D I P L O M A R B E I T

Stochastische Schadenreservierung

Ausgeführt am Institut für
Wirtschaftsmathematik
der Technischen Universität Wien

unter der Anleitung von
Privatdoz. Dipl.-Ing. Dr.techn. Stefan Gerhold

durch

Michael Schützenhofer
Engerthstraße 249/7/10
1020 Wien

Wien,7. Januar 2015 ______________________

Inhaltsverzeichnis

Inhaltsverzeichnis **ii**

Tabellenverzeichnis **iv**

1 Einleitung **1**

2 Das Chain-Ladder Verfahren **2**
2.1 Allgemeine Definitionen und Annahmen 2
2.2 Chain-Ladder Verfahren 3
2.2.1 Chain-Ladder Faktoren, Chain-Ladder Schätzer, Chain-Ladder Reserve 4
2.2.2 Begründung des Chain-Ladder Verfahrens 4
2.2.3 Standardabweichung und Konfidenzintervall der Reserve 5
2.2.4 Beispiele Chain-Ladder Verfahren 11
2.2.4.1 KFZ-Kasko 11
2.2.4.2 Rechtsschutz 14

3 Das lineare Modell **17**
3.1 Grundlagen 17
3.1.1 Das elementare lineare Modell 17
3.1.2 Das erweiterte lineare Modell 19
3.2 Das lineare Modell in der Schadenreservierung 23
3.2.1 Das lineare Modell von Mack 26
3.2.2 Beispiele 30
3.2.2.1 KFZ-Kasko 30
3.2.2.2 Rechtsschutz 33

4 Das lognormale loglineare Modell **36**
4.1 Grundlagen 36
4.1.1 Das elementare lognormale loglineare Modell 36
4.1.2 Das erweiterte lognormale loglineare Modell 38
4.2 Das lognormale loglineare Modell in der Schadenreservierung 40
4.2.1 Das lognormale logadditive Modell für Zuwächse 41
4.2.2 Beispiele 43
4.2.2.1 KFZ-Kasko 44
4.2.2.2 Rechtsschutz 46

5 Stochastische Abwicklungsfaktoren **49**
5.1 Das lognormale Modell . . . 50
5.1.1 Grundlagen des Modells . . . 50
5.1.2 Simulationsalgorithmus zur Berechnung der Endschadenstände . . . 52
5.2 Das loggamma Modell . . . 53
5.2.1 Grundlagen des Modells . . . 53
5.2.2 Simulationsalgorithmus zur Berechnung der Endschadenstände . . . 56
5.3 Das loginverse Gauß Modell . . . 57
5.3.1 Grundlagen des Modells . . . 57
5.3.2 Simulationsalgorithmus zur Berechnung der Endschadenstände . . . 59
5.4 Beispiele . . . 60
5.4.1 KFZ-Kasko . . . 61
5.4.1.1 Lognormales Modell . . . 62
5.4.1.2 Loggamma Modell . . . 63
5.4.1.3 Loginverses Gauß Modell . . . 65
5.4.2 Rechtsschutz . . . 67
5.4.2.1 Lognormales Modell . . . 68
5.4.2.2 Loggamma Modell . . . 69
5.4.2.3 Loginverses Gauß Modell . . . 71

6 Vergleich der Verfahren **73**
6.1 KFZ-Kasko . . . 73
6.2 Rechtsschutz . . . 75

7 Schlusswort **78**

Literaturverzeichnis **80**

Tabellenverzeichnis

2.1 Allgemeines Abwicklungsquadrat für Schadenstände 2
2.2 Allgemeines Abwicklungsquadrat für inkrementelle Zuwächse 3
2.3 Kumuliertes KFZ-Kasko Dreieck 11
2.4 KFZ-Kasko Chain Ladder 12
2.5 Werte KFZ-Kasko Chain Ladder 13
2.6 Tatsächliches KFZ-Kasko Quadrat 13
2.7 Kumuliertes Rechtsschutz Dreieck 14
2.8 Rechtsschutz Chain Ladder 15
2.9 Tatsächliches Rechtsschutz Quadrat 16

3.1 Inkrementelles Dreieck KFZ-Kasko 30
3.2 KFZ-Kasko Lineares Modell 31
3.3 Inkrementelles Dreieck Rechtsschutz 33
3.4 Rechtsschutz Lineares Modell 34

4.1 KFZ-Kasko Loglineares Modell 45
4.2 Rechtsschutz Loglineares Modell 47

5.1 Dreieck der individuellen Abwicklungsfaktoren 49
5.2 Abwicklungsquadrat bei stochastischen Faktoren 50
5.3 Berechnungsalgorithmus der Digammafunktion 55
5.4 Logarithmierte individuelle Abwicklungsfakoren KFZ-Kasko 61
5.5 KFZ-Kasko Lognormales Modell 62
5.6 KFZ-Kasko Loggamma Modell 64
5.7 KFZ-Kasko Loginverses Gauß Modell 66
5.8 Logarithmierte individuelle Abwicklungsfaktoren Rechtsschutz 67
5.9 Rechtsschutz Lognormales Modell 68
5.10 Rechtsschutz Loggamma Modell 70
5.11 Rechtsschutz Loginverses Gauß Modell 71

6.1 Vergleich KFZ-Kasko 73
6.2 Vergleich Rechtsschutz 75

Kapitel 1

Einleitung

Aufgrund verschiedener Vorgaben, wie den International Financial Reporting Standards (IFRS), ist es bei der Schadenreservierung von Versicherungsunternehmen nicht mehr ausreichend, nur die Schadenreserven der Fachabteilungen zu verwenden, sondern die Schadenreserve muss anhand versicherungsmathematischer Verfahren berechnet werden. Dabei wird unter Zuhilfenahme der vergangenen Schadenzahlungen beziehungsweise der vergangenen Schadenaufwendungen (Schadenzahlungen + Schadenreserven) die zukünftig benötigte Schadenreserve berechnet. Wichtig ist hierbei zu erwähnen, dass für diese Verfahren nur gleichartige Risiken zusammengefasst werden dürfen, da ansonsten keine korrekte Schadenreserve berechnet werden kann.
Unterschieden wird hierbei, ob es sich um ein Verfahren basierend auf Schadenstände oder inkrementelle Zuwächse handelt.

Eines der bekanntesten Schadenreservierungsverfahren in dieser Hinsicht ist das Chain-Ladder Verfahren. Dieses Verfahren wird in Kapitel 2 vorgestellt.
Die nächste Methode, die in dieser Arbeit betrachtet wird, ist das lineare Modell. Dabei wird zunächst auf die Grundlagen des linearen Modells eingegangen, um diese Theorie dann in der Schadenreservierung anzuwenden.
Eine weitere Schadenreservierungsmethode, die in Kapitel 4 erklärt wird, ist das loglineare lognormale Modell. Dieses Modell knüpft stark an das lineare Modell an und ist das erste Modell dieser Arbeit, in dem bereits bei der Reserveberechnung eine Verteilungsannahme an die Schadenstände getroffen wird.
Danach wird in Kapitel 5 das Chain-Ladder Verfahren erweitert.Dieses Mal mit der Annahme, dass die individuellen Abwicklungsfakoren einer bestimmten Verteilung folgen.

Für all diese Verfahren werden die Reserven auch anhand zweier realer Versicherungsbestände berechnet. Neben der Reserve wird auch die Standardabweichung des Schätzers bestimmt sowie die Standardabweichung und die Konfidenzintervalle der tatsächlichen Reserve die sich aus den Modellen ergeben würden. Zum Abschluss der Verfahren werden die berechneten Werte mit den Daten verglichen, die sich bei den realen Versicherungsbeständen ergeben haben.
Zuletzt werden die Ergebnisse der verschiedenen Verfahren miteinander verglichen.

Kapitel 2

Das Chain-Ladder Verfahren

2.1 Allgemeine Definitionen und Annahmen

Bevor wir uns nun mit dem Chain-Ladder Verfahren befassen werden, folgen noch einige Definitonen, die sich auch auf die weiteren Schadenreservierungsverfahren beziehen werden.
Mit n bezeichnen wir die Anzahl der betrachteten Schadenjahre und somit auch der betrachteten Abwicklungsjahre. Dabei wird die Annahme getroffen, dass die betrachteten Schadenfälle nach n Jahren abgewickelt sind und es hier zu keinen weiteren Schadenbearbeitungen mehr kommt.
Mit $P_{i,j}$(Zahlungs-Entwicklung) bezeichnen wir die getätigten Schadenzahlungen des i-ten Schadenjahres(SJ) bis zum Ende des j-ten Abwicklungsjahres(AJ),
mit $I_{i,j}$(Schadenaufwands-Entwicklung) die Schadenzahlungen zuzüglich der vorhandenen Reserve für Schäden aus dem i-ten Schadenjahr am Ende des j-ten Abwicklungsjahres , wobei $i, j \in \{1, \ldots, n\}$ gilt. Weiters setzen wir voraus, dass die $P_{i,j}$ und $I_{i,j}$ für $i+j \leq n+1$ beobachtbar sind und für $i + j > n + 1$ nicht.
Das heißt die Datenstruktur hat folgende Form

SJ \ AJ	1	2	$\cdots$	n+1-i	$\cdots$	n-1	n
1	$S_{1,1}$	$S_{1,2}$	$\cdots$	$S_{1,n+1-i}$	$\cdots$	$S_{1,n-1}$	$S_{1,n}$
2	$S_{2,1}$	$S_{2,2}$	$\cdots$	$S_{2,n+1-i}$	$\cdots$	$S_{2,n-1}$	$S_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$S_{i,1}$	$S_{i,2}$	$\cdots$	$S_{i,n+1-i}$	$\cdots$	$S_{i,n-1}$	$S_{i,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n-1	$S_{n-1,1}$	$S_{n-1,2}$	$\cdots$	$S_{n-1,n+1-i}$	$\cdots$	$S_{n-1,n-1}$	$S_{n-1,n}$
n	$S_{n,1}$	$S_{n,2}$	$\cdots$	$S_{n,n+1-i}$	$\cdots$	$S_{n,n-1}$	$S_{n,n}$

Tabelle 2.1: Allgemeines Abwicklungsquadrat für Schadenstände

wobei die grau eingefärbten Felder noch nicht beobachtbar sind und daher geschätzt werden müssen.$S_{i,j}$ bezeichne dabei sowohl die Zahlungs- als auch die Schadenaufwands-Entwicklung.

Nun können wir die inkrementellen Schadenzahlungen $Z_{i,j}$ durch

$$Z_{i,j} := \begin{cases} P_{i,1} & \text{falls j=1} \\ P_{i,j} - P_{i,j-1} & \text{sonst} \end{cases}$$

und den inkrementellen Schadenaufwand $Y_{i,j}$ durch

$$Y_{i,j} := \begin{cases} I_{i,1} & \text{falls j=1} \\ I_{i,j} - I_{i,j-1} & \text{sonst} \end{cases}$$

definieren. Somit können wir auch ein Abwicklungsquadrat für inkrementelle Zuwächse betrachten.

SJ \ AJ	1	2	$\cdots$	n+1-i	$\cdots$	n-1	n
1	$T_{1,1}$	$T_{1,2}$	$\cdots$	$T_{1,n+1-i}$	$\cdots$	$T_{1,n-1}$	$T_{1,n}$
2	$T_{2,1}$	$T_{2,2}$	$\cdots$	$T_{2,n+1-i}$	$\cdots$	$T_{2,n-1}$	$T_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$T_{i,1}$	$T_{i,2}$	$\cdots$	$T_{\tilde{i},n+1-\tilde{i}}$	$\cdots$	$T_{i,n-1}$	$T_{i,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n-1	$T_{n-1,1}$	$T_{n-1,2}$	$\cdots$	$T_{n-1,n+1-i}$	$\cdots$	$T_{n-1,n-1}$	$T_{n-1,n}$
n	$T_{n,1}$	$T_{n,2}$	$\cdots$	$T_{n,n+1-i}$	$\cdots$	$T_{n,n-1}$	$T_{n,n}$

Tabelle 2.2: Allgemeines Abwicklungsquadrat für inkrementelle Zuwächse

Dabei steht $T_{i,j}$ sowohl für die inkrementellen Schadenzahlungen als auch für den inkrementellen Schadenaufwand.

2.2 Chain-Ladder Verfahren

Das Chain-Ladder Verfahren ist ein Verfahren, das auf dem Abwicklungsquadrat für Schadenstände beruht. Die Grundlage des Chain-Ladder Verfahren ist das Abwicklungsmuster für Faktoren. Dabei orientieren sich die Annahmen dieser Arbeit an denen, die in Thomas Mack [6] getroffen werden. Diese Annahmen führen dazu, dass wir unter anderem die Varianz der zukünftigen Reserve und die Varianz der Schätzer der Reserve berechnen können.

Annahme 2.1 Abwicklungsmuster für Faktoren.
Es existieren Parameter $\varphi_1, ..., \varphi_{n-1}$ so, dass für alle $i \in \{1, ..., n\}$ und $j \in \{1, ..., n-1\}$

$$\mathbb{E}[S_{i,j+1}|S_{i,1}, .., S_{i,j}] = S_{i,j} \cdot \varphi_j$$

gilt. Daraus ergibt sich

$$\mathbb{E}[S_{i,j+1}] = \mathbb{E}[S_{i,j}]\varphi_j$$

Dabei steht $S_{i,1}, .., S_{i,j}$ für die von diesen Schadenständen erzeugte Sigma-Algebra.
Um das Chain-Ladder Verfahren nun anwenden zu können, müssen wir vorraussetzen, dass ein Abwicklungsmuster für Faktoren vorliegt und diese Abwicklungsfaktoren unbekannt sind.

2.2.1 Chain-Ladder Faktoren, Chain-Ladder Schätzer, Chain-Ladder Reserve

Das Chain-Ladder Verfahren besteht nun im wesentlichen aus zwei Schritten. Im ersten Schritt wird für alle $j \in \{1, .., n-1\}$ der Abwicklungsfaktor φ_j durch den sogenannten Chain-Ladder Faktor geschätzt. Dabei ist der Chain-Ladder Faktor als

$$F_j := \frac{\sum_{k=1}^{n-j} S_{k,j+1}}{\sum_{k=1}^{n-j} S_{k,j}}$$

definiert. Im zweiten Schritt kann nun für alle $i, j \in \{1, ..., n\}$ mit $i + j > n + 1$ der zu erwartende Schadenstand $\mathbb{E}[S_{i,j}]$ durch den Chain-Ladder Schätzer $\hat{S}_{i,j}^{CL}$, mit

$$\hat{S}_{i,j}^{CL} := S_{i,n+1-i} \prod_{k=n+1-i}^{j-1} F_k,$$

geschätzt werden.
Wenn wir nun noch $\hat{S}_{i,n+1-i}^{CL} := S_{i,n+1-i}$ definieren, gilt offensichtlich die Rekursion

$$\hat{S}_{i,j}^{CL} = \hat{S}_{i,j-1}^{CL} F_{j-1}.$$

Nach der Berechnung der Chain-Ladder Schätzer kann auch schon die Chain-Ladder Reserve berechnet werden. Diese ist für die Schadenjahre $i \in \{2, ..., n\}$ durch $\hat{R}_i^{CL}$ gegeben, wobei

$$\hat{R}_i^{CL} := \hat{S}_{i,n}^{CL} - S_{i,n+1-i}$$

gilt. Für das erste beobachtete Schadenjahr hat die erwartete Reserve den Wert 0, da wir angenommen haben, dass alle Schadenfälle nach n Jahren abgewickelt sind.

Anhand der Reserveberechnung ist auch ein Schwachpunkt des Chain-Ladder Verfahrens zu erkennen. Zur Berechnung der Endschadenstände $\hat{S}_{i,n}^{CL}$ wird nur der Schadenstand des aktuellen Kalenderjahres herangezogen und mit den Faktoren multipliziert. Ist dieser Schadenstand nun überdurchscnittlich hoch oder niedrig, setzt sich dies bis zum Ende des Beobachtungszeitraums fort. Dies hat vorallem im aktuellten Schadenjahr große Auswirkungen, da hier erst ein Wert beobachtet wurde und alle anderen geschätzt werden müssen.

2.2.2 Begründung des Chain-Ladder Verfahrens

In diesem Unterkapitel wollen wir kurz erläutern, wieso die oben getroffene Definiton der Chain-Ladder Faktoren sinnvoll ist.
Für alle Schadenjahre $i \in \{1, ..., n-1\}$ und alle Abwicklungsjahre $j \in \{1, ..., n-1\}$ definieren wir den individuellen Abwicklungsfaktor durch

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

Hier gilt, dass die individuellen Abwicklungsfaktoren $F_{i,j}$ für $i+j<n+1$ beobachtbar sind und für $i+j \geq n+1$ nicht. Unter Zuhilfenahme der nicht beobachtbaren individuellen Abwicklungsfaktoren können die nicht beobachtbaren Schadenstände als

$$S_{i,j} = S_{i,n+1-i} \prod_{k=n+1-i}^{j-1} F_{i,k}$$

berechnet werden.
Ersetzt man in obiger Gleichung die individuellen Abwicklungsfaktoren durch die Chain-Ladder Faktoren, erhält man den Chain-Ladder Schätzer. Da für die Chain-Ladder Faktoren folgende Gleichheit gilt

$$F_j = \frac{\sum_{k=1}^{n-j} S_{k,j+1}}{\sum_{k=1}^{n-j} S_{k,j}} = \sum_{k=1}^{n-j} \frac{S_{k,j+1}}{\sum_{l=1}^{n-j} S_{l,j}} = \sum_{k=1}^{n-j} \frac{S_{k,j+1}\frac{S_{k,j}}{S_{k,j}}}{\sum_{l=1}^{n-j} S_{l,j}} = \sum_{k=1}^{n-j} \frac{S_{k,j}}{\sum_{l=1}^{n-j} S_{l,j}} F_{j,k}$$

können sie auch als gewichtetes Mittel der beobachtbaren individuellen Abwicklungsfaktoren geschrieben werden, was die getroffene Wahl der Chain-Ladder Faktoren bestätigt.

2.2.3 Standardabweichung und Konfidenzintervall der Reserve

Nun, da wir einen Schätzer für die zukünftige Reserve gefunden haben, interessieren wir uns auch für die Standardabweichung der Schadenjahrreserven und der Gesamtreserve. Mit Hilfe der Standardabweichung werden wir dann ein Konfidenzintervall der Reserven berechnen.
Dafür muss zunächst eine Annahme an die bedingte Varianz der Schadenstände getroffen werden und an die Schadenstände verschiedener Schadenjahre.

Annahme 2.2 Varianz der bedingten Schadenstände.
Es existieren Faktoren $\alpha_1^2, ..., \alpha_{n-1}^2$, sodass für alle $i \in \{1, ..., n\}$ und alle $j \in \{1, ..., n-1\}$ gilt

$$Var(S_{i,j+1}|S_{i,1}, ..., S_{i,j}) = S_{i,j}\alpha_j^2$$

Annahme 2.3 Unabhängigkeit der Schadenjahre.
Die Variablen $\{S_{i,1}, ..., S_{i,n}\}$ und $\{S_{j,1}, ..., S_{j,n}\}$ sind für $i \neq j$ unnabhängig.

Satz 2.1.
Für alle $j \in \{1, ..., n-2\}$ ist

$$\hat{\alpha}_j^2 := \frac{1}{n-j-1} \sum_{k=1}^{n-j} S_{k,j} (\frac{S_{k,j+1}}{S_{k,j}} - F_j)^2$$

ein erwartungstreuer Schätzer für α_j^2, das heißt

$$\mathbb{E}[\hat{\alpha}_j^2] = \alpha_j^2$$

Beweis.
Bevor wir die Gleichheit zeigen können, definieren wir $B_j := \{S_{i,k} | i + k \leq n+1, k \leq j\}$. Nun gilt mit der Gleichheit $\sum_{k=1}^{n-j} S_{k,j+1} = F_j \sum_{k=1}^{n-j} S_{k,j}$

$$(n-j-1)\hat{\alpha}_j^2 = \sum_{k=1}^{n-j} (\frac{S_{k,j+1}^2}{S_{k,j}} - 2 \cdot S_{k,j+1} F_j + S_{k,j} F_j^2) = \sum_{k=1}^{n-j} \frac{S_{k,j+1}^2}{S_{k,j}} - \sum_{k=1}^{n-j} S_{k,j} F_j^2$$

Mit B_j gilt nun

$$\mathbb{E}[(n-j-1)\hat{\alpha}_j^2 | B_j] = \sum_{k=1}^{n-j} \mathbb{E}[S_{k,j+1}^2 | B_j] \frac{1}{S_{k,j}} - \sum_{k=1}^{n-j} S_{k,j} \mathbb{E}[F_j^2 | B_j]$$

Aufgrund der Unabhängigkeit der verschiedenen Schadenjahre gilt

$$\begin{aligned} \mathbb{E}[S_{k,j+1}^2 | B_j] = \mathbb{E}[S_{k,j+1}^2 | S_{k,1}, .., S_{k,j}] &= Var(S_{k,j+1} | S_{k,1}, .., S_{k,j}) + \mathbb{E}[S_{k,j+1} | S_{k,1}, .., S_{k,j}]^2 \\ &= S_{k,j}\alpha_j + (S_{k,j}\varphi_j)^2 \end{aligned}$$

Weiter gilt

$$Var(F_j | B_j) = \frac{\sum_{k=1}^{n-j} Var(S_{k,j+1} | B_j)}{(\sum_{k=1}^{n-j} S_{k,j})^2} = \frac{\sum_{k=1}^{n-j} S_{k,j}\alpha_j^2}{(\sum_{k=1}^{n-j} S_{k,j})^2} = \frac{\alpha_j^2}{\sum_{k=1}^{n-j} S_{k,j}} \tag{2.1}$$

und

$$\mathbb{E}[F_j | B_j] = \frac{\sum_{k=1}^{n-j} \mathbb{E}[S_{k,j+1} | B_j]}{\sum_{k=1}^{n-j} S_{k,j}} = \frac{\sum_{k=1}^{n-j} \mathbb{E}[S_{k,j+1} | S_{k,1}, .., S_{k,j}]}{\sum_{k=1}^{n-j} S_{k,j}} = \frac{\sum_{k=1}^{n-j} S_{k,j}\varphi_j}{\sum_{k=1}^{n-j} S_{k,j}} = \varphi_j \tag{2.2}$$

Damit folgt

$$\mathbb{E}[F_j^2 | B_j] = Var(F_j | B_j) + \mathbb{E}[F_j | B_j]^2 = \varphi_j^2 + \frac{\alpha_j^2}{\sum_{k=1}^{n-j} S_{k,j}}$$

und schlussendlich

$$\begin{aligned} \mathbb{E}[(n-j-1)\hat{\alpha}_j^2 | B_j] &= \sum_{k=1}^{n-j} \alpha_j + S_{k,j}\varphi_j^2 - \sum_{k=1}^{n-j} (S_{k,j}\varphi_j^2 + \frac{S_{k,j}\alpha_j^2}{\sum_{k=1}^{n-j} S_{k,j}} \\ &= (n-j)\alpha_j - \alpha_j = (n-j-1)\alpha_j \end{aligned}$$

Damit ist alles gezeigt, denn es gilt

$$\mathbb{E}[\hat{\alpha}_j^2] = \mathbb{E}[\mathbb{E}[\hat{\alpha}_j^2 | B_j]] = \mathbb{E}[\alpha_j^2] = \alpha_j^2$$

□

Nun benötigen wir noch einen Schätzer für α_{n-1}^2. Da die Folge $\alpha_1, ..., \alpha_{n-2}$ normalerweise monoton fallend ist, kann man vereinfachend annehmen, dass folgende Gleichung gilt

$$\frac{\hat{\alpha}_{n-3}^2}{\hat{\alpha}_{n-2}^2} = \frac{\hat{\alpha}_{n-2}^2}{\hat{\alpha}_{n-1}^2}$$

Dies führt zu

$$\hat{\alpha}_{n-1}^2 = min(\frac{\hat{\alpha}_{n-3}^2}{\hat{\alpha}_{n-2}^2}, min(\hat{\alpha}_{n-2}^2, \hat{\alpha}_{n-3}^2))$$

Nun können wir die Varianz der zukünftigen Endschadenstände schätzen. Mit ihrer Hilfe erhalten wir die Varianz der Reserve, da

$$Var(S_{i,n}) = Var(R_i + S_{i,n+1-i}) = Var(R_i)$$

aufgrund der Tatsache, dass $S_{i,n+1-i}$ konstant ist, gilt. Dabei wurde der von Braun [2] präsentierte Algorithmus zur Berechnung der Varianz verwendet.

