Dies kann z.B. erfolgen, indem die Anzahl der Großschäden wie in den vorangegangenen Kapiteln beschrieben, z.B. mithilfe des Chain-Ladder-Verfahrens, und separat die durchschnittliche Schadenhöhe eines Großschadens prognostiziert wird.

Weiters können die Chain-Ladder-Faktoren mithilfe einer Regression geglättet werden. Dabei muss jedoch beachtet werden, dass die Abwicklungsfaktoren aus den ersten Abwicklungsjahren oft nicht dem selben Trend wie die späteren unterliegen. In diesem Fall sollten sie nicht in die Regression miteinbezogen werden. Folgende Funktionen bieten sich für die Kurvenanpassung an:

Exponentialfunktion	$\hat{f}_k = 1 + a \cdot e^{-b \cdot k}$
Weibull-Funktion	$\hat{f}_k = \frac{1}{1 - \exp(-a \cdot k^b)}$
Potenzfunktion	$\hat{f}_k = a^{b^k}$
Inverse Potenzfunktion	$\hat{f}_k = 1 + a \cdot (c + k)^b$

Welcher Typ verwendet werden soll, kann durch einen grafischen Vergleich oder durch Betrachtung der quadratischen Approximationsfehler entschieden werden.

12.2 Tail-Schätzung

Da das Abwicklungsdreieck nur Zahlungen bis zum Jahr n enthält, in der Praxis der Schaden aber oft noch nicht nach n Jahren abgewickelt ist, kann eine weitere Abwicklung vom Jahr n bis zum Endjahr $n + d$, wobei $d > 0$ ist, ein sogenannter **Tail**, geschätzt werden.

Allgemein unterscheidet man zwischen lang- und kurzabwickelnden Sparten, d.h. Sparten, in denen die Abwicklung länger bzw. kürzer dauert. Langabwickelnde Sparten sind z.B. KFZ-Haftpflicht, Allgemeine Haftpflicht oder Rechtsschutz. Beispiele für kurzabwickelnde Sparten

sind KFZ-Kasko oder Haushalt.

Ein Tail muss geschätzt werden, wenn

- der letzte Abwicklungsfaktor noch nicht nahe bei Eins ist, d.h. die Abwicklung noch nicht beendet ist,
- aus vergleichbaren Schadenfällen bekannt ist, dass die Abwicklung länger als n Jahre dauert oder
- zu wenige zuverlässige Daten für ein größeres Dreieck vorhanden sind.

Deshalb verwenden Aktuarien oft einen Tail-Faktor $f_\infty > 1$, um den Endschaten $S_{i,\infty}$ zu prognostizieren:

$$\widehat{S}_{i,\infty}^{\text{CL}} = \widehat{S}_{i,n}^{\text{CL}} \widehat{f}_\infty$$

Es gibt verschiedene Möglichkeiten, diesen Tail-Faktor zu schätzen:

- **Marktfaktoren:** Grundsätzlich sollte jedes Versicherungsunternehmen eigene Zahlungsmuster berechnen. In manchen Fällen, wenn z.B. zu wenige historische Daten zur Verfügung stehen, kann es jedoch sinnvoll sein, markteinheitliche Abwicklungsmuster in Form von Markt-Abwicklungsfaktoren zu verwenden. Es sollte aber sichergestellt sein, dass die Verwendung der Marktinformation nicht zu einer systematischen Über- oder Unterschätzung der Reserve führt.
- **Extrapolation:** Dabei werden die Chain-Ladder-Faktoren \widehat{f}_k mit geeigneten Funktionen extrapoliert. Es wird eine Funktion an die geschätzten Chain-Ladder-Faktoren wie in Kapitel 12.1 beschrieben angepasst und an den Stellen $k \geq n + 1$ ausgewertet. Der Tail-Faktor berechnet sich dann folgendermaßen:

$$\widehat{f}_\infty = \prod_{k=n+1}^{n+d} \widehat{f}_k$$

- **Verteilungsanpassung:** Dabei betrachtet man die Zuwächse der Schadenzahlung $Z_{i,k}$ mit $k \in \{0, \dots, n\}$ als eine Aufteilung der Schadenzahlung auf die Abwicklungsjahre $k \in \{0, \dots, n\}$. Diese wird mittels einer parametrischen Verteilungsfunktion direkt modelliert. Eine einfache Anpassungsmöglichkeit stellt die Exponentialverteilung mit folgender Verteilungsfunktion dar

$$F(t) = 1 - \exp\left(-\frac{t}{m}\right),$$

wobei m die durchschnittliche Abwicklungsdauer ist. $F(t)$ gibt an, welcher Anteil nach t Jahren bereits abgewickelt ist. Die „geglätteten“ Chain-Ladder-Faktoren für $k \in \{1, \dots, n\}$ werden folgendermaßen ermittelt:

$$\widetilde{f}_k = \frac{F(k)}{F(k-1)}$$

Diese müssen nicht verwendet werden, jedoch kann nun der Tail-Faktor bestimmt werden:

$$\widehat{f}_\infty = \frac{1}{F(n)}$$

Tabelle 12.1: Chain-Ladder-Abwicklungsfaktoren \hat{f}_k

k	\hat{f}_k	k	\hat{f}_k
1	1,05569	9	1,00018
2	1,02242	10	1,00030
3	1,01148	11	1,00377
4	1,00954	12	1,00036
5	1,00498	13	1,00005
6	1,00163	14	1,00021
7	1,00401	15	1,00015
8	1,00191		

12.3 Numerisches Beispiel

Angenommen man erhält aus dem Abwicklungsdreieck die Abwicklungsfaktoren \hat{f}_k , die in Tabelle 12.1 dargestellt sind.

Auffällig ist die Werte in den Jahren 7 und 11. In Abbildung 12.1 wurde versucht, die Chain-Ladder-Abwicklungsfaktoren \hat{f}_k mit verschiedenen Kurven wie in Kapitel 12.1 beschrieben zu glätten.

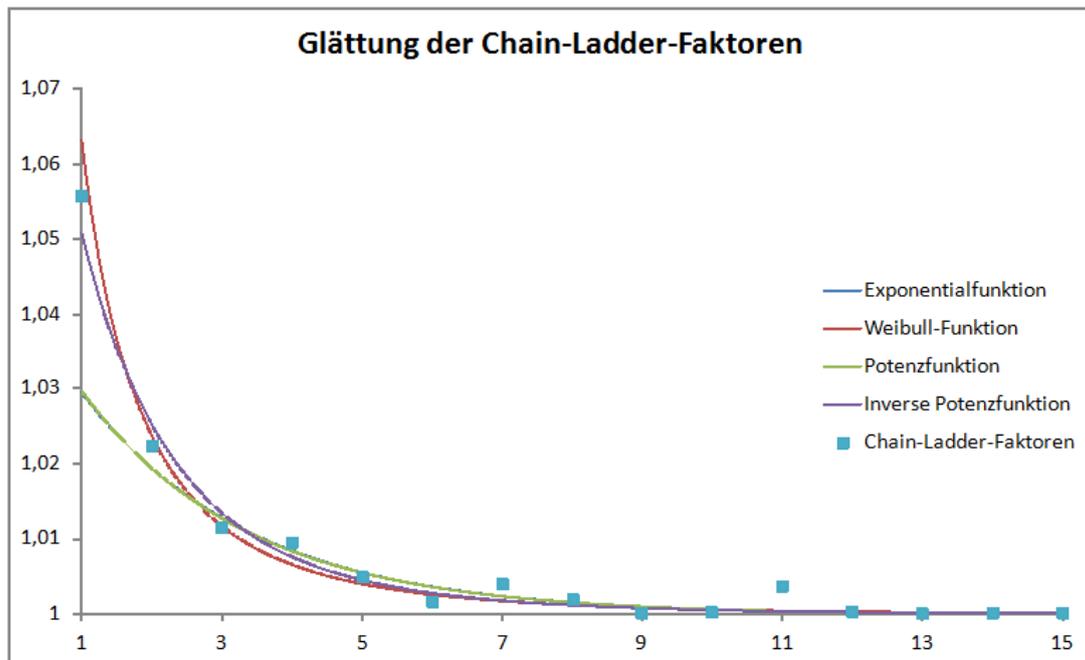


Abbildung 12.1: Glättung der Chain-Ladder-Faktoren

Kapitel 13

Implementierung in Excel

In Excel wurde mithilfe von VBA ein Programm erstellt, das für gegebene Daten das Chain-Ladder-, Bornhuetter-Ferguson-, Benktander-Hovinen-, Cape-Cod-, Additive und Munich-Chain-Ladder-Verfahren wie in den obigen Kapiteln beschrieben rechnet. Das Excel-Dokument besteht aus mehreren Tabellenblättern:

- Input Jahresdaten
- Input Quartalsdaten
- Daten
- CL-Faktoren
- Chain Ladder
- Bornhuetter-Ferguson
- Benktander-Hovinen
- Cape-Cod
- Additiv
- Munich Chain Ladder

Dabei sind die Tabellenblätter „Chain Ladder“, „Bornhuetter-Ferguson“, „Benktander-Hovinen“, „Cape-Cod“, „Additiv“ und „Munich Chain Ladder“ zu Beginn ganz leer und werden durch die Makros befüllt.

Das Tabellenblatt „Input Jahresdaten“ und auch das Tabellenblatt „Input Quartalsdaten“ hat zu Beginn den in Abbildung 13.1 dargestellten Inhalt.

	A	B	C	D	E	F	G
1	Anzahl Jahre					kum. Daten	
2	1. Anfallsjahr					Form	
3	letztes Anf.jahr						
4							
5							
6	Neues Schadendreieck eingeben		Daten für Bornhuetter-Ferguson, Benktander-Hovinen, Cape-Cod und Additiv eingeben				
7							
8	Daten einlesen						
9							
10							

Abbildung 13.1: Sheet „Input Jahresdaten“

Das Abwicklungsdreieck kann entweder auf Jahresbasis oder auf Quartalsbasis eingegeben werden. Im Folgenden wird das Programm für die Eingabe von Jahresdaten erklärt. Es funktioniert

analog für die Eingabe von Quartalsdaten.

Zuerst müssen zwei der Zellen B1-B3 ausgefüllt werden. Klickt man dann auf den Button „Neues Schadendreieck eingeben“, so erscheint zuerst das Formular aus Abbildung 13.2. Es muss angegeben werden, ob kumulierte Daten oder Zuwächse ins Abwicklungsdreieck eingegeben werden.

Abbildung 13.2: Formular zur Dateneingabe

Danach erscheint das Formular aus Abbildung 13.3. Man kann auswählen, in welcher Form das Abwicklungsdreieck eingegeben wird.

Abbildung 13.3: Formular zur Dateneingabe

Hat man im Formular aus Abbildung 13.3 die erste Möglichkeit ausgewählt, so wird die Vorlage zur Dateneingabe wie in Abbildung 13.4 erstellt.

PAID		Abwicklungsjahr					INCURRED		Abwicklungsjahr						
Anfallsjahr		0	1	2	3	4	5	Anfallsjahr		0	1	2	3	4	5
vor 2007								vor 2007							
2007								2007							
2008								2008							
2009								2009							
2010								2010							
2011								2011							
2012								2012							

Abbildung 13.4: Dateneingabe mit Abwicklungsdreieck mit Abwicklungsjahren 0, 1, 2, ...

Hat man im Formular aus Abbildung 13.3 die zweite Möglichkeit ausgewählt, so wird die Vorlage zur Dateneingabe wie in Abbildung 13.5 erstellt.

PAID								INCURRED							
Anfallsjahr	Abwicklungsjahr							Anfallsjahr	Abwicklungsjahr						
	vor 2007	2007	2008	2009	2010	2011	2012		vor 2007	2007	2008	2009	2010	2011	2012
vor 2007								vor 2007							
2007								2007							
2008								2008							
2009								2009							
2010								2010							
2011								2011							
2012								2012							

Abbildung 13.5: Dateneingabe mit Abwicklungsdreieck mit Abwicklungsjahren 2007, 2008, ...

