Schadenversicherungsmathematik

Prof. Dr. Michael Kohler

Sommersemester 2009

Vorlesung orientiert sich an:

Thomas Mack:
Schadenversicherungsmathematik.
Gesellschaft für Versicherungsmathematik,
Verlag Versicherungswirtschaft, 2002

1 Einführung

1.1 Schadenversicherung

Im Versicherungsvertrag (Police) verpflichtet sich das Versicherungsunternehmen gegen Erhalt eines vereinbarten, im Voraus fälligen Geldbetrags (Prämie), bei Eintritt von im Vertrag näher definierten ungewissen Ereignissen (Schäden) bestimmte, in ihrer Höhe meist vom betreffenden Ereignis abhängende Zahlungen an den Vertragspartner (Versicherungsnehmer) zu leisten, die den aus dem Ereignis resultierenden wirtschaftlichen Nachteil des Versicherungsnehmers reduzieren oder auszugleichen sollen.

Wir behandeln Schadenversicherungen, wie z. B.

- Kfz-Versicherungen
- Hausrat-Versicherung
- Versicherungen gegen Feuerschäden oder Naturgewalten

Dazu gehören **nicht** die Lebensversicherungen.

Auftretende Fragestellungen:

1) Wie groß ist der Schaden im Mittel?
2) Welchen Zuschlag für Schwankungen soll die Prämie zum Schaden im Mittel enthalten?
3) Wieviel Geld soll die Versicherung beiseite legen bei lang andauernden Schadenabwicklungen (z. B. in der Haftpflicht-Versicherung - angerichteter Schaden eines Architekten beim Hausbau)?
4) Wie verändern wir die Prämie, wenn wir einen Teil des Schadens an Versicherungsnehmer oder Rückversicherer abgeben (z. B. Selbstbeteiligung oder Deckungssumme oder Rückversicherung)?

1.2 Ausgleich im Kollektiv

Abschluss einer Versicherung statt selbständiges beiseite legen von Geld lohnt sich wegen des sogenannten **Ausgleich im Kollektiv**:
Versicherung muss (sofern die Schäden nicht vollständig positiv korreliert sind) zur Absicherung von Schwankungen weniger Geld pro Versicherten zurücklegen, als wenn das jeder einzeln machen würde.

Einfaches Beispiel (idealisiert):

Sind die Schadenhöhen $X_1, \dots X_n$ der n Versicherten unabhängig und identisch verteilt mit Erwartungswert $\mu = EX_1$ und Varianz $\sigma^2 = V(X_1) < \infty$ so gilt für $c > 0$:

$$
\begin{aligned}
&P\left[\sum_{i=1}^{n} X_i > n\cdot(\mu+c)\right]\\
=&P\left[\sum_{i=1}^{n} X_i - n\cdot\mu > n\cdot c\right]\\
\overset{\text{Tschebyscheff}}{\leq}&\ \frac{V(\sum_{i=1}^{n} X_i)}{(n\cdot c)^2} \overset{\text{Unabhängigkeit}}{=} \frac{n\cdot V(X_1)}{n^2\cdot C^2} = \frac{\sigma^2}{n\cdot c^2}.
\end{aligned}
$$

Diese Wahrscheinlichkeit ist kleiner oder gleich wie $\varepsilon > 0$, wenn gilt

$$
\frac{\sigma^2}{n\cdot c^2} \leq \varepsilon \Rightarrow c \geq \frac{\sigma}{\sqrt{n}}\cdot\frac{1}{\sqrt{\varepsilon}}.
$$

Folgerung:
Je größer die Anzahl n der Versicherten ist, umso kleiner muss der Anteil c pro Versicherten sein, so dass der Gesamtschaden nur mit kleiner Wahrscheinlichkeit den Wert

$$
n\cdot(\mu+c)
$$

übersteigt.

1.3 Bestandteile der Prämie

Wir betrachten versichertes Kollektiv mit einem jährlich (zufälligen) Gesamtschaden S. G sei die (im Moment als bekannt) vorausgesetzte Verteilungsfunktion von S, also

$$
G(x) = P[S \leq x] \quad (x\in\mathbb{R}).
$$

Dann setzen wir

$$
\text{Prämie für Kollektiv} = \frac{E(S) + \text{Schwankungszuschlag} + k_0}{1-p},
$$

wobei

p = Anteil der zur Bruttoprämie proportionalen Betriebskosten wie Provisionen oder Steuern (z. B. $p = 0.2$).

k_0 = Prämienunabhängige Kosten für das Kollektiv

Sei c das insgesamt vorhandene Sicherheits- oder Eigenkapital. Dann beschreibt die sogenannte **Illiquiditätswahrscheinlichkeit**

$$
\varepsilon = P[S > E(S) + c] = 1 - G(E(S)+c)
$$

die Wahrscheinlichkeit, dass der Gesamtschaden das Kapital des Unternehmens übersteigt (und damit die Versicherung zahlungsunfähig ist) (z. B. $\varepsilon =$