Satz 2.2 Schätzer für die Varianz der Endschadenstände.
Für $i \in \{2, .., n\}$ erhalten wir folgenden Schätzer der Varianz der Endschadenstände

$$V\hat{a}r(S_{i,n}) = (S_{i,n}^{CL})^2 \sum_{k=n+1-i}^{n-1} \frac{\frac{(\hat{\alpha}_k)^2}{F_k^2}}{S_{i,k}^{CL}}$$

Beweis.
Da für die Varianz einer Zufallsvariable

$$Var(X) = \mathbb{E}[Var(X|F)] + Var(\mathbb{E}[X|F])$$

gilt, wobei F eine Sigma-Algebra ist, erhalten wir mit $D_{i,k} := \{S_{i,1}, .., S_{i,k}\}$ und
$i \in \{2, .., n\}$,$k \in \{1, .., n-1\}$

$$\begin{aligned} Var(S_{i,k+1}) &= \mathbb{E}[Var(S_{i,k+1}|D_{i,k})] + Var(\mathbb{E}[S_{i,k+1}|D_{i,k}]) = \mathbb{E}[S_{i,k}]\alpha_k^2 + Var(S_{i,k}\varphi_k) \\ &= \mathbb{E}[S_{i,k}]\alpha_k^2 + \varphi_k^2 Var(S_{i,k}) = S_{i,n+1-i}\varphi_{n+1-i}...\varphi_{k-1}\alpha_k^2 + \varphi_k^2 Var(S_{i,k}) \quad . \end{aligned}$$

Dabei ist zu beachten, dass

$$Var(S_{i,k}) = 0 \quad k \leq n+1-i$$

gilt.
Somit erhalten wir

$$Var(S_{i,n}) = S_{i,n+1-i} \sum_{k=n+1-i}^{n-1} \varphi_{n+1-i}..\varphi_{k-1}\alpha_k^2\varphi_{k+1}^2..\varphi_{n-1}^2 \quad .$$

Um nun einen Schätzer für die Varianz zu erhalten, ersetzen wir φ_k,α_k durch deren Schätzer F_k und $\hat{\alpha}_k$. Damit ergibt sich

$$
\begin{aligned}
\hat{Var}(S_{i,n}) &= S_{i,n+1-i} \sum_{k=n+1-i}^{n-1} F_{n+1-i}..F_{k-1}\hat{\alpha}_k^2 F_{k+1}^2..F_{n-1}^2 \\
&= S_{i,n+1-i}^2 F_{n+1-i}^2..F_{n-1}^2 \sum_{k=n+1-i}^{n-1} \frac{\hat{\alpha}_k^2}{S_{i,n+1-i}F_k^2 F_{n+1-i}..F_{k-1}} \\
&= (S_{i,n}^{CL})^2 \sum_{k=n+1-i}^{n-1} \frac{\hat{\alpha}_k^2}{S_{i,k}F_k^2} \quad .
\end{aligned}
$$

□

Um nun ein Konfidenzintervall für die Reserve zu erhalten, muss eine Verteilungsannahme an die tatsächlichen Reserven R_i getroffen werden. Aufgrund des Zentralen Grenzwertsatzes kann man bei einer entsprechenden Anzahl an Schadenreserven annehmen, dass diese Reserven normalverteilt sind mit Erwartungswert R_i^{CL} und Varianz $Var(R_i)$. Mit diesen Annahmen erhält man ein symmetrisches 95% Konfidenzintervall für die zukünftige Reserve R_i in der Form von

$$
(R_i^{CL} - 2 \cdot \sqrt{\hat{Var}(R_i)}, R_i^{CL} + 2 \cdot \sqrt{\hat{Var}(R_i)}) \quad .
$$

Wie man hier schön sehen kann, wird das Konfidenzintervall negativ, sobald $\sqrt{\hat{Var}(R_i)}$ größer als 50% von R_i^{CL} ist. Dies ist jedoch für eine Reserve nicht möglich.
Wenn dieser Fall eintritt, sollte als Verteilungsannahme die Log-Normalverteilung getroffen werden. In diesem Fall würden für die Parameter der Verteilung die Gleichungen

$$
\begin{aligned}
exp(\mu_i + \frac{\sigma_i^2}{2}) &= R_i^{CL} \\
exp(2\mu_i + \sigma_i^2)(exp(\sigma_i^2) - 1) &= \hat{Var}(R_i)
\end{aligned}
$$

gelten. Diese Gleichungen führen zu den Parametern

$$
\sigma_i^2 = ln(1 + \frac{\hat{Var}(R_i)}{(R_i^{CL})^2})
$$

und

$$
\mu_i = ln(R_i^{CL}) - \frac{\sigma_i^2}{2} \quad .
$$

Möchten wir nun ein bestimmtes Quantil der Verteilung von R_i berechnen, können wir wie folgt vorgehen. Zunächst wird das entsprechende Quantil der Standardnormalverteilung berechnet, welches wir mit x_q bezeichnen. Nun erhalten wir mit $exp(x_q \cdot \sigma_i + \mu_i)$ das entsprechende Quantil der Log-Nnormalverteilung mit den oben definierten Parametern.
Nun wollen wir noch die Varianz der Gesamtreserve $R = R_2 + ... + R_n$ schätzen. Damit können wir dann auf gleiche Weise ein Konfidenzintervall angeben. Da die Unabhängigkeit der verschiedenen Schadenjahre angenommen wurde, erhalten wir diesen Schätzer indem

die Schätzer der Varianzen der Endschadenstände aufsummiert werden.

Satz 2.3 Schätzer für die Varianz der Gesamtreserve.
Der Schätzer der Varianz der Gesamtreserve ist aufgrund der Unabhängigkeit der einzelnen Schadenjahre durch

$$\hat{Var}(\sum_{k=2}^{n} R_k) = \sum_{k=2}^{n} \hat{Var}(R_k)$$

gegeben.

Von Interesse ist nun noch ein Schätzer für die Varianz des Chain Ladder Schätzers. Dieser wird in Satz 2.4 vorgestellt und ist mit Hilfe des Algorithmus aus [2] berechnet worden.

Satz 2.4 Schätzer für die Varianz der Chain Ladder Endschadenstände.
Für $i \in \{2, .., n\}$ können wir die Varianz der Chain Ladder Endschadenstände durch

$$\hat{Var}(S_{i,n}^{CL}) = \sum_{k=n+1-i}^{n-1} \hat{\alpha_k}^2 \frac{(S_{i,k}^{CL})^2}{\sum_{j=1}^{n-k} S_{j,k}} F_{k+1}^2 \dots F_{n-1}^2$$

schätzen.

Beweis.
Da für die Varianz einer Zufallsvariable

$$Var(X) = \mathbb{E}[Var(X|F)] + Var(\mathbb{E}[X|F])$$

gilt, wobei F eine Sigma-Algebra ist, erhalten wir mit (2.1) und (2.2) und für $k > n+1-i$

$$\begin{aligned} Var(S_{i,k}^{CL}) &= \mathbb{E}[Var(S_{i,k}^{CL}|B_{k-1})] + Var(\mathbb{E}[S_{i,k}^{CL}|B_{k-1}]) \\ &= \mathbb{E}[(S_{i,k-1}^{CL})^2 Var(F_{k-1}|B_{k-1})] + Var(S_{i,k-1}^{CL}\mathbb{E}[F_{k-1}|B_{k-1}]) \\ &= \frac{\alpha_{k-1}^2}{\sum_{j=1}^{n+1-k} S_{j,k-1}} \mathbb{E}[(S_{i,k-1}^{CL})^2] + \varphi_{k-1}^2 Var(S_{i,k-1}^{CL}) \end{aligned}$$

Dabei ist zu beachten, dass

$$Var(S_{i,k}^{CL}) = 0 \quad k \leq n+1-i$$

gilt.
Wenden wir diese Gleichung nun iterativ an für $n+1-i < k \leq n$ und ersetzen φ_{k-1}, α_{k-1}, $\mathbb{E}[S_{i,k-1}^2]$ durch deren Schätzer F_{k-1},$\hat{\alpha}_{k-1}$,$(S_{i,k-1}^{CL})^2$, so erhalten wir die obige Behauptung.
□

Da unter der Annahme, dass $S_{i,n+1-i}$ konstant ist,

$$Var(S_{i,n}^{CL}) = Var(S_{i,n+1-i} + R_i^{CL}) = Var(R_i^{CL})$$

gilt, erhalten wir auch einen Schätzer für die Varianz der Chain Ladder Reserve.

Zum Abschluss wollen wir noch einen Schätzer für die Varianz des Gesamtreserveschätzers betrachten. Dabei können wir nicht so einfach vorgehen wie bei der tatsächlichen Reserve, da die Schätzer der Reserve der einzelnen Schadenjahre nicht unkorreliert sind.

Satz 2.5 Schätzer für die Varianz der Chain Ladder Gesamtschadenstände. Ein Schätzer für die Varianz der Chain Ladder Gesamtschadenstände ist durch

$$\hat{Var}(\sum_{k=2}^{n} S_{k,n}^{CL}) = \sum_{k=2}^{n}(\hat{Var}(S_{k,n}^{CL}) + 2 \sum_{\substack{j\in\{3,..,n\} \\ k<j}} \hat{Cov}(S_{k,n}^{CL}, S_{j,n}^{CL}))$$

gegeben, wobei

$$\begin{aligned}\hat{Cov}(S_{k,n}^{CL}, S_{j,n}^{CL}) =& S_{k,n+1-k} S_{j,n+1-j} F_{n+1-j}..F_{n-k} * \\ & (F_{n+1-k}^2 + \frac{\hat{\alpha}_{n+1-k}^2}{\sum_{l=1}^{k-1} S_{l,n+1-k}})..(F_{n-1}^2 + \frac{\hat{\alpha}_{n-1}^2}{S_{1,n-1}}) \\ & - S_{k,n+1-k} S_{j,n+1-j} F_{n+1-j}..F_{n-k} F_{n+1-k}^2..F_{n-1}^2\end{aligned}$$

gilt.

Beweis. Da für die Varianz einer Summe von Zufallsvariablen $X_1,..,X_n$

$$Var(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n}(Var(X_k) + 2 \sum_{\substack{j\in\{2,..,n\} \\ k<j}} Cov(X_k, X_j))$$

gilt, müssen wir noch die Kovarianzen der Schätzer der Endschadenstände betrachten. Da für die Kovarianz

$$Cov(X_j, X_k) = \mathbb{E}[X_j X_k] - \mathbb{E}[X_j]\mathbb{E}[X_k]$$

gilt, werden wir diese Erwartungswerte getrennt betrachten. Sei nun $k < j$. Mit (2.2) erhalten wir

$$\begin{aligned}\mathbb{E}[S_{k,n}^{CL}] &= \mathbb{E}[S_{k,n+1-k} F_{n+1-k}..F_{n-1}] = \mathbb{E}[S_{k,n+1-k} F_{n+1-k}..F_{n-2} \mathbb{E}[F_{n-1}|B_{n-1}]] \\ &= S_{k,n+1-k} \varphi_{n-1} \mathbb{E}[F_{n+1-k}..F_{n-2}] = .. = S_{k,n+1-k} \varphi_{n+1-k}..\varphi_{n-1}\end{aligned}$$

und damit

$$\mathbb{E}[S_{k,n}^{CL}]\mathbb{E}[S_{j,n}^{CL}] = S_{j,n+1-j} S_{k,n+1-k} \varphi_{n+1-j}..\varphi_{n-k} \varphi_{n+1-k}^2..\varphi_{n-1}^2 \quad .$$

Unter Zuhilfenahme von (2.2) und (2.1) erhalten wir

$$\begin{aligned}\mathbb{E}[S_{k,n}^{CL} S_{j,n}^{CL}] &= S_{j,n+1-j} S_{k,n+1-k} \mathbb{E}[F_{n+1-j}..F_{n-k} F_{n+1-k}^2..F_{n-1}^2] \\ &= S_{j,n+1-j} S_{k,n+1-k} \mathbb{E}[F_{n+1-j}..F_{n-k} F_{n+1-k}^2..F_{n-2}^2 \mathbb{E}[F_{n-1}^2|B_{n-1}]] \\ &= S_{j,n+1-j} S_{k,n+1-k} (\varphi_{n-1}^2 + \frac{\alpha_{n-1}^2}{S_{1,n-1}}) \mathbb{E}[F_{n+1-j}..F_{n-k} F_{n+1-k}^2..F_{n-2}^2] \\ &= .. = S_{j,n+1-j} S_{k,n+1-k} \varphi_{n+1-j}..\varphi_{n-k} * \\ &\quad (\varphi_{n+1-k}^2 + \frac{\alpha_{n+1-k}^2}{\sum_{l=1}^{k-1} S_{l,n+1-k}})..(\varphi_{n-1}^2 + \frac{\alpha_{n-1}^2}{S_{1,n-1}}) \quad .\end{aligned}$$

Somit erhalten wir den oben angegebenen Schätzer, wenn wir anstatt φ_k und α_k^2 deren Schätzer F_k und $\hat{\alpha}_k^2$ verwenden sowie die bereits oben berechneten Schätzer für Varianz der Jahresendschadenstände.

$\square$

Somit haben wír auch einen Schätzer für die Varianz der Gesamtschadenreserve, denn es gilt

$$Var(\sum_{k=2}^{n} S_{k,n}^{CL}) = Var(\sum_{k=2}^{n} R_k^{CL})$$

2.2.4 Beispiele Chain-Ladder Verfahren

Nun werden wir dieses Verfahren anhand zweier realer Versicherungsbestände betrachten. Dabei werden wir zunächst das Verfahren durchführen und dann dieses Ergebnis mit dem realen Datenbestand abgleichen um die Aussagekraft der Schätzung zu erhalten.
Dabei ist zu beachten, dass diese einmalige Durchführung der Plausibiliesierung nur zur Veranschaulichung dient und öfters durchgeführt werden muss, um ein Verfahren zu verwerfen oder beizubehalten, da eine einmalige Durchführung von Ausreißern beeinflusst werden kann.

Die zwei realen Versicherungsbestände die wir betrachten werden, sind KFZ-Kasko und Rechtsschutz. Dabei sind die Daten die wir zur Modellierung betrachten die Schadenzahlungen dieser zwei Sparten.

2.2.4.1 KFZ-Kasko

Als Ausgangspunkt des Modells steht das kumulierte Zahlungsdreieck dieser Sparte, welches folgende Form hat.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47		
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93			
5	15.452.148,36	18.356.911,24	18.489.553,46				
6	15.463.013,72	18.430.481,19					
7	13.768.695,59						

Tabelle 2.3: Kumuliertes KFZ-Kasko Dreieck

Aus diesem Dreieck erhält man nun die individuellen Abwicklungsfaktoren

SJ \ AJ	1	2	3	4	5	6	7
1	1,245699	1,000098	1,000859	1,000211	1,000084	1,000041	
2	1,214976	1,004405	1,001121	1,000003	1,000007		
3	1,193434	1,007916	1,005709	1,000115			
4	1,163302	1,008444	1,002708				
5	1,187984	1,007226					
6	1,191907						
7							

beziehungsweise die Chain-Ladder Faktoren

	1	2	3	4	5	6
F_j	1,195747	1,006096	1,002760	1,000103	1,000041	1,000041

Betrachtet man die individuellen Abwicklungsfaktoren erkennen wir, dass sie kaum streuen und daher die Annahmen des Chain-Ladder Modells erfüllt sind. Wenden wir nun diese Faktoren auf unser Zahlungsdreieck an erhalten wir das geschätzte Zahlungsquadrat, welches folgendende Form hat.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	15.498.922,06
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47	19.723.504,97	19.724.312,26
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93	18.922.119,72	18.922.896,36	18.923.670,88
5	15.452.148,36	18.356.911,24	18.489.553,46	18.540.585,85	18.542.500,45	18.543.261,51	18.544.020,49
6	15.463.013,72	18.430.481,19	18.542.827,66	18.594.007,10	18.595.927,20	18.596.690,46	18.597.451,63
7	13.768.695,59	16.463.875,47	16.564.234,13	16.609.952,51	16.611.667,74	16.612.349,55	16.613.029,50

Tabelle 2.4: KFZ-Kasko Chain Ladder

Dabei sind die grau eingefärbten Werte die Schätzer. Bevor wir nun die Reserve angeben, berechnen wir noch die $\hat{\alpha}_j, j \in \{1, .., n-1\}$, um für die Reserven auch die Standardabweichung und die Konfidenzintervalle angeben zu können. Die Schätzer $\hat{\alpha}_j$ haben folgende Form.

	1	2	3	4	5	6
$\hat{\alpha}_j$	9518,308575	169,7710971	97,47988015	0,151758688	0,040176823	0,040176823

Somit erhalten wir für die Reserven folgende Werte

SJ	R_i^{CL}	Stdabw. $S_{i,n}$	2,5%-Quantil	97,5%-Quantil	Stdabw. $S_{i,n}^{CL}$
2	634,35	789,10	57,46	2.749,15	883,96
3	1.616,79	1.258,92	321,94	5.054,79	1.351,58
4	3.504,95	2.095,79	995,59	9.089,32	1.680,42
5	54.467,03	42.512,72	10.807,44	170.580,91	22.483,71
6	166.970,44	70.427,01	68.487,20	345.587,33	34.593,84
7	2.844.333,91	371.309,49	2.174.709,18	3.657.810,84	149.482,42
Gesamt	3.071.527,48	380.321,75	2.211.476,74	4.160.855,56	219.992,53

Tabelle 2.5: Werte KFZ-Kasko Chain Ladder

Dabei wurde zur Berechnung der Konfidenzintervalle die Annahme getroffen, dass die Reserven logarithmisch normalverteilt sind, da wir ansonsten für die Schadenjahre 2-5 als untere Intervallgrenze eine negative Reserve erhalten würden.
Nun betrachten wir noch das tatsächlich beobachtete Quadrat. Dieses hat folgende Form

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	15.499.202,02
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47	19.722.695,47	19.722.939,17
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93	18.927.186,83	18.931.158,56	18.931.978,64
5	15.452.148,36	18.356.911,24	18.489.553,46	18.489.986,41	18.490.702,29	18.491.202,29	18.491.373,02
6	15.463.013,72	18.430.481,19	18.582.598,10	18.589.779,85	18.599.426,35	18.600.784,49	18.601.256,49
7	13.768.695,59	16.330.431,63	16.439.713,95	16.471.855,69	16.472.920,60	16.473.930,60	16.473.930,60

Tabelle 2.6: Tatsächliches KFZ-Kasko Quadrat

Daher ergeben sich folgende Differenzen zwischen den beobachteten und geschätzten Werten.

SJ	R_i	$R_i^{CL} - R_i$
2	914,31	- 279,96
3	243,70	1.373,09
4	11.812,71	- 8.307,76
5	1.819,56	52.647,47
6	170.775,30	- 3.804,86
7	2.705.235,01	139.098,90
Gesamt	2.890.800,59	180.726,89

Hier ist gut zu erkennen, dass, bis auf das 5. Schadenjahr, die beobachteten Werte und die geschätzten sehr ähnlich sind. Außerdem befindet sich die tatsächlich beobachtete Reserve unter Ausnahme des 5. Schadenjahres für alle Schadenjahre im jeweiligen Konfidenzintervall. Im 5. Schadenjahr wird die Reserve deutlich überschätzt was dazu führt, dass die tatsächliche Reserve geringer ist als die untere Grenze des Konfidenzintervalls.

2.2.4.2 Rechtsschutz

Auch hier beginnen wir das Modell mit der Betrachtung des kumulierten Zahlungsdreiecks. Dieses hat folgende Form.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78		
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44			
5	363.636,21	1.116.438,19	1.623.695,82				
6	419.610,32	1.433.715,17					
7	448.107,48						

Tabelle 2.7: Kumuliertes Rechtsschutz Dreieck

Betrachten wir nun bei dieser Sparte die individuellen Abwicklungsfaktoren

SJ \ AJ	1	2	3	4	5	6	7
1	4,61648	1,44610	1,21078	1,13350	1,04399	1,09286	
2	4,94934	1,67280	1,20454	1,19958	1,05156		
3	4,19430	1,45477	1,16057	1,12968			
4	3,92608	1,46158	1,20176				
5	3,07021	1,45435					
6	3,41678						
7							

ist schön zu erkennen, dass vorallem im ersten Abwicklungsjahr die Faktoren stark streuen und daher die Chain Ladder Annahmen nicht ganz erfüllt sind.Weiter sehen wir, dass diese Sparte länger abwickelt als KFZ-Kasko. Daher ist bei den letzten Abwicklungsjahren die Schätzung schwerer, da hier die Datenbasis nicht mehr so groß ist, aber noch größere Zahlungen fließen können.
Die Chain-Ladder Faktoren haben bei dieser Sparte folgende Form.

j	1	2	3	4	5	6
F_j	3,71423	1,48462	1,19247	1,15391	1,04842	1,09286

Damit erhalten wir das geschätzte Zahlungsquadrat. Auch hier sind wieder die grau gefärbten Werte die Schätzer.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	1.435.739,53
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78	1.547.618,79	1.691.330,48
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44	2.043.600,51	2.142.556,18	2.341.513,68
5	363.636,21	1.116.438,19	1.623.695,82	1.936.208,23	2.234.205,00	2.342.390,17	2.559.904,23
6	419.610,32	1.433.715,17	2.128.517,05	2.538.192,30	2.928.838,87	3.070.659,75	3.355.800,84
7	448.107,48	1.664.373,56	2.470.956,27	2.946.540,73	3.400.035,15	3.564.672,40	3.895.687,44

Tabelle 2.8: Rechtsschutz Chain Ladder

Die Schätzer $\hat{\alpha}_j, j \in \{1, .., n-1\}$ haben bei der Sparte Rechtsschutz folgende Form.

j	1	2	3	4	5	6
$\hat{\alpha}_j$	95.966,59	5.303,19	538,36	1.631,90	29,71	29,71

Somit erhalten wir für die Reserve

SJ	R_i^{CL}	Stdabw. $S_{i,n}$	2,5%-Quantil	97,5%-Quantil	Stdabw. $S_{i,n}^{CL}$
2	121.994,23	6.247,11	109.500,00	134.488,45	7.454,65
3	215.189,70	9.916,98	195.355,74	235.023,66	10.647,16
4	570.487,24	62.692,49	445.102,27	695.872,21	48.637,63
5	936.208,41	76.321,23	783.565,96	1.088.850,86	58.578,16
6	1.922.085,67	162.895,40	1.596.294,87	2.247.876,46	113.643,85
7	3.447.579,96	516.139,90	2.415.300,15	4.479.859,76	302.433,65
Gesamt	7.213.545,20	550.298,07	6.112.949,07	8.314.141,33	604.212,59

Hier wurde zur Berechung der Konfidenzintervalle angenommen, dass die Reserven normalverteilt sind. Auch hier wollen wir nun das tatächlich beobachtete Zahlungsquadrat betrachten und mit dem geschätzten vergleichen. Das tatsächliche Quadrat hat folgende Form.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	1.358.927,95
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78	1.539.980,31	1.628.371,44
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44	1.932.512,08	2.061.823,12	2.215.163,34
5	363.636,21	1.116.438,19	1.623.695,82	1.928.900,99	2.251.362,16	2.567.299,12	2.859.606,91
6	419.610,32	1.433.715,17	2.091.419,97	2.711.636,97	3.216.827,36	3.621.598,68	3.822.963,90
7	448.107,48	1.593.877,88	2.344.880,19	3.065.084,79	3.601.092,25	3.960.953,39	4.116.655,97

Tabelle 2.9: Tatsächliches Rechtsschutz Quadrat

Daher ergeben sich für die Reserven folgende Differenzen.

SJ	R_i	$R_i^{CL} - R_i$
2	45.182,65	76.811,58
3	152.230,66	62.959,04
4	444.136,90	126.350,34
5	1.235.911,09	- 299.702,68
6	2.389.248,73	- 467.163,06
7	3.668.548,49	- 220.968,53
Gesamt	7.935.258,52	- 721.713,32

Hier ist gut zu erkennen, dass die Reserven in jedem Abwicklungsjahr entweder deutlich überschätzt beziehungsweise unterschätzt werden. Außerdem sind für alle Schadenjahre mit Ausnahme des 7. die tatsächlich beobachteten Reserven außerhalb des Konfidenzintervalls. Daher ist die getroffene Annahme, dass die Reserven normalverteilt sind mit den angegebenen Parametern zu verwerfen.
Zu erwähnen ist, dass die tatsächliche Reserve auch nicht im Konfidenzintervall liegen würde wenn wir die Annahme treffen würden, dass die tatsächliche Reserve logarithmisch normalverteilt ist.

Kapitel 3

Das lineare Modell

3.1 Grundlagen

Lineare Modelle sind in der mathematischen Statistik weit verbreitet. Die Grundannahme dieser Modelle ist, dass die Erwartungswerte aller Zufallsvariablen mittels einer bekannten Matrix von einem unbekannten Parameter, welcher ein Vektor ist, abhängen.
Zunächst werden wir das elementare lineare Modell beschreiben. Hier sind alle Zufallsvariablen beobachtbar und das Problem besteht in der Schätzung des Parameters.
Danach betrachten wir das erweiterte lineare Modell. Hier sind nicht mehr alle Zufallsvariablen beobachtbar. Daher ist das Problem nicht mehr nur die Schätzung des Parameters, sondern auch die Prognose der nicht beobachtbaren Zufallsvariablen mittels der beobachtbaren.
Das erweiterte lineare Modell werden wir auch in der Schadenreservierung heranziehen. Dabei werden wir uns stark an Radtke und Schmidt [4] orientieren. Zu erwähnen ist, dass die Beweise, die in diesem Kapitel verwendet werden, nicht in [4] vorkommen und daher selbst geführt wurden.

3.1.1 Das elementare lineare Modell

Wir betrachten hier einen Zufallsvektor $\mathbf{X} \in \mathbb{R}^m$ und definieren mit

$$\Sigma := Var(\mathbf{X})$$

die Kovarianzmatrix von $\mathbf{X}$. Das elementare Modell besteht aus folgenden Annahmen.

Annahme 3.1 Elementares lineares Modell.

- Es existiert ein Parameter $\beta \in \mathbb{R}^s$ und eine Matrix $A \in \mathbb{R}^{m \times s}$ mit $rang(A) = s$ und $\mathbb{E}[\mathbf{X}] = A\beta$
- Die Matrix Σ ist positiv definit.

Wir setzen nun voraus, dass diese Annahmen erfüllt sind und die Matrix Σ bekannt ist. Wir wollen nun den unbekannten Parameter β schätzen. Für einen Schätzer $\hat{\beta}$ von β

bezeichnen wir die Differenz $\hat{\beta} - \beta$ als Schätzfehler und den Erwartungswert

$$\mathbb{E}[(\hat{\beta} - \beta)(\hat{\beta} - \beta)]$$

als erwarteten quadratischen Schätzfehler. Dies bildet die Grundlage für den Vergleich von Schätzern für β. Aufgrund folgender Gleichung

$$\begin{aligned}\mathbb{E}[(\hat{\beta} - \beta)(\hat{\beta} - \beta)] &= spur(Var(\hat{\beta} - \beta) + \mathbb{E}[\hat{\beta} - \beta]^T\mathbb{E}[\hat{\beta} - \beta]) \\ &= spur(Var(\hat{\beta} - \beta)) + \mathbb{E}[\hat{\beta} - \beta]^T\mathbb{E}[\hat{\beta} - \beta]\end{aligned}$$

ist der erwartete quadratische Schätzfehler durch den Erwartungswert und die Varianz des Schätzfehlers bestimmt.

Wir betrachten nun Eigenschaften die ein Schätzer $\hat{\beta}$ besitzen kann.
Ein Zufallsvektor $\hat{\beta}$ heißt linearer Schätzer, wenn eine Matrix B existiert mit

$$\hat{\beta} = B\mathbf{X}$$

Ein Zufallsvektor $\hat{\beta}$ heißt erwartungstreuer Schätzer, wenn

$$\mathbb{E}[\hat{\beta}] = \beta$$

gilt.
Ein Zufallsvektor $\hat{\beta}$ heißt Gauss-Markov Schätzer, wenn er ein erwartungstreuer, linearer Schätzer ist und unter allen erwartungstreuen, linearen Schätzern den kleinsten erwarteten quadratischen Schätzfehler besitzt.
Für jeden erwartungstreuen Schätzer $\hat{\beta}$ für β gilt $\mathbb{E}[(\hat{\beta} - \beta)(\hat{\beta} - \beta)] = spur(Var(\hat{\beta} - \beta))$.

Satz 3.1 Gauss-Markov Theorem des elementaren linearen Modells.
Es existiert ein eindeutig bestimmter Gauss Markov Schätzer $\beta^\star$ für β und es gilt

$$\beta^\star = (A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1}\mathbf{X}$$

und

$$Var(\beta^\star) = (A^T\Sigma^{-1}A)^{-1}$$

Beweis.
Das der Schätzer $\beta^\star$ linear ist, ist offensichtlich. Aufgrund der Annahmen des linearen Modells gilt

$$\mathbb{E}[\beta^\star] = (A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1}\mathbb{E}[\mathbf{X}] = (A^T\Sigma^{-1}A)^{-1}(A^T\Sigma^{-1}A)\beta = \beta \quad .$$

Daher handelt es sich bei $\beta^\star$ um einen erwartungstreuen, linearen Schätzer.
Sei $\hat{\beta}$ ein weiterer erwartungstreuer linearer Schätzer. Dann gilt

$$\beta = \mathbb{E}[\hat{\beta}] = \mathbb{E}[B\mathbf{X}] = BA\beta \quad .$$

Daher muss $BA = I$ gelten, wobei I die Einheitsmatrix bezeichnet. Unter Zuhilfenahme dieser Gleichung folgt

$$\begin{aligned}
0 \leq Var(\hat{\beta} - \beta^\star) &= Var(\hat{\beta}) + Var(\beta^\star) - Cov(\beta^\star, \hat{\beta}) - Cov(\hat{\beta}, \beta^\star) \\
&= Var(\hat{\beta}) + Var(\beta^\star) - Cov(B\mathbf{X}, \beta^\star) - Cov(\beta^\star, B\mathbf{X}) \\
&= Var(\hat{\beta}) + Var(\beta^\star) - B \cdot Cov(X, X)((A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1})^T \\
&- (A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1} \cdot Cov(X, X)B^T \\
&= Var(\hat{\beta}) + Var(\beta^\star) - BA(A^T\Sigma^{-1}A)^{-1} - (A^T\Sigma^{-1}A)^{-1}(BA)^T \\
&= Var(\hat{\beta}) + Var(\beta^\star) - 2Var(\beta^\star) \quad .
\end{aligned}$$

Daher folgt $Var(\beta^\star) \leq Var(\hat{\beta})$. Schließlich fehlt noch die Varianz von $\beta^\star$.

$$\begin{aligned}
Var(\beta^\star) &= (A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1} \cdot Var(\mathbf{X})((A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1})^T \\
&= (A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1}AA^{-1}\Sigma(A^T)^{-1} = (A^T\Sigma^{-1}A)^{-1}
\end{aligned}$$

□

3.1.2 Das erweiterte lineare Modell

In diesem Kapitel betrachten wir wieder einen Vektor $\mathbf{X} \in \mathbb{R}^m$, bei dem nur noch die ersten m_1 Einträge beobachtbar sind und die Übrigen $m_2 := m - m_1$ nicht. Das heißt, der Vektor $\mathbf{X}$ lässt sich schreiben als

$$\mathbf{X} = \begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{pmatrix}$$

wobei wir annehmen, dass $\mathbf{X_1}$ beobachtbar ist und $\mathbf{X_2}$ nicht. Definiert man nun die Kovarianzmatrizen

$$\begin{aligned}
\Sigma_{11} &:= Var(\mathbf{X_1}) \\
\Sigma_{12} &:= Cov(\mathbf{X_1}, \mathbf{X_2}) \\
\Sigma_{21} &:= Cov(\mathbf{X_2}, \mathbf{X_1}) \\
\Sigma_{22} &:= Var(\mathbf{X_2})
\end{aligned}$$

so erhalten wir die Kovarianzmatrix von $\mathbf{X}$ durch

$$Var(\mathbf{X}) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad .$$

Kommen wir nun zu den Annahmen die in diesem Modell getroffen werden.

Annahme 3.2 Erweitertes lineares Modell.

- Es existiert ein Parameter $\beta \in \mathbb{R}^s$ und Matrizen $A_1 \in \mathbb{R}^{m_1 \times s}$, $A_2 \in \mathbb{R}^{m_2 \times s}$ mit $rang(A_1) = s$ und
$$\mathbb{E}\left[\begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{pmatrix}\right] = \begin{pmatrix} \mathbf{A_1} \\ \mathbf{A_2} \end{pmatrix} \beta$$

- Σ_{11} ist positiv definit.

Wir setzen in diesem Abschnitt voraus, dass die Annahmen des erweiterten linearen Modells erfüllt sind.

Dabei interessieren wir uns nicht nur für die Schätzung von β und die Prognose von $\mathbf{X_2}$, sondern für die Schätzung von $C\beta$ mit $C \in \mathbb{R}^{r\times s}$ und die Prognose von $D\mathbf{X_2}$ mit $D \in \mathbb{R}^{r\times m_2}$. Hier ist dann die Schätzung von β beziehungsweise $\mathbf{X_2}$ ein Spezialfall. Von besonderem Interesse sind die Fälle $C = c^T$ mit $c \in \mathbb{R}^s$ und $D = d^T$ mit $d \in \mathbb{R}^{m_2}$, da wir auf diese Weise Linearkombinationen der Schätzungen bzw. Prognosen erhalten.
Als erstes betrachten wir die Schätzung von $C\beta$. Dafür kommt jeder Vektor $\hat{Y}$ in Frage, der nur von $\mathbf{X_1}$ abhängt. Dabei bezeichnen wir als Schätzfehler die Differenz

$$\hat{Y} - C\beta$$

und den Erwartungswert

$$\mathbb{E}[(\hat{Y} - C\beta)^T(\hat{Y} - C\beta)]$$

als erwarteten quadratischen Schätzfehler. Auch hier bezeichnen wir einen Schätzer $\hat{Y}$ für $C\beta$ als Gauss-Markov-Schätzer, wenn er ein erwartungstreuer, linearer Schätzer ist und den kleinsten quadratischen Schätzfehler besitzt.

Satz 3.2 Gauss-Markov Theorem des erweiterten linearen Modells.
Es existiert ein eindeutig bestimmter Gauss-Markov Schätzer $Y^\star(C\beta)$ für $C\beta$ und es gilt

$$Y^\star(C\beta) = C(A_1^T\Sigma_{11}^{-1}A_1)^{-1}A_1^T\Sigma_{11}^{-1}\mathbf{X_1}$$

und

$$Var(Y^\star) = C(A^T\Sigma_{11}^{-1}A)^{-1}C^T \quad .$$

Im folgenden werden wir $C\beta^\star$ für $Y^\star(C\beta)$ schreiben.
Der Beweis hierzu ist analog zum elementaren linearen Modell.

Nun werden wir die Prognose von $D\mathbf{X_2}$ mit $D \in \mathbb{R}^{r\times m_2}$ betrachten. Als Prädiktor kommt auch hier jeder Zufallsvektor $\hat{Y}$ in Frage der nur von $\mathbf{X_1}$ abhängt. Nun betrachten wir Eigenschaften, die dieser Prädiktor besitzen kann. Auch hier heißt ein Prädiktor linear, wenn es eine Matrix Q gibt mit

$$\hat{Y} = Q\mathbf{X_1} \quad .$$

Ein Prädiktor für $D\mathbf{X_2}$ heißt erwartungstreu wenn

$$\mathbb{E}[\hat{Y}] = \mathbb{E}[D\mathbf{X_2}]$$

gilt.
Ein Prädiktor heißt Gauss-Markov Prädiktor, wenn er unter allen erwartungstreuen linearen Prädiktoren den kleinsten erwarteten quadratischen Prognosefehler besitzt.
Dabei gilt hier für den quadratischen Prognosefehler $\hat{Y}$ von $D\mathbf{X_2}$

$$\mathbb{E}[(\hat{Y} - D\mathbf{X_2})^T(\hat{Y} - D\mathbf{X_2})] = spur(Var(\hat{Y} - D\mathbf{X_2})) + \mathbb{E}[\hat{Y} - D\mathbf{X_2}]^T\mathbb{E}[\hat{Y} - D\mathbf{X_2}]$$

und somit für den quadratischen Prognosefehlers des Gauss-Markov Schätzers

$$\mathbb{E}[(\hat{Y} - D\mathbf{X_2})^T(\hat{Y} - D\mathbf{X_2})] = spur(Var(\hat{Y} - D\mathbf{X_2})) \quad .$$

Satz 3.3 Gauss-Markov Theorem für Prädiktoren.
Es existiert ein eindeutig bestimmter Gauss-Markov Prädiktor $Y^\star(D\mathbf{X_2})$ für $D\mathbf{X_2}$ und es gilt

$$Y^\star(D\mathbf{X_2}) = D(A_2\beta^\star + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{X_1} - A_1\beta^\star))$$

und

$$\begin{aligned} Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) = D(&\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})D^T \\ &+ D(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T D^T \end{aligned}$$

Beweis.
Zunächst ist anzumerken, dass für die Kovarianzmatrix einer Summe von Zufallsvariablen

$$Var(X + Y) = Var(X) + Var(Y) + Cov(X, Y) + Cov(Y, X)$$

und

$$Cov(X, Y) = Cov(Y, X)^T$$

gilt. Daher gilt

$$\Sigma_{11} = \Sigma_{11}^T, \Sigma_{12} = \Sigma_{21}^T, \Sigma_{21} = \Sigma_{12}^T \quad .$$

Aufgrund der Definition von $\beta^\star$ ist die Linearität erfüllt. Die Erwartungstreue ist ebenfalls erfüllt, denn es gilt

$$\begin{aligned} \mathbb{E}[Y^\star(D\mathbf{X_2})] &= D(A_2\mathbb{E}[\beta^\star] + \Sigma_{21}\Sigma_{11}^{-1}(\mathbb{E}[\mathbf{X_1}] - A_1\mathbb{E}[\beta^\star])) \\ &= D(A_2\beta + \Sigma_{21}\Sigma_{11}^{-1}(A_1\beta - A_1\beta)) = DA_2\beta = D\mathbb{E}[\mathbf{X_2}] = \mathbb{E}[D\mathbf{X_2}] \quad . \end{aligned}$$

Nun betrachten wir $Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2})$. Einerseits gilt

$$Var(D\mathbf{X_2}) = D\Sigma_{22}D^T$$

und andererseits

$$\begin{aligned} Var(Y^\star(D\mathbf{X_2})) &= D((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)Var(\beta^\star)(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T \\ &+ \Sigma_{21}\Sigma_{11}^{-1}Var(\mathbf{X_1})(\Sigma_{21}\Sigma_{11}^{-1})^T \\ &+ (A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}A_1^T\Sigma_{11}^{-1}Cov(\mathbf{X_1}, \mathbf{X_1})(\Sigma_{21}\Sigma_{11}^{-1})^T \\ &+ ((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}A_1^T\Sigma_{11}^{-1}Cov(\mathbf{X_1}, \mathbf{X_1})(\Sigma_{21}\Sigma_{11}^{-1})^T)^T)D^T \\ &= D((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \\ &+ (A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)A_1^{-1}\Sigma_{12} + ((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)A_1^{-1}\Sigma_{12})^T)D^T \quad . \end{aligned}$$

Für die $Cov(Y^\star(D\mathbf{X_2}), D\mathbf{X_2})$ gilt

$$\begin{aligned}
Cov(Y^\star(D\mathbf{X_2}), D\mathbf{X_2}) &= D((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)Cov(\beta^\star, \mathbf{X_2}) + \Sigma_{21}\Sigma_{11}^{-1}Cov(\mathbf{X_1}, \mathbf{X_2}))D^T \\
&= D((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}A_1^T\Sigma_{11}^{-1}Cov(\mathbf{X_1}, \mathbf{X_2}) \\
&+ \Sigma_{21}\Sigma_{11}^{-1}Cov(\mathbf{X_1}, \mathbf{X_2}))D^T \\
&= D((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)A_1^{-1}\Sigma_{12} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})D^T \quad .
\end{aligned}$$

Damit erhalten wir

$$\begin{aligned}
Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) &= Var(Y^\star(D\mathbf{X_2})) + Var(D\mathbf{X_2}) \\
&- Cov(Y^\star(D\mathbf{X_2}), D\mathbf{X_2}) - Cov(D\mathbf{X_2}, Y^\star(D\mathbf{X_2})) \\
&= D((\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}) \\
&+ (A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T)D^T \quad .
\end{aligned}$$

Nun zeigen wir noch die Minimierung des quadratischen Prognosefehlers. Dazu sei $\hat{Y} = Q\mathbf{X_1}$ ein weiterer linearer erwartungstreuer Schätzer für $D\mathbf{X_2}$. Dann gilt

$$\begin{aligned}
Cov(\hat{Y} - D\mathbf{X_2}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) &= Cov(\hat{Y}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) \\
&- Cov(D\mathbf{X_2}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) \quad .
\end{aligned}$$

Betrachten wir zunächst $Cov(\hat{Y}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2})$. Hier gilt

$$\begin{aligned}
Cov(\hat{Y}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) &= Q(Cov(\mathbf{X_1}, (A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)\beta^\star + \Sigma_{21}\Sigma_{11}^{-1}\mathbf{X_1} - \mathbf{X_2})D^T \\
&= Q(Cov(\mathbf{X_1}, (A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)\beta^\star) \\
&+ Cov(\mathbf{X_1}, \Sigma_{21}\Sigma_{11}^{-1}\mathbf{X_1}) - Cov(\mathbf{X_1}, \mathbf{X_2}))D^T \\
&= Q(\Sigma_{11}((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}A_1^T\Sigma_{11}^{-1})^T \\
&+ \Sigma_{11}(\Sigma_{21}\Sigma_{11}^{-1})^T - \Sigma_{12})D^T \\
&= Q(A_1(A_1^T\Sigma_{11}^{-1}A_1)^{-1}(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T)D^T \quad .
\end{aligned}$$

Unter Zuhilfenahme folgender Gleichheit

$$DA_2\beta = D\mathbb{E}[\mathbf{X_2}] = \mathbb{E}[D\mathbf{X_2}] = \mathbb{E}[\hat{Y}] = \mathbb{E}[Q\mathbf{X_1}] = QA_1\beta$$

folgt

$$Q = DA_2A_1^{-1} \quad .$$

Daher gilt

$$Cov(\hat{Y}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) = DA_2((A_1^T\Sigma_{11}^{-1}A_1)^{-1}(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T)D^T \quad .$$

Für $Cov(D\mathbf{X_2}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2})$ gilt indes

$$\begin{aligned}
Cov(D\mathbf{X_2}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) &= D(Cov(\mathbf{X_2}, Y^\star(D\mathbf{X_2})) - Cov(\mathbf{X_2}, D\mathbf{X_2})) \\
&= D(Cov(\mathbf{X_2}, (A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)\beta^\star) \\
&+ Cov(\mathbf{X_2}, \Sigma_{21}\Sigma_{11}^{-1}\mathbf{X_1}) - Cov(\mathbf{X_2}, \mathbf{X_2}))D^T \\
&= D(\Sigma_{21}((A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}A_1^T\Sigma_{11}^{-1})^T \\
&+ \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \Sigma_{22})D^T \\
&= D(\Sigma_{21}\Sigma_{11}^{-1}A_1(A_1^T\Sigma_{11}^{-1}A_1)^{-1}(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T \\
&+ \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \Sigma_{22})D^T \quad .
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
Cov(\hat{Y} - D\mathbf{X_2}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) &= D((\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}) \\
&\quad + (A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)(A_1^T\Sigma_{11}^{-1}A_1)^{-1}(A_2 - \Sigma_{21}\Sigma_{11}^{-1}A_1)^T)D^T \\
&= Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2})
\end{aligned}$$

und damit

$$\begin{aligned}
0 &\leq Var((\hat{Y} - D\mathbf{X_2}) - (Y^\star(D\mathbf{X_2}) - D\mathbf{X_2})) \\
&= Var(\hat{Y} - D\mathbf{X_2}) + Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) \\
&\quad - Cov(\hat{Y} - D\mathbf{X_2}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) - (Cov(\hat{Y} - D\mathbf{X_2}, Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}))^T \\
&= Var(\hat{Y} - D\mathbf{X_2}) - Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) \quad .
\end{aligned}$$

Somit erhalten wir

$$Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) \leq Var(\hat{Y} - D\mathbf{X_2}) \quad .$$

□

Am Ende dieses Abschnitts wollen wir noch ein Lemma betrachten, dass ein Spezialfall des Gauss-Markov Theorems ist und welches wir für die Schadenreservierung benötigen werden.

Lemma 3.1.
Sei $\Sigma_{21} = O$, wobei O jene Matrix ist, deren Einträge gleich Null sind. Dann gilt

$$Y^\star(D\mathbf{X_2}) = DA_2\beta^\star$$

und

$$Var(Y^\star(D\mathbf{X_2}) - D\mathbf{X_2}) = D(\Sigma_{22} + A_2(A_1^T\Sigma_{11}^{-1}A_1)^{-1}A_2^T)D^T \quad .$$

3.2 Das lineare Modell in der Schadenreservierung

Das lineare Modell ist ein Modell basierend auf Schadenzuwächsen. Dabei wird vorausgesetzt, dass die $T_{i,j}$ für $i+j \leq n+1$ beobachtbar sind und für $i+j > n+1$ nicht. Unser Ziel ist es nun diese Zuwächse so zu strukturieren, dass wir die Theorie des erweiterten linearen Modells anwenden können.

Satz 3.4.
Sei $m := n^2, m_1 := \frac{n(n+1)}{2}, m_2 := \frac{n(n-1)}{2}$. Durch die Zuordnung

$$\mathbf{X}_{h(i,j)} := T_{i,j}$$

mit

$$h(i,j) = \begin{cases} j-(i-1)+\frac{(2n+3-(i-1))(i-1)}{2} & \text{falls } i+j \leq n+1 \\ ((i+j)-(n+2))+\frac{n(n+1)+(i-2)^2+i}{2} & \text{falls } i+j > n+1 \end{cases}$$

erhält man eine Anordnung der Zuwächse zu einem Zufallsvektor

$$\mathbf{X} = \begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{pmatrix}$$

wobei $\mathbf{X_1}$ aus den m_1 beobachtbaren Zuwächsen besteht und $\mathbf{X_2}$ aus den m_2 nicht beobachtbaren.

Beweis. Zunächst werden wir zeigen, dass die Funktion $h_{i,j}$ für $i+j \leq n+1$ jeden Wert der Menge $\{1,..,m_1\}$ genau einmal annimmt.
Wenn wir i festhalten sehen wir, dass $h_{i,j}$ streng monoton wachsend in j ist. Weiter gilt

$$h_{1,1} = 1$$

und

$$\begin{aligned} h_{n,1} &= 1-(n-1)+\frac{(2n+3-(n-1))(n-1)}{2} = 2-n+\frac{(n+4)(n-1)}{2} \\ &= \frac{4-2n+n^2+4n-n-4}{2} = \frac{n^2+n}{2} = m_1 \end{aligned}$$

Abschließend bleibt folgende Gleichheit zu zeigen

$$h_{i,n+1-i} = h_{i+1,1} - 1$$

Es gilt

$$h_{i+1,1} - 1 = 1 - i + \frac{(2n+3-i)i}{2} - 1 = -i + \frac{(2n+3-i)i}{2}$$

und

$$\begin{aligned} h_{i,n+1-i} &= n+1-i-i+1+\frac{(2n+3-(i-1))(i-1)}{2} \\ &= n+2-2i+\frac{(2n+3-i+1)(i-1)}{2} \\ &= n+2-2i+\frac{(2n+3-i)i+i-2n-4+i}{2} = -i+\frac{(2n+3-i)i}{2} \end{aligned}$$

Somit ist für $i+j \leq n+1$ die Annahme gezeigt.

Nun betrachten wir die Funktion $h_{i,j}$ für $i+j > n+1$ und zeigen, dass diese jeden Wert der Menge $\{m_1+1,..,m\}$ genau einmal annimmt. Auch hier ist leicht zu erkennen, dass die Funktion streng monoton wachsend ist in j für festes i. Es gilt

$$h_{2,n} = (n+2-(n+2)) + \frac{n(n+1)+(2-2)^2+2}{2} = m_1 + 1$$

Weiter gilt

$$h_{n,n} = 2n - (n+2) + \frac{n(n+1) + (n-2)^2 + n}{2} = n - 2 + \frac{n(n+1) + n^2 - 4n + 4 + n}{2}$$
$$= \frac{n(n+1) - n + n^2}{2} = n^2 = m$$

Nun bleibt noch zu zeigen, dass

$$h_{i,n} = h_{i+1,n+1-i} - 1$$

gilt. Betrachtet man $h_{i,n}$, so gilt

$$h_{i,n} = i + n - (n+2) + \frac{n(n+1) + i^2 - 4i + 4 + i}{2} = \frac{n(n+1) - i + i^2}{2}$$

Andererseits haben wir

$$h_{i+1,n+1-i} - 1 = (n + 2 - (n+2)) + \frac{n(n+1) + (i-1)^2 + i + 1}{2} - 1$$
$$= \frac{n(n+1) + i^2 - i + 2}{2} - 1 = \frac{n(n+1) + i^2 - i}{2}$$

und somit gilt auch für $i + j > n + 1$ die Behauptung gezeigt. □

Um nun zu einem linearen Modell zu kommen, nehmen wir an, dass es einen Parameter $\beta \in \mathbb{R}^s$ gibt mit $s \leq m$ und für alle $i, j \in \{1, .., n\}$ einen Vektor $\mathbf{g_{i,j}} \in \mathbb{R}^s$ gibt, sodass für alle $i, j \in \{1, .., n\}$

$$\mathbb{E}[T_{i,j}] = \mathbf{g_{i,j}}^T \beta$$

gilt. Bezeichnet man mit $\mathbf{a_j}$ die j-te Zeile einer Matrix A, so erhalten wir durch die Zuordnung

$$\mathbf{a_{h_{i,j}}} := g_{i,j}^T$$

eine Matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{A_1} \\ \mathbf{A_2} \end{pmatrix}$$

sodass

$$\mathbb{E}[\begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{pmatrix}] = \begin{pmatrix} \mathbf{A_1} \\ \mathbf{A_2} \end{pmatrix} \beta$$

gilt.

Wenn außerdem Σ_{11} positiv definit ist und $rang(A_1) = s$ gilt, sind die Annahmen des erweiterten, linearen Modells erfüllt.

Unser Interesse bei der Schadenreservierung gilt dem Gauss Markov Prädiktor

$$Y^\star(\sum_{i=2}^{n} \sum_{j=n+2-i}^{n} T_{i,j})$$

für die globale Reserve, und dem Gauss-Markov Prädiktor

$$Y^\star(\sum_{j=n+2-i}^{n} T_{i,j})$$

für die Reserve des Schadenjahres $i \in \{2, .., n\}$.

In all diesen Fällen ist es also notwendig den Gauss-Markov Prädiktor für $d^T \mathbf{X_2}$ mit $d \in \mathbb{R}^{m_2}$ zu bestimmen. Da

$$Y^\star(d^T \mathbf{X_2}) = d^T Y^\star(\mathbf{X_2})$$

gilt, folgt daraus

$$\begin{aligned} Y^\star(\sum_{i=2}^{n} \sum_{j=n+2-i}^{n} T_{i,j}) &= \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} Y^\star(T_{i,j}) \\ Y^\star(\sum_{j=n+2-i}^{n} T_{i,j}) &= \sum_{j=n+2-i}^{n} Y^\star(T_{i,j}) \quad . \end{aligned}$$

Daher ist es also ausreichend die Gauss-Markov Prädiktoren der einzelnen nicht beobachtbaren Schadenstände zu bestimmen.

3.2.1 Das lineare Modell von Mack

Als ein spezielles lineares Modell wollen wir das von Mack betrachten. Dieses Modell besteht aus folgenden Annahmen.

Annahme 3.3 Lineares Modell von Mack.
Es existieren Parameter $\upsilon_1, ..., \upsilon_n \in (0, \infty)$ sowie $\zeta_1, ..., \zeta_n \in \mathbb{R}$ und $\sigma_1^2, ..., \sigma_n^2 \in (0, \infty)$ mit

$$\mathbb{E}[\frac{T_{i,j}}{\upsilon_i}] = \zeta_j$$

und

$$Cov(\frac{T_{i,j}}{\upsilon_i}, \frac{T_{k,l}}{\upsilon_k}) = \frac{\sigma_j^2}{\upsilon_i} \delta_{i,k} \delta_{j,l}$$

für alle $i, j, k, l \in \{1, .., n\}$.

In diesem Abschnitt wird vorausgesetzt, dass die Annahmen des linearen Modells von Mack erfüllt sind, die Parameter $\upsilon_1, .., \upsilon_n$ und $\sigma_1^2, .., \sigma_n^2$ bekannt sind und die Parameter $\zeta_1, ..\zeta_n$ unbekannt sind.
Dabei stellen die Parameter $\upsilon_1, .., \upsilon_n$ Volumsmaße dar. Diese können entweder die Prämie des Schadenjahres i oder die Größe des Bestandes im Schadenjahr i sein. Sollten die Varianzen $\sigma_1^2, .., \sigma_n^2$ nicht bekannt sein, so müssen diese genauso wie die Parameter $\zeta_1, ..\zeta_n$ geschätzt werden.

Setzen wir nun $s := n$ und

$$\beta := \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix}$$

Dann gilt für alle $i, j \in \{1, ..., n\}$

$$\upsilon_i \zeta_j = g_{i,j}^T \beta \quad ,$$

wobei

$$g_{i,j} := \upsilon_i e_j \quad .$$

Daher gibt es Matrizen $A_1 \in \mathbb{R}^{m_1 \times s}$ und $A_2 \in \mathbb{R}^{m_2 \times s}$ mit

$$\mathbb{E}[\begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{pmatrix}] = \begin{pmatrix} \mathbf{A_1} \\ \mathbf{A_2} \end{pmatrix} \beta$$

Da alle Zuwächse des Schadenjahres 1 beobachtbar sind, gilt $rang(A_1) = n$. Außerdem ist nach den Annahmen Σ_{11} eine invertierbare Diagonalmatrix. Daher erfüllt das lineare Modell von Mack die Annahmen des erweiterten linearen Modells. Des Weiteren gilt $\Sigma_{12} = \Sigma_{21}^T = O$.

Aufgrund der Definitionen von Σ_{11}, A_1, A_2 folgen für den Gauss-Markov Schätzer $\beta^\star$ von β,

$$\beta^\star = \begin{pmatrix} \zeta_1^\star \\ \zeta_2^\star \\ \vdots \\ \zeta_n^\star \end{pmatrix} \quad ,$$

wobei

$$\zeta_j^\star = \frac{\sum_{i=1}^{n+1-j} T_{i,j}}{\sum_{i=1}^{n+1-j} \upsilon_i}$$

gilt. Aufgrund der Kovarianzannahme der Inkremente gilt

$$Cov(\zeta_j^\star, \zeta_k^\star) = \frac{\sigma_j^2}{\sum_{l=1}^{n+1-j} \upsilon_l} \delta_{j,k} \quad .$$

Satz 3.5 Erwartungstreue der Gauss-Markov Schätzer.
Für alle $j \in \{1, ..., n\}$ ist $\zeta_j^\star$ ein erwartungstreuer Schätzer für ζ_j, das heißt

$$\mathbb{E}[\zeta_j^\star] = \zeta_j$$

Beweis.
Sei $j \in \{1, ..., n\}$ beliebig. Dann gilt

$$\begin{aligned} \mathbb{E}[\zeta_j^\star] = \mathbb{E}[\frac{\sum_{i=1}^{n+1-j} T_{i,j}}{\sum_{i=1}^{n+1-j} \upsilon_i}] &= \frac{\mathbb{E}[\sum_{i=1}^{n+1-j} T_{i,j}]}{\sum_{i=1}^{n+1-j} \upsilon_i} \\ = \frac{\sum_{i=1}^{n+1-j} \upsilon_i \zeta_j}{\sum_{i=1}^{n+1-j} \upsilon_i} &= \zeta_j \end{aligned}$$

$\square$

Nun betrachten wir noch die Schätzer für die Varianz. Dieser ist [8] entnommen.