Hat man im Formular aus Abbildung 13.3 die dritte Möglichkeit ausgewählt, so wird die Vorlage zur Dateneingabe wie in Abbildung 13.6 erstellt. Hierbei kann das Abwicklungsjahr entweder als Differenz (0, 1, ...) oder als Jahreszahl (z.B. 2007) eingegeben werden.

PAID			INCURRED		
Anfallsjahr	Abwicklungsjahr	Betrag	Anfallsjahr	Abwicklungsjahr	Betrag

Abbildung 13.6: Dateneingabe in Listenform

Es können nun die Daten für Paid und/oder Inurred eingegeben werden. (Man beachte, dass für das Munich-Chain-Ladder-Verfahren beide Dreiecke benötigt werden.) Danach kann man auf den Button „Daten für Bornhuetter-Ferguson, Benktander-Hovinen, Cape-Cod und Additiv eingeben“ klicken. Es erscheint dann eine Vorlage, wo die Schätzungen der erwarteten Endschäden $\mathbb{E}[S_{i,n}]$, die Prämien π_i , die Schätzungen der Abwicklungsfaktoren $\hat{\alpha}_k$ und ein Variationskoeffizient $Vco(\hat{\mu}_i)$ eingegeben werden kann.

Danach können die Daten eingelesen werden. Falls Quartalsdaten eingegeben werden, kann man auswählen, ob mit Quartals- oder Jahresdaten gerechnet werden soll. Im Sheet „Daten“ werden die Inkremente der Paid- und/oder Incurred-Daten angezeigt. Eventuell können in diesem Sheet noch Großschäden entfernt werden.

Danach werden die individuellen und die Chain-Ladder-Abwicklungsfaktoren in einem Sheet berechnet. Diese können ebenfalls noch manuell geändert werden.

Dann kann man mit einem Klick auf einen Button die einzelnen Verfahren rechnen. Dabei kann man auswählen, ob die Verfahren für die Paid- oder Incurred-Daten gerechnet werden sollen. In den einzelnen Sheets werden die Prognosen der Reserve und die Schätzung für die mittlere quadratische Abweichung dieser Prognosen wie in den obigen Kapiteln beschrieben berechnet.

Anhang A

Mittlere quadratische Abweichung einer Prognose

Dieses Kapitel lehnt sich an [6] an.

Die mittlere quadratische Abweichung soll helfen, die Qualität einer Prognose zu beurteilen.

Sei X eine quadratisch integrierbare Zufallsvariable und \mathcal{D} eine σ -Algebra, die aus einer Menge von Beobachtungen erzeugt wird. Wir nehmen an, dass \hat{X} ein \mathcal{D} -messbarer Schätzer für $\mathbb{E}[X|\mathcal{D}]$ bzw. eine \mathcal{D} -messbare Prognose für X ist. Dann ist die **bedingte mittlere quadratische Abweichung** der Prognose \hat{X} für X folgendermaßen definiert:

$$\text{MSE}_{X|\mathcal{D}}(\hat{X}) = \mathbb{E} \left[(\hat{X} - X)^2 \middle| \mathcal{D} \right]$$

Für eine \mathcal{D} -messbare Prognose \hat{X} gilt also

$$\begin{aligned} \text{MSE}_{X|\mathcal{D}}(\hat{X}) &= \mathbb{E} \left[\hat{X}^2 - 2\hat{X}X + X^2 \middle| \mathcal{D} \right] = \hat{X}^2 - 2\hat{X}\mathbb{E}[X|\mathcal{D}] + \mathbb{E}[X^2|\mathcal{D}] = \\ &= \text{Var}(X|\mathcal{D}) + \mathbb{E}[X|\mathcal{D}]^2 + \hat{X}^2 - 2\hat{X}\mathbb{E}[X|\mathcal{D}] = \\ &= \text{Var}(X|\mathcal{D}) + (\hat{X} - \mathbb{E}[X|\mathcal{D}])^2. \end{aligned} \tag{A.1}$$

Der erste Term in (A.1) wird als **stochastischer Fehler** bezeichnet und beschreibt die Varianz innerhalb des Modells, welche nicht eliminiert werden kann. Der zweite Term in (A.1) ist der **Schätzfehler**, der die Unsicherheit in der Schätzung von $\mathbb{E}[X|\mathcal{D}]$ widerspiegelt. Dieser wird im Allgemeinen kleiner, wenn mehr Beobachtungen vorhanden sind. In der Praxis verschwindet er jedoch nicht ganz, nachdem zukünftiges bedingtes Verhalten durch vergangene Information geschätzt wird.

Für die **unbedingte mittlere quadratische Abweichung** der \mathcal{D} -messbaren Prognose \hat{X} für X gilt

$$\text{MSE}_X(\hat{X}) = \mathbb{E} \left[\text{MSE}_{X|\mathcal{D}}(\hat{X}) \right] = \mathbb{E}[\text{Var}(X|\mathcal{D})] + \mathbb{E} \left[(\hat{X} - \mathbb{E}[X|\mathcal{D}])^2 \right].$$

Es werden nun die beiden Fälle X von \mathcal{D} unabhängig bzw. X von \mathcal{D} abhängig betrachtet.

- **Fall 1:** X ist von \mathcal{D} unabhängig.

Das ist z.B. der Fall, wenn wir eine i.i.d. Stichprobe haben und X prognostizieren wollen. Es gilt dann

$$\mathbb{E}[X|\mathcal{D}] = \mathbb{E}[X] \quad \text{und} \quad \text{Var}(X|\mathcal{D}) = \text{Var}(X).$$

Für die unbedingte mittlere quadratische Abweichung erhält man dann

$$\text{MSE}_X(\hat{X}) = \text{Var}(X) + \mathbb{E}\left[\left(\hat{X} - \mathbb{E}[X]\right)^2\right].$$

Falls \hat{X} ein unverzerrter Schätzer für $\mathbb{E}[X]$ ist, d.h. $\mathbb{E}[\hat{X}] = \mathbb{E}[X]$, gilt

$$\text{MSE}_X(\hat{X}) = \text{Var}(X) + \text{Var}(\hat{X}).$$

Beispiel: Seien X und X_1, \dots, X_n i.i.d. mit Mittelwert μ und Varianz $\sigma^2 < \infty$. Dann gilt für den Schätzer $\hat{X} = \frac{1}{n} \sum_{i=1}^n X_i$

$$\text{MSE}_{X|\mathcal{D}}(\hat{X}) = \sigma^2 + \left(\frac{1}{n} \sum_{i=1}^n X_i - \mu\right)^2$$

Nach dem Gesetz der großen Zahlen konvergiert der letzte Term gegen Null. In der Praxis ist man jedoch an einer Schätzung der mittleren quadratischen Abweichung für ein konkretes n interessiert. Nachdem μ aber nicht bekannt ist, kann nur der Erwartungswert für diese Abweichung berechnet werden, die unbedingte mittlere quadratische Abweichung:

$$\text{MSE}_X(\hat{X}) = \text{Var}(X) + \text{Var}(\hat{X}) = \sigma^2 + \frac{\sigma^2}{n}$$

Man sieht also, dass die Schwankung von $\frac{1}{n} \sum_{i=1}^n X_i$ um μ im Mittel von der Ordnung σ/\sqrt{n} ist.

- **Fall 2:** X ist von der σ -Algebra \mathcal{D} , die von den Beobachtungen erzeugt wird, abhängig.

Hier ist die Situation komplizierter. Es gilt

$$\begin{aligned} \text{MSE}_X(\hat{X}) &= \mathbb{E}[\text{Var}(X|\mathcal{D})] + \mathbb{E}\left[\left(\hat{X} - \mathbb{E}[X|\mathcal{D}]\right)^2\right] = \\ &= \text{Var}(X) - \text{Var}(\mathbb{E}[X|\mathcal{D}]) + \mathbb{E}\left[\left(\hat{X} - \mathbb{E}[X|\mathcal{D}]\right)^2\right] = \\ &= \text{Var}(X) - \mathbb{E}\left[\left(\mathbb{E}[X|\mathcal{D}] - \mathbb{E}[X]\right)^2\right] + \mathbb{E}\left[\left(\hat{X} - \mathbb{E}[X|\mathcal{D}]\right)^2\right] = \\ &= \text{Var}(X) + \mathbb{E}\left[-\left(\mathbb{E}[X|\mathcal{D}] - \mathbb{E}[X]\right)^2 + \left(\hat{X} - \mathbb{E}[X|\mathcal{D}]\right)^2\right] = \\ &= \text{Var}(X) + \mathbb{E}\left[\left(\hat{X} - \mathbb{E}[X|\mathcal{D}] + \mathbb{E}[X|\mathcal{D}] - \mathbb{E}[X]\right)^2\right] \\ &\quad - 2 \mathbb{E}\left[\left(\mathbb{E}[X|\mathcal{D}] - \mathbb{E}[X]\right)^2\right] \end{aligned}$$

$$\begin{aligned} & - 2 \mathbb{E} \left[\left(\hat{X} - \mathbb{E}[X | \mathcal{D}] \right) \left(\mathbb{E}[X | \mathcal{D}] - \mathbb{E}[X] \right) \right] = \\ & = \text{Var}(X) + \mathbb{E} \left[\left(\hat{X} - \mathbb{E}[X] \right)^2 \right] \\ & - 2 \mathbb{E} \left[\left(\mathbb{E}[X | \mathcal{D}] - \mathbb{E}(X) \right) \left(\hat{X} - \mathbb{E}(X) \right) \right]. \end{aligned}$$

Für einen unverzerrten Schätzer \hat{X} , d.h. $\mathbb{E}[\hat{X}] = \mathbb{E}[X]$, gilt

$$\text{MSE}_X(\hat{X}) = \text{Var}(X) + \text{Var}(\hat{X}) - 2 \text{Cov}(\mathbb{E}[X | \mathcal{D}], \hat{X}).$$

Dieser Werte gibt wieder nur Auskunft über die mittlere quadratische Abweichung, gemittelt über alle möglichen Daten, man erhält aber keine Information über die Qualität der Prognose \hat{X} für gegebene Daten.

Anhang B

Numerisches Beispiel - Vergleich der Methoden

In Tabelle B.1 bzw. Abbildung B.1 werden die Reserve-Prognosen der unterschiedlichen Methoden pro Jahr und als Summe dargestellt. In Abbildung B.2 ist die Prognose \hat{R} für die Gesamtreserve $R = \sum_{i=1}^n R_i$ und der Standardfehler, d.h. die Wurzel aus der mittleren quadratischen Abweichung, für verschiedene Verfahren dargestellt. In Abbildung B.3 sieht man die Reserveprognosen des Chain-Ladder-Verfahrens im Vergleich zu jenen des Munich-Chain-Ladder-Verfahrens.

Tabelle B.1: Vergleich der Reserve-Prognosen der verschiedenen Verfahren

i	R_i^{CL}	R_i^{BF}	R_i^{BH}	R_i^{CC}	R_i^{AD}	R_i^{MCL}
1	143	155	143	136	168	188
2	846	849	846	810	936	1.679
3	2.998	3.268	3.001	3.000	3.352	3.190
4	3.996	4.281	4.000	4.223	4.665	3.787
5	5.836	6.524	5.851	6.222	6.744	7.204
6	7.281	8.598	7.320	7.720	8.258	8.784
7	22.965	25.642	23.212	23.068	23.664	22.544
8	90.166	92.966	91.214	95.653	96.573	95.017
	R^{CL}	R^{BF}	R^{BH}	R^{CC}	R^{AD}	R^{MCL}
	134.230	142.282	135.587	140.832	144.360	142.393