1% oder $\varepsilon = 0,1\%$ oder...) Ist z der risikolose Zins am Kapitalmarkt, dann muss (in Abhängigkeit von ε) für das Sicherheitskapital am Kapitalmarkt der Zinssatz

$$r_\varepsilon > z$$

gezahlt werden. Der Aufschlag

$$(r_\varepsilon - z) \cdot c$$

muss vom Schwankungszuschlag bezahlt werden. Damit benötigt die Berechnung des Schwankungszuschlags primär eine Schätzung der Verteilungsfunktion des Gesamtschadens, die wir im nächsten Kapitel behandeln.
Der Schwankungszuschlag für das versicherte Kollektiv muss auf die einzelnen Versicherten (mit Schadenhöhen $S_1, \ldots S_n$, also $S = S_1 + \ldots + S_n$) umgelegt werden. Sind die Schadenhöhen unabhängig, bietet sich eine Aufteilung entsprechend der Varianzen an, d. h. Versicherter i zahlt Schwankungszuschlag

$$(r_\varepsilon - z) \cdot c \cdot \frac{V(S_i)}{V(S)}$$

Vorteil dabei ist, das der Schwankungszuschlag pro Versicherter unverändert bleibt, wenn sich einige Versicherte zu einem Kollektiv zusammenschließen. Ist die Voraussetzung der Unabhängigkeit nicht erfüllt, so kann man

$$(r_{\varepsilon - z}) \cdot c \cdot \frac{\mathrm{Cov}(S_i, S)}{V(S)}$$

als Schwankungszuschlag für den i-ten Versicherten verwenden.

1.4 Schwierigkeit bei der Prämienkalkulation

Oft wird der Gesamtschaden durch einige wenige Großschäden dominiert.

Beispiel Feuerversicherungen:
Hier liegen 85% der Schäden unter dem Mittelwert, tragen aber nur 15% zur Gesamtschadenlast bei. Da von diesen Großschäden aber nur wenige Beobachtungen vorliegen, sind diese ohne starke (und evt. unrealistische) Annahmen an die zugrunde liegende Verteilung nur schlecht schätzbar.

2 Schätzung der Verteilung des Gesamtschadens

2.1 Das kollektive Modell

Bisher haben wir den Gesamtschaden im sogenannten **individuellen Modell** dargestellt als

$$S = S_2 + S_2 + \ldots + S_n,$$

also als Summe der Schäden der einzelnen Versicherten. Problem dabei ist, dass wir aufgrund der Inhomogenität der Versicherten nicht annehmen können, dass die Schäden der einzelnen Versicherten identisch verteilt sind. Statt dessen müssten wir viele Verteilungen für verschiedene homogene Untergruppen der Versicherten einzeln schätzen.

Statt dessen betrachten wir zur Schätzung des Gesamtschadens das sogenannte **kollektive Modell**, bei dem vernachlässigt wird, wer welchen Schaden verursacht hat. Statt dessen wird

$$S = \sum_{n=1}^{N} X_1$$

gesetzt, wobei $X_1, X_2, \ldots$ die (zufälligen) Höhen des 1. Schadens, 2. Schadens etc. sind, und N die **zufällige** Anzahl der Schäden ist. Hierbei wir vorausgesetzt, dass

1.) $X_1, X_2, X_3, \ldots$ unabhängig identisch verteilt sind

und dass

2.) $N, X_1, X_2, \ldots$ unabhängig sind.

Dabei kann die Annahme der identischen Verteiltheit der $X_1, X_2, \ldots$ durch Verwendung einer aus mehreren Verteilungen „zusammengesetzten"Verteilung erreicht werden, während 2. primär der Vereinfachung dient.

Es stellt sich nun heraus, dass man sowohl die Schadenzahl N als auch die Schadenhöhen X_1 pro Schadenfall gut modellieren bzw. schätzen kann, und daraus mit den Formeln von Panjer auch die Verteilung des Gesamtschaden

$$S = \sum_{i=1}^{N} X_1$$

zumindest approximativ ermitteln kann.

2.2 Modellierung der Schadenzahl

Die Anzahl von Schäden kann z. B. durch eine Poisson-Verteilung modelliert werden, d. h.

$$P[N = n] = \frac{\vartheta^n}{n!} \cdot e^{-\vartheta} \quad (n \in \mathbb{N}_0)$$

für ein $\vartheta > 0$.

Man kann zeigen: Diese tritt als Zählprozess zwischen $\exp(\lambda)$-verteilten unabhängigen Wartezeiten im Intervall $[0, T]$ auf, wobei dann $\vartheta = \lambda \cdot T$. D. h. wenn die Wartezeiten zwischen zwei Schäden unabhängig $\exp(\lambda)$-verteilt ist, ist die Schadenzahl $\pi(\lambda \cdot T)$-verteilt.
Für eine $\pi(\vartheta)$-verteilte $ZV \quad N$ gilt:

$$E(N) = \vartheta, \quad V(N) = \vartheta.$$

Bei Vorliegen von Beobachtungen $n_1 \dots n_k$ von N kann ϑ mit Hilfe des Maximum-Likelihood-Prinzips geschätzt werden:

$$\begin{aligned}\hat{\vartheta} &= \operatorname{argmax}_{\vartheta>0} \prod_{i=1}^{k} P_\vartheta[N = \vartheta_n] = \operatorname{argmax}_{\vartheta>0} \prod_{i=1}^{k} \frac{\vartheta^{n_i}}{n_i!} \cdot e^{-\vartheta} \\ &\overset{(!)}{=} \frac{1}{k} \sum_{i=1}^{k} n_i.\end{aligned}$$

Erweiterungen

a) Poisson-Verteilung volumenabhängig definieren.
Da die Summe von K unabhängigen $\pi(\vartheta)$-verteilten Zufallsvariablen $\pi(K \cdot \vartheta)$-verteilt ist, bietet es sich an, die jährliche Schadenzahl durch eine

$$\pi(K_t \cdot \vartheta)\text{-verteilt } ZV$$

zu modellieren, wobei K_t die Anzahl der Versicherten im Jahr t ist.

b) Einbeziehung von oszillatorischen jährlichen Einflüsse.
Häufig schwanken die äußeren Bedingungen pro Jahr, z. B. Winter mit oder ohne Glatteis in der Kfz-Haftpflichtversicherung. Hier kann man die oszillatorischