
Seminar aus Finanz- und Versicherungsmathematik

**Schadenreservierung bei lang andauernder
Schadenabwicklung**

Christina Pommer

Inhaltsverzeichnis

1	Problemstellung	3
1.1	Lang andauernde Schadenabwicklung und ihre Ursachen	3
1.2	Grundsätzliches zu den mathematischen Verfahren	4
1.3	Das Abwicklungsdreieck	5
1.4	Bedeutung der Schätzgenauigkeit	6
2	Das Chain – Ladder – Verfahren	8
2.1	Allgemeines	8
2.2	Das Chain – Ladder – Verfahren	8
2.3	Sensitivität und Genauigkeit	11
2.4	Beispiel	14
3	Kreuzklassifizierte parametrische Verfahren	17
3.1	Allgemeines	17
3.2	Ein auf der Gammaverteilung beruhendes Verfahren	18
3.3	Weitere Verfahren	19
4	Modifikationen der Verfahren	20
4.1	Trennung von IBNR- und IBNER- Schäden	20
4.2	Separation von Kalenderjahreffekten	22
5	Zusammenfassung	24
6	Anhang	25

1 Problemstellung

1.1 Lang andauernde Schadenabwicklung und ihre Ursachen

Zwischen dem Eintritt eines Schadenfalles und dessen vollständiger Regulierung besteht immer eine zeitliche Differenz, da einerseits der Schaden dem Versicherer gemeldet und von ihm überprüft werden muss, und andererseits die Abwicklung selbst einige Zeit, zum Beispiel für die Reparatur, in Anspruch nehmen kann. Im Gegensatz zu den meisten Versicherungsbranchen, wo die Regulierungszeit meist nur ein bis zwei Monate dauert, kann sie in der Haftpflichtversicherung durchaus mehrere Jahre betragen. Zwei wesentlichen Gründe dafür sind:

1. *Späte Manifestation*: Manche Schäden werden erst lange Zeit nach ihrer Verursachung bemerkt, weil sie nur unter bestimmten Gegebenheiten auftreten:
z.B. Fehler eines Architekten, die erst bei besonderer Belastung eines Bauwerks bemerkbar werden, oder Fehler eines Notars, die erst bei der Eröffnung eines Testaments auftreten. Die Haftpflichtversicherung deckt auch solche Schäden, sofern die Polizza zur Zeit der Verursachung des Fehlers in Kraft war.
2. *Lange Regulierungsdauer*: Selbst nach dem Schadeneintritt und Meldung an den Versicherer kann es lange Zeit dauern, bis die endgültige Schadenhöhe bekannt ist, z.B. wenn sie vom Ausgang eines Gerichtsprozesses oder von Kosten einer längeren ärztlichen Behandlung abhängt.

Es gibt also zwei Arten von Schäden zum Ende einer Abrechnungsperiode, deren Höhe noch nicht vollständig bekannt ist:

- Schäden nach Nummer 1, die zwar schon verursacht bzw. eingetreten, dem Versicherer aber noch nicht bekannt sind, werden als IBNR- Schäden bezeichnet (IBNR = incurred but not reported). Der Versicherer ist verpflichtet eine IBNR-Schadenreserve zu bilden, falls solche Schäden erfahrungsgemäß zu erwarten sind.
- Schäden nach Nummer 2, die zwar dem Versicherer gemeldet, aber noch nicht vollständig reguliert sind, werden IBNER- Schäden (IBNER = incurred but not enough reserved), manchmal auch RBNS- Schäden (RBNS = reported but not settled) genannt. Der zuständige Schadensachbearbeiter muss, basierend auf der verfügbaren Information des Einzelschadens und seiner Erfahrung, eine vermutlich ausreichende Einzelfallreserve festsetzen. Die später tatsächlich anfallenden

Kosten können dann, gemessen an der Einzelfallreserve, zu Abwicklungsgewinnen oder -verlusten führen. Nimmt die Schadenhöhe im Lauf der Abwicklungszeit stärker zu als angenommen, so tritt ein Abwicklungsverlust ein. Die Reserve für ein erfahrungsgemäß negatives Abwicklungsergebnis eines ganzen Bestandes wird als IBNER-Schadenreserve bezeichnet.

Im Folgenden wird für die zwei eben genannten Reservearten der Überbegriff (Spät-) Schadenreserve verwendet.

Eine möglichst genaue Schätzung der Schadenreserve ist einerseits zur Erstellung der externen Rechnungslegung und andererseits zur Prämienkalkulation, welche von dem vergangenen Schadenverlauf abhängt, erforderlich. Die IBNR- Reserve wird in beiden Fällen benötigt, anders als die IBNER-Reserve, welche nur für die Prämienkalkulation und die intere Rechnungslegung verwendet wird.

Die Spätschadenreservierungsproblematik betrifft den Rückversicherer in der Haftpflichtversicherung bei den häufig vorkommenden Schadenexzedentenrückversicherungsverträgen stark. Hier übernimmt der Rückversicherer jenen Teil der Schadenhöhe eines Schadenfalles, der einen zuvor festgesetzten Betrag, die Priorität übersteigt. Oft ist dem Erstversicherer bei der Meldung eines Schadens noch nicht bekannt, ob die Schadenhöhe die Priorität übersteigen wird, und er kann deshalb dem Rückversicherer am Ende eines Vertragsjahres nicht definitiv mitteilen, welche Beträge er übernehmen muss. Für den Rückversicherer steigt daher die Spätschadenreservierungsproblematik enorm, da für ihn auch solche Schäden als IBNR-Schäden zählen.

1.2 Grundsätzliches zu den mathematischen Verfahren

Aufgrund der entscheidenden Bedeutung der Schätzung der Spätschadenreserve für die Rechnungslegung und die Prämienkalkulation sind einige mathematische Reserveschätzverfahren entwickelt worden. Sie alle versuchen, die Erfahrungen und Informationen früherer Schadeneintrittsjahre auf spätere Anfalljahre zu übertragen. Somit kann es jedoch auch passieren, dass sie wegen Trend- oder Strukturänderungen, die sich im Lauf der Zeit ereignen (z.B. in der Schadenregulierung, Zeichnungspolitik oder Rechtsprechung) versagen. Für die nachstehend dargestellten Verfahren nehmen wir deshalb an, dass im Verlauf der betrachteten Jahre keine solchen Trend- oder Strukturbrüche stattfinden.

1.3 Das Abwicklungsdreieck

Den meisten mathematischen Verfahren zur Berechnung der Spätschadenreserve liegt das sogenannte Abwicklungsdreieck zugrunde, das sich in seiner Grundform folgendermaßen darstellen lässt:

Sei S_{ik} , $1 \leq i \leq I$, $k = 1, 2, \dots$, der im Abwicklungsjahr k aufgewendete Betrag für im Anfalljahr i eingetretene Schäden. Das erste Abwicklungsjahr, $k = 1$, ist das Anfalljahr selbst, das zweite Abwicklungsjahr, $k = 2$, ist das dem Anfalljahr folgende Kalenderjahr, usw.

Das jüngste Anfalljahr $i = I$ ist das dem aktuellen Jahr vorangehende Jahr, von dem nur die Schäden bekannt sind, welche in diesem Jahr eingetreten und gemeldet wurden. Von dem am längsten zurückliegenden Anfalljahr $i = 1$ sind bereits $k = I$ Abwicklungsjahre bekannt. Das erste Anfalljahr wird wenn möglich so gewählt, dass es nach fachmännischer Einschätzung nahezu vollständig abgewickelt ist, so dass die Beträge $S_{1,I+1}$, $S_{1,I+2}, \dots$, der künftigen Abwicklungsjahre vermutlich alle gleich Null sind.

Im Abwicklungsdreieck (Tabelle 1) sind nur die Beträge S_{ik} mit $i + k \leq I + 1$ bekannt, wobei die Zeilen die Anfalljahre i und die Spalten die Abwicklungsjahre k darstellen.

		Abwicklungsjahre k							
Schadenjahre i	S_{11}	S_{12}	\dots	S_{1k}	\dots	$S_{1,I+1-i}$	\dots	$S_{1,I-1}$	S_{1I}
	S_{21}	S_{22}	\dots	S_{2k}	\dots	$S_{2,I+1-i}$	\dots	$S_{2,I-1}$	
	\vdots	\vdots	\dots	\vdots	\dots	\vdots			
	S_{i1}	S_{i2}	\dots	S_{ik}	\dots	$S_{i,I+1-i}$			
	\vdots	\vdots	\dots	\vdots					
	$S_{I+1-k,1}$	$S_{I+1-k,2}$	\dots	$S_{I+1-k,k}$					
	\vdots	\vdots							
	$S_{I-1,1}$	$S_{I-1,2}$							
	S_{I1}								

Tabelle 1: Das Abwicklungsdreieck

Die Parallelen zur Hypotenuse des Dreiecks entsprechen den einzelnen Kalenderjahren. Zur entsprechenden Situation vor einem Jahr sind nur die Werte S_{1I} , $S_{I-1,2}, \dots$, $S_{I+1-k,k}, \dots$, S_{1I} der Hypotenuse (Diagonale) dazugekommen, die das letzte Kalenderjahr betreffen.

Nach einem möglichen kurzen Anstieg zu Beginn der Schadenabwicklung nehmen die Beträge S_{ik} für jedes Anfalljahr i mit wachsendem k tendenziell ab und werden

schlussendlich gleich Null sein, da Änderungen im Schadenstand nach langer Abwicklungsdauer immer unwahrscheinlicher werden.

Unter der Annahme, dass nach I Abwicklungsjahren alle Schäden eines Anfalljahres bekannt und vollständig reguliert sind, bezeichnet $S_{i+} = S_{i1} + S_{i2} + \dots + S_{iI}$ den Gesamtschaden für Anfalljahr i . Von S_{i+} ist bis jetzt jedoch nur der Betrag $S_{i1} + S_{i2} + \dots + S_{i,I+1-i}$ bekannt, der noch unbekannt Teil ist

$$R_i := S_{i,I+2-i} + S_{i,I+3-i} + \dots + S_{iI}.$$

Ziel der mathematischen Verfahren ist, den Betrag R_i für jedes Anfalljahr zu schätzen.

Es gibt verschiedene Abwandlungen des Abwicklungsdreiecks, so steht häufig an einer Stelle (i, k) der kumulierte Schadenstand $C_{ik} = S_{i1} + S_{i2} + \dots + S_{ik}$ und nicht der Zuwachs S_{ik} . Die Zuwächse sind durch $S_{ik} = C_{ik} - C_{i,k-1}$ mit $C_{i0} = 0$ einfach zu bestimmen.

Die Beträge S_{ik} beziehen sich im Allgemeinen auf die Schäden eines Bestandes von Risiken. In der Schadenreservierung ist es wichtig, möglichst große und homogene Teilbestände zu bilden, da einerseits die Abwicklung eines einzelnen Schadenfalles wenig über andere Schadenabwicklungen aussagt und deshalb Informationen nach dem Prinzip vom Ausgleich im Kollektiv immer zuverlässiger werden, wenn sie auf mehr Schäden beruhen. Andererseits kann es passieren, dass Informationen verloren gehen, wenn Teilbestände zusammengefasst werden, die sich in ihrem Abwicklungsverhalten unterscheiden, weshalb auf die Bildung homogener Teilbestände geachtet werden muss. Die Schätzung der Spätschadenreserve ist daher keine rein mathematische Aufgabe, sondern ist stark von der Wahl der Teilbestände abhängig.

1.4 Bedeutung der Schätzgenauigkeit

Nicht zu vergessen ist, dass es sich bei der zu schätzenden Spätschadenreserve um eine Zufallsvariable handelt, bei der man sich natürlich für ihren Erwartungswert interessiert. Hier gilt jedoch nicht, wie bei der Prämienkalkulation, dass die Abweichung des Gesamtschadens vom Erwartungswert relativ gering ist, da die Schäden mit langer Abwicklungsdauer nur einen kleinen Teil der Gesamtzahl aller Schäden bilden. Da also die Schätzung des Erwartungswertes mit großen Unsicherheiten verbunden ist, ist neben der Punktschätzung für den Erwartungswert der Reserve eine Angabe über die Genauigkeit dieser Schätzung wichtig.

Außerdem gilt, dass die Spätschadenreserve eine Prognose über mehrere Jahre beinhaltet, was nochmals zu höherer Unsicherheit führt. Mit der Reserveschätzung ist die Bindung entsprechender Geldmittel verbunden, die später nicht einfach aus Prämien

des damaligen Portefeuilles aufgestockt werden kann, falls die Reserve zu niedrig angesetzt wurde.

Ferner können zwei verschiedene Schätzverfahren nur dann verglichen werden, wenn zumindest von einem die Genauigkeit der Punktschätzung bekannt ist.

Der Angabe zur Schätzgenauigkeit kommt bei der Schadenreservierung also eine hohe Bedeutung zu. Sie sagt aus, welchen Informationswert die Punktschätzung hat.

Wir werden im Weiteren ein paar Schätzverfahren kennenlernen, darunter auch das bedeutende Chain-Ladder-Verfahren. Hierfür werde ich eine Möglichkeit der Berechnung zur Schätzung und Genauigkeitsmessung der Spätschadenreserve anführen. Ziel ist, die Spätschadenreserve R_i für Anfalljahr i möglichst gut durch einen Schätzer \hat{R}_i zu prognostizieren, wobei wir mit der zur Varianz analogen bedingten mittleren quadratischen Abweichung arbeiten.

2 Das Chain – Ladder – Verfahren

2.1 Allgemeines

Das Chain-Ladder Verfahren ist eines der ältesten und bekanntesten mathematischen Verfahren zur Schätzung der Schadenreserve, das heute noch eine wichtige Rolle spielt. Das Verfahren ist verteilungsfrei¹ und wird normalerweise in rein deterministischer Form dargestellt, kann aber dennoch stochastisch modelliert werden. Da das stochastische Modell einerseits die Grenzen bzw. die erforderlichen Voraussetzungen zeigt, und andererseits die Angabe der Genauigkeit des Reserveschätzers ermöglicht, werden im weiteren die Verfahren immer stochastisch modelliert.

2.2 Das Chain – Ladder – Verfahren

Wir lassen für jedes Anfalljahr i einen individuellen Erwartungswert der Reserve zu, nehmen aber an, dass die Aufteilung des Endschadens auf die einzelnen Abwicklungsjahre im Schnitt für jedes Anfalljahr gleich ist.

Mit den Bezeichnungen aus Kapitel 1.3 schreiben wir den Endschaden des Anfalljahres i

$$C_{iI} = S_{i1} + S_{i2} + \dots + S_{iI}$$

in der multiplikativen Form

$$C_{iI} = C_{i1} \cdot F_{i1} \cdot F_{i2} \cdot \dots \cdot F_{i,I-1}, \quad \text{wobei} \quad F_{ik} = \frac{C_{i,k+1}}{C_{ik}}$$

die multiplikative Zunahme des akkumulierten Schadenstandes C_{ik} von Abwicklungsjahr k zu Abwicklungsjahr $k + 1$ ist. Für diese Darstellung sollte $C_{ik} > 0$ für alle Schadenstände sein, andernfalls wird die multiplikative Darstellung statt C_{i1} mit dem ersten $C_{ik} > 0$ geschrieben.

Wir haben, wie vorhin erwähnt, die Annahme getroffen, dass die Aufteilung des Endschadens auf die einzelnen Abwicklungsjahre im Schnitt für jedes Anfalljahr gleich ist. Dies lässt sich so formalisieren, dass wir für die Zufallsvariablen F_{ik} einen vom Anfalljahr i unabhängigen Erwartungswert

$$\mathbb{E}(F_{ik}) = f_k \quad 1 \leq i \leq I, \quad 1 \leq k \leq I - 1$$

¹Verteilungsfreien Modellen liegt keine spezielle Verteilungsform zu Grunde.

annehmen.

Die f_k geben die durchschnittliche Steigerung des Schadenstands von Abwicklungsjahr k auf Abwicklungsjahr $k + 1$ an und werden mittels des C_{ik} -gewichteten arithmetischen Mittels

$$\hat{f}_k = \frac{\sum_{i=1}^{I-k} C_{ik} F_{ik}}{\sum_{i=1}^{I-k} C_{ik}} = \frac{\sum_{i=1}^{I-k} C_{i,k+1}}{\sum_{i=1}^{I-k} C_{ik}} \quad 1 \leq k \leq I-1 \quad (1)$$

aus den bis jetzt vorliegenden Werten der C_{ik} geschätzt. Eine Besonderheit des Chain-Ladder-Verfahrens ist, dass der Endschaden C_{iI} durch

$$\hat{C}_{iI} = C_{i,I+1-i} \cdot \hat{f}_{I+1-i} \cdot \dots \cdot \hat{f}_{I-1}, \quad 2 \leq i \leq I, \quad (2)$$

bzw. die Reserve $R_i = C_{iI} - C_{i,I+1-i}$ der einzelnen Anfalljahre i durch

$$\hat{R}_i = C_{i,I+1-i} \cdot (\hat{f}_{I+1-i} \cdot \dots \cdot \hat{f}_{I-1} - 1) \quad (3)$$

geschätzt werden kann. Dass diese Darstellung möglich ist, wird später gezeigt.

Die Prognose beruht lediglich auf dem aktuellen Schadenstand $C_{i,I+1-i}$ des Anfalljahres i , wobei die vergangenen Schadenstände $C_{i1}, \dots, C_{i,I-i}$ nicht berücksichtigt werden, was verwundernswert ist. Daher treffen wir die dem Chain-Ladder-Verfahren zu Grunde liegenden Modellannahmen:

(CL1) Es gibt Abwicklungsfaktoren f_1, \dots, f_{I-1} mit

$$\mathbb{E} \left(\frac{C_{i,k+1}}{C_{ik}} \mid C_{i1}, \dots, C_{ik} \right) = f_k, \quad 1 \leq i \leq I, \quad 1 \leq k \leq I-1, \quad \text{für } C_{ik} > 0$$

oder gleichbedeutend $\mathbb{E}(C_{i,k+1} \mid C_{i1}, \dots, C_{ik}) = C_{ik} f_k$.

Diese Annahme besagt, dass der bedingte Erwartungswert von $C_{i,k+1}$ nur von C_{ik} abhängt. Die vorigen Werte $C_{i1}, \dots, C_{i,k-1}$ verbessern den Informationsgehalt des jeweils aktuellsten Schadenstandes nicht.

Unter Verwendung des Satzes vom iterierten Erwartungswert (siehe Anhang 1.) folgt aus CL1

$$\mathbb{E} \left(\frac{C_{i,k+1}}{C_{ik}} \right) = \mathbb{E} \left(\mathbb{E} \left(\frac{C_{i,k+1}}{C_{ik}} \mid C_{i1}, \dots, C_{ik} \right) \right) = f_k,$$

was wieder unser eigentliches Modell $\mathbb{E}(F_{ik}) = f_k$ darstellt.

(CL2) Die Anfalljahre $\{C_{i1}, \dots, C_{iI}\}$, $1 \leq i \leq I$, sind unabhängig.

Die Formulierung der Modellannahme CL1 als bedingter Erwartungswert zeigt, dass wir nicht $\mathbb{E}(C_{iI})$ bzw. $\mathbb{E}(R_i)$, sondern vielmehr $\mathbb{E}(C_{iI}|D)$ bzw. $\mathbb{E}(R_i|D)$ schätzen wollen, wobei

$D = \{C_{ik} \mid i + k \leq I + 1\}$ die bisher bekannten Daten sind.

Dass die Chain-Ladder-Projektion (2) tatsächlich auf dem Modell aus CL1 und CL2 basiert, zeigt folgender Satz:

Satz: Unter den Annahmen CL1 und CL2 gilt

$$\mathbb{E}(C_{iI}|D) = C_{i,I+1-i} \cdot f_{I+1-i} \cdot \dots \cdot f_{I-1}, \quad 2 \leq i \leq I.$$

Beweis: Mit der Abkürzung $D_i = \{C_{i1}, \dots, C_{i,I+1-i}\}$ gilt wegen CL2 und durch wiederholte Anwendung von CL1

$$\begin{aligned} \mathbb{E}(C_{iI}|D) &= \mathbb{E}(C_{iI}|D_i) \\ &= \mathbb{E}(\mathbb{E}(C_{iI}|C_{i1}, \dots, C_{i,I-1})|D_i) \\ &= \mathbb{E}(C_{i,I-1}f_{I-1}|D_i) \\ &= \mathbb{E}(C_{i,I-1}|D_i) f_{I-1} \\ &= \dots \\ &= \mathbb{E}(C_{i,I+1-i}|D_i) f_{I+1-i} \cdot \dots \cdot f_{I-1} \\ &= C_{i,I+1-i} f_{I+1-i} \cdot \dots \cdot f_{I-1}. \end{aligned}$$

□

Ein weiterer Satz zeigt, dass die benutzten Schätzer für $f_{I+1-i} \cdot \dots \cdot f_{I-1}$ vernünftig sind:

Satz: Unter den Annahmen CL1 und CL2 sind die gemäß (1) und (2) berechneten Schätzer \hat{f}_k erwartungstreu (siehe Anhang 2.) und unkorreliert mit

$$\mathbb{E}(\hat{f}_{I+1-i} \cdot \dots \cdot \hat{f}_{I-1}) = f_{I+1-i} \cdot \dots \cdot f_{I-1}.$$

Beweis: Siehe [Ma], Kapitel 3.2.4

Angeichts der Tatsache, dass f_{k-1} und f_k von denselben Daten $C_{1k}, \dots, C_{I-k,k}$ abhängen, ist die Unkorreliertheit der \hat{f}_k eher überraschend. Weiters gilt auch, dass die Faktoren F_{ik} für jedes feste Anfalljahr i unkorreliert sind, insbesondere sind also auch aufeinander folgende Abwicklungsfaktoren $F_{i,k-1}$ und F_{ik} unkorreliert. Deshalb sollte das Chain-Ladder-Verfahren nicht bei Portefeuilles angewandt werden, bei denen im Allgemeinen auf einen eher hohen Abwicklungsfaktor ein niedriger folgt, oder umgekehrt.

Da $F_{ik} = \frac{C_{i,k+1}}{C_{ik}}$ für $C_{ik} = 0$ undefiniert ist, wird häufig

$$\hat{f}_k = \frac{\sum_{i=1}^{I-k} C_{ik} F_{ik}}{\sum_{i=1}^{I-k} C_{ik}} \quad \text{mit } F_{ik} := 0 \text{ falls } C_{ik} = 0$$

an Stelle von

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

verwendet, da dies f_k überschätzt, falls ein Paar $(C_{ik}, C_{i,k+1})$ mit $C_{ik} = 0$ und $C_{i,k+1} > 0$ enthalten ist.

Nachdem man sich die Spätschadenreserve \hat{R}_i der einzelnen Anfalljahre berechnet hat, kann die Gesamtreserve R mit

$$\hat{R} = \hat{R}_2 + \dots + \hat{R}_I$$

geschätzt werden.

2.3 Sensitivität und Genauigkeit

Es gibt zwei heikle Stellen im Abwicklungsdreieck, die auch beim Chain-Ladder-Verfahren zu beachten sind. Die erste sensible Situation stellt das rechte obere Eck dar: Hier wird der Schätzer des letzten Abwicklungsfaktors f_{I-1} berechnet, der nur von einem Beobachtungswert S_{1I} abhängt. Die im ältesten Anfalljahr beobachtete Veränderung $\hat{f}_{I-1} = (C_{1,I-1} + S_{1I})/C_{1,I-1}$ wird auf alle anderen Anfalljahre übertragen, selbst wenn der Wert untypisch ist.

Das andere spitze Eck bildet die zweite kritische Stelle: Der Faktor C_{1I} , im jüngsten Anfalljahr, beruht nur auf den Daten aus einem Beobachtungsjahr. In manchen Versicherungsbereichen mit langer Abwicklungsdauer (vor allem bei Schadenexzedenten-Rückversicherungen) sagt der Zahlungsstand C_{1I} selbst am Ende des Anfalljahres noch wenig aus. Der Reserveschätzer $\hat{R}_I = C_{1I} \cdot (\hat{f}_1 \cdot \dots \cdot \hat{f}_{I-1} - 1)$ für das jüngste Anfalljahr beruht nun auf diesem unsicheren Wert. Betrachten wir den Extremfall $C_{1I} = 0$, so wird die ganze Reserve R_I zu $\hat{R}_I = 0$ geschätzt, was keinen Sinn macht.

Es ist möglich ein unplausibles C_{1I} zu korrigieren, indem man es durch den volumengewichteten Durchschnitt aller bisher beobachteten Anfalljahre

$$C_{1I} := v_I \sum_{i=1}^I C_{i1} / \sum_{i=1}^I v_i$$

ersetzt, wobei v_i das Volumen (d.h. die Anzahl der Risiken bzw. deren Gesamtversicherungssumme oder das Prämienvolumen) von Anfalljahr i ist.

Sind mehrere der Erstjahresstände C_{i1} ebenfalls unplausibel, kann man sie durch Rückwärtsprojektion

$$\hat{C}_{i1} = \frac{C_{i,I+1-i}}{\hat{f}_1 \cdot \dots \cdot \hat{f}_{I-i}}$$

aus den jeweils aktuellen Schadenständen schätzen.

Man kann nun C_{I1} im Chain-Ladder-Reserveschätzer \hat{R}_I aus (3) durch $v_I \hat{q}_1$ ersetzen, wobei \hat{q}_1 der Schätzer für die mittlere Erstjahresschadenquote ist:

$$\hat{q}_1 = \sum_{i=1}^I C_{i,I+1-i} / \sum_{i=1}^I (\hat{f}_1 \cdot \dots \cdot \hat{f}_{I-i} \cdot v_i)$$

Diese Methode wird als Cape-Cod-Verfahren bezeichnet, mit dem natürlich auch die anderen aktuellen Schadenstände $C_{I-1,2}, C_{I-2,3}, \dots$ angepasst werden können.

Eine andere einfache Methode besteht darin, zunächst die gesamt Nettoprämieinnahme des Anfalljahrs I abzüglich C_{I1} zu reservieren, bis nach dem folgenden Jahr zuverlässigere Daten vorliegen.

Angesichts dieser Sensitivitäten wollen wir nun eine Formel für die Genauigkeit des Reserveschätzers \hat{R}_i herleiten.

Die Schätzung der Genauigkeitsmessung der Spätschadenreserve basiert auf folgendem Modell: Ziel ist, die Spätschadenreserve R_i für Anfalljahr i möglichst gut durch einen Schätzer \hat{R}_i zu prognostizieren. Gesucht ist also der bedingte mittlere quadratische Fehler

$$\text{mse}(\hat{R}_i) = \mathbb{E}((R_i - \hat{R}_i)^2 | D)$$

zwischen Schätzer und Prognoseziel, gegeben die Daten des Abwicklungsdreiecks

$$D = \{C_{ik} | i + k \leq I + 1\}.$$

Zur Berechnung der mittleren quadratischen Abweichung benutzen wir den für beliebige Skalare a gültigen Verschiebungssatz

$$\mathbb{E}(X - a)^2 = \text{Var}(X) + (\mathbb{E}(X) - a)^2$$

bzw. dessen Verallgemeinerung auf bedingte Erwartungswerte

$$\mathbb{E}((X - h(Y))^2 | Y) = \text{Var}(X|Y) + (\mathbb{E}(X|Y) - h(Y))^2.$$

Da \hat{R}_i aus den Daten D berechnet wird und deshalb eine Funktion von D ist, $\hat{R}_i = h_i(D)$, gilt

$$\text{mse}(\hat{R}_i) = \text{Var}(R_i|D) + (\mathbb{E}(R_i|D) - \hat{R}_i)^2. \quad (4)$$

Da wegen $R_i = C_{iI} - C_{i,I+1-i}$ und $\hat{R}_i = \hat{C}_{iI} - C_{i,I+1-i}$

$$\text{mse}(\hat{R}_i) = \mathbb{E}((R_i - \hat{R}_i)^2 | D) = \mathbb{E}((C_{iI} - \hat{C}_{iI})^2 | D) = \text{mse}(\hat{C}_{iI})$$

gilt, ersetzen wir in obiger Gleichung \hat{R}_i durch \hat{C}_{iI} und erhalten:

$$\text{mse}(\hat{C}_{iI}) = \underbrace{\text{Var}(C_{iI}|D)}_{\text{Zufallsfehler}} + \underbrace{(\mathbb{E}(C_{iI}|D) - \hat{C}_{iI})^2}_{\text{Schätzfehler}}$$

Damit liefert der Verschiebungssatz die Zerlegung in die zwei Komponenten Zufalls- und Schätzfehler.

Betrachtet man Gleichung (4), so kann man erkennen, dass $\mathbb{E}(R_i|D)$ die beste Prognose von R_i im Quadratmittel ist, da $\hat{R}_i = \mathbb{E}(R_i|D)$ den quadratischen Fehler minimiert.

Damit wir diese zwei Komponenten schätzen können, treffen wir eine weitere Modellannahme über die Varianz der C_{ik} :

(CL3) Es gibt Proportionalitätskonstanten $\sigma_1^2, \dots, \sigma_{I-1}^2$ mit

$$\text{Var}\left(\frac{C_{i,k+1}}{C_{ik}} \mid C_{i1}, \dots, C_{ik}\right) = \frac{\sigma_k^2}{C_{ik}}, \quad 1 \leq i \leq I, \quad 1 \leq k \leq I-1, \quad \text{für } C_{ik} > 0$$

oder gleichbedeutend $\text{Var}(C_{i,k+1} \mid C_{i1}, \dots, C_{ik}) = C_{ik} \sigma_k^2$.

Zur konkreten Berechnung für σ_k^2 wählen wir als Schätzer

$$\hat{\sigma}_k^2 = \frac{1}{I-k-1} \sum_{j=1}^{I-k} C_{jk} \left(\frac{C_{j,k+1}}{C_{jk}} - \hat{f}_k \right)^2, \quad 1 \leq k \leq I-2$$

Wie man sehen kann, erhalten wir durch die Formel jedoch keinen Schätzer für σ_{I-1} . Falls $f_{I-1} = 1$ und unter der Annahme, dass nach $I-1$ Abwicklungsjahren keine weiteren Schadenzahlungen mehr anfallen, können wir $\sigma_{I-1} = 0$ setzen. Ansonsten muss die Folge $\sigma_1, \dots, \sigma_{I-3}, \sigma_{I-2}$ um einen weiteren Term extrapoliert werden oder mit Hilfe der Forderung

$$\frac{\hat{\sigma}_{I-3}}{\hat{\sigma}_{I-2}} = \frac{\hat{\sigma}_{I-2}}{\hat{\sigma}_{I-1}},$$

solange $\hat{\sigma}_{I-3} > \hat{\sigma}_{I-2}$ erfüllt ist, das heißt

$$\hat{\sigma}_{I-1}^2 = \min\left(\frac{\hat{\sigma}_{I-2}^4}{\hat{\sigma}_{I-3}^2}, \hat{\sigma}_{I-3}^2\right),$$

berechnet werden.