Satz 3.6 Erwartungstreuer Schätzer der Varianz.
Für alle $j \in \{1, .., n-1\}$ ist

$$\hat{\sigma_j}^2 := \frac{1}{n-j} \sum_{k=1}^{n+1-j} \upsilon_k (\frac{T_{k,j}}{\upsilon_k} - \zeta_j^\star)^2$$

ein erwartungstreuer Schätzer für σ_j^2.

Beweis.
Sei $j \in \{1, ..., n-1\}$ beliebig. Dann gilt

$$\begin{aligned}
\mathbb{E}[(n-j)\hat{\sigma_j}^2] &= \sum_{k=1}^{n+1-j} (\frac{1}{\upsilon_k}\mathbb{E}[T_{k,j}^2] - 2\mathbb{E}[T_{k,j}\zeta_j^\star] + \upsilon_k \mathbb{E}[(\zeta_j^\star)^2]) \\
&= \sum_{k=1}^{n+1-j} (\frac{1}{\upsilon_k}(\mathbb{E}[T_{k,j}]^2 + Var(T_{k,j})) - 2(\mathbb{E}[T_{k,j}\zeta_j^\star] - \mathbb{E}[T_{k,j}]\mathbb{E}[\zeta_j^\star])) \\
&- 2\sum_{k=1}^{n+1-j} \mathbb{E}[T_{k,j}]\mathbb{E}[\zeta_j^\star] + \sum_{k=1}^{n+1-j} (\upsilon_k (Var(\zeta_j^\star) + \mathbb{E}[\zeta_j^\star]^2)) \\
&= \sum_{k=1}^{n+1-j} (\frac{1}{\upsilon_k}(\upsilon_k^2 \zeta_j^2 + \upsilon_k \sigma_j^2) - 2(Cov(T_{k,j}, \zeta_j^\star) + \upsilon_k \zeta_j^2) + \upsilon_k (\frac{\sigma_j^2}{\sum_{l=1}^{n+1-j} \upsilon_l} + \zeta_j^2)) \\
&= \sum_{k=1}^{n+1-j} (\upsilon_k \zeta_j^2 + \sigma_j^2 - 2\upsilon_k \zeta_j^2 - 2\frac{\upsilon_k \sigma_j^2}{\sum_{l=1}^{n+1-j} \upsilon_l} + \upsilon_k \frac{\sigma_j^2}{\sum_{l=1}^{n+1-j} \upsilon_l} + \upsilon_k \zeta_j^2) \\
&= (n+2-j-1)\sigma_j^2 - \sigma_j^2 \frac{\sum_{k=1}^{n+1-j} \upsilon_k}{\sum_{k=1}^{n+1-j} \upsilon_k} = (n-j)\sigma_j^2
\end{aligned}$$

□

Nun brauchen wir noch einen Schätzer $\hat{\sigma_n^2}$ für σ_n^2. Dafür nehmen wir an, dass

$$ln(\hat{\sigma_k^2}) = \alpha_0 + k\alpha_1 \quad \forall k \in \{1, .., n-1\}$$

gilt. Mittels der Methode der kleinsten Quadrate können Schätzer $\hat{\alpha}_0, \hat{\alpha}_1$ von α_0, α_1 bestimmt werden und damit

$$\hat{\sigma}_n^2 := exp(\hat{\alpha}_0 + n\hat{\alpha}_1)$$

definieren.

Nun wollen wir die Gauss-Markov Prädiktoren bestimmen. Da für alle $i, j \in \{1, .., n\}$ mit $i + j > n + 1$ ein Einheitsvektor $e_{i,j} \in \mathbb{R}^{m_2}$ existiert mit

$$\begin{aligned}
T_{i,j} &= e_{i,j}^T \mathbf{X_2} \\
g_{i,j}^T &= e_{i,j}^T A_2 \quad ,
\end{aligned}$$

folgt aufgrund der oben angesprochenen Linearität des Gauss-Markov Prädiktors

$$
\begin{aligned}
Y^\star(T_{i,j}) &= Y^\star(e_{i,j}^T \mathbf{X_2}) = e_{i,j}^T Y^\star(\mathbf{X_2}) = e_{i,j}^T A_2 \beta^\star \\
&= g_{i,j}^T \beta^\star = \upsilon_i e_j \beta^\star = \upsilon_i \zeta_j^\star \quad .
\end{aligned}
$$

Damit gilt für den Schätzer der globalen Reserve

$$
Y^\star(\sum_{i=2}^{n} \sum_{j=n+2-i}^{n} T_{i,j}) = \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} \upsilon_i \zeta_j^\star
$$

und für die Varianz

$$
\begin{aligned}
&Var(Y^\star(\sum_{i=2}^{n} \sum_{j=n+2-i}^{n} T_{i,j}) - \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} T_{i,j}) = \\
&\sum_{i=2}^{n} \Big(\frac{1}{\sum_{j=1}^{n+1-i} \upsilon_j} + \frac{1}{\sum_{j=n-i+2}^{n} \upsilon_j} \Big) (\sum_{j=n-i+2}^{n} \upsilon_j)^2 \sigma_i^2 \quad .
\end{aligned}
$$

Abschließend erhalten wir für die Reserven der Schadenjahre $i \in \{2, ..., n\}$ folgende Schätzer

$$
Y^\star(\sum_{j=n+2-i}^{n} T_{i,j}) = \sum_{j=n+2-i}^{n} \upsilon_i \zeta_j^\star \quad ,
$$

deren Varianz durch

$$
Var(Y^\star(\sum_{j=n+2-i}^{n} T_{i,j}) - \sum_{j=n+2-i}^{n} T_{i,j}) = \sum_{j=n+2-i}^{n} \left(\frac{1}{\sum_{k=1}^{n+1-j} \upsilon_k} + \frac{1}{\upsilon_i} \right) \upsilon_i^2 \sigma_j^2
$$

gegeben ist.

3.2.2 Beispiele

Auch bei diesem Modell werden wir die Zahlungsdreiecke der realen Versicherungsbestände KFZ-Kasko und Rechtsschutz betrachten. Jedoch werden bei diesem Modell die inkrementellen Zahlungen zur Berechnung herangezogen.

3.2.2.1 KFZ-Kasko

Am Anfang dieses Modells betrachten wir das inkrementelle Zahlungsdreieck, welches folgende Form hat.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	2.434.458,98	1.212,06	2.594,12	2.610,44	1.032,34	505,50
2	12.685.757,26	2.727.135,16	67.889,88	17.352,06	42,72	110,63	
3	16.301.266,73	3.153.221,70	154.001,39	111.947,00	2.258,65		
4	16.084.447,76	2.626.624,89	157.987,68	51.105,60			
5	15.452.148,36	2.904.762,88	132.642,22				
6	15.463.013,72	2.967.467,47					
7	13.768.695,59						

Tabelle 3.1: Inkrementelles Dreieck KFZ-Kasko

Als Volumsmaß betrachten wir die verdiente Prämie v_i des Schadenjahres i. Dabei gilt

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{pmatrix} = \begin{pmatrix} 13.713.457,77 \\ 18.514.732,78 \\ 24.031.112,54 \\ 28.720.787,59 \\ 30.055.066,86 \\ 29.777.398,96 \\ 27.921.354,56 \end{pmatrix}$$

Bevor wir die Schätzer $\zeta^\star$ betrachten, betrachten wir noch $\frac{T_{i,j}}{v_i}$.

SJ \ AJ	1	2	3	4	5	6	7
1	0,722524	0,177523	0,000088	0,000189	0,000190	0,000075	0,000037
2	0,685171	0,147295	0,003667	0,000937	0,000002	0,000006	
3	0,678340	0,131214	0,006408	0,004658	0,000094		
4	0,560028	0,091454	0,005501	0,001779			
5	0,514128	0,096648	0,004413				
6	0,519287	0,099655					
7	0,493124						

Vergleichen wir diese mit den Schätzern

j	1	2	3	4	5	6	7
$\zeta_j^\star$	0,576978	0,116106	0,004466	0,002153	0,000087	0,000035	0,000037

sehen wir, dass die individuellen Faktoren sehr stark streuen und daher die Annahme des Modells nicht ganz erfüllt ist. Mit Hilfe dieser Schätzer erhalten wir das geschätze Abwicklungsquadrat, wobei auch hier die grau eingefärbten Werte die Schätzer sind.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	2.434.458,98	1.212,06	2.594,12	2.610,44	1.032,34	505,50
2	12.685.757,26	2.727.135,16	67.889,88	17.352,06	42,72	110,63	682,48
3	16.301.266,73	3.153.221,70	154.001,39	111.947,00	2.258,65	852,26	885,83
4	16.084.447,76	2.626.624,89	157.987,68	51.105,60	2.507,52	1.018,58	1.058,69
5	15.452.148,36	2.904.762,88	132.642,22	64.721,52	2.624,01	1.065,90	1.107,88
6	15.463.013,72	2.967.467,47	132.982,30	64.123,58	2.599,76	1.056,05	1.097,64
7	13.768.695,59	3.241.849,21	124.693,42	60.126,72	2.437,72	990,23	1.029,23

Tabelle 3.2: KFZ-Kasko Lineares Modell

Die Schätzer der Varianz haben folgende Werte

j	1	2	3	4	5	6	7
$\hat{\sigma}_j$	196.090,1337	2.2423,3902	99,03286621	78,37030318	0,140232706	0,037839579	0,000886627

Somit erhalten wir für die Reserve

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw $R_i^\star$
2	682,48	128,12	462,29	973,24	148,87
3	1.738,09	964,69	539,07	4.284,22	845,80
4	4.584,79	2.267,12	1.612,78	10.472,79	1.754,41
5	69.519,30	48.588,10	16.141,06	201.157,00	28.920,89
6	201.859,34	72.718,18	94.437,79	381.909,28	39.804,30
7	3.431.126,52	794.386,41	2.116.502,59	5.279.315,43	349.442,29
Gesamt	3.709.510,52	799.189,96	2.368.378,53	5.552.361,92	361.584,45

wobei $R_i^\star = \sum_{j=n+1-i}^{n} Y^\star(T_{i,j})$ und $R_i = \sum_{j=n+1-i}^{n} T_{i,j}$ gilt.
Dabei wurde zur Berechnung der Konfidenzintervalle angenommen, dass die Reserven der Log-Normalverteilung folgen, da ansonsten im 3. und 5. Schadenjahr die untere Grenze negativ werden würde.

Betrachten wir nun das tatsächlich beobachtete Quadrat

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	2.434.458,98	1.212,06	2.594,12	2.610,44	1.032,34	505,50
2	12.685.757,26	2.727.135,16	67.889,88	17.352,06	42,72	110,63	914,31
3	16.301.266,73	3.153.221,70	154.001,39	111.947,00	2.258,65	0	243,70
4	16.084.447,76	2.626.624,89	157.987,68	51.105,60	7.020,90	3.971,73	820,08
5	15.452.148,36	2.904.762,88	132.642,22	432,95	715,88	500,00	170,73
6	15.463.013,72	2.967.467,47	152.116,91	7.181,75	9.646,50	1.358,14	472,00
7	13.768.695,59	2.561.736,04	109.282,32	32.141,74	1.064,91	1.010,00	0

erhalten wir als Differenzen zwischen den tatsächlichen und den geschätzten Reserven folgende Werte.

SJ	R_i	$R_i^\star - R_i$
2	914,31	- 231,83
3	243,70	1.494,39
4	11.812,71	- 7.227,92
5	1.819,56	67.699,74
6	170.775,30	31.084,04
7	2.705.235,01	725.891,51
Gesamt	2.890.800,59	818.709,93

Hier können wir sehen, dass die Schadenjahre 5-7 überschätzt worden sind. Vor allem die Reserve für das 7. Schadenjahr ist deutlich höher als die später tatsächlich beobachtete. Weiter ist zu erwähnen, dass die tatsächliche Reserve für die Schadenjahre 2-4 nicht im angegebenen Konfidenzintervall liegt.

3.2.2.2 Rechtsschutz

Auch hier betrachten wir zunächst das inkrementelle Zahlungsdreieck. Dieses hat folgende Form.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	348.829,14	198.640,90	135.725,98	104.081,37	38.871,09	85.672,89
2	104.432,33	412.438,75	347.753,32	176.849,21	207.855,20	64.416,49	
3	184.521,83	589.417,16	351.963,73	180.781,62	169.456,44		
4	256.817,88	751.470,36	465.401,63	297.336,57			
5	363.636,21	752.801,98	507.257,63				
6	419.610,32	1.014.104,85					
7	448.107,48						

Tabelle 3.3: Inkrementelles Dreieck Rechtsschutz

So wie bei der Sparte KFZ-Kasko wird auch hier als Volumsmaß die verdiente Prämie υ_i für das i-te Schadenjahr herangezogen.

$$\begin{pmatrix} \upsilon_1 \\ \upsilon_2 \\ \upsilon_3 \\ \upsilon_4 \\ \upsilon_5 \\ \upsilon_6 \\ \upsilon_7 \end{pmatrix} = \begin{pmatrix} 1.490.954,27 \\ 2.111.251,05 \\ 2.908.105,48 \\ 3.890.261,79 \\ 4.784.506,22 \\ 5.723.593,69 \\ 6.722.810,89 \end{pmatrix}$$

Betrachten wir bei diesem Beispiel die Werte $\frac{T_{i,j}}{\upsilon_i}$ um die Annahme des Modells zu überprüfen erhalten wir

SJ \ AJ	1	2	3	4	5	6	7
1	0,064694	0,233964	0,133231	0,091033	0,069809	0,026071	0,057462
2	0,049465	0,195353	0,164714	0,083765	0,098451	0,030511	
3	0,063451	0,202681	0,121029	0,062165	0,058270		
4	0,066016	0,193167	0,119632	0,076431			
5	0,076003	0,157342	0,106021				
6	0,073312	0,177180					
7	0,066655						

Auch hier ist eine Streuung der individuellen Faktoren zu erkennen. Vor allem im zweiten und dritten Abwicklungsjahr ist diese Streuung sehr stark. Daher ist die Annahme des Modells nicht ganz erfüllt.
Als Schätzer $\zeta_j^\star$ erhalten wir

j	1	2	3	4	5	6	7
$\zeta_j^\star$	0,067806042	0,18504581	0,12321419	0,076024024	0,073943169	0,02867343	0,057461782

Mit deren Hilfe erhalten wir das geschätzte Dreieck welches folgende Gestalt hat. Dabei sind auch hier die geschätzten Wert grau eingefärbt.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	348.829,14	198.640,90	135.725,98	104.081,37	38.871,09	85.672,89
2	104.432,33	412.438,75	347.753,32	176.849,21	207.855,20	64.416,49	121.316,25
3	184.521,83	589.417,16	351.963,73	180.781,62	169.456,44	83.385,36	167.104,92
4	256.817,88	751.470,36	465.401,63	297.336,57	287.658,28	111.547,15	223.541,38
5	363.636,21	752.801,98	507.257,63	363.737,42	353.781,55	137.188,20	274.926,26
6	419.610,32	1.014.104,85	705.227,97	435.130,62	423.220,65	164.115,06	328.887,90
7	448.107,48	1.244.027,99	828.345,71	511.095,14	497.105,94	192.766,04	386.304,70

Tabelle 3.4: Rechtsschutz Lineares Modell

Bevor wir nun die Konfidenzintervalle der Reserven berechnen können benötigen wir noch die Schätzer $\hat{\sigma}_j$. Diese haben folgende Werte.

j	1	2	3	4	5	6	7
σ_j	216,04	1.795,88	1.315,97	340,54	1.003,97	17,22	78,71

Mit deren Hilfe erhalten wir für die Standardabweichung der Reserve und für die Konfidenzintervalle folgende Werte.

SJ	$R_i^\star$	Stadabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw. $R_i^\star$
2	121.316,25	12.890,81	95.534,62	147.097,87	15.339,74
3	250.490,28	16.702,81	217.084,66	283.895,91	22.065,67
4	622.746,81	65.413,28	491.920,25	753.573,37	56.614,21
5	1.129.633,42	83.016,75	963.599,92	1.295.666,93	74.816,93
6	2.056.582,20	125.604,65	1.805.372,90	2.307.791,50	104.161,23
7	3.659.645,52	174.940,44	3.309.764,64	4.009.526,40	137.296,82
Gesamt	7.840.414,48	240.824,67	7.358.765,14	8.322.063,82	366.956,45

Hier wurde zur Berechnung der Konfidenzintervalle genauso wie im Chain-Ladder Modell angenommen, dass die Reserven normalverteilt sind.
Zum Abschluss wollen wir auch hier die geschätzten Werte mit den tatsächlich beobachteten Werten vergleichen. Hierfür benötigen wir zunächst das tatsächlich beobachtete Zahlungsquadrat.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	348.829,14	198.640,90	135.725,98	104.081,37	38.871,09	85.672,89
2	104.432,33	412.438,75	347.753,32	176.849,21	207.855,20	64.416,49	45.182,65
3	184.521,83	589.417,16	351.963,73	180.781,62	169.456,44	63.839,53	88.391,13
4	256.817,88	751.470,36	465.401,63	297.336,57	161.485,64	129.311,04	153.340,22
5	363.636,21	752.801,98	507.257,63	305.205,17	322.461,17	315.936,96	292.307,79
6	419.610,32	1.014.104,85	657.704,80	620.217,00	505.190,39	404.771,32	201.365,22
7	448.107,48	1.145.770,40	751.002,31	720.204,60	536.007,46	359.861,14	155.702,58

Somit erhalten wir folgende Differenzen zwischen den tatsächlichen und beobachteten Reserven.

SJ	R_i	$R_i^\star - R_i$
2	45.182,65	76.133,60
3	152.230,66	98.259,62
4	444.136,90	178.609,91
5	1.235.911,09	- 106.277,67
6	2.389.248,73	- 332.666,53
7	3.668.548,49	- 8.902,97
Gesamt	7.935.258,52	- 94.844,04

Hier ist gut zu erkennen, dass die Schadenjahre 2-4 deutlich überschätzt werden und die Schadenjahre 5 und 6 unterschätzt werden. Außerdem ist zu erwähnen, dass bis auf die Schadenjahre 5 und 7 die tatsächlichen Reserven nicht im angegebenen Konfidenzintervall liegen. Somit ist hier die Annahme der Normalverteilung zu verwerfen.

Kapitel 4

Das lognormale loglineare Modell

4.1 Grundlagen

In diesem Modell wird angenommen, dass der Logarithmus eines Zufallsvektors normalverteilt ist und den Annahmen des linearen Modells genügt. Dabei werden wir wie im linearen Modell zunächst das elementare lognormale loglineare Modell betrachten und danach das erweiterte Modell. Auch hier werden wir zeigen, dass die ersten zwei Momente von Linearkombinationen der nicht beobachtbaren Koordinaten erwartungstreu geschätzt werden können, und mit deren Hilfe Konfidenzintervalle der Reserven bestimmen. Dabei werden wir uns auch in diesem Kapitel an Radtke und Schmidt [4] orientieren. Auch hier wurden die Beweise selbst geführt, da sie schon wie im linearen Modell nicht in [4] vorkommen wurden.

4.1.1 Das elementare lognormale loglineare Modell

Sei nun $\mathbf{Z}$ ein Zufallsvektor mit Werten in $(0, \infty)^m$. Im elementaren lognormalen loglinearen Modell werden folgende Annahmen getroffen.

Annahme 4.1 Elementares lognormales loglineares Modell.
Es gibt einen Parameter $\beta \in \mathbb{R}^s$ und eine Matrix $A \in \mathbb{R}^{m \times s}$ mit $rang(A) = s$ derart, dass

$$\mathbf{Z} \sim LN(A\beta, \Sigma)$$

gilt. Dabei steht LN für die logarithmische Normalverteilung.

Wir setzen in diesem Abschnitt voraus, dass die Annahmen des Modells erfüllt sind und dass die Matrix Σ bekannt ist. Durch diese Annahmen erhalten wir für $ln(\mathbf{Z})$ ein elementares lineare Modell. Der Gauss-Markov Schätzer ist daher durch

$$\beta^\star = (A^T \Sigma A)^{-1} A^T \Sigma^{-1} ln(\mathbf{Z})$$

gegeben und da

$$ln(\mathbf{Z}) \sim N(A\beta, \Sigma)$$

gilt, können wir die Verteilung von $\beta^\star$ bestimmen.

Lemma 4.1.
Die Verteilung von $\beta^\star$ ist durch

$$\beta^\star \sim N(\beta, (A^T\Sigma^{-1}A)^{-1})$$

gegeben.

Beweis.
Aus der Wahrscheinlichkeitstheorie wissen wir, dass für $Y := AX$, wobei $X \sim N(\mu, \Sigma)$,

$$Y \sim N(A\mu, A\Sigma A^T)$$

gilt. Daher gilt

$$(A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1}\Sigma((A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1})^T = A^{-1}\Sigma(A^T)^{-1} = (A^T\Sigma^{-1}A)^{-1}$$

und

$$(A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1}A\beta = A^{-1}\Sigma(A^T)^{-1}A^T\Sigma^{-1}A\beta = \beta$$

Daraus folgt

$$\beta^\star \sim N(\beta, (A^T\Sigma^{-1}A)^{-1})$$

und somit ist alles bewiesen. □

Nun wollen wir die Momente des Vektors **Z** schätzen. Dazu sei $c \in \mathbb{R}^m$ und $C \in \mathbb{R}^{m\times m}$ und eine Transformation von β durch

$$h(\beta, c, C) := exp(c^T A\beta + \frac{1}{2}spur(\Sigma C))$$

definiert.

Satz 4.1 Erwartungstreuer Schätzer der Transformation.
Durch die Funktion

$$H(\beta, c, C) := exp(c^T A\beta^\star + \frac{1}{2}spur(\Sigma C) - \frac{1}{2}c^T A(A^T\Sigma^{-1}A)^{-1}A^T c)$$

ist ein erwartungstreuer Schätzer für $h(\beta, c, C)$ gegeben. Außerdem ist $H(\beta, c_1, C_1)H(\beta, c_2, C_2) - H(\beta, c_1 + c_2, C_1 + C_2)$ ein erwartungstreuer Schätzer für die Kovarianz von $H(\beta, c_1, C_1)$ und $H(\beta, c_2, C_2)$.

Beweis.
Zunächst beweisen wir die Erwartungstreue des Schätzers für die Funktion $h(\beta, c, C)$.Da

$$\mathbb{E}[exp(c^T A\beta^\star)] = exp(c^T A\beta + \frac{1}{2}c^T A(A^T\Sigma^{-1}A)^{-1}A^T c)$$

gilt, folgt

$$\mathbb{E}[H(\beta, c, C)] = h(\beta, c, C) \quad .$$

Wenden wir uns nun dem erwartungstreuen Schätzer der Kovarianz zu. Hier gilt

$$\begin{aligned}
Cov(H(\beta, c_1, C_1), H(\beta, c_2, C_2)) &= \mathbb{E}[H(\beta, c_1, C_1)H(\beta, c_2, C_2)] \\
&- \mathbb{E}[H(\beta, c_1, C_1)]\mathbb{E}[H(\beta, c_1, C_2)] \\
&= \mathbb{E}[H(\beta, c_1, C_1)H(\beta, c_2, C_2)] - h(\beta, c_1, C_1)h(\beta, c_2, C_2) \\
&= \mathbb{E}[H(\beta, c_1, C_1)H(\beta, c_2, C_2)] \\
&- exp((c_1 + c_2)^T A\beta + \frac{1}{2}spur(\Sigma(C_1 + C_2))) \\
&= \mathbb{E}[H(\beta, c_1, C_1)H(\beta, c_2, C_2)] - \mathbb{E}[H(\beta, c_1 + c_2, C_1 + C_2)] \\
&= \mathbb{E}[H(\beta, c_1, C_1)H(\beta, c_2, C_2) - H(\beta, c_1 + c_2, C_1 + C_2)]
\end{aligned}$$

Somit ist die Behauptung gezeigt. □

Definieren wir nun für alle $i \in \{1, .., m\}$

$$E_i := e_i e_i^T \quad ,$$

dann erhalten wir

$$\mathbb{E}[\mathbf{Z_i}] = exp(e_i^T A\beta + \frac{1}{2}e_i^T \Sigma e_i) = h(\beta, e_i, E_i) = \mathbb{E}[H(\beta, e_i, E_i)] \quad .$$

Also ist $H(\beta, e_i, E_i)$ ein erwartungstreuer Schätzer für $\mathbb{E}[\mathbf{Z_i}]$. Für einen Vektor $d \in \mathbb{R}^m$ gilt dann

$$\mathbb{E}[d^Z \mathbf{Z}] = \mathbb{E}[\sum_{i=1}^{m} d_i \mathbf{Z_i}] = \mathbb{E}[\sum_{i=1}^{m} d_i H(\beta, e_i, E_i)] \quad .$$

Weiter gilt aufgrund der Erwartungstreue des Schätzers der Kovarianzen

$$\begin{aligned}
Var(\sum_{i=1}^{n} d_i H(\beta, e_i, E_i)) &= \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(d_i H(\beta, e_i, E_i), d_j H(\beta, e_j, E_j)) \\
&= \mathbb{E}[\sum_{i=1}^{n} \sum_{j=1}^{n} d_i d_j H(\beta, e_i, E_i)H(\beta, e_j, E_j)] \\
&- \mathbb{E}[\sum_{i=1}^{n} \sum_{j=1}^{n} H(\beta, e_i + e_j, E_i + E_j)] \quad .
\end{aligned}$$

Somit können wir den Erwartungswert einer Linearkombination von Koordinaten von $\mathbf{Z}$ erwartungstreu schätzen und Konfidenzintervalle für diesen Schätzer angeben.