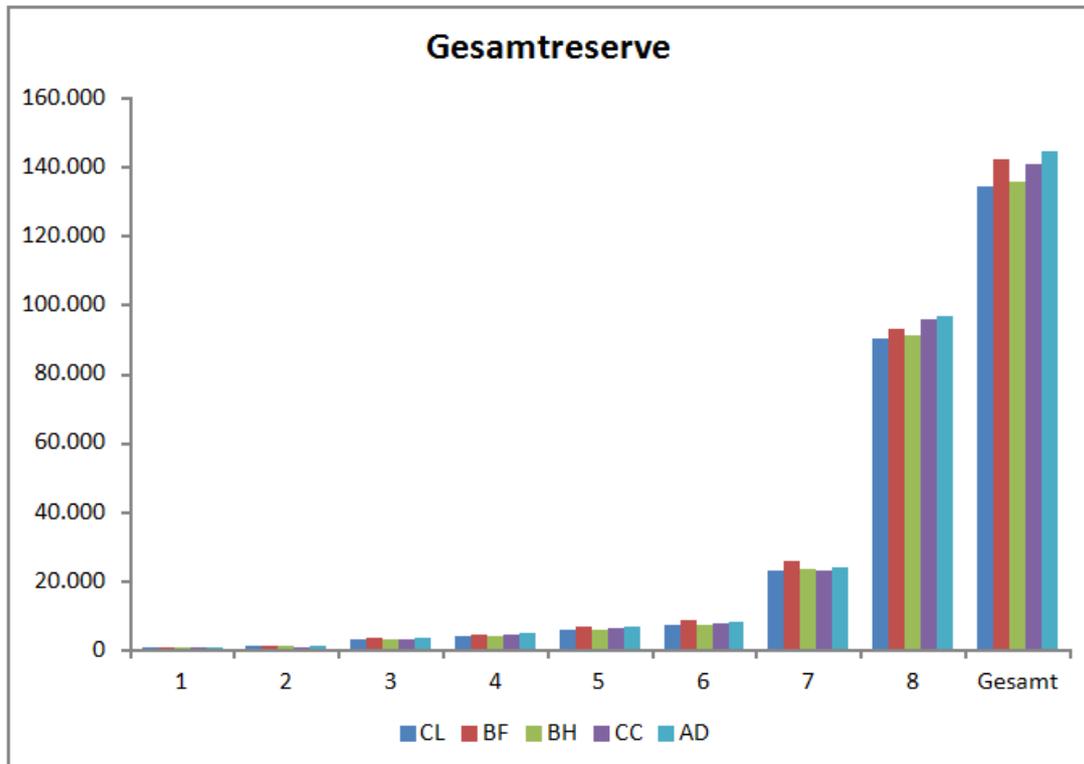


Abbildung B.1: Vergleich der Reserve-Prognosen der verschiedenen Verfahren

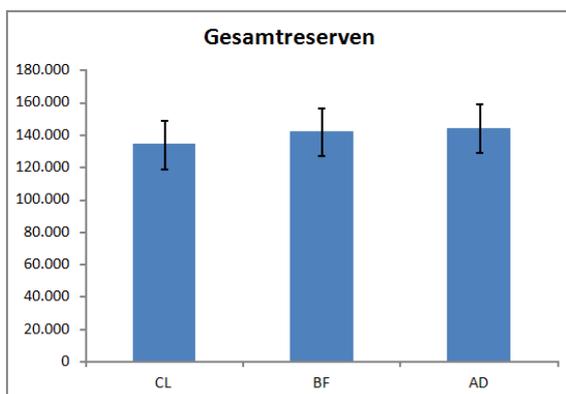


Abbildung B.2: Vergleich der Prognosen der Gesamtreserve der verschiedenen Verfahren

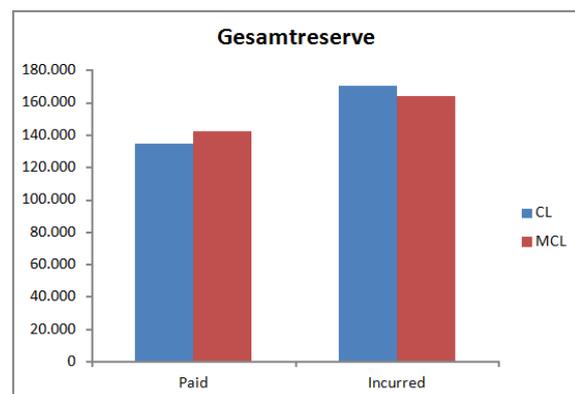


Abbildung B.3: Vergleich der Prognosen der Gesamtreserve des Chain-Ladder- und des Munich-Chain-Ladder-Verfahrens