Dadurch können wir nun einen Schätzer für $\text{mse}(\hat{R}_i)$ herleiten.

Satz: Unter den Annahmen (CL1), (CL2) und (CL3) kann der mittlere quadratische Fehler des Reserveschätzers \hat{R}_i des Chain-Ladder-Verfahrens durch

$$(\text{s.e.}(\hat{R}_i))^2 = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

geschätzt werden.

Beweis: Siehe [Ma], Kapitel 3.2.5

Der Standardfehler $s.e.(\hat{R}_i)$ ist die Quadratwurzel des Schätzers $mse(\hat{R}_i)$.

Wir haben also jetzt in \hat{R}_i und $s.e.(\hat{R}_i)$ Schätzer für Erwartungswert und Standardfehler der Schadenreserve R_i für Anfalljahr i .

Die Gesamtreserve $R = R_2 + \dots + R_I$ wird durch $\hat{R} = \hat{R}_2 + \dots + \hat{R}_I$ geschätzt. Da die \hat{R}_i nicht voneinander unabhängig sind, weil sie über die Schätzer $\hat{f}_k, \hat{\sigma}_k$ positiv korrelieren, können wir nicht einfach die quadrierten Standardfehler der einzelnen Anfalljahre \hat{R}_i aufaddieren, um den Standardfehler von \hat{R} zu berechnen.

Satz: Unter denselben Annahmen wie im vorigen Satz erhalten wir für den mittleren quadratischen Fehler der Gesamtreserve $\hat{R} = \hat{R}_2 + \dots + \hat{R}_I$ den Schätzer

$$(s.e.(\hat{R}))^2 = \sum_{i=2}^I \left((s.e.(\hat{R}_i))^2 + \hat{C}_{iI} \left(\sum_{j=i+1}^I \hat{C}_{jI} \right) \sum_{k=I+1-i}^{I-1} \frac{2 \hat{\sigma}_k^2 / \hat{f}_k^2}{\sum_{n=1}^{I-k} C_{nk}} \right)$$

Beweis: Siehe [Ma], Kapitel 3.2.5

2.4 Beispiel

Bekannt sind die kumulierten Beträge C_{ik} für $i + k \leq I + 1$, wobei die Zeilen $i = 1, \dots, 6$ die Anfalljahre und die Spalten $k = 1, \dots, 6$ die Abwicklungsjahre darstellen:

	Jahr 1	Jahr 2	Jahr 3	Jahr 4	Jahr 5	Jahr 6
2000	4370	6293	10292	12460	13660	14307
2001	2701	5291	7162	8945	9338	
2002	4483	6729	10074	11142		
2003	3254	5804	8351			
2004	8010	12118				
2005	5582					

Tabelle 2: kumulierte Schadenstände

Wir berechnen nun die Schätzer $\hat{f}_k = \sum_{i=1}^{I-k} C_{i,k+1} / \sum_{i=1}^{I-k} C_{ik}$ für die Übergangsfaktoren:

\hat{f}_1	\hat{f}_2	\hat{f}_3	\hat{f}_4	\hat{f}_5
1,588	1,488	1,182	1,074	1,047

Tabelle 3: Übergangsfaktoren

Um die fehlenden Beträge $\hat{C}_{i,k+1}$ für $i+k > I+1$ zu berechnen, multiplizieren wir C_{ik} mit \hat{f}_k und erhalten:

	Jahr 1	Jahr 2	Jahr 3	Jahr 4	Jahr 5	Jahr 6
2000	4370	6293	10292	12460	13660	14307
2001	2701	5291	7162	8945	9338	9780
2002	4483	6729	10074	11142	11971	12538
2003	3254	5804	8351	9874	10608	11111
2004	8010	12118	18028	21315	22901	23986
2005	5582	8864	13187	15592	16752	17546

Tabelle 4: kumulierte Schadenstände für alle Zeiten

Schneller erhält man die Endschäden mit $\hat{C}_{iI} = C_{i,I+1-i} \cdot \hat{f}_{I+1-i} \cdot \dots \cdot \hat{f}_{I-1}$ für $2 \leq i \leq I$, womit die Reserve für jedes Anfalljahr bestimmt werden kann durch

$$\hat{R}_i = C_{i,I+1-i} \cdot (\hat{f}_{I+1-i} \cdot \dots \cdot \hat{f}_{I-1} - 1) :$$

\hat{R}_1	\hat{R}_2	\hat{R}_3	\hat{R}_4	\hat{R}_5	\hat{R}_6
0	442	1396	2760	11868	11964

Tabelle 5: Spätschadenreserven

Die Gesamtreserve R wird mit $\hat{R} = \hat{R}_2 + \dots + \hat{R}_I$ geschätzt:

$$\hat{R} = 28430$$

Zur Berechnung der Schätzgenauigkeit berechnen wir zunächst $\hat{\sigma}_k$ durch

$$\hat{\sigma}_k^2 = \frac{1}{I-k-1} \sum_{j=1}^{I-k} C_{jk} \left(\frac{C_{j,k+1}}{C_{jk}} - \hat{f}_k \right)^2, \quad 1 \leq k \leq I-2 \quad \text{und} \quad \hat{\sigma}_{I-1}^2 = \min \left(\frac{\hat{\sigma}_{I-2}^4}{\hat{\sigma}_{I-3}^2}, \hat{\sigma}_{I-3}^2 \right),$$

$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_4$	$\hat{\sigma}_5$
12,95	9,073	7,025	3,779	2,033

Tabelle 6: Schätzer der Parameter σ_k

um nun $s.e.(\hat{R}_i)$ mittels

$$(s.e.(\hat{R}_i))^2 = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

zu berechnen:

$s.e.(\hat{R}_2)$	$s.e.(\hat{R}_3)$	$s.e.(\hat{R}_4)$	$s.e.(\hat{R}_5)$	$s.e.(\hat{R}_6)$
255	599	992	2332	2851

Tabelle 7: mittlerer quadratischer Fehler von \hat{R}_i

Schließlich interessiert uns noch der mittlere quadratische Fehler der Gesamtreserve \hat{R} , den wir durch

$$(s.e.(\hat{R}))^2 = \sum_{i=2}^I \left((s.e.(\hat{R}_i))^2 + \hat{C}_{iI} \left(\sum_{j=i+1}^I \hat{C}_{jI} \right) \sum_{k=I+1-i}^{I-1} \frac{2 \hat{\sigma}_k^2 / \hat{f}_k^2}{\sum_{n=1}^{I-k} C_{nk}} \right)$$

erhalten:

$$s.e.(\hat{R}) = 4639$$

3 Kreuzklassifizierte parametrische Verfahren

Da es eine Vielzahl an mathematischen Verfahren zur Berechnung der Spätschadenreserve gibt, möchte ich in diesem Kapitel noch weitere Methoden erwähnen, allerdings nur bei dem auf der Gammaverteilung basierenden Verfahren näher ins Detail gehen.

3.1 Allgemeines

Eine der Modellannahmen beim Chain-Ladder-Verfahren war

$$(CL1) \quad \mathbb{E}(C_{i,k+1} \mid C_{i1}, \dots, C_{ik}) = C_{ik} f_k, \quad 1 \leq i \leq I, \quad 1 \leq k \leq I-1.$$

Wenden wir darauf den Erwartungswert-Operator an und benutzen den Satz vom iterierten Erwartungswert auf Seite 9, so sehen wir, dass das Chain-Ladder-Verfahren ein Spezialfall einer Modellklasse ist, für die

$$(A) \quad \mathbb{E}(C_{i,k+1}) = \mathbb{E}(C_{ik}) f_k, \quad 1 \leq i \leq I, \quad 1 \leq k \leq I-1$$

mit unbekanntem positiven Parametern f_1, \dots, f_{I-1} gilt.

Satz: Jedes Modell, für das (A) gilt, erfüllt auch

$$(B) \quad \mathbb{E}(S_{ik}) = x_i y_k, \quad 1 \leq i, \quad k \leq I,$$

mit unbekanntem Parametern $x_i, y_k > 0$ mit $y_1 + \dots + y_I = 1$. Umgekehrt ist auch jedes Modell der Art (B) ein Modell der Art (A).

Beweis: Siehe [Ma], Kapitel 3.3.1

Hierbei handelt es sich um die kreuzklassifizierte Modelle.

Der Unterschied zwischen Chain-Ladder-Modell (CL1) und den Modellklassen (A) und (B) wird deutlich, wenn wir annehmen, wir kennen die Modellparameter f_1, \dots, f_{I-1} von (CL1) und x_i, y_k von (B). Im Modell (B) wäre der Reserveschätzer (bei zusätzlich angenommener Unabhängigkeit der Abwicklungsjahre)

$$\hat{R}_i = x_i (y_{I+2-i} + \dots + y_I)$$

unabhängig von den beobachteten Daten², dh. \hat{R}_i ändert sich bei Verwendung anderer Daten D nicht. Dagegen würde sich der Wert der Chain-Ladder-Reserve

$$\hat{R}_i = C_{i,I+1-i} (f_{I+1-i} \cdot \dots \cdot f_{I-1} - 1)$$

unter anderen Daten sehr wohl verändern.

²Dies gilt nur bei festen x_i und y_k . im Allgemeinen sind die Schätzer für x_i und y_k von den Daten abhängig.

Ziel ist es, die Parameter x_i, y_k zu schätzen, womit wir in \hat{x}_i, \hat{y}_k automatisch auch Schätzer für die eigentlich interessanten Werte $\mathbb{E}(S_{ik}), i + k > I + 1$ der künftigen Schadenabwicklung erhalten.

Im folgenden Kapitel werden wir sehen, dass x_i, y_i trotz der fehlenden Beobachtungen von $S_{ik}, i + k > I + 1$ geschätzt werden können. Wir nehmen hierzu die Unabhängigkeit aller S_{ik} und die Positivität $S_{ik} > 0$ an, wir betrachten also nur Verteilungen, die auf der positiven reellen Achse definiert sind. Im Vergleich zum Chain-Ladder-Verfahren sind diese Annahmen deutlich strenger, da wir dort nur die Unabhängigkeit der Anfalljahre vorausgesetzt und negative Zuwächse nicht ausgeschlossen haben.

3.2 Ein auf der Gammaverteilung beruhendes Verfahren

In diesem Kapitel verwenden wir das individuelle Modell für die Zufallsvariable S_{ik} der Änderung des Schadenstands von Anfalljahr i in Abwicklungsjahr k :

$$S_{ik} = \sum_{n=1}^{v_i} R_{ikn},$$

wobei wir die Polizzenzahl v_i als Volumenmaß von Anfalljahr i benutzen. Möglich sind auch andere Volumenmaße, wie z.B. das Bedarfsprämienvolumen oder die Anzahl der im ersten Abwicklungsjahr gemeldeten Schäden. R_{ikn} bezeichnet den von der n -ten Police stammenden Änderungsbetrag.

Wir suchen nun eine Verteilung, die wir diesem Modell zu Grunde legen können. Es ist nicht sinnvoll, irgendein Verteilungsmodell zu nehmen, denn dabei könnte sich eine ganz falsche Gestalt für die Verteilung ergeben. Da die R_{ikn} für die meisten Policen gleich Null sind, liegt die Hauptmasse der Wahrscheinlichkeit im Punkt 0. Um auch für die vielen Fälle kleinerer Risikogruppen eine brauchbare Approximation der Gesamtschadenverteilung zu finden, verwenden wir stetige Verteilungen, bei denen die Faltungspotenzen (und damit auch die Verteilung) berechnet werden können. Die bekannteste Verteilung auf $(0, \infty)$, bei der die Faltungspotenzen berechnet werden können, ist die Gammaverteilung. (siehe Anhang 3.)

Wir nehmen an, dass sie in allen Zellen (i, k) und für alle Policen $n = 1, \dots, v_i$ denselben Formparameter $\alpha < 1$ hat, um die Parameterzahl nicht zu groß werden zu lassen.

Wir gehen vom allgemeinen Ansatz

$$\mathbb{E}(R_{ikn}) = x_i y_k, \quad \text{und} \quad \text{Var}(R_{ikn}) = \frac{(x_i y_k)^2}{\alpha}, \quad 1 \leq n \leq v_i$$

aus und erhalten insgesamt für S_{ik} eine Gammaverteilung mit

$$\mathbb{E}(S_{ik}) = v_i x_i y_k \quad \text{und} \quad \text{Var}(S_{ik}) = \frac{v_i (x_i y_k)^2}{\alpha}$$

mit Formparameter v_i α .

Nehmen wir für die S_{ik} globale Unabhängigkeit an, so erhalten wir als Likelihoodfunktion (siehe Anhang 4.)

$$L = \prod_{i,k} \exp\left(-\frac{\alpha S_{ik}}{x_i y_k}\right) \cdot \left(\frac{\alpha S_{ik}}{x_i y_k}\right)^{\alpha v_{ik}} / (S_{ik} \Gamma(\alpha v_{ik})).$$

Die Likelihoodschätzer sind die Werte \hat{x}_i, \hat{y}_k , die L maximieren. Durch Nullsetzen der partiellen Ableitungen von L nach den Parametern erhalten wir die Maximum-Likelihood-Schätzer:

$$\hat{x}_i = \sum_{k=1}^{I+1-i} \frac{S_{ik}}{\hat{y}_k} / ((I+1-i)v_i), \quad 1 \leq i \leq I,$$

$$\hat{y}_k = \sum_{i=1}^{I+1-k} \frac{S_{ik}}{\hat{x}_i} / \sum_{i=1}^{I+1-k} v_i, \quad 1 \leq k \leq I.$$

Sie ergeben sich durch abwechselnde Berechnung von \hat{x}_i und \hat{y}_k , ausgehend von den Startwerten $\hat{y}_1 = \dots = \hat{y}_I = 1$.

Nun kann der Endschaten

$$\hat{C}_{iI} = C_{i,I+1-i} + \hat{R}_i$$

mit dem Reserveschätzer

$$\hat{R}_i = \sum_{k=I+2-i}^I v_i \hat{x}_i \hat{y}_k$$

berechnet werden.

Zur Berechnung des Likelihoodschätzers $\hat{\alpha}$ verwendet man ebenfalls die Likelihoodmethode, auf die ich hier jedoch nicht weiter eingehen möchte.

3.3 Weitere Verfahren

Es gibt noch viele andere Verfahren, wie zum Beispiel:

- ein Verfahren mittels der Methode der kleinsten Quadrate, das auf der Lognormalverteilung aufbaut. Es ist das Basismodell einer verbreiteten Schadenreservierungs-Software (ICRFS von Ben Zehnwirth).
- ein auf der Inversen Gaußverteilung beruhendes Verfahren. Hier kann durch einen zusätzlichen Parameter die Varianzstruktur verbessert werden.
- ein Modell mit der Poissonverteilung, aus dem das Chain-Ladder-Verfahren hergeleitet werden kann.