4.1.2 Das erweiterte lognormale loglineare Modell

Hier betrachten wir einen Zufallsvektor $\mathbf{Z}$ mit Werten in $(0, \infty)^m$ von dem die ersten m_1 Koordinaten beobachtbar und die restlichen $m_2 := m - m_1$ Koordinaten nicht beobachtbar sind. Wir können $\mathbf{Z}$ daher wieder als

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z_1} \\ \mathbf{Z_2} \end{pmatrix}$$

schreiben, wobei wir annehmen, dass $\mathbf{Z_1}$ beobachtbar ist und $\mathbf{Z_2}$ nicht.
Definieren wir nun für $i, j \in \{1, 2\}$

$$\Sigma_{i,j} := Cov(log(\mathbf{Z_i}), log(\mathbf{Z_j}))$$

und

$$\Sigma := \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

dann gilt

$$Var(log(\mathbf{Z})) = \Sigma \quad .$$

Nun betrachten wir die Annahmen in diesem Abschnitt

Annahme 4.2 Erweitertes lognormales loglineares Modell.
Es gibt einen Parameter $\beta \in \mathbb{R}^s$ und eine Matrix

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$$

mit $A_1 \in \mathbb{R}^{m_1 \times s}$ und $A_2 \in \mathbb{R}^{m_2 \times s}$ und $rang(A_1) = s$, sodass

$$\mathbf{Z} \sim LN(A\beta, \Sigma)$$

gilt.

Wir setzen in diesem Abschnitt voraus, dass die Annahmen des erweiterten lognormalen loglinearen Modells erfüllt sind und die Matrix Σ_{11} bekannt ist. Da sich aus dem erweiterten lognormalen loglinearen Modell für $\mathbf{Z}$ ein erweitertes lineares Modell für $log(\mathbf{Z})$ ergibt, erhält man den Gauss-Markov Schätzer $\beta^\star$ für β durch

$$\beta^\star = (A_1^T \Sigma_{11}^{-1} A_1)^{-1} A_1^T \Sigma_{11}^{-1} log(\mathbf{Z_1}) \quad .$$

Da weiter

$$log(\mathbf{Z_1}) \sim N(A_1\beta, \Sigma_{11})$$

gilt, folgt

$$\beta^\star \sim N(\beta, (A_1^T \Sigma_{11}^{-1} A_1)^{-1}) \quad .$$

Dadurch erhalten wir, dass für $c \in \mathbb{R}^m$ und $C \in \mathbb{R}^{m \times m}$ die Funktion

$$H(\beta, c, C) := exp(c^T A \beta^\star + \frac{1}{2} spur(\Sigma C) - \frac{1}{2} c^T A (A_1^T \Sigma_{11}^{-1} A_1)^{-1} A^T c)$$

ein erwartungstreuer Schätzer für

$$h(\beta, c, C) := exp(c^T A\beta + \frac{1}{2} spur(\Sigma C))$$

ist. Mit Hilfe der Funktion $H(\beta, c, C)$ kann man genauso wie im elementaren lognormalen loglinearen Modell erwartungstreue Schätzer für Linearkombinationen von Erwartungswerten der Koordinaten von $\mathbf{Z}$ und erwartungstreue Schätzer für die Varianz dieser Schätzer konstruieren. Von besonderem Interesse sind dabei Schätzer des nicht beobachtbaren Vektors $\mathbf{Z_2}$.

4.2 Das lognormale loglineare Modell in der Schadenreservierung

Dieses Modell basiert auf Zuwächsen. Dabei wird wieder angenommen, dass die Zuwächse $T_{i,j}$ für $i + j \leq n + 1$ beobachtbar sind und für $i + j > n + 1$ nicht. Außerdem wird vorausgesetzt, dass für alle $i, j \in \{1, .., n\}$

$$\mathbb{P}(T_{i,j} \in (0, \infty)) = 1$$

gilt. Diese Annahme ist erforderlich, da sonst die Logarithmen der Zuwächse nicht gebildet werden können.
Diese Annahme ist nicht in jedem Abwicklungsdreieck erfüllt. In den meisten Abwicklungsdreiecken für Schadenaufwände ist diese Annahme verletzt, da in den ersten Abwicklungsjahren eine Überreservierung stattfindet. Daher ist das Modell auf diese Abwicklungsdreiecke nicht anwendbar.

Sei nun $m := n^2, m_1 := \frac{n(n+1)}{2}$ und $m_2 := \frac{(n-1)n}{2}$. Genauso wie im linearen Modell erhält man durch die Zuordnung

$$\mathbf{Z_{h(i,j)}} := T_{i,j}$$

mit

$$h(i,j) = \begin{cases} j - (i-1) + \frac{(2n+3-(i-1))(i-1)}{2} & \text{falls } i + j \leq n + 1 \\ ((i+j) - (n+2)) + \frac{n(n+1)+(i-2)^2+i}{2} & \text{falls } i + j > n + 1 \end{cases}$$

eine Anordnung der Zuwächse $T_{i,j}$ zu einem Zufallsvektor

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z_1} \\ \mathbf{Z_2} \end{pmatrix}$$

bei dem die Koordinaten von $\mathbf{Z_1}$ aus den m_1 beobachtbaren Zuwächsen bestehen und die von $\mathbf{Z_2}$ aus den m_2 nicht beobachtbaren. Weiter definieren wir für $i, j \in \{1, 2\}$

$$\Sigma_{i,j} := Cov(log(\mathbf{Z_i}), log(\mathbf{Z_j}))$$

und

$$\Sigma := \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Um nun zu einem lognormalen loglinearen Modell zu gelangen, nehmen wir an, dass

$$\mathbf{Z} \sim LN(\mathbb{E}[log(\mathbf{Z})], \Sigma)$$

gilt und dass ein Parameter $\beta \in \mathbb{R}^s$ mit $s \leq m$ und für alle $i, j \in \{1, .., n\}$ ein Vektor $g_{i,j} \in \mathbb{R}^s$ existiert mit

$$\mathbb{E}[log(T_{i,j})] = g_{i,j}^T \beta \quad .$$

Durch die Zuordnung

$$a_{h(i,j)} := g_{i,j}^T$$

erhalten wir nun eine Matrix

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$$

für die

$$\mathbb{E}[log(\mathbf{Z})] = A\beta$$

gilt. Wenn außerdem Σ_{11} positiv definit ist und $rang(A_1) = s$ gilt, sind die Annahmen des erweiterten lognormalen loglinearen Modells erfüllt.

4.2.1 Das lognormale logadditive Modell für Zuwächse

Dieses Modell besteht aus folgenden Annahmen.

Annahme 4.3 Das lognormale logadditive Modell für Zuwächse.

- Die Verteilung des Zufallsvektors $\mathbf{Z}$ ist durch

  $$\mathbf{Z} \sim LN(\mathbb{E}[log(\mathbf{Z})], \Sigma)$$

  gegeben.

- Es existieren Parameter $\mu \in (0, \infty), \zeta_1, .., \zeta_{n-1} \in \mathbb{R}$ und $\xi_1, .., \xi_{n-1} \in \mathbb{R}$ so, dass für alle $i, j \in \{1, .., n\}$

  $$\mathbb{E}[log(T_{i,j})] = \mu + (1 - \delta_{1,i})\xi_{i-1} + (1 - \delta_{1,j})\zeta_{j-1}$$

  gilt.

Wir setzen voraus, dass diese Annahmen erfüllt sind und dass die Matrix Σ_{11} bekannt ist. Definieren wir nun

$$\beta := \begin{pmatrix} \mu \\ \xi_1 \\ \vdots \\ \xi_{n-1} \\ \zeta_1 \\ \vdots \\ \zeta_{n-1} \end{pmatrix}$$

und

$$g_{i,j} := e_1 + (1 - \delta_{1,i})e_i + (1 - \delta_{1,j})e_{n-1+j}$$

erhalten wir

$$\mathbb{E}[log(T_{i,j})] = g_{i,j}^T\beta \quad .$$

Für alle $i, j \in \{1, .., n\}$ mit $i + j > n + 1$ existiert ein Einheitsvektor $e_{i,j} \in \mathbb{R}^m$ mit

$$T_{i,j} = e_{i,j}^T \mathbf{Z}$$

und

$$g_{i,j} = e_{i,j}^T A \quad .$$

Somit erhalten wir

$$\mathbb{E}[T_{i,j}] = exp(g_{i,j}^T \beta + \frac{1}{2} e_{i,j}^T \Sigma e_{i,j}) = \mathbb{E}[exp(g_{i,j}^T \beta^\star + \frac{1}{2} e_{i,j}^T \Sigma e_{i,j} - \frac{1}{2} g_{i,j}^T (A_1^T \Sigma_{11}^{-1} A_1)^{-1} g_{i,j})] \quad ,$$

wobei $\beta^\star$ der Gauss-Markov Schätzer von β ist. Daher erhalten wir einen erwartungstreuen Schätzer für den Erwartungswert des nicht beobachtbaren Zuwachses $T_{i,j}$. Damit können wir den erwartungstreuen Schätzer des Erwartungswertes einer Linearkombination von nicht beobachtbaren Zuwächsen bilden und einen Schätzer für die Varianz dieses Schätzers. Nun betrachten wir noch einen Spezialfall der in der Schadenreservierung von großer Bedeutung ist.

Lemma 4.2. Falls für die Kovarianzmatrix $\Sigma_{12} = O$ gilt, folgt für die Varianz

$$Var(\sum_{i=m_1+1}^{m_2} d_i H(\beta, e_i, E_i) - \sum_{i=m_1+1}^{m_2} d_i Z_i) = Var(\sum_{i=m_1+1}^{m_2} d_i H(\beta, e_i, E_i)) + Var(\sum_{i=m_1+1}^{m_2} d_i Z_i)$$

wobei $d_i \in \{0, 1\}$ ist und Z_i die i-te Koordinate des Vektors $\mathbf{Z}$ bezeichnet.

Beweis.
Zunächst gilt

$$Var(\sum_{i=m_1+1}^{m_2} d_i H(\beta, e_i, E_i) - \sum_{i=m_1+1}^{m_2} d_i Z_i) = Var(\sum_{i=m_1+1}^{m_2} d_i H(\beta, e_i, E_i)) + Var(\sum_{i=m_1+1}^{m_2} d_i Z_i)$$
$$- 2 * Cov(\sum_{i=m_1+1}^{m_2} d_i H(\beta, e_i, E_i), \sum_{i=m_1+1}^{m_2} d_i Z_i)$$

und da die Kovarianz linear ist, reicht es

$$Cov(H(\beta, e_i, E_i), Z_j) = 0 \text{ für } i, j \in \{m_1 + 1, ..., m_2\}$$

zu zeigen.
Da für zwei unabhängige Zufallsvariablen X, Y und messbare Funktionen $f, g : \mathbb{R} \to \mathbb{R}$ auch $f(X), g(Y)$ unabhängig sind und bei Normalverteilung aus Unkorreliertheit bereits Unahbhängigkeit folgt erhalten wir

$$Cov(H(\beta, e_i, E_i), Z_i) = Cov(exp(\sum_{k=1}^{m_1} b_{k,i} log(Z_k) + c_i), exp(log(Z_j))) = 0$$

wobei $b_{k,i}, c_i \in \mathbb{R}$. $\square$

4.2.2 Beispiele

Auch in diesem Modell betrachten wir für die Sparten KFZ-Kasko und Rechtsschutz die inkrementellen Zahlungsdreiecke. Bevor wir aber die konkreten Sparten betrachten, kommen wir noch zu den Werten und Matrizen, die für beide Sparten gleich sind.
Da wir für beide Sparten $n = 7$ Schadenjahre betrachten, gilt $m = 49, m_1 = 28, m_2 = 21$ und $s = 13$. Somit erhalten wir

$$A_1 = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\end{pmatrix}$$

beziehungsweise

$$A_2 = \begin{pmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

Desweiteren wird die Annahme getroffen, dass

$$Cov(ln(\mathbf{Z_i}), ln(\mathbf{Z_j})) = 0$$

gilt für $i, j \in \{1, .., m\}, i \neq j$.

4.2.2.1 KFZ-Kasko

Genauso wie beim linearen Modell ist der Ausgangspunkt das inkrementelle Zahlungsdreieck

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	2.434.458,98	1.212,06	2.594,12	2.610,44	1.032,34	505,50
2	12.685.757,26	2.727.135,16	67.889,88	17.352,06	42,72	110,63	
3	16.301.266,73	3.153.221,70	154.001,39	111.947,00	2.258,65		
4	16.084.447,76	2.626.624,89	157.987,68	51.105,60			
5	15.452.148,36	2.904.762,88	132.642,22				
6	15.463.013,72	2.967.467,47					
7	13.768.695,59						

Aus den Daten der Einzelschäden erhalten wir für die Gesamtschadenzahlungen folgende Parameter

$\sigma_{i,j}$	1	2	3	4	5	6	7
1	0,00026654	0,00151731	6,33122560	3,35759558	1,37477122	1,27066449	0,81748229
2	0,00024095	0,00159811	0,16678285	0,70188528	8,68007637	6,95742763	
3	0,00015332	0,00116524	0,03981319	0,06852133	1,81566048		
4	0,00014792	0,00129015	0,05272335	0,05181053			
5	0,00015920	0,00142288	0,01529713				
6	0,00014072	0,00202793					
7	0,00015750						

wobei $\sigma_{i,j}$ den Parameter der Schadenzahlung $T_{i,j}$ bezeichnet. Mit der oben getroffenen Annahme bezüglich der Kovarianz erhalten wir die Varianz-Kovarianz-Matrix Σ_{11} der beobachtbaren logarithmierten Schadenstände.
Die Parameter der unbekannten Schadenzahlungen erhalten wir, indem der Mittelwert des jeweiligen Abwicklungsjahres gebildet wird. Dadurch ergibt sich folgendes Bild

$\sigma_{i,j}$	1	2	3	4	5	6	7
1							
2							0,81748229
3						4,11404606	0,81748229
4					3,956836023	4,11404606	0,81748229
5				1,044953178	3,956836023	4,11404606	0,81748229
6			1,321168425	1,044953178	3,956836023	4,11404606	0,81748229
7		0,00150360	1,32116843	1,04495318	3,95683602	4,11404606	0,81748229

und wir erhalten wieder mit der Annahme über die Kovarianz der logarithmierten Schadenstände die Varianz-Kovarianz-Matrix Σ_{22} der logarithmierten, nicht beobachtbaren Schadenstände.

Somit sind nun alle Parameter gegeben, die für die Berechnung von $\beta^\star$ notwendig sind un wir erhalten

$$\beta^\star = \begin{pmatrix} 16,14231279 \\ 0,224525665 \\ 0,464037233 \\ 0,431863563 \\ 0,406510314 \\ 0,410209021 \\ 0,295595351 \\ -1,628554677 \\ -4,740101312 \\ -5,48064667 \\ -8,87428206 \\ -9,582305738 \\ -9,916764748 \end{pmatrix}$$

Damit ergibt sich mit den gegebenen Daten das nachstehende geschätze Zahlungsdreieck, wobei die geschätzten Werte wieder grau eingefärbt sind.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	2.434.458,98	1.212,06	2.594,12	2.610,44	1.032,34	505,50
2	12.685.757,26	2.727.135,16	67.889,88	17.352,06	42,72	110,63	632,61
3	16.301.266,73	3.153.221,70	154.001,39	111.947,00	2.258,65	5.134,30	803,84
4	16.084.447,76	2.626.624,89	157.987,68	51.105,60	11.151,84	4.971,74	778,39
5	15.452.148,36	2.904.762,88	132.642,22	106.565,27	10.872,61	4.847,25	758,90
6	15.463.013,72	2.967.467,47	260.038,68	106.960,69	10.912,95	4.865,24	761,72
7	13.768.695,59	2.703.045,28	231.875,94	95.376,61	9.731,05	4.338,32	679,22

Tabelle 4.1: KFZ-Kasko Loglineares Modell

Daher liefert uns das Modell die folgenden Werte für die Reserven

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw $R_i^\star$
2	632,61	711,45	68,91	2.564,22	472,83
3	5.938,14	39.844,54	17,48	43.836,37	4.210,15
4	16.901,98	88.698,72	81,31	123.111,68	8.964,01
5	123.044,04	168.555,07	9.285,26	566.830,91	19.839,21
6	383.539,28	463.064,20	36.718,85	1.630.065,45	31.643,17
7	3.045.046,42	426.018,41	2.282.720,75	3.983.974,05	65.549,26
Gesamt	3.575.102,47	658.624,71	2.439.822,17	5.066.684,91	98.165,36

wobei auch hier $R_i = \sum_{j=n+1-i}^{n} T_{i,j}$ gilt und $R_i^\star$ die Summe der jeweiligen Schätzer ist.

Auch in diesem Modell wollen wir die Schätzer mit den tatsächlich beobachteten Werten vergleichen, welche folgende Form haben.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	2.434.458,98	1.212,06	2.594,12	2.610,44	1.032,34	505,50
2	12.685.757,26	2.727.135,16	67.889,88	17.352,06	42,72	110,63	914,31
3	16.301.266,73	3.153.221,70	154.001,39	111.947,00	2.258,65	0	243,70
4	16.084.447,76	2.626.624,89	157.987,68	51.105,60	7.020,90	3.971,73	820,08
5	15.452.148,36	2.904.762,88	132.642,22	432,95	715,88	500,00	170,73
6	15.463.013,72	2.967.467,47	152.116,91	7.181,75	9.646,50	1.358,14	472,00
7	13.768.695,59	2.561.736,04	109.282,32	32.141,74	1.064,91	1.010,00	0

Somit erhalten wir im lognormalen loglinearen Modell als Differenzen zwischen der tatsächlich benötigten Reserve und den Schätzern

SJ	R_i	$R_i^\star - R_i$
2	914,31	- 281,70
3	243,70	5.694,44
4	11.812,71	5.089,27
5	1.819,56	121.224,48
6	170.775,30	212.763,98
7	2.705.235,01	339.811,41
Gesamt	2.890.800,59	684.301,88

Hier ist zu erwähnen, dass die Reserven der ersten drei Schadenjahre sehr gut geschätzt werden, die Reserven der letzten drei Schadenjahre jedoch deutlich überschätzt werden. Gut zu erkennen ist, dass mit Ausnahme des 4. Schadenjahres alle tatsächlichen Reserven im angegebenen Konfidenzintervall liegen.

4.2.2.2 Rechtsschutz

Zunächst betrachten wir das inkrementelle Zahlungsdreieck.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	348.829,14	198.640,90	135.725,98	104.081,37	38.871,09	85.672,89
2	104.432,33	412.438,75	347.753,32	176.849,21	207.855,20	64.416,49	
3	184.521,83	589.417,16	351.963,73	180.781,62	169.456,44		
4	256.817,88	751.470,36	465.401,63	297.336,57			
5	363.636,21	752.801,98	507.257,63				
6	419.610,32	1.014.104,85					
7	448.107,48						

Auch hier erhalten wir aus den Einzelschadendaten die geschätzten Parameter für die Verteilung.

$\sigma_{i,j}$	1	2	3	4	5	6	8
1	0,005887705	0,004420324	0,013582138	0,07684008	0,139389663	0,118784826	0,179381666
2	0,005731153	0,004250075	0,009242454	0,051347268	0,09695653	0,266794889	
3	0,003461622	0,002091881	0,008950169	0,02899561	0,036989306		
4	0,002939904	0,002262736	0,006241943	0,026734087			
5	0,002237127	0,002106584	0,006348592				
6	0,0025403	0,002444385					
7	0,002454249						

Wie schon bei KFZ-Kasko bezeichnet $\sigma_{i,j}$ die Varianz von $ln(T_{i,j})$. Die Varianz der unbekannten logarithmierten Schadenstände erhalten wir wieder, indem wir für jedes Abwicklungsjahr den Mittelwert bilden.

$\sigma_{i,j}$	1	2	3	4	5	6	7
1							
2							0,179381666
3						0,192789857	0,179381666
4					0,091111833	0,192789857	0,179381666
5				0,045979261	0,091111833	0,192789857	0,179381666
6			0,008873059	0,045979261	0,091111833	0,192789857	0,179381666
7		0,002929331	0,008873059	0,045979261	0,091111833	0,192789857	0,179381666

So können wir mit der oben getroffenen Annahme bezüglich der Kovarianz die Varianz-Kovarianz-Matrix der logarithmierten Schadenstände bilden. Nun ist es auch hier möglich $\beta^\star$ zu berechnen, wobei wir folgende Werte erhalten

$$\beta^\star = \begin{pmatrix} 11,62162425 \\ 0,232326358 \\ 0,572813004 \\ 0,857104107 \\ 1,011816615 \\ 1,254517812 \\ 1,391164143 \\ 1,021635986 \\ 0,614844903 \\ 0,082138546 \\ -0,014137428 \\ -0,969579503 \\ -0,263332532 \end{pmatrix}$$

und damit für die geschätzten Schadenzahlungen.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	348.829,14	198.640,90	135.725,98	104.081,37	38.871,09	85.672,89
2	104.432,33	412.438,75	347.753,32	176.849,21	207.855,20	64.416,49	107.872,03
3	184.521,83	589.417,16	351.963,73	180.781,62	169.456,44	79.145,73	151.683,76
4	256.817,88	751.470,36	465.401,63	297.336,57	267.835,20	105.172,83	201.564,95
5	363.636,21	752.801,98	507.257,63	338.789,16	312.662,12	122.775,64	235.301,35
6	419.610,32	1.014.104,85	724.925,46	431.770,23	398.473,66	156.471,10	299.880,49
7	448.107,48	1.244.398,01	830.437,57	494.609,28	456.463,55	179.244,72	343.526,79

Tabelle 4.2: Rechtsschutz Loglineares Modell

Für die Reserven ergeben sich dadurch folgende Werte.

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw $R_i^\star$
2	107.872,03	47.815,10	42.274,55	230.056,34	44.136,90
3	230.829,49	76.501,33	114.884,34	417.889,88	66.365,26
4	574.572,98	131.062,91	357.033,06	878.927,55	98.499,52
5	1.009.528,27	169.733,00	712.927,09	1.390.226,35	122.444,86
6	2.011.520,93	226.884,52	1.596.314,77	2.502.881,40	171.868,42
7	3.548.679,93	268.502,44	3.042.294,76	4.115.789,89	263.722,34
Gesamt	7.483.003,62	421.539,63	6.675.695,67	8.361.407,18	672.142,84

Betrachten wir in diesem Modell die tasächlich beobachteten Zahlungen

	1	2	3	4	5	6	7
1	96.455,48	348.829,14	198.640,90	135.725,98	104.081,37	38.871,09	85.672,89
2	104.432,33	412.438,75	347.753,32	176.849,21	207.855,20	64.416,49	45.182,65
3	184.521,83	589.417,16	351.963,73	180.781,62	169.456,44	63.839,53	88.391,13
4	256.817,88	751.470,36	465.401,63	297.336,57	161.485,64	129.311,04	153.340,22
5	363.636,21	752.801,98	507.257,63	305.205,17	322.461,17	315.936,96	292.307,79
6	419.610,32	1.014.104,85	657.704,80	620.217,00	505.190,39	404.771,32	201.365,22
7	448.107,48	1.145.770,40	751.002,31	720.204,60	536.007,46	359.861,14	155.702,58

erhalten wir folgende Differenzen zwischen den tatsächlich Reserven und deren Schätzern.

SJ	R_i	$R_i^\star - R_i$
2	45.182,65	62.689,38
3	152.230,66	78.598,83
4	444.136,90	130.436,08
5	1.235.911,09	- 226.382,82
6	2.389.248,73	- 377.727,80
7	3.668.548,49	- 119.868,56
Gesamt	7.935.258,52	- 452.254,90

Hier ist sehr gut zu erkennen, dass die Reserven der ersten drei Schadenjahre deutlich überschätzt werden, die Reserven der letzten drei Schadenjahre deutlich unterschätzt werden. Zu erwähnen ist, dass für alle Schadenjahre die tatsächliche Schadenreserve im angegebenen Konfidenzintervall liegt, jedoch manchmal sehr knapp an den Intervallgrenzen.

Kapitel 5

Stochastische Abwicklungsfaktoren

In diesem Kapitel werden wir nochmals das Chain-Ladder Verfahren betrachten, jedoch werden wir dieses Mal annehmen, dass die Abwicklungsfaktoren einer Verteilung folgen. Dazu werden wir als Verteilungen die Log-Nomralverteilung, die logarithmische Gammaverteilung und die loginverse Gauß-Verteilung betrachten. Dabei wird die wichtigste Eigenschaft dieser drei Verteilungen sein, dass sie bezüglich der Multiplikation invariant sind. Weiter werden wir Algorithmen sehen, mittels derer wir die Endschadenstände und somit die Reserven und weitere Kennzahlen schätzen können. Dabei werden wir uns in diesem Kapitel an Mary Kelly [5] orientieren, wobei wir die Endschadenstände in etwas abgewandelter Form berechnen werden.

Genauso wie im klassischen Chain-Ladder Verfahren betrachten wir hier ein Schadendreieck basierend auf kumulierten Schadenständen. Anhand dieses Dreiecks können wir, wie bereits im Chain-Ladder Verfahren erwähnt, individuelle Abwicklungsfaktoren bilden. Diese haben dann folgende Form.

SJ \ AJ	1	2	$\cdots$	n-i	$\cdots$	n-2	n-1
1	$F_{1,1}$	$F_{1,2}$	$\cdots$	$F_{1,n-i}$	$\cdots$	$F_{1,n-2}$	$F_{1,n-1}$
2	$F_{2,1}$	$F_{2,2}$	$\cdots$	$F_{2,n-i}$	$\cdots$	$F_{2,n-2}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\iddots$		
i	$F_{i,1}$	$F_{i,2}$	$\cdots$	$F_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$	$\iddots$				
n-2	$F_{n-2,1}$	$F_{n-2,2}$					
n-1	$F_{n-1,1}$						

Tabelle 5.1: Dreieck der individuellen Abwicklungsfaktoren

In diesen Modellen wird nun angenommen, dass sich die Endschadenstände wie folgt berechnen lassen.

SJ \ AJ	1	2	$\cdots$	n+1-i	$\cdots$	n-1	n
1	I_1	$I_1 D_{1,1}$	$\cdots$	$I_1 \prod_{k=1}^{n-i} D_{1,k}$	$\cdots$	$I_1 \prod_{k=1}^{n-2} D_{1,k}$	$I_1 \prod_{k=1}^{n-1} D_{1,k}$
2	I_2	$I_2 D_{2,1}$	$\cdots$	$I_2 \prod_{k=1}^{n-i} D_{2,k}$	$\cdots$	$I_2 \prod_{k=1}^{n-2} D_{2,k}$	$I_2 \prod_{k=1}^{n-1} D_{2,k}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	I_i	$I_i D_{i,1}$	$\cdots$	$I_i \prod_{k=1}^{n-i} D_{i,k}$	$\cdots$	$I_i \prod_{k=1}^{n-2} D_{i,k}$	$I_i \prod_{k=1}^{n-1} D_{i,k}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n-1	I_{n-1}	$I_{n-1} D_{n-1,1}$	$\cdots$	$I_{n-1} \prod_{k=1}^{n-i} D_{n-1,k}$	$\cdots$	$I_{n-1} \prod_{k=1}^{n-2} D_{n-1,k}$	$I_{n-1} \prod_{k=1}^{n-1} D_{n-1,k}$
n	I_n	$I_n D_{n,1}$	$\cdots$	$I_n \prod_{k=1}^{n-i} D_{n,k}$	$\cdots$	$I_n \prod_{k=1}^{n-2} D_{n,k}$	$I_n \prod_{k=1}^{n-1} D_{n,k}$

Tabelle 5.2: Abwicklungsquadrat bei stochastischen Faktoren

Definieren wir nun $S_{i,j} := I_i \prod_{k=1}^{j-1} D_{i,k}$ erhalten wir für das Schadenjahr i den Endschadenstand $U_i | S_{i,n+1-i} := S_{i,n} = I_i \prod_{k=1}^{n-1} D_{i,k}$ und den Gesamtendschadenstand durch $S_n := \sum_{i=1}^{n} U_i$. Dabei sind sowohl die I_i als auch die $D_{i,j}$ Zufallsvariablen. Wir werden als Vereinfachung annehmen, dass die Zufallsvariablen I_i und $D_{i,j}$ für $i + j \leq n$ deterministisch sind und durch das bereits beobachtbare Schadendreieck gegeben sind.
Weiter nehmen wir an, dass die $D_{i,j}$ für $i \in \{1, .., n\}$ unabhängig und identisch verteilt sind. Desweiteren wird angenommen, dass für festes $i \in \{1, .., n\}$ die $D_{i,j}$ und $D_{i,k}$ unabhängig sind für $j, k \in \{1, .., n\}, j \neq k$.

5.1 Das lognormale Modell

5.1.1 Grundlagen des Modells

Die Log-Normalverteilung findet häufig Anwendung bei natürlichen Phänomenen. Jedoch kann sie auch zur Berechnung von Aktienkursen und Endschadenständen verwendet werden.
Eine Zufallsvariable $Y \sim LN(\mu, \sigma^2)$ folgt einer Log-Normalverteilung, wenn ihre Dichtefunktion durch

$$g(y) = \frac{1}{(2\pi)^{\frac{1}{2}} \sigma y} exp\Big(\frac{-\big(ln(y) - \mu\big)^2}{2\sigma^2} \Big), y \geq 0$$

gegeben ist.
Eine Eigenschaft der Log-Normalverteilung ist, dass $Y = exp(X) \sim LN(\mu, \sigma^2)$ ist, wenn $X \sim N(\mu, \sigma)$.
Nun wollen wir zeigen, dass die Lognormalverteilung unter der Multiplikation abgeschlossen ist. Dafür seien zunächst $X_1, .., X_n$ unabhängige Zufallsvariablen mit $X_i \sim N(\mu_i, \sigma_i^2)$. Dann gilt $T_n := \sum_{k=1}^{n} (X_k + b_k) \sim N(\mu, \sigma)$, wobei $\mu = \sum_{k=1}^{n} (\mu_k + b_k)$ und $\sigma^2 = \sum_{k=1}^{n} \sigma_k^2$.
Sind nun $Y_1, .., Y_n$ unabhängig und logarithmisch normalverteilt, dann gilt für $\prod_{k=1}^{n} (\alpha_k Y_k)$ mit $\alpha_j \geq 0$

$$\prod_{k=1}^{n} (\alpha_k Y_k) = \prod_{k=1}^{n} exp(log(\alpha_k) + (X_k)) = exp(\sum_{k=1}^{n} (log(\alpha_k) + (X_k)))$$

und ist daher logarithmisch normalverteilt mit $\mu = \sum_{k=1}^{n} (\mu_k + log(\alpha_k))$ und $\sigma^2 = \sum_{k=1}^{n} \sigma_k^2$.

Nun wollen wir annehmen, dass die Abwicklungsfaktoren $D_{i,j}$ für $i \in \{1,..,n\}$ und $j \in \{1,..,n-1\}$ einer Log-Normalverteilung mit Parametern μ_j und σ_j^2 folgen.
Dabei vereinfacht die Annahme der Unabhängigkeit die Berechnungen der Reserve. Somit erhalten wir, dass der Endschadenstand $U_i|S_{i,n+1-i} = I_i \prod_{k=1}^{n-1} D_{i,k}$ einer Log-Normalverteilung folgt mit $\mu = log(S_{i,n+1-i}) + \sum_{k=n+1-i}^{n-1} \mu_k$ und $\sigma^2 = \sum_{k=n+1-i}^{n-1} \sigma_k^2$. Nun müssen wir noch die Parameter μ_k und σ_k^2 schätzen. Da wir angenommen haben, dass die Abwicklungsfaktoren unabhängig sind, stellen die individuellen Abwicklungsfaktoren $F_{i,j}$ für $i = 1,..,n-j$ eine Zufallsstichprobe der Log-Normalverteilung dar.

Lemma 5.1.
Mit Hilfe der Maximum-Likelihood Methode erhält man

$$\hat{\mu_j} = \frac{1}{n-j} \sum_{k=1}^{n-j} log(F_{k,j})$$

als Schätzer für μ_j, $j \in \{1,..,n-1\}$ und

$$\hat{\sigma_j^2} = \frac{\sum_{k=1}^{n-j} (log(F_{k,j} - \hat{\mu_j})^2)}{n-j}$$

für σ_j^2, wobei $j \in \{1,..,n-2\}$. Der Schätzer für σ_{n-1}^2 wird genauso wie im linearen Modell 3.2.1 mit Hilfe der Kleinstquadratmethode berechnet.

Beweis. Die Likelihood Funktion ist gegeben durch

$$f(F_{1,j},..,F_{n-j,j},\mu,\sigma_j^2) = \prod_{k=1}^{n-j} g(F_{k,j})$$

wobei g die Dichtefunktion einer Log-Normalverteilung mit Parametern μ_j und σ_j ist. Durch den Übergang zur Log-Likelihood Funktion erhalten wir

$$\begin{aligned} log(f) := log(f(F_{1,j},..,F_{n-j,j},\mu,\sigma_j^2)) &= \sum_{k=1}^{n-j} log(g(F_{k,j})) \\ &= \sum_{k=1}^{n-j} \Big(log(\frac{1}{(2\pi)^{\frac{1}{2}}\sigma_j y}) - \frac{(ln(F_{k,j}) - \mu_j)^2}{2\sigma_j^2} \Big) \quad . \end{aligned}$$

Damit erhalten wir

$$\frac{\partial log(f)}{\partial \mu_j} = \sum_{k=1}^{n-j} \frac{-(ln(F_{k,j}) - \mu_j)(-1)}{\sigma_j^2} = \sum_{k=1}^{n-j} \frac{(ln(F_{k,j}) - \mu_j)}{\sigma_j^2} = 0$$

und es gilt

$$\sum_{k=1}^{n-j} ln(F_{k,j}) = \sum_{k=1}^{n-j} \hat{\mu}_j = \hat{\mu}_j(n-j) \quad .$$

Nun wollen wir den Parameter σ_j^2 schätzen. Hier gilt

$$\frac{\partial log(f)}{\partial \sigma_j} = \sum_{k=1}^{n-j} \Big(-\frac{(2\pi)^{\frac{1}{2}} \sigma_j F_{k,j}}{(2\pi)^{\frac{1}{2}} \sigma_j^2 F_{k,j}} - \frac{(log(F_{k,j}) - \hat{\mu_j})^2}{2\sigma_j^3}(-2) \Big)$$
$$= \sum_{k=1}^{n-j} (-\frac{\sigma_j}{\sigma_j^2} + \frac{(log(F_{k,j}) - \hat{\mu_j})^2}{\sigma_j^3}) = 0$$

und daraus folgt

$$\sum_{k=1}^{n-j} (log(F_{k,j}) - \hat{\mu_j})^2 = \sum_{k=1}^{n-j} \hat{\sigma}_j^2 = \hat{\sigma}_j^2 (n-j) \quad .$$

□

Die Schätzer $\hat{\mu}_j$ und $\hat{\sigma}_j^2$ sind notwendig um den Endschadenstand zu berechnen.

5.1.2 Simulationsalgorithmus zur Berechnung der Endschadenstände

Da wir in diesem Modell mit Zufallsvariablen arbeiten, muss die Berechnung der Endschadenstände öfters durchgeführt werden, damit eine statistische Aussage getroffen werden kann.
Daher wird nun ein Algorithmus für diese Berechnung vorgestellt, der aus folgenden Schritten besteht.

1. Erzeuge zwei Zufallsvariablen U_1, U_2, die auf dem Intervall $[0,1]$ gleichverteilt sind, für die also $U_1, U_2 \sim U(0,1)$ gilt.

2. Mit Hilfe der Box-Muller Methode können wir nun zwei unabhängige, standardnormalverteilte Zufallsvariablen erzeugen. Diese Variablen berechnen sich durch

$$Z_1 = (-2log(U_1))^{\frac{1}{2}} \cos(2\pi U_2) \quad \text{und} \quad Z_2 = (-2log(U_1))^{\frac{1}{2}} \sin(2\pi U_2) \quad .$$

3. Mit der Transformation

$$Y_1 = exp(\hat{\mu} + \hat{\sigma} Z_1) \quad \text{und} \quad Y_2 = exp(\hat{\mu} + \hat{\sigma} Z_2)$$

   erhalten wir zwei logarithmisch normalverteilte Zufallsvariablen mit Paramatern $\hat{\mu}$ und $\hat{\sigma}$.

4. Für jedes Schadenjahr $i \in \{2,..,n\}$ und Abwicklungsjahr $j \in \{2,..,n\}$ mit $i+j > n+1$ wiederholen wir Schritt 1 bis 3 500 Mal um 1000 Beobachtungen $S_{i,j;1},..,S_{i,j;1000}$ des Schadenstandes $S_{i,j}$ zu erhalten, wobei $\hat{\mu} = log(S_{i,n+1-i}) + \sum_{k=n+1-i}^{j-1} \hat{\mu_k}$ und $\hat{\sigma^2} = \sum_{k=n+1-i}^{j-1} \hat{\sigma_k^2}$

5. Damit erhalten wir unter anderem 1000 Beobachtungen $S_1^\star,..,S_{1000}^\star$ des Gesamtschadenstandes, wobei $S_k^\star = \sum_{l=2}^{n} S_{l,n;k}$ gilt.

6. Unter Zuhilfenahme dieser 1000 Beobachtungen können wir, sowohl für die Endschadenstände der einzelnen Schadenjahre, als auch für die Gesamtendschadenstände und somit auch für die Reserve, die empirische Verteilungsfunktion, die Mittelwerte, Varianzen und weitere statistische Größen berechnen.

Eine weitere Möglichkeit, die auch in dieser Arbeit verwendet wird, besteht darin die logarithmisch normalverteilten Zufallsvariablen mit R Statistic [9] zu erzeugen. Dabei ist zu beachten, dass hier die Parameter der Verteilung der Erwartungswert und die Standardabweichung sind.

5.2 Das loggamma Modell

5.2.1 Grundlagen des Modells

In diesem Modell werden wir die Annahme treffen, dass die individuellen Abwicklungsfaktoren logarithmisch gammaverteilt sind. Diese Verteilungsannahme hat nicht mehr so viele wünschenswerte Eigenschaften wie noch die Log-Normalverteilung.
So ist die logarithmische Gammaverteilung nur für Werte größer eins sinnvoll definiert. Somit kann dieses Modell nur für Dreiecke verwendet werden, bei denen es für jedes Schadenjahr in jedem Abwicklungsjahr positive Zahlungen gab. Es besteht die Möglichkeit, die beobachteten Abwicklungsfaktoren zu verschieben, sodass alle Faktoren größer eins sind und dieses Modell auf alle Dreiecke angewandt werden kann, jedoch werden wir in dieser Arbeit nicht näher darauf eingehen.
Ein weiterer Nachteil besteht darin, dass die logarithmsiche Gammaverteilung nur dann bezüglich der Multiplikation abgeschlossen ist, wenn der Skalierungsparameter konstant ist für alle Abwicklungsjahre. Dabei ist die logarithmische Gammaverteilung bezüglich der Multiplikation abgeschlossen, da die Gammaverteilung bezüglich der Addition abgeschlossen ist. Somit müssen zur Berechnung der Schätzer alle bekannten Daten herangezogen werden und wir können uns nicht wie beim lognormalen Modell auf die einzelnen Spalten des Dreiecks einschränken.
Jedoch ist diese Verteilung trotzdem von großer Interesse, da es Sparten gibt, bei denen die Verteilung der Abwicklungsfaktoren der logarithmischen Gammaverteilung näher kommmt als die Log-Normalverteilung, wie wir später bei der Sparte Rechtsschutz sehen werden.

Die Dichte der Gammaverteilung mit Lageparameter α und Skalierungsparameter λ, wir schreiben dann $X \sim Gamma(\alpha, \lambda)$, ist durch

$$f(x) = \frac{x^{\alpha-1} exp(-\lambda x)\lambda^{\alpha}}{\Gamma(\alpha)} \quad x \geq 0$$

gegeben, wobei

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} exp(-t) dt$$

die Gammafunktion ist.
Die momenterzeugende Funktion ist gegeben durch

$$M_X(t) = \mathbb{E}[exp(tX)] = (\frac{\lambda}{\lambda - t})^{\alpha} \quad \lambda < t \quad .$$

Hier sehen wir sehr schön, wieso der Skalierungsparameter für die Abgeschlossenheit bezüglich der Addition konstant sein muss. Wie wir ebenfalls an der momenterzeugenden Funktion sehen, gilt für $X_1, ..., X_n$ mit $X_i \sim Gamma(\alpha_i, \lambda)$,
dass $X = \sum_{i=1}^{n} X_i \sim Gamma(\alpha = \sum_{i=1}^{n} \alpha_i, \lambda)$. Weiter ist die Gammaverteilung nicht abgeschlossen bezüglich Verschiebungen des Lageparameters und daher die logarithmische Gammaverteilung nicht bezüglich der Multiplikation mit Skalaren. Wenn $X \sim Gamma(\alpha, \lambda)$ und wir $Z = X + \delta$ betrachten, so folgt Z einer verschobenen Gammaverteilung mit Dichtefunktion

$$h(z) = \frac{(z-\delta)^{\alpha-1} epx(-\lambda(z-\delta))\lambda^{\alpha}}{\Gamma(\alpha)} \quad z \geq \delta$$

und momenterzeugenden Funktion

$$M_Z(t) = \mathbb{E}[exp(tX + t\delta)] = exp(t\delta)(\frac{\lambda}{\lambda - t})^{\alpha} \quad \lambda < t \quad .$$

Betrachten wir nun $Y = exp(X)$, so ist Y logarithmisch gammaverteilt mit Dichtefunktion

$$g(y) = f(log(y))\frac{1}{y} = \frac{(log(y))^{\alpha-1} y^{-(\lambda+1)} \lambda^{\alpha}}{\Gamma(\alpha)} \quad y \geq 1 \quad .$$

Wir nehmen nun an, dass die individuellen Abwicklungsfaktoren $D_{i,j}$ für $i \in \{1, .., n\}$ und $j \in \{1, .., n-1\}$ einer logarithmischen Gammaverteilung folgen mit Lageparameter α_j und Skalierungsparameter λ. Somit ist der Endschadenstand $U_i | S_{i,n+1-i} = I_i \prod_{k=1}^{n-1} D_{i,k}$ verschoben logarithmisch gammaverteilt und hat die Dichtefunktion

$$g_i(u_i) = \frac{(log(u_i) - log(S_{i,n+1-i}))^{\alpha-1} S_{i,n+1-i}^{\lambda} u_i^{(-\lambda-1)} \lambda^{\alpha}}{\Gamma(\alpha)} \quad u_i \geq S_{i,n+1-i}$$

mit $\alpha = \sum_{j=n+1-i}^{n-1} \alpha_j$.
Nun wollen wir die Parameter des Modells schätzen. Dies ist weitaus aufwendiger als noch im lognormalen Modell. Die Schätzung erfolgt wieder unter Zuhilfenahme der Likelihood-Methode und wir erhalten als Loglikelihood-Funktion

$$\begin{aligned} Log(L(\alpha_{\{j\}}, \lambda; F_{\{i,j\}})) = \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} &((\alpha_j - 1) log(log(F_{i,j})) + \alpha_j log(\lambda) \\ &- log(\Gamma(\alpha_j)) - (\lambda + 1) log(F_{i,j})) \end{aligned}$$

wobei hier nur noch jene Terme stehen, die von α_j oder λ abhängen. DIe Maximierung dieser Funktion führt zu folgenden n Gleichungen

$$\begin{aligned} \hat{\lambda} &= \frac{\sum_{j=1}^{n-1} (n-j) \hat{\alpha}_j}{\sum_{j=1}^{n-1} \sum_{i=1}^{n-j} log(F_{i,j})} \\ \psi(\hat{\alpha}_j) &= log(\hat{\lambda}) + \frac{\sum_{i=1}^{n-j} log(log(F_{i,j}))}{n-j} \quad \text{für} \quad j = 1, .., n-1 \end{aligned}$$

wobei $\psi(x) = \frac{\partial log(\Gamma(x))}{\partial x}$ die Digammafunktion ist. Eine Annäherung der Digammafunktion ist gegeben durch

$$\begin{aligned}\psi(x) \sim log(x) &- \frac{1}{2x} - \frac{1}{12x^2} + \frac{1}{120x^4} - \frac{1}{256x^6} + \frac{1}{240x^8} \\ &- \frac{1}{132x^{10}} + \frac{691}{32760x^{12}} - \frac{1}{12x^{14}} + \frac{255}{28936x^{16}} + - \ldots \quad ,\end{aligned} \tag{5.1}$$

wobei diese Annäherung für $x \to \infty$ gilt. Da für die Gammafunktion

$$\Gamma(x+1) = x\Gamma(x)$$

gilt, erhalten wir für die Digammafunktion die Beziehung

$$\psi(x+1) = \psi(x) + \frac{1}{x} \tag{5.2}$$

Mit Hilfe der Approximation (5.1), der Beziehung (5.2) und der Tabelle 5.3 können wir nun für $x \in [0,1;4000]$ die Funktion $\psi(x)$ sehr gut schätzen.

Wertbereich von x	Berechnungsmethodik
[0,1;40]	Verwendung der ersten 10 Terme von 5.1 mit $t = 40 + (x - \lfloor x \rfloor)$. $\psi(x)$ kann dann mittels 5.2 durch $\psi(t)$ berechnet werden.
(40;60]	Verwendung der ersten 10 Terme von 5.1
(60;200]	Verwendung der ersten 8 Terme von 5.1
(200;500]	Verwendung der ersten 6 Terme von 5.1
(500;4000]	Verwendung der ersten 5 Terme von 5.1

Tabelle 5.3: Berechnungsalgorithmus der Digammafunktion

Bezeichnen wir mit $l := 40 - \lfloor x \rfloor$, so erhalten wir $\psi(x)$ für $x \in [0,1;40]$ durch

$$\psi(x) = \psi(t) - \sum_{j=1}^{l} \frac{1}{x+l-j} \quad .$$

Nun können wir die Parameter $\hat{\alpha}_1, ..., \hat{\alpha}_{n-1}$ berechnen. Dazu definieren wir den Vektor

$$\mathbb{F}(\hat{\alpha}) := (f_1(\hat{\alpha}), .., f_{n-1}(\hat{\alpha}))^T$$

wobei für $k \in \{1, .., n-1\}$

$$f_k(\hat{\alpha}) = f_k(\hat{\alpha}_1, ..., \hat{\alpha}_{n-1}) := \psi(\hat{\alpha}_k) - log(\frac{\sum_{j=1}^{n-1}(n-j)\hat{\alpha}_j}{\sum_{j=1}^{n-1}\sum_{i=1}^{n-j} log(F_{i,j})}) - \frac{\sum_{i=1}^{n-k} log(log(F_{i,k}))}{n-k} \quad .$$

Die Jacobi-Matrix $\mathbb{J}(\hat{\alpha})$ des Vektors $\mathbb{F}(\hat{\alpha})$ ist durch

$$\mathbb{J}(\hat{\alpha}) = \psi'(\hat{\alpha}) - \frac{1}{\sum_{j=1}^{n-1}(n-j)\hat{\alpha}_j} D$$

gegeben, wobei

$$\psi'(\hat{\alpha}) := \begin{pmatrix} \psi'(\hat{\alpha_1}) & 0 & \cdots & 0 \\ 0 & \psi'(\hat{\alpha_2}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi'(\hat{\alpha_{n-1}}) \end{pmatrix}$$

und

$$D := \begin{pmatrix} n-1 & n-2 & \cdots & 1 \\ n-1 & n-2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ n-1 & n-2 & \cdots & 1 \end{pmatrix}$$

gilt. Die Werte von $\psi'(\hat{\alpha_k})$ erhalten wir, indem wir (5.1) und (5.2) differenzieren und den Algorithmus verwenden. Dabei gilt für $x \in [0,1;40]$ und $l := 40 - \lfloor x \rfloor$ die Gleichung

$$\psi'(x) = \psi'(40 + (x - \lfloor x \rfloor)) + \sum_{k=1}^{l} \frac{1}{(x+l-1)^2} \quad .$$

Somit erhalten wir unter Zuhilfenahme des mehrdimensionalen Newton-Verfahrens

$$(\hat{\alpha}_{1,m}, ..., \hat{\alpha}_{n-1,m})^T = (\hat{\alpha}_{1,m-1}, ..., \hat{\alpha}_{n-1,m-1})^T - \mathbb{F}(\hat{\alpha}_{m-1})\mathbb{J}^{-1}(\hat{\alpha}_{m-1}) \quad ,$$

welches wir durchführen müssen bis Kovergenz beobachtbar ist. Eine Möglichkeit zur Berechnung der Startwerte für das Newton-Verfahren ist die Momentenmethode. Dabei gilt für $j \in \{1, .., n-1\}$

$$\hat{\alpha}_{j,0} = \frac{\frac{1}{n-j}(\sum_{i=1}^{n-j} log(d_{i,j}))^2}{\sum_{i=1}^{n-j}(log(d_{i,j}))^2 - \frac{1}{n-j}(\sum_{i=1}^{n-j} log(d_{i,j}))^2} \quad .$$

Bessere Startwerte erhalten wir, wenn wir die α_j aus den Werten berechnen, die wir im loginversen Modell erhalten. Nun können wir auch $\hat{\lambda}$ berechnen. Somit können wir mit den geschätzten Parametern die Schadenstände simulieren und anhand des Simulationsergebnisses Erwartungswert und Varianz des Schätzers berechnen.

5.2.2 Simulationsalgorithmus zur Berechnung der Endschadenstände

Auch für dieses Modell werden wir einen Algorithmus zur Berechnung der Endschadenstände angeben. Zunächst präsentieren wir einen Algorithmus, der in allen Programmen verwendet werden kann, die gleichverteilte Zufallsvariablen erzeugen können.

1. Zunächst wird eine Zufallsvariable $X \sim Gamma(\hat{\alpha}, 1)$ erzeugt. Dabei gilt es zu unterscheiden, ob $\hat{\alpha} \geq 1$ oder $\hat{\alpha} < 1$ ist. Dieser Algorithmus zur Simulierung einer gammaverteilten Zufallsvariable wurde aus Devroye [3] entnommen.

   (a) Algorithmus für $\hat{\alpha} \geq 1$:
   Definiere $b := \hat{\alpha} - 1$ und $c := 3\hat{\alpha} - \frac{3}{4}$

i. Erzeuge zwei unabhängige auf $[0,1]$ gleichverteilte Zufallsvariablen U, V
ii. Definiere $W := U(1-U), Y := \sqrt{\frac{c}{W}}(U - \frac{1}{2}), X := b + Y$
iii. Falls $X < 0$, gehe zu i., sonst gehe zu iv..
iv. Definiere $Z := 64W^3V^2$
v. Falls $Z \leq 1 - \frac{2Y^2}{X}$ gehe zu vii. , sonst gehe zu vi..
vi. Falls $log(Z) \leq 2(blog(\frac{X}{b}) - Y)$ gehe zu vii., sonst gehe zu i..
vii. Gib X zurück.

(b) Algorithmus für $\hat{\alpha} < 1$
Definiere $c := \frac{1}{\hat{\alpha}}, d = \hat{\alpha}^{\frac{\hat{\alpha}}{1-\hat{\alpha}}}(1-\hat{\alpha})$
i. Erzeuge zwei unabhängige auf $[0,1]$ gleichverteilte Zufallsvariablen U, V
ii. Damit erzeuge zwei exponentialverteilte Zufallsvariable
$X = log(\frac{1}{1-U}), Y = log(\frac{1}{1-V})$.
iii. Definiere $Z = X^c$. Damit ist Z weibullverteil mit Parameter a
iv. Falls $X + Y \leq d + Z$, gib Z zurück, sonst gehe zu i.

2. Eine verschobene logarithmische Gammaverteilung mit Parametern $\hat{\alpha}$, $\hat{\lambda}$ und Verschiebung δ erhält man durch
$$Y = exp(\frac{1}{\hat{\lambda}}X + \delta) \quad .$$

3. Für jedes Schadenjahr $i \in \{2, .., n\}$ und jedes Abwicklungsjahr $j \in \{2, .., n\}$ mit $i + j > n + 1$ wiederhole die Schritte eins und zwei 1000 Mal um Beobachtungen $S_{i,j;1}, S_{i,j;2}, .., S_{i,j;1000}$ zu erzeugen, wobei $\delta = log(S_{i,n+1-1})$ gilt. Dabei sind $\hat{\alpha} = \sum_{k=n+1-i}^{j-1} \hat{\alpha_k}$ und $\hat{\lambda}$ jene Werte, die mittels der Likelihood-Methode berechnet wurden.

4. Somit erhält man auch hier 1000 Beobachtungen der jeweiligen Endschadenstände und des Gesamtendschadenstandes.

5. Diese 1000 Beobachtungen können dafür verwendet werden, um den Erwartungswert, die Varianz und andere statistische Kennzahlen der Endschadenstände und der Reserven zu berechnen.

In dieser Arbeit werden wir jedoch die logarithmisch gammaverteilte Zufallsvariable nicht mit dem im 1. und 2. Schritt beschriebenen Algorithmus bestimmen, sondern mit dem Statistikprogramm R Statistics [9]. Alle anderen Schritte werden wie oben beschrieben durchgeführt.

5.3 Das loginverse Gauß Modell

5.3.1 Grundlagen des Modells

In diesem Modell werden wir die Annahme treffen, dass die individuellen Abwicklungsfaktoren der loginversen Gauß-Verteilung folgen. Genauso wie die logarithmische Gammaverteilung ist diese Verteilung nur für Werte größer eins definiert. Ebenso ist die loginverse Gauß-Verteilung nur dann bezüglich der Multiplikation abgeschlossen, wenn die

Ereignisrate konstant ist. Daher müssen auch in diesem Modell alle bekannten Daten zur Schätzung der Parameter herangezogen werden.
Die Dichte der inversen Gauß-Verteilung mit Ereignisrate $\lambda > 0$ und Mittelwert $\mu > 0$, wir schreiben dann $X \sim IG(\mu, \lambda)$, ist durch

$$f(x) = (\frac{\lambda}{2\pi x^3})^{\frac{1}{2}} exp(\frac{-\lambda(x-\mu)^2}{2\mu^2 x}) \text{ für } x > 0 \tag{5.3}$$

gegeben. Durch die Umparametrisierung $\lambda := \mu^2\beta$ erhalten wir folgende Dichte

$$f(x) = \mu(\frac{\beta}{2\pi x^3})^{\frac{1}{2}} exp(\frac{-\beta(x-\mu)^2}{2x}) \text{ für } x > 0 \tag{5.4}$$

und schreiben dann $X \sim IG(\mu, \beta)$. Die momenterzeugende Funktion dieser Verteilung ist durch

$$M_X(t) = exp(\mu\beta - \mu\beta(1 - \frac{2}{\beta}t)^{\frac{1}{2}}) \quad t \leq \frac{\beta}{2}$$

gegeben. Anhand der Darstellung der momenterzeugenden Funktion ist zu erkennen, dass eine Summe unabhängiger, invers gaußverteilter Zufallsvariablen wieder invers gaußverteilt ist, wenn β konstant ist.
Definieren wir $Z := X + \delta$, wobei $X \sim IG(\beta, \mu)$ und $\delta > 0$, so ist Z nicht invers gaußverteilt sondern verschoben invers gaußverteilt. Diese Verteilung hat die Dichtefunktion

$$h(z) = \mu(\frac{\beta}{2\pi(z-\delta)^3})^{\frac{1}{2}} exp(\frac{-\beta(z-\delta-\mu)^2}{2(z-\delta)}) \quad z \geq \delta$$

und die momenterzeugende Funktion

$$M_Z(t) = exp(\mu\beta + t\delta - \mu\beta(1 - \frac{2}{\beta}t)^{\frac{1}{2}}) \quad .$$

Somit erhalten wir für unabhängige Zufallsvariablen $X_1, ..., X_n$, wobei $X_i \sim IG(\mu_i, \beta)$, dass $T_n := \sum_{i=1}^n (X_i + \delta_i)$ einer verschobenen inversen Gauß-Verteilung folgt mit Parametern $\mu = \sum_{i=1}^n \mu_i$ und β und einer Verschiebung $\delta = \sum_{i=1}^n \delta_i$.
Ist $X \sim IG(\mu, \beta)$ und defineren wir $Y := exp(X)$, dann ist Y logarithmisch invers gaußverteilt mit Dichtefunktion

$$g(y) = f(log(y))\frac{1}{y} = \mu(\frac{\beta}{2\pi log(y)^3})^{\frac{1}{2}} \frac{1}{y} exp(\frac{-\beta(log(y)-\mu)^2}{2log(y)}) \text{ für } y > 1 \quad .$$

Damit erhalten wir

$$g(y) = \mu(\frac{\beta}{2\pi log(y)^3})^{\frac{1}{2}} y^{-(1+\frac{1}{2}\beta)} exp(\beta\mu - \frac{\beta\mu^2}{2log(y)}) \quad .$$

In diesem Modell wollen wir nun annehmen, dass die Abwicklungsfaktoren $D_{i,j}$ für $i \in \{1, .., n\}$ und $j \in \{1, .., n-1\}$ logarithmisch invers gaußverteilt sind mit Parametern μ_j und β. Dann ist die Verteilung von $U_i | S_{i,n+1-i} = I_i \prod_{j=1}^{n-1} D_{i,j}$ eine verschobene logarithmische inverse Gauß-Verteilung mit Dichtefunktion

$$\begin{aligned} g_i(u_i) = \mu(&\frac{\beta}{2\pi(log(u_i) - log(S_{i,n+1-i}))^3})^{\frac{1}{2}} (S_{i,n+1-i})^{\frac{1}{2}\beta} u_i^{-(1+\frac{1}{2}\beta)} \\ &* exp(\beta\mu - \frac{\beta\mu^2}{2(log(u_i) - log(S_{i,n+1-i}))}) \end{aligned}$$

wobei $u_i \geq S_{i,n+1-i}$ und $\mu = \sum_{j=n+1-i}^{n-1} \mu_j$ gilt.
Genauso wie im loggamma Modell ist die Schätzung der Parameter sehr rechenintensiv. Unter Zuhilfenahme der beobachtbaren, individuellen Abwicklungsfaktoren $F_{i,j}$ erhalten wir die Loglikelihoodfunktion

$$log(L(\mu_{\{j\}}, \beta, d_{\{i,j\}})) = \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} (log(\mu_j) + \frac{1}{2} log(\beta) - \frac{1}{2} \beta \frac{(log(F_{i,j}) - \mu_j)^2}{log(F_{i,j})} \quad ,$$

wobei jene Terme, die unabhängig von β und μ_j sind, bereits weggelassen wurden. Für die Maximum-Likelihoodschätzer ergeben sich somit folgende n Gleichungen

$$\frac{1}{\hat{\beta}} = \frac{\sum_{j=1}^{n-1} \sum_{i=1}^{n-j} (\frac{(log(F_{i,j}) - \hat{\mu}_j)^2}{log(F_{i,j})})}{\frac{(n-1)n}{2}}$$

$$\frac{n-j}{\hat{\beta}} = \hat{\mu}_j^2 \sum_{i=1}^{n-j} (\frac{1}{log(F_{i,j})}) - (n-j)\hat{\mu}_j \quad ,$$

wobei $j \in \{1, .., n-1\}$. Diese Gleichungen können wir mittels einem zweischrittigen iterativen Verfahren lösen.
Als Startwert dieses Verfahrens wählen wir $\mu_{j,0} = \frac{1}{n-j} \sum_{i=1}^{n-j} log(F_{i,j})$. Dieser Algorithmus wird solange ausgeführt bis Konvergenz der Parameter eintritt.

1. Unter Zuhilfenahme der Werte $\hat{\mu}_{1,m-1}, .., \hat{\mu}_{n-1,m-1}$ wird der Maximumlikelihoodschätzer $\hat{\beta}_m$ berechnet.

2. Nachdem $\hat{\beta}_m$ berechnet wurde, kann damit $\hat{\mu}_{j,m}$ für $j \in \{1, .., n-1\}$ mit Hilfe des Maximumlikelihoodschätzers berechnet werden. Dabei gilt

$$\hat{\mu}_{j,m} = \frac{n-j + ((n-j)(n-j+4\hat{\beta}_m^{-1} \sum_{i=1}^{n-j} \frac{1}{log(F_{i,j})}))^{\frac{1}{2}}}{2 \sum_{i=1}^{n-j} \frac{1}{log(F_{i,j})}} \quad .$$

Nun können wir mit Hilfe des Algorithmus, der im nächsten Unterkapitel vorgestellt wird, Endschadenstände simulieren und mit Hilfe dieser Stichprobe Erwartungswert, Varianz und weitere statistische Kennzahlen berechnen.

5.3.2 Simulationsalgorithmus zur Berechnung der Endschadenstände

Auch für dieses Modell werden wir zunächst einen allgemeinen Algorithmus betrachten, der mit allen Programmen durchgeführt werden kann, die uniformverteilte Zufallsvariablen erzeugen kann. Dabei ist zu beachten, dass die Zufallsvariablen mit einer Parametrisierung, wie sie bei (5.4) der Fall ist, erzeugt werden.

1. Erzeuge auf dem Intervall $[0, 1]$ gleichverteilte Zufallsvariablen U_1 und U_2.

2. Mit der Box-Müller Methode werden zwei unabhängige chi-quadratverteilte Zufallsvariablen Z_1, Z_2 mit einem Freiheitsgrad erzeugt, wobei

$$Z_1 = ((-2log(U_1)^{\frac{1}{2}} cos(2\pi U_2))^2, Z_2 = ((-2log(U_1)^{\frac{1}{2}} sin(2\pi U_2))^2$$

gilt.

3. Nun definieren wir $X_{1,2} := \hat{\mu} + \frac{Z_{1,2} - (Z_{1,2}^2 + 4\hat{\beta}\hat{\mu}Z_{1,2})^{\frac{1}{2}}}{2\hat{\beta}}$

4. Für jede der beiden Zufallsvariablen X_1, X_2 erzeugen wir eine auf $[0,1]$ gleichverteilte Zufallsvariable $V_{1,2}$

5. Wenn $V_i \leq \frac{\hat{\mu}}{\hat{\mu}+X_i}$ für $i \in \{1,2\}$ definiere $Y_i := exp(X_i + \delta)$, sonst definere $Y_i := exp(\frac{\hat{\mu}^2}{X_i} + \delta)$. Dann folgt Y_i einer verschobenen logarithmisch inversen Gauß-Verteilung mit Parametern $\hat{\mu}, \hat{\beta}$ und Verschiebung δ.

6. Für jedes Schadenjahr $i \in \{2, .., n\}$ und Abwicklungsjahr $j \in \{2, .., n\}$ mit $i + j > n+1$ wiederholen wir die oben genannten Schritte 500 Mal um 1000 Beobachtungen $S_{i,j;1}, ..., S_{i,j;1000}$ zu erzeugen wobei $\delta = log(S_{i,n+1-i})$ und $\hat{\beta}$ und $\hat{\mu} = \sum_{k=n+1-i}^{j-1} \hat{\mu}_j$ die oben berechneten Werte sind.

7. Somit erhält man auch hier 1000 Beobachtungen der jeweiligen Endschadenstände und des Gesamtendschadenstandes. Mit Hilfe dieser Stichprobe können wir statistische Kennzahlen der Schadenstände und der Reserven berechnen.

Für diese Arbeit werden wir die invers gaußverteilten Zufallsvariablen mit Hilfe des Statistikprogramms R [9] berechnen. Dabei ist zu beachten, dass in diesem Programm die Dichte mit der Parametrisierung wie in (5.3) hinterlegt ist. Daher müssen wir in diesen Fall mit den Parametern $\hat{\mu}$ und $\hat{\lambda} = \hat{\mu}^2\hat{\beta}$ arbeiten.

5.4 Beispiele

Auch bei diesen Modellen werden wir die Reserve anhand der kumulierten Zahlungsdreiecke der KFZ-Kasko und Rechtsschutz Sparte berechnen, wobei auch hier nochmals zu erwähnen ist, dass die verwendeten Daten aus der Praxis stammen. Dabei werden wir zuerst die verschiedenen Verteilungsannahmen auf KFZ-Kasko und danach auf Rechtsschutz anwenden.

Dabei wird bei den nachfolgenden Beispielen R_i für die tatsächliche benötigte Reserve stehen, das heißt für $i \in \{2, .., 7\}$ gilt

$$R_i = S_{i,n} - S_{i,n+1-i}$$

und $R_i^\star$ bezeichnet für $i \in \{2, .., 7\}$ den Schätzer der Reserve, es gilt also

$$R_i^\star = \hat{S}_{i,n} - S_{i,n+1-i} \quad .$$

Dabei erhalten wir die Schätzer $\hat{S}_{i,j}$ für $i + j > n + 1$ durch das empirische Mittel der jeweiligen Stichprobe,

$$\hat{S}_{i,j} = \frac{\sum_{k=1}^{1000} S_{i,j;k}}{1000} \quad .$$

Die Standardabweichung der tatsächlichen Reserve erhalten wir durch

$$s.d.(R_i) = \sqrt{\frac{1}{1000-1} \sum_{k=1}^{1000} (S_{i,n;k} - \hat{S}_{i,n})^2} \quad .$$

da auch hier aufgrund der Tatsache, dass $S_{i,n+1-i}$ eine Konstante ist,

$$Var(R_i) = Var(S_{i,n})$$

gilt. Somit erhalten wir

$$s.d(R_i^\star) = \frac{s.d.(R_i)}{\sqrt{1000}} \quad .$$

Die Quantile der Konfidenzintervalle erhalten wir aus der Stichprobe.

5.4.1 KFZ-Kasko

Wir werden hier die Beispiele ebenfalls mit den bereits beobachtbaren kumulierten Zahlungsdreieck beginnen, welches bereits aus dem Chain-Ladder Modell bekannt ist.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47		
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93			
5	15.452.148,36	18.356.911,24	18.489.553,46				
6	15.463.013,72	18.430.481,19					
7	13.768.695,59						

Aus diesen Daten erhalten wir die logarithmierten, individuellen Abwicklungsfaktoren

SJ \ AJ	1	2	3	4	5	6	7
1	0,21970	0,00010	0,00021	0,00021	0,00008	0,00004	
2	0,19472	0,00440	0,00112	0,000003	0,00001		
3	0,17683	0,00788	0,00569	0,00011			
4	0,15126	0,00841	0,00270				
5	0,17226	0,00720					
6	0,17555						
7							

Tabelle 5.4: Logarithmierte individuelle Abwicklungsfakoren KFZ-Kasko

Um später unser Ergebnis mit den tatsächlich beobachteten Werten vergleichen zu können, benötigen wir noch das tatsächliche Abwicklungsquadrat welches wie folgt aussieht

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	15.499.202,02
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47	19.722.695,47	19.722.939,17
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93	18.927.186,83	18.931.158,56	18.931.978,64
5	15.452.148,36	18.356.911,24	18.489.553,46	18.489.986,41	18.490.702,29	18.491.202,29	18.491.373,02
6	15.463.013,72	18.430.481,19	18.582.598,10	18.589.779,85	18.599.426,35	18.600.784,49	18.601.256,49
7	13.768.695,59	16.330.431,63	16.439.713,95	16.471.855,69	16.472.920,60	16.473.930,60	16.473.930,60

Nun ist alles vorbereitet um die Reserven unter den verschiedenen Verteilungsannahmen zu berechnen.

5.4.1.1 Lognormales Modell

Der erste Schritt, der in diesem Modell durchgeführt werden muss, ist die Schätzung der Parameter der Log-Normalverteilung. Diese Schätzer sehen wie folgt aus

j	1	2	3	4	5	6
$\hat{\mu_j}$	0,18172	0,00560	0,00243	0,00011	0,00005	0,00004
$\hat{\sigma_j^2}$	0,000448398259	0,000009477330	0,000004341273	0,000000007268	0,000000001461	0,000000000043

Nun können wir mithilfe des Algorithmus aus 5.1.2 die nicht beobachtbaren Schadenstände simulieren. Die Stichproben die wir dadurch erhalten, können wir nutzen, um die zukünftigen Schadenstände zu schätzen.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	15.498.918,79
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47	19.723.576,57	19.724.384,56
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93	18.922.248,83	18.923.099,34	18.923.872,21
5	15.452.148,36	18.356.911,24	18.489.553,46	18.533.378,76	18.535.545,61	18.536.399,49	18.537.151,10
6	15.463.013,72	18.430.481,19	18.532.973,25	18.580.889,63	18.582.913,54	18.583.772,66	18.584.533,74
7	13.768.695,59	16.504.277,86	16.596.446,31	16.636.732,47	16.638.543,84	16.639.294,48	16.639.979,53

Tabelle 5.5: KFZ-Kasko Lognormales Modell

Für die Reserven und deren Kennzahlen erhalten wir folgendes Bild.

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw. $R_i^\star$
2	631,08	110,64	403,25	852,49	3,50
3	1.689,09	766,10	234,26	3.157,88	24,23
4	3.706,28	1.814,66	225,02	7.370,34	57,38
5	47.597,64	39.572,22	6.111,34	125.437,09	1.251,38
6	154.052,55	69.055,49	16.937,80	289.007,99	2.183,73
7	2.871.283,94	361.967,23	2.158.993,84	3.574.043,60	11.446,41
Gesamt	3.078.960,58	371.430,25	2.306.445,26	3.751.638,74	11.745,66

Auffallend ist hier, dass die Standardabweichung des Schätzers sehr gering ist im Vergleich zu den anderen Modellen, da hier auf mehr Daten zurückgegriffen werden kann, um den Schadenstand zu schätzen.
Vergleichen wir nun unsere Schätzer mit den tatsächlichen Werten, erhalten wir folgende Differenzen.

SJ	R_i	$R_i^\star - R_i$
2	914,31	- 283,23
3	243,70	1.445,39
4	11.812,71	- 8.106,43
5	1.819,56	45.778,08
6	170.775,30	- 16.722,75
7	2.705.235,01	166.048,93
Gesamt	2.890.800,59	188.159,99

Hier ist gut zu erkennen, dass mit Ausnahme des 5. Schadenjahres die Reserve gut geschätzt wird. Anzumerken ist, dass die Schadenjahre 2 bis 5 nicht oder nur sehr knapp im angegebenen Konfidenzintervall liegen. Bei mehrmaligem Auftreten dieser Situation wäre die Verteilungsannahme nochmals zu überdenken.

5.4.1.2 Loggamma Modell

Auch in diesem Modell müssen wir als ersten Schritt die Parameter der Verteilung schätzen. Da diese Schätzung unter Zuhilfenahme des Newton-Verfahrens funktioniert, müssen wir zunächst die Startwerte berechnen.
Würden wir zur Berechnung der Startwerte die Momentenmethode verwenden, könnten wir für unsere Daten keine Konvergenz beobachten. Daher müssen wir, wie in 5.2.1 beschrieben, auf die Werte des loginversen Gauß Modell zurückgreifen, um die Startwerte $\hat{\alpha}_{j,0}$ zu berechnen.

j	1	2	3	4	5	6
$\hat{\alpha}_{j,0}$	33,1395352	0,33273305	0,403037	0,03853953	0,04976607	0,0894312

Mit deren Hilfe wurde das Newton-Verfahren durchgeführt bis für alle $j \in \{1,..,6\}$

$$|\hat{\alpha}_{j,m} - \hat{\alpha}_{j,m-1}| < 10^{-6}$$

gilt. Somit erhalten wir für die Parameter der Verteilung die Schätzer

j	1	2	3	4	5	6
$\hat{\alpha}_j$	60,542543	1,4258571	0,88769409	0,24552795	0,2198458	0,246035

und

$$\hat{\lambda} = 332,61464 \quad .$$

Nun können wir den Algorithmus 5.2.2 anwenden um die Schadenstände zu schätzen. Dieser liefert uns folgende Werte

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	15.509.761,41
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47	19.736.361,84	19.751.287,89
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93	18.934.881,07	18.948.603,66	18.962.990,47
5	15.452.148,36	18.356.911,24	18.489.553,46	18.540.029,05	18.553.799,13	18.566.695,23	18.581.554,44
6	15.463.013,72	18.430.481,19	18.510.697,79	18.557.872,51	18.570.056,93	18.583.210,46	18.597.397,19
7	13.768.695,59	16.522.811,00	16.590.905,75	16.633.900,87	16.645.641,68	16.657.203,19	16.668.325,57

Tabelle 5.6: KFZ-Kasko Loggamma Modell

Somit erhalten wir für die Reserven folgende Werte

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw. $R_i^\star$
2	11.473,70	22.496,46	0,02	84.779,52	711,40
3	28.592,42	39.931,12	31,70	142.764,51	1.262,73
4	42.824,54	49.687,38	418,16	189.268,52	1.571,25
5	92.000,98	73.443,57	24.875,82	282.577,78	2.322,49
6	166.916,00	94.259,03	40.981,12	413.268,55	2.980,73
7	2.899.629,98	402.373,20	2.165.174,60	3.802.950,72	12.724,16
Gesamt	3.241.437,60979	428.388,16	2.470.479,74	4.158.080,39	13.546,82

Vergleichen wir nun die geschätzten Reserven mit den tatsächlich benötigten, erhalten wir

SJ	R_i	$R_i^\star - R_i$
2	914,31	10.559,39
3	243,70	28.348,72
4	11.812,71	31.011,83
5	1.819,56	90.181,42
6	170.775,30	- 3.859,30
7	2.705.235,01	194.394,97
Gesamt	2.890.800,59	350.637,02

In diesem Modell werden vor allem die Schadenjahre 2 bis 5 stark überschätzt. Daher erhalten wir auch für die Gesamtreserve eine relativ große Abweichung. Hier befinden sich bis auf das 5. Schadenjahr alle tatsächlichen Reserven in deren Konfidenzintervall. Dies liegt auch daran, dass die Reserven für die Schadenjahre 2 bis 4 sehr stark streuen und daher die Konfidenzintervalle deutlich größer werden.

5.4.1.3 Loginverses Gauß Modell

Um bei diesem Modell die Endschadenstände berechnen zu können, ist es zunächst notwendig die Parameter zu schätzen. Da diese sowie im loggamma Modell über einen Algorithmus berechnet werden müssen, ist es notwendig, Startwerte für diesen Algorithmus anzugeben.

j	1	2	3	4	5	6
$\hat{\mu}_{j,0}$	0,1817220	0,0055972	0,0024320	0,0001096	0,0000454	0,0000409

Mit Hilfe dieser Startwerte und des Algorithmus im Abschnitt 5.3.1 können wir nun die Parameter schätzen, wobei der Algorithmus durchgeführt wurde bis für alle $j \in \{1, .., 6\}$

$$|\hat{\mu}_{j,m} - \hat{\mu}_{j,m}| < 10^{-6}$$

gilt. Daher erhalten wir

j	1	2	3	4	5	6
$\hat{\mu}_j$	0,1847591	0,0018550	0,0022470	0,0002149	0,0002775	0,0004986

und

$$\hat{\beta} = 179,3662264 \quad .$$

Unter Verwendung des Algorithmus 5.3.2 kann nun auch in diesem Modell das Zahlungsquadrat erzeugt werden.

SJ \ AJ	1	2	3	4	5	6	7
1	9.908.307,89	12.342.766,87	12.343.978,93	12.346.573,05	12.349.183,49	12.350.215,83	12.350.721,33
2	12.685.757,26	15.412.892,42	15.480.782,30	15.498.134,36	15.498.177,08	15.498.287,71	15.505.573,43
3	16.301.266,73	19.454.488,43	19.608.489,82	19.720.436,82	19.722.695,47	19.728.027,93	19.738.270,19
4	16.084.447,76	18.711.072,65	18.869.060,33	18.920.165,93	18.925.937,32	18.931.582,92	18.942.849,08
5	15.452.148,36	18.356.911,24	18.489.553,46	18.533.573,50	18.536.758,64	18.541.787,84	18.549.946,86
6	15.463.013,72	18.430.481,19	18.468.055,45	18.506.842,99	18.510.254,37	18.514.642,51	18.523.990,39
7	13.768.695,59	16.558.979,09	16.590.735,27	16.628.723,38	16.631.895,32	16.636.351,30	16.644.313,55

Tabelle 5.7: KFZ-Kasko Loginverses Gauß Modell

Für die Reserven und deren Kennzahlen erhalten wir aus den Stichproben folgende Werte

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw. $R_i^\star$
2	7.285,72	21.540,29	127,12	63.262,73	681,16
3	15.574,72	40.404,23	467,02	114.133,62	1.277,69
4	22.683,15	57.875,47	567,67	148.693,62	1.830,18
5	60.393,40	76.444,40	12.559,67	289.672,25	2.417,38
6	93.509,20	92.983,33	12.748,58	364.601,98	2.940,39
7	2.875.617,96	553.899,80	1.976.903,01	4.191.569,83	17.515,85
Gesamt	3.075.064,15	577.757,06	2.178.990,64	4.339.849,38	18.270,28

und folgende Abweichung zur tatsächlichen Reserve

SJ	R_i	$R_i^\star - R_i$
2	914,31	6.371,41
3	243,70	15.331,02
4	11.812,71	10.870,44
5	1.819,56	58.573,84
6	170.775,30	- 77.266,10
7	2.705.235,01	170.382,95
Gesamt	2.890.800,59	184.263,56

Auch hier wird die Reserve des 5. Schadenjahres deutlich überschätzt. Betrachten wir hier die jeweiligen Konfidenzintervalle der Reserven ist gut zu erkennen, dass für das 3. und 5. Schadenjahr die tatsächliche Reserve außerhalb des Konfidenzintervalls liegt.

5.4.2 Rechtsschutz

Auch bei der Sparte Rechtsschutz ist das Erste, das wir benötigen, das beobachtbare Zahlungsdreieck, um die individuellen Abwicklungsfaktoren zu berechnen. Das Zahlungsdreieck hat folgende Form

SJ \ AbwJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78		
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44			
5	363.636,21	1.116.438,19	1.623.695,82				
6	419.610,32	1.433.715,17					
7	448.107,48						

und daraus erhalten wir die logarithmierten, individuellen Abwicklungsfaktoren

SJ \ AJ	1	2	3	4	5	6	7
1	1,52963212	0,36886939	0,19126396	0,12530781	0,04304523	0,088798	
2	1,59925418	0,51450171	0,18609673	0,18196981	0,05027561		
3	1,43372527	0,37484736	0,14890776	0,12193821			
4	1,36764216	0,37951529	0,18378992				
5	1,12174476	0,37456149					
6	1,2286979						
7							

Tabelle 5.8: Logarithmierte individuelle Abwicklungsfaktoren Rechtsschutz

Um auch hier unsere Ergebnisse mit den tatsächlich eingetretenen Schadenzahlungen zu vergleichen, benötigen wir noch das Abwicklungsquadrat

SJ \ AbwJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	1.358.927,95
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78	1.539.980,31	1.628.371,44
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44	1.932.512,08	2.061.823,12	2.215.163,34
5	363.636,21	1.116.438,19	1.623.695,82	1.928.900,99	2.251.362,16	2.567.299,12	2.859.606,91
6	419.610,32	1.433.715,17	2.091.419,97	2.711.636,97	3.216.827,36	3.621.598,68	3.822.963,90
7	448.107,48	1.593.877,88	2.344.880,19	3.065.084,79	3.601.092,25	3.960.953,39	4.116.655,97

5.4.2.1 Lognormales Modell

Um die zukünftigen Schadenstände schätzen zu können, benötigen wir auch hier Schätzer für die Parameter der Verteilung. Diese erhalten wir aus den logarithmierten, individuellen Abwicklungsfaktoren und haben folgende Form

j	1	2	3	4	5	6
$\hat{\mu_j}$	1,38012	0,40246	0,17751	0,14307	0,04666	0,088798
$\hat{\sigma_j^2}$	0,027181549512	0,003149783605	0,000280107184	0,000758414447	0,000013069602	0,000004999082

Somit erhalten wir unter Verwendung des Algorithmus 5.1.2 folgendes geschätztes Zahlungsquadrat.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	1.435.753,61
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78	1.546.380,34	1.690.092,48
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44	2.048.736,96	2.146.453,06	2.345.804,13
5	363.636,21	1.116.438,19	1.623.695,82	1.940.113,32	2.237.966,95	2.344.949,02	2.562.441,12
6	419.610,32	1.433.715,17	2.144.554,80	2.560.518,09	2.953.566,36	3.094.347,87	3.381.641,77
7	448.107,48	1.799.704,72	2.683.354,10	3.204.145,59	3.695.074,77	3.871.649,44	4.231.343,54

Tabelle 5.9: Rechtsschutz Lognormales Modell

Nun da die Endschadenstände berechnet sind, können wir die Reserve und deren Kennzahlen unter Zuhilfenahme der Stichproben berechnen. Dabei erhalten wir folgende Werte

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw. $R_i^\star$
2	122.008,31	3.193,90	116.162,80	128.292,82	101,00
3	213.951,70	7.322,16	200.121,07	228.312,37	231,55
4	574.777,69	66.874,16	452.480,68	711.121,43	2.114,75
5	938.745,30	82.798,66	774.923,29	1.099.162,41	2.618,32
6	1.947.926,60	222.955,16	1.560.643,14	2.409.596,49	7.050,46
7	3.783.236,06	744.637,00	2.466.207,54	5.328.397,81	23.547,49
Gesamt	7.580.645,66	778.776,88	6.171.778,13	9.227.171,42	24.627,09

Natürlich gilt auch hier unser Interesse der Abweichung von unseren Schätzern zu den tatsächlichen Werten

SJ	R_i	$R_i^\star - R_i$
2	45.182,65	76.825,66
3	152.230,66	61.721,04
4	444.136,90	130.640,79
5	1.235.911,09	- 297.165,79
6	2.389.248,73	- 441.322,13
7	3.668.548,49	114.687,57
Gesamt	7.935.258,52	- 354.612,86

Betrachtet man die Schadenjahre ist gut zu erkennen, dass mit Ausnahme des 7. Schadenjahres die Reserven deutlich über- beziehungsweise unterschätzt werden. Betrachtet man hingegen die Gesamtreserve ist diese Fehlschätzung nicht so stark zu erkennen, da sich diese ausgleichen. Vergleicht man die tatsächliche Reserve mit den Konfidenzintervallen, ist zu erkennen, dass die Schadenjahre 2 bis 5 nicht im Konfidenzintervall liegen. Daher wäre auch hier die Verteilungsannahme nochmals zu überdenken.