Literaturverzeichnis

- [1] Klaus D. Schmidt. *Versicherungsmathematik, 3. Auflage*. Springer-Verlag, Berlin, 2009.
- [2] Thomas Mack. *Schadenversicherungsmathematik, 2. Auflage*. Verlag Versicherungswirtschaft, Karlsruhe, 2002.
- [3] RICHTLINIE 2009/138/EG des Europäischen Parlaments und des Rates vom 25. November 2009 betreffend die Aufnahme und Ausübung der Versicherungs- und der Rückversicherungstätigkeit (Solvabilität II) (Neufassung). Amtsblatt der Europäischen Union, 17. Dezember 2009.
- [4] Michael Merz und Mario V. Wüthrich. A credibility approach to the Munich Chain-Ladder Method. *Blätter der Deutschen der Gesellschaft für Versicherungs- und Finanzmathematik*, 27(4):619–628, 2006.
- [5] Klaus Th. Hess und Klaus D. Schmidt. A Comparison of Models for the Chain-Ladder Method. *Insurance: Mathematics and Economics*, 31(3):351–364, 2002.
- [6] Mario V. Wüthrich und Michael Merz. *Stochastic Claims Reserving Methods in Insurance*. John Wiley & Sons Ltd, Chichester, 2008.
- [7] Thomas Mack. Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *Astin Bulletin*, 23(2):93–109, 1993.
- [8] Klaus D. Schmidt. Methods and Models of Loss Reserving Based on Run-Off Triangles: A Unifying Survey. *Casualty Actuarial Society Forum*, pages 269–317, Fall 2006.
- [9] Erich L. Lehmann. *Theory of Point Estimation*. John Wiley & Sons, Inc., New York, 1983.
- [10] Hermann Witting. *Mathematische Statistik I*. Teubner, Stuttgart, 1985.
- [11] Johann Pfanzagl. *Parametric Statistical Theory*. De Gruyter, Berlin, 1994.
- [12] Schweizerische Eidgenossenschaft. Technical document on the Swiss Solvency Test. <http://www.finma.ch>, 2. Oktober 2006.
- [13] Thomas Mack. Credible Claims Reserves: The Benktander Method. *Astin Bulletin*, 30(2):333–347, 2000.
- [14] Michael Radtke und Klaus D. Schmidt. *Handbuch zur Schadenreservierung*. Verlag Versicherungswirtschaft, Karlsruhe, 2004.
- [15] Michael Merz und Mario V. Wüthrich. Prediction Error of the Multivariate Additive Loss Reserving Method for Dependent Lines of Business. *Variance*, 3(1):131–151, 2009.

- [16] Gerhard Quarg und Thomas Mack. Munich Chain Ladder: A Reserving Method that Reduces the Gap between IBNR Projections based on Paid Losses and IBNR Projections based on Incurred Losses. *Blätter der Deutschen der Gesellschaft für Versicherungs- und Finanzmathematik*, 26(4):597–630, 2004.
- [17] Peter J. Brockwell und Richard A. Davis. *Time Series: Theory and Methods*. Springer-Verlag, Berlin, 1991.
- [18] Hans Bühlmann und Alois Gisler. *A Course in Credibility Theory and its Applications*. Springer-Verlag, Berlin, 2005.
- [19] Gesamtverband der Deutschen Versicherungswirtschaft e. V. *Methoden zur Schätzung von Schaden- und Prämienrückstellungen in der Kompositversicherung. Überarbeitete Fassung*. Berlin, 2001.
- [20] Thomas Mack. The Standard Error of Chain Ladder Estimates: Recursive Calculation and Inclusion of a Tail Factor. *Astin Bulletin*, 29(2):361–366, 1999.
- [21] Peter Marchert. Spaltenbuchstabe ermitteln. *Office-Loesung.de*: http://www.office-loesung.de/ftopic63712-0-0_asc.php, 14. Jänner 2006.