4 Modifikationen der Verfahren

Es gibt viele Varianten und Verfeinerungen der Verfahren. Man kann z.B. durch Ausnutzung der Zusatzinformation über die Entwicklung der Zahl der Schäden eine Trennung von Schadenzahl und Schadenhöhe vornehmen. Im Folgenden werde ich jedoch nur den Unterschied von IBNR- und IBNER-Schäden und die Separation von Kalenderjahreffekten näher beschreiben.

4.1 Trennung von IBNR- und IBNER- Schäden

In Kapitel 1 haben wir zwar den Unterschied zwischen IBNR- und IBNER- Schäden erklärt, haben die zwei Arten jedoch zu einem Begriff zusammengefasst und sind auch in den besprochenen Verfahren nicht weiter auf den Unterschied eingegangen. In der Realität bestehen beide Unsicherheitsquellen gleichzeitig, oft liegen die Daten jedoch nur in einer Form vor, sodass eine Unterscheidung zwischen IBNR- und IBNER-Schäden nicht möglich ist. Allerdings sind Versicherungsunternehmen in der Lage, die zur Trennung erforderlichen Informationen bereitzustellen. Wir wollen nun die Reserve mit getrennten Schadenarten berechnen.

Wir zerlegen den Änderungsbetrag $S_{ik} = C_{ik} - C_{i,k-1}$ des Schadenstands von Anfalljahr i in die Bestandteile U_{ik} , der sich nur auf Schäden bezieht, die vor Abwicklungsjahr k noch nicht gemeldet waren (IBNR-Schäden), und T_{ik} , der sich nur auf Schäden bezieht, die vor Abwicklungsjahr k schon gemeldet waren (IBNER-Schäden):

$$S_{ik} = T_{ik} + U_{ik}.$$

Es gilt $T_{i1} = 0$, d.h. $C_{i1} = S_{i1} = U_{i1}$.

Wir nehmen an, dass der echte Spätschaden U_{ik} sowohl von $U_{i1}, \dots, U_{i,k-1}$ als auch von $T_{i1}, \dots, T_{i,k-1}$ unabhängig ist und können somit den unbedingten Erwartungswert von U_{ik} , der nur vom Abwicklungsjahr k und dem bekannten Volumen v_i des Portefeuilles im Anfalljahr i abhängt, modellieren:

$$\begin{aligned}\mathbb{E}(U_{ik}) &= v_i m_k, & 1 \leq i, k \leq I, \\ \text{Var}(U_{ik}) &= v_i s_k^2, & 1 \leq i, k \leq I,\end{aligned}$$

mit unbekanntem Parametern m_k, s_k .

Da T_{ik} zwar normalerweise unabhängig von U_{ik} nicht aber von der vergangenen Schadenabwicklung ist, müssen wir bei der Modellierung des Änderungsbetrags T_{ik} der offenen Schäden anders vorgehen. Da alle hinter T_{ik} stehenden Einzelschäden im vorangegangenen Abwicklungsjahr bereits bekannt waren, nehmen wir an, dass der Erwartungswert von T_{ik} von der Gesamthöhe aller noch offenen Schäden abhängt. Da die

Änderungsrate vom Abwicklungsjahr abhängen wird, verwenden wir die Abkürzung $D_{i,k-1}$ für die Gesamtheit aller Daten von Anfalljahr i bis einschließlich Abwicklungsjahr $k - 1$ und schreiben

$$\begin{aligned}\mathbb{E}(T_{ik} | D_{i,k-1}) &= C_{i,k-1} h_{k-1}, \\ \text{Var}(T_{ik} | D_{i,k-1}) &= C_{i,k-1} t_{k-1}^2,\end{aligned}$$

wobei h_k und t_k unbekannte Parameter sind. Wir verwenden $C_{i,k-1}$ in obiger Formel, um das Modell möglichst einfach zu halten, da wir korrekterweise die Summe aller Einzelfallreserven zum Ende von Abwicklungsjahr $k - 1$ einbeziehen müssten. Da die Einzelfallreserven einen Anteil b_k des Gesamtschadenstandes C_{ik} ausmachen, ist die Vereinfachung kein Problem, wenn wir den Faktor b_k in h_k miteinbeziehen.

Für C_{ik} erhalten wir folgendes Modell:

$$\mathbb{E}(C_{ik} | D_{i,k-1}) = \mathbb{E}(C_{i,k-1} + T_{ik} + U_{ik} | D_{i,k-1}) = C_{i,k-1} (1 + h_{k-1}) + v_i m_k.$$

Für die unbekannt Parameter haben wir folgende erwartungstreue Schätzer:

$$\begin{aligned}\hat{m}_k &= \frac{\sum_{i=1}^{I+1-k} U_{ik}}{\sum_{i=1}^{I+1-k} v_i}, \quad 1 \leq k \leq I, \\ \hat{s}_k^2 &= \frac{1}{I-k} \sum_{i=1}^{I+1-k} v_i \left(\frac{U_{ik}}{v_i} - \hat{m}_k \right)^2, \quad 1 \leq k \leq I-1, \\ \hat{s}_I^2 &= \min \{ \hat{s}_k^2 | 1 \leq k \leq I-1 \} \\ \hat{h}_k &= \frac{\sum_{i=1}^{I-k} T_{i,k+1}}{\sum_{i=1}^{I-k} C_{ik}}, \quad 1 \leq k \leq I-1, \\ \hat{t}_k^2 &= \frac{1}{I-k-1} \sum_{i=1}^{I-k} C_{ik} \left(\frac{T_{i,k+1}}{C_{ik}} - \hat{h}_k \right)^2, \quad 1 \leq k \leq I-2, \\ \hat{t}_{I-1}^2 &= \min \left\{ \frac{\hat{t}_{I-2}^4}{\hat{t}_{I-3}^2}, \hat{t}_{I-3}^2 \right\}\end{aligned}$$

Mit den Abkürzungen

$$\begin{aligned}D &= \{U_{ik}, T_{ik} | i+k \leq I+1\}, \\ D_i &= \{U_{ik}, T_{ik} | 1 \leq k \leq I+1-i\}\end{aligned}$$

für die bekannten Daten und unter der Annahme der Unabhängigkeit der Anfalljahre gilt für $k > I+1-i$

$$\mathbb{E}(C_{ik} | D) = \mathbb{E}(C_{ik} | D_i) = \mathbb{E}(\mathbb{E}(C_{ik} | D_{i,k-1}) | D_i) = v_i m_k + \mathbb{E}(C_{i,k-1} | D_i)(1 + h_{k-1}).$$

Durch sukzessives Anwenden dieser Beziehung und mit der Abkürzung $g_k = 1 + h_k$ erhalten wir den mittleren Endschaten:

$$\mathbb{E}(C_{iI}|D) = C_{i,I+1-i} g_{I+1-i} \cdots g_{I-1} + \sum_{k=I+2-i}^I v_i m_k g_k \cdots g_{I-1}$$

Durch Verwendung unserer Schätzer erhalten wir den Endschaten durch

$$\hat{C}_{iI} = C_{i,I+1-i} \hat{g}_{I+1-i} \cdots \hat{g}_{I-1} + \sum_{k=I+2-i}^I v_i \hat{m}_k \hat{g}_k \cdots \hat{g}_{I-1}$$

und die Reserve R_i durch

$$\hat{R}_i = \hat{C}_{iI} - C_{i,I+1-i}.$$

Falls die Daten nur auf wenigen Schäden beruhen, kann die Trennung von IBNR- und IBNER- Schäden dazu führen, dass das Abwicklungsmuster der Daten schlechter zu erkennen ist als ohne Trennung. Man sollte das Verfahren der Trennung jedoch auf jeden Fall anwenden, wenn die Datendetaillierung dies zulässt.

4.2 Separation von Kalenderjahreffekten

In all den erwähnten Verfahren wurde die Unabhängigkeit der Anfalljahre angenommen. Diese Annahme ist jedoch problematisch, da jeder kalenderjahrweise wirkende Einfluss meist mehrere Anfalljahre zugleich betrifft. Dies gilt außer für deutliche Änderungen der Inflationsrate insbesondere für Änderungen in der Reservierungs- und Regulierungspraxis, sowie in der Rechtsprechung. So kann z.B. eine von der Versicherungsunternehmen durchgeführte Aktion zur Klärung offener Schadenfälle zu einem stärkeren Anstieg der Einzelfallreserven in einem Kalenderjahr führen, was sich diagonal auf mehrere Anfall- und Abwicklungsjahre auswirkt.

Bis jetzt wurde die monetäre Inflation noch nicht erwähnt, da wir die Beträge S_{ik} als um die monetäre Inflation bereinigte Daten vorausgesetzt haben. Trotzdem ist das Inflationsproblem hiermit nicht gelöst, da in der Haftpflichtversicherung häufig eine sogenannte „superimposed inflation“ wirkt. Dies bedeutet, dass oft nach der üblichen Inflationsbereinigung noch ein trendmäßiger Anstieg der jährlichen Schadenlast erkennbar ist, der wiederum auf Veränderungen in der Rechtsprechung, steigendes Anspruchdenken der Geschädigten oder Ähnliches zurückzuführen ist.

Es ist daher sinnvoll, das Abwicklungsdreieck zu Beginn der Spätschadenreserve-schätzung auf das Vorhandensein von solchen Kalenderjahreffekten zu untersuchen, und gegebenenfalls die Auswirkung des Effekts auf die Höhe der Schadenreserve durch Weglassen oder Glätten einzelner Daten zu eliminieren.

Wie wir bereits wissen, bilden die Zeilen die Anfalljahre und die Spalten die Abwicklungsjahre. Die Kalenderjahre stellen eine dritte Richtung dar, sie entsprechen den Parallelen zur Hypotenuse (S_{ik} mit konstantem $i + k$). Im Gegensatz zur Unabhängigkeit der Anfall- und Abwicklungsjahre, wirkt der Kalenderjahreffekt zugleich auf Anfall- und Abwicklungsjahre.

Wir verdeutlichen dies anhand des kreuzklassifizierten Verfahrens $\mathbb{E}(S_{ik}) = x_i y_k, 1 \leq I, k \leq I$, indem wir auf dieses Modell einen konstanten jährlichen Inflationsfaktor $u > 1$ wirken lassen:

$$\mathbb{E}(\underline{S}_{ik}) = x_i y_k u^{i+k-2}, 1 \leq I, k \leq I$$

Dies können wir wieder in der kreuzklassifizierten Form

$$\mathbb{E}(S_{ik}) = (x_i u^{i-1}) (y_k u^{k-1})$$

schreiben.

Man sieht, dass die Inflation sowohl auf die Anfalljahr- als auch auf die Abwicklungsjahrparameter wirkt. Deshalb ist es nicht möglich, mit Hilfe eines kreuzklassifizierten Modells zu erkennen, ob in den Daten ein konstanter Kalenderjahreffekt enthalten ist. Außerdem wird ein konstanter Kalenderjahreffekt (der bereits in den Daten des Abwicklungsdreiecks enthalten ist) durch ein kreuzklassifiziertes Verfahren auf die zu berechnenden Werte $S_{ik}, i + k > I + 1$, extrapoliert. Die Tatsache, dass eine konstante Inflation nicht erkannt werden kann, ist jedoch bei kreuzklassifizierten Modellen nicht weiter störend, solange man mangels besserem Wissen annehmen muss, dass dieselbe konstante Inflation auch in den nächsten Jahren weiterwirkt.

Nichtkonstante Kalenderjahreffekte u_j bilden ein Modell der Form

$$\underline{S}_{ik} = S_{ik} u_{i+k},$$

die jedoch nicht in multiplikativer Form $u_{ik} = t_i w_k$ aus Spalten- und Zeileneffekten zusammengesetzt werden können, da die Anwendung eines kreuzklassifizierten Modells die Abweichungen zwischen u_{ik} und $t_i w_k$ mit den Zufallsschwankungen der Schadenvariablen S_{ik} vermischt und dies dann auf die Zukunft überträgt.

Eine nichtkonstante Inflation kann beschrieben werden, wenn es gelingt, die Daten S_{ik} so zu transformieren, dass ohne Inflation entweder keine systematischen Anfalljahrunterschiede oder keine systematischen Abwicklungsjahrunterschiede vorkommen, denn dann können etwa doch vorhandene Anfalljahr- bzw. Abwicklungsjahrunterschiede direkt auf die Inflation zurückgeführt werden.

5 Zusammenfassung

Die Berechnung der Schadenreserve ist für Versicherungsunternehmen unerlässlich und wird in Zukunft an Bedeutung gewinnen. In vielen Ländern (vor allem USA und Kanada) wird von den Aufsichtsbehörden verlangt, dass die Höhe der Spätschadenreserve (ähnlich wie bei der Deckungsrückstellung in der Lebensversicherung) von einem Aktuar überprüft wird.

Wir haben nun ein paar mathematische Verfahren kennengelernt, es gibt jedoch eine Vielzahl von anderen Methoden, Verfeinerungen und Erweiterungen.

Das Chain-Ladder-Verfahren ist das wichtigste Verfahren und wird auch seine zentrale Stellung innerhalb der verschiedenen Reservierungsverfahren beibehalten. Andere werden nur dann angewandt, wenn ihr Standardfehler deutlich unter dem des Chain-Ladder-Verfahrens liegt.

Die Einsatzmöglichkeit hängt natürlich auch von den gegebenen Daten ab: Besser ist, einen großen Teilbestand zu verwenden, da dies den Standardfehler aufgrund des Ausgleichs im Kollektiv verkleinert.

Die Berechnung des Standardfehlers sollte nicht außer Acht gelassen werden, da dieser die Genauigkeit der geschätzten Reserve widerspiegelt.

6 Anhang

1. Satz vom iteriertem Erwartungswert

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Z))$$

$$\mathbb{E}(\mathbb{E}(X|Y, Z)|Y) = \mathbb{E}(X|Y)$$

2. erwartungstreu

\hat{f}_k ist erwartungstreu für f_k , wenn $\mathbb{E}(\hat{f}_k) = f_k$ gilt.

3. Die Dichtefunktion der Gammaverteilung der Form

$$g(x) = \exp\left(-\frac{x}{\mu}\right) \cdot x^{\alpha-1} \cdot \frac{\left(\frac{\alpha}{\mu}\right)^\alpha}{\Gamma(\alpha)} \quad x > 0$$

ist eine Abweichung von der üblichen Darstellung. Hier kommt der Erwartungswert μ direkt als Parameter vor. Der Formparameter $\alpha > 0$ bestimmt die Form der Dichtefunktion. Die Varianz ist $\frac{\mu^2}{\alpha}$.

4. Likelihoodfunktion

Für kontinuierliche Verteilungen ist die Likelihoodfunktion die gemeinsame Dichte der Stichprobe (X_1, \dots, X_n) , aufgefasst als Funktion des Parameters θ

$$l(\theta; x_1, \dots, x_n) := \prod_{i=1}^n f(x_i|\theta).$$

Literatur

- [Ma] Thomas Mack: Schadenversicherungsmathematik, 2002
Schriftenreihe Angewandte Versicherungsmathematik, Heft 28
- [Vi] Skriptum zur Vorlesung „Angewandte Statistik“, R. Viertl, 2008
- [1] <http://www.mi.uni-koeln.de/jost/schmidt.pdf>
- [2] <http://www.actuaries.ch>