5.4.2.2 Loggamma Modell

Genauso wie bei KFZ-Kasko müssen wir zur Schätzung der Parameter die Werte verwenden, die wir durch das logarithmische inverse Gauß Modell erhalten, da auch hier die Momentenmethode keine Konvergenz liefert. Daher erhalten wir als Startwerte des Newton-Verfahrens

j	1	2	3	4	5	6
$\hat{\alpha}_{j,0}$	155,601746	46,007127	20,9409153	16,6814338	6,13399122	11,014484

Nach Verwendung des Newton-Verfahrens, wobei auch hier die Abbruchbedingung

$$|\hat{\alpha}_{j,m} - \hat{\alpha}_{j,m-1}| < 10^{-6}$$

für alle $j \in \{1, .., 6\}$ ist, erhalten wir für die Verteilung als Parameter

j	1	2	3	4	5	6
$\hat{\alpha}_j$	161,210378	47,3053983	21,222471	16,9931483	5,94962162	10,912799

und

$$\hat{\lambda} = 117,3088946 \quad .$$

Nun können wir die Schadenstände berechnen, indem wir den Algorithmus 5.2.2 verwenden und so für die Schadenstände Stichproben erzeugen. So erhalten wir

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	1.442.389,93
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78	1.553.390,20	1.703.248,29
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44	2.052.574,20	2.160.125,46	2.371.828,46
5	363.636,21	1.116.438,19	1.623.695,82	1.947.589,07	2.249.453,80	2.367.252,43	2.602.933,33
6	419.610,32	1.433.715,17	2.154.794,74	2.586.528,94	2.988.745,19	3.146.088,42	3.455.462,06
7	448.107,48	1.779.799,25	2.669.588,22	3.202.384,76	3.703.632,88	3.898.049,50	4.275.235,80

Tabelle 5.10: Rechtsschutz Loggamma Modell

und für die Reserve folgende Werte

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw. $R_i^\star$
2	128.644,63	41.624,83	61.742,30	224.142,04	1.316,29
3	227.107,51	58.933,13	128.413,94	358.438,83	1.863,63
4	600.802,02	122.223,10	394.591,81	866.756,80	3.865,03
5	979.237,51	169.496,09	684.818,89	1.347.257,57	5.359,94
6	2.021.746,89	305.229,94	1.473.163,66	2.653.524,05	9.652,22
7	3.827.128,32	608.509,62	2.794.394,42	5.212.812,22	19.242,76
Gesamt	7.784.666,87	718.630,44	6.535.906,93	9.270.056,68	22.725,09

Für die Abweichung der soeben berechneten Reserve zu der tatsächlich benötigten Reserve ergibt sich folgende Tabelle

SJ	R_i	$R_i^\star - R_i$
2	45.182,65	83.461,98
3	152.230,66	74.876,85
4	444.136,90	156.665,12
5	1.235.911,09	- 256.673,58
6	2.389.248,73	- 367.501,84
7	3.668.548,49	158.579,83
Gesamt	7.935.258,52	- 150.591,65

Auch hier werden die Reserven der Schadenjahre überschätzt beziehungsweise unterschätzt, jedoch liegen hier mit Ausnahme des 2. Schadenjahres alle Reserven im angegeben Konfidenzintervall.

5.4.2.3 Loginverses Gauß Modell

So wie bei KFZ-Kasko müssen zur Berechnung der Parameter zunächst die Startwerte berechnet werden, bevor wir den Algorithmus, der im Abschnitt 5.3.1 beschrieben ist, anwenden können.
Dabei haben die Startwerte folgende Form

j	1	2	3	4	5	6
$\hat{\mu}_{j,0}$	1,380116067	0,40245905	0,177514591	0,143071944	0,04666042	0,088798

und daraus erhalten wir die optimalen Werte

j	1	2	3	4	5	6
$\hat{\mu}_j$	1,36834142	0,404580663	0,184151673	0,1466944	0,0539415	0,0968599

und

$$\hat{\beta} = 113,715585544 \quad ,$$

wobei auch hier der Algorithmus abgegbrochen wurde, sobald für alle $j \in \{1,..,6\}$

$$|\hat{\mu}_{j,m} - \hat{\mu}_{j,m-1}| < 10^{-6}$$

gilt.
Da wir nun die Parameter der Verteilung geschätzt haben, können wir den Algorithmus 5.3.2 verwenden um Stichproben für unsere Schadenstände zu erzeugen und damit das Zahlungsquadrat schätzen.

SJ \ AJ	1	2	3	4	5	6	7
1	96.455,48	445.284,62	643.925,52	779.651,50	883.732,87	922.603,96	1.008.276,85
2	104.432,33	516.871,08	864.624,40	1.041.473,61	1.249.328,81	1.313.745,30	1.447.096,97
3	184.521,83	773.938,99	1.125.902,72	1.306.684,34	1.476.140,78	1.558.122,92	1.717.873,73
4	256.817,88	1.008.288,24	1.473.689,87	1.771.026,44	2.052.481,14	2.167.568,64	2.388.825,16
5	363.636,21	1.116.438,19	1.623.695,82	1.955.099,58	2.265.398,22	2.392.613,75	2.631.728,73
6	419.610,32	1.433.715,17	2.146.908,40	2.581.678,41	2.992.369,65	3.159.374,06	3.482.386,93
7	448.107,48	1.772.741,38	2.657.300,64	3.199.618,65	3.703.390,36	3.907.929,05	4.301.477,91

Tabelle 5.11: Rechtsschutz Loginverses Gauß Modell

Betrachten wir nun die Reserven und deren Kennzahlen erhalten wir folgendes Bild

SJ	$R_i^\star$	Stdabw. R_i	2,5%-Quantil	97,5%-Quantil	Stdabw. $R_i^\star$
2	133.351,67	42.372,59	71.297,92	232.384,23	1.339,94
3	241.732,95	62.692,16	143.103,50	384.800,91	1.982,50
4	617.798,72	119.763,93	432.194,53	900.769,33	3.787,27
5	1.008.032,91	179.181,91	703.585,14	1.411.756,41	5.666,23
6	2.048.671,76	312.804,30	1.523.544,81	2.722.281,48	9.891,74
7	3.853.370,43	592.634,82	2.835.496,43	5.049.369,38	18.740,76
Gesamt	7.902.958,43	710.259,74	6.619.691,46	9.367.359,20	22.460,39

Zuletzt betrachten wir auch hier die Abweichung der geschätzten Reserven zu den tatsächlich benötigten Reserven

SJ	R_i	$R_i^\star - R_i$
2	45.182,65	88.169,02
3	152.230,66	89.502,29
4	444.136,90	173.661,82
5	1.235.911,09	- 227.878,18
6	2.389.248,73	- 340.576,97
7	3.668.548,49	184.821,94
Gesamt	7.935.258,52	- 32.300,09

Hier ist zwar die geschätzte Gesamtreserve annähernd gleich wie die tatsächliche Reserve, jedoch sind die einzelnen Schadenjahre stark über- beziehungsweise unterschätzt. Auch hier liegen die Reserven mit Ausnahme des 2. Schadenjahres immer in dem angegebenen Konfidenzintervall, jedoch teilweise sehr knapp.

Kapitel 6

Vergleich der Verfahren

In diesem Kapitel wollen wir die verschiedenen Verfahren miteinander vergleichen. Dabei ist zu erwähnen, dass die Verfahren auf unterschiedlichen Annahmen beruhen und es daher schwer ist die Differenzen zwischen den Modellen zu erklären. Jedoch können wir die Modelle mit der tatsächlichen Reserve vergleichen und so ein Bild bekommen, welche Verfahren sich anbieten würden. Dazu werden wir natürlich auch die Standardabweichung und die Konfidenzintervalle heranziehen.

6.1 KFZ-Kasko

Betrachten wir für diese Sparte die tatsächliche Reserve R_i und deren Schätzer erhalten wir

SJ	R_i	Chain Ladder	Lineares Modell	Loglineares Modell	CL Lognormal	CL Loggamma	CL Loginvers
2	914,31	634,35	682,48	632,61	631,08	11.473,70	7.285,72
3	243,70	1.616,79	1.738,09	5.938,14	1.689,09	28.592,42	15.574,72
4	11.812,71	3.504,95	4.584,79	16.901,98	3.706,28	42.824,54	22.683,15
5	1.819,56	54.467,03	69.519,30	123.044,04	47.597,64	92.000,98	60.393,40
6	170.775,30	166.970,44	201.859,34	383.539,28	154.052,55	166.916,00	93.509,20
7	2.705.235,01	2.844.333,91	3.431.126,52	3.045.046,42	2.871.283,94	2.899.629,98	2.875.617,96
Gesamt	2.890.800,59	3.071.527,48	3.709.510,52	3.575.102,47	3.078.960,58	3.241.437,61	3.075.064,15

Tabelle 6.1: Vergleich KFZ-Kasko

Hier ist gut zu erkennen, dass sowohl das klassische Chain Ladder Modell als auch die Modelle mit den stochastischen Abwicklungsfaktoren die Reserven sehr gut schätzen. Wie schon bei der Betrachtung der Modelle erwähnt, wird die Reserve beim linearen und loglinearen Modell deutlich überschätzt. Dieses Bild ist nicht verwunderlich, da wir schon bei der Betrachtung der Modelle festgestellt haben, dass die Chain Ladder Annahmen erfüllt sind und die des linearen Modells nicht zur Gänze.

Die Reserve des 5. Schadenjahres wird in allen Modellen deutlich überschätzt. Dies liegt daran, dass selbst für eine kurzabwickelnde Sparte, wie KFZ-Kasko, die nicht beobachtbaren Schadenzahlungen sehr gering sind. Betrachten wir die Standardabweichung der Reserven erhalten wir folgende Werte

SJ	Chain Ladder	Lineares Modell	Loglineares Modell	CL Lognormal	CL Loggamma	CL Loginvers
2	789,10	128,12	711,45	110,64	22.496,46	21.540,29
3	1.258,92	964,69	39.844,54	766,10	39.931,12	40.404,23
4	2.095,79	2.267,12	88.698,72	1.814,66	49.687,38	57.875,47
5	42.512,72	48.588,10	168.555,07	39.572,22	73.443,57	76.444,40
6	70.427,01	72.718,18	463.064,20	69.055,49	94.259,03	92.983,33
7	371.309,49	794.386,41	426.018,41	361.967,23	402.373,20	553.899,80
Gesamt	380.321,75	799.189,96	658.624,71	371.430,25	428.388,16	577.757,06

Da wir uns für Modelle interessieren, deren Zufallsfehler klein ist, ist auch hier das klassische Chain Ladder Modell und das Chain Ladder Modell mit den stochastischen Abwicklungsfaktoren, hier vorallem das lognormale Modell, zu wählen. Ebenfalls von Interesse ist für uns noch die Standardabweichung der Schätzer, die ebenfalls niedrig sein soll.

SJ	Chain Ladder	Lineares Modell	Loglineares Modell	CL Lognormal	CL Loggamma	CL Loginvers
2	883,96	148,87	472,83	3,50	711,40	681,16
3	1.351,58	845,80	4.210,15	24,23	1.262,73	1.277,69
4	1.680,42	1.754,41	8.964,01	57,38	1.571,25	1.830,18
5	22.483,71	28.920,89	19.839,21	1.251,38	2.322,49	2.417,38
6	34.593,84	39.804,30	31.643,17	2.183,73	2.980,73	2.940,39
7	149.482,42	349.442,29	65.549,26	11.446,41	12.724,16	17.515,85
Gesamt	219.992,53	361.584,45	98.165,36	11.745,66	13.546,82	18.270,28

Hier sind vorallem die Schätzer der stochastischen Abwicklungsfaktoren zu beachten, die aufgrund der Menge an Daten aus der die Stichprobe besteht eine geringe Standardabweichung haben. Zum Abschluss wollen wir noch die Konfidenzintervalle der Verfahren betrachten, die eine Aussage treffen sollen, in welchen Bereichen sich die Reserven befinden werden.

SJ	Chain Ladder	Lineares Modell	Loglineares Modell
2	(57,46;2.749,15)	(462,29;973,24)	(68,91;2.564,21)
3	(321,94;5.054,78)	(539,06;4.284,22)	(17,47;43.836,37)
4	(995,58;9.089,32)	(1.612,77;10.472,79)	(81,3;123.111,68)
5	(10.807,44;170.580,91)	(16.141,05;201.156,99)	(9.285,26;566.830,91)
6	(68.487,19;345.587,33)	(94.437,79;381.909,27)	(36.718,84;1.630.065,45)
7	(2.174.709,18;3.657.810,84)	(2.116.502,58;5.279.315,42)	(2.282.720,75;3.983.974,05)
Gesamt	(2.381.820,85;3.901.141,67)	(2.368.378,53;5.552.361,91)	(2.439.822,17;5.066.684,91)

SJ	CL Lognormal	CL Loggamma	CL Loginvers
2	(403,25;852,48)	(0,01;84.779,52)	(127,11;63.262,72)
3	(234,25;3.157,88)	(31,7;142.764,51)	(467,02;114.133,61)
4	(225,01;7.370,33)	(418,15;189.268,51)	(567,67;148.693,62)
5	(6.111,33;125.437,08)	(24.875,82;282.577,78)	(12.559,67;289.672,24)
6	(16.937,8;289.007,99)	(40.981,11;413.268,54)	(12.748,58;364.601,97)
7	(2.158.993,83;3.574.043,6)	(2.165.174,6;3.802.950,72)	(1.976.903;4.191.569,83)
Gesamt	(2.306.445,25;3.751.638,74)	(2.470.479,73;4.158.080,39)	(2.178.990,64;4.339.849,37)

Auch hier ist gut zu erkennen, dass das klassische Chain Ladder Verfahren sowie das stochastische Verfahren mit den logarithmisch normalverteilten Abwicklungsfaktoren, die geringste Streuung erwarten lassen würden.
Insgesamt fällt die Wahl entweder auf das Chain Ladder Modell und auf die stochastischen Abwicklungsfaktoren, hier vorallem auf das lognormale Modell. Bei der Entscheidung, welches der beiden Modelle man wählt, ist zu beachten, dass die Standardabweichung der Schätzer des klassischen Chain Ladder Modells deutlich größer sind, jedoch hier die Berechnung der Reserven und deren Kennzahlen deutlich weniger aufwendig ist.

6.2 Rechtsschutz

Auch hier wollen wir zunächst die tatsächliche Reserve R_i mit ihren Schätzern vergleichen. Dabei ist nochmals zu erwähnen, dass dies nur nachträglich möglich ist und daher nur einen Eindruck schafft, wie gut die getroffene Wahl war, jedoch keine Aussage liefert, für welches Verfahren man sich entscheidet.

SJ	R_i	Chain Ladder	Lineares Modell	Loglineares Modell	CL Lognormal	CL Loggamma	CL Loginvers
2	45.182,65	121.994,23	121.316,25	107.872,03	122.008,31	128.644,63	133.351,67
3	152.230,66	215.189,70	250.490,28	230.829,49	213.951,70	227.107,51	241.732,95
4	444.136,90	570.487,24	622.746,81	574.572,98	574.777,69	600.802,02	617.798,72
5	1.235.911,09	936.208,41	1.129.633,42	1.009.528,27	938.745,30	979.237,51	1.008.032,91
6	2.389.248,73	1.922.085,67	2.056.582,20	2.011.520,93	1.947.926,60	2.021.746,89	2.048.671,76
7	3.668.548,49	3.447.579,96	3.659.645,52	3.548.679,93	3.783.236,06	3.827.128,32	3.853.370,43
Gesamt	7.935.258,52	7.213.545,20	7.840.414,48	7.483.003,62	7.580.645,66	7.784.666,87	7.902.958,43

Tabelle 6.2: Vergleich Rechtsschutz

Hier ist gut zu erkennen, dass sich die Reserven der einzelnen Verfahren doch teilweise deutlich unterscheiden. Betrachten wir die Gesamtreserve ist gut zu erkennen, dass das lineare Modell und das Modell der stochastischen Abwicklungsfaktoren, hier vorallem das loginverse Modell, diese relativ genau schätzen. Betrachten wir jedoch die Einzelschadenreserven ist gut zu erkennen, dass diese meist über- beziehungsweise unterschätzt werden. Um nun eine Aussage treffen zu können, welches Verfahren wir wählen, betrachten wir die Standardabweichung der tatsächlichen Reserve und deren Schätzer. Dabei erhalten wir für die Standardabweichung der tatsächliche Reserve R_i folgendes Bild.

SJ	Chain Ladder	Lineares Modell	Loglineares Modell	CL Lognormal	CL Loggamma	CL Loginvers
2	6.247,11	12.890,81	47.815,10	3.193,90	41.624,83	42.372,59
3	9.916,98	16.702,81	76.501,33	7.322,16	58.933,13	62.692,16
4	62.692,49	65.413,28	131.062,91	66.874,16	122.223,10	119.763,93
5	76.321,23	83.016,75	169.733,00	82.798,66	169.496,09	179.181,91
6	162.895,40	125.604,65	226.884,52	222.955,16	305.229,94	312.804,30
7	516.139,90	174.940,44	268.502,44	744.637,00	608.509,62	592.634,82
Gesamt	550.298,07	240.824,67	421.539,63	778.776,88	718.630,44	710.259,74

Hier ist gut zu erkennen, dass das lineare und loglineare Modell die geringste Standardabweichung hat. Die stochastischen Abwicklungsfaktoren hingegen haben eine sehr hohe Standardabweichung, die zumindest teilweise auf die große Schwankung der individuellen Abwicklungsfaktoren zurückzuführen ist.
Für die Schätzer der Reserven erhalten wir als Standardabweichung folgenden Tabelle.

SJ	Chain Ladder	Lineares Modell	Loglineares Modell	CL Lognormal	CL Loggamma	CL Loginvers
2	7.454,65	15.339,74	44.136,90	101,00	1.316,29	1.339,94
3	10.647,16	22.065,67	66.365,26	231,55	1.863,63	1.982,50
4	48.637,63	56.614,21	98.499,52	2.114,75	3.865,03	3.787,27
5	58.578,16	74.816,93	122.444,86	2.618,32	5.359,94	5.666,23
6	113.643,85	104.161,23	171.868,42	7.050,46	9.652,22	9.891,74
7	302.433,65	137.296,82	263.722,34	23.547,49	19.242,76	18.740,76
Gesamt	604.212,59	366.956,45	672.142,84	24.627,09	22.725,09	22.460,39

Hier sind die Modelle mit den stochastischen Abwicklungsfaktoren deutlich niedriger als die restlichen Verfahren. Dies ist wieder auf die Verwendung der Stichproben zurückzuführen. Weiter ist zu erwähnen, dass der Schätzfehler des linearen Modells nochmal deutlich niedriger ist als der, der zwei übrigen Verfahren.
Zuletzt wollen wir auch hier die Konfidenzintervalle der Verfahre betrachten.

SJ	Chain Ladder	Lineares Modell	Loglineares Modell
2	(109.500;134.488,44)	(95.534,62;147.097,87)	(42.274,55;230.056,34)
3	(195.355,74;235.023,66)	(217.084,65;283.895,9)	(114.884,34;417.889,88)
4	(445.102,27;695.872,21)	(491.920,24;753.573,36)	(357.033,05;878.927,55)
5	(783.565,95;1.088.850,86)	(963.599,91;1.295.666,92)	(712.927,09;1.390.226,35)
6	(1.596.294,87;2.247.876,45)	(1.805.372,9;2.307.791,5)	(1.596.314,77;2.502.881,39)
7	(2.415.300,14;4.479.859,76)	(3.309.764,63;4.009.526,39)	(3.042.294,76;4.115.789,88)
Gesamt	(6.112.949,07;8.314.141,33)	(7.358.765,14;8.322.063,81)	(6.675.695,66;8.361.407,18)

SJ	CL Lognormal	CL Loggamma	CL Loginvers
2	(116.162,79;128.292,81)	(61.742,29;224.142,04)	(71.297,92;232.384,23)
3	(200.121,06;228.312,36)	(128.413,93;358.438,82)	(143.103,49;384.800,9)
4	(452.480,67;711.121,43)	(394.591,8;866.756,79)	(432.194,52;900.769,32)
5	(774.923,29;1.099.162,41)	(684.818,89;1.347.257,56)	(703.585,14;1.411.756,41)
6	(1.560.643,14;2.409.596,49)	(1.473.163,65;2.653.524,05)	(1.523.544,81;2.722.281,47)
7	(2.466.207,54;5.328.397,81)	(2.794.394,42;5.212.812,22)	(2.835.496,42;5.049.369,38)
Gesamt	(6.171.778,13;9.227.171,42)	(6.535.906,92;9.270.056,67)	(6.619.691,45;9.367.359,2)

Bei Betrachtung der Konfidenzintervalle ist zu erkennen, dass das lineare Modell jenes Verfahren ist, dass die geringste Abweichung bei der zu erwarteten Reserve liefern würde. Ebenfalls gut zu erkennen ist, dass die oberen Intervallgrenzen der stochastischen Abwicklungsfaktoren teilweise deutlich höher sind als die der anderen Verfahren.

Insgesamt wird die Wahl bei dieser Sparte auf das lineare Modell fallen, da hier sowohl der Zufallsfehler als auch der Schätzfehler gering ist. Würden unsere Wahl allein auf dem Schätzfehler basieren, der bei Schätzungen gering sein soll, würde unsere Wahl auf die stochastischen Abwicklungsfaktoren fallen. Jedoch ist bei diesen Verfahren der Zufallsfehler sehr hoch.

Kapitel 7

Schlusswort

Vergleicht man die verschiedenen Verfahren ist gut zu erkennen, dass die Daten sehr unterschiedliche Annahmen erfüllen müssen. Daher sind die Ergebnisse nur bedingt vergleichbar. Wie wir aber an den zwei Versicherungsbeständen gesehen haben, ist es notwendig, Verfahren mit verschiedenen Voraussetzungen auf die Daten anzuwenden, um einen guten Schätzer zu erhalten, da bei Verwendung von Verfahren deren Annahmen nicht erfüllt sind, die Reserven deutlich über- beziehungsweise unterschätzt werden.
Ein gute Indikator für die richtige Wahl eines Verfahrens ist die Standardabweichung des Schätzers und der tatsächlichen Reserve.

Wie bereits bei den einzelnen Verfahren erwähnt, ist der Vergleich der Schätzer mit den tatsächlichen Reserven notwendig, um herauszufinden, ob die getroffenen Annahmen korrekt gewesen sind. Jedoch ist die einmalige Durchführung nicht ausreichend, um eine gesicherte Aussage treffen zu können, da die Daten von Ausreißern beeinflusst werden können. Jedoch ist zu erwähnen, dass bei mehrmaliger Verschätzung die Wahl des Verfahrens zu überdenken ist.

Wie wir ebenfalls beobachten haben können, liefern ähnliche Verfahren teilweise auch ähnliche Schätzer und ähnliche Standardabweichungen der tatsächlichen Reserve. Hier haben wir als weiteres Entscheidungskriterium noch die Standardabweichung des Schätzers. Jedoch besteht die Möglichkeit, dass dieser Schätzfehler nur aufgrund eines erhöhten Rechenaufwandes vermindert wird. Daher gilt es in solchen Situationen abzuschätzen, ob sich dieser erhöhte Rechenaufwand lohnt.

Danksagung

Mein besonderer Dank gilt Dr. Stefan Gerhold, der meine Diplomarbeit betreut hat und mich bei Fragen schnellstmöglich unterstützt hat.
Weiter möchte ich mich sehr herzlich bei der HDI Versicherung AG bedanken, die mir die verwendeten Daten zur Verfügung gestellt hat und bei meinen Kollegen die mir stets mit Rat und Tat zur Verfügung gestanden sind.

Literaturverzeichnis

[1] Gurker W. *Angewandte Statistik*, Vorlesungsskriptum, (S 2011)

[2] Braun,C.(2006) *Der Prognosefehler des Chain-Ladder-Verfahrens bei korrelierten Abwicklungsdreiecken*,Präsentation der Munich Re Group
unter : http://www.math.tu-dresden.de/sto/schmidt/verein/
2006_Dresdner_Forum/2006_christian_braun.pdf

[3] Devroye, L.(1986) *Non-Uniform Random Variate Generation*,Springer-Verlag, New York

[4] Radtke M.,Schmidt K.D.(Hrsg.)(2004): *Handbuch zur Schadenreservierung*, 1.Auflage,Verlag Versicherungswirtschaft, Karlsruhe

[5] Kelly, M.V.(1992): *Practical Loss Reserving Method with Stochastic Development Factors*,Casualty Actuarial Society Discussion Paper Program,Seiten 355-381

[6] Mack T.(1994): *Measuring the variability of Chain Ladder reserve estimates*,Casualty Actuarial Society Forum,Spring, Vol 1, Seiten 101-182

[7] Mack T.(1993): *Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates*,ASTIN Bulletin,23;2, Seiten 213-225

[8] Schmidt, K.D.(2009): *Linear Models in Loss Reserving*,Technische Universität Dresden,
Präsentation in Rennes beim Journées d'économétrie et d'économie de l'assurance
unter: http://www.math.tu-dresden.de/sto/schmidt/talks/2009_10_23_Rennes.pdf

[9] R Core Team(2014): *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria
unter: http://www.R-project.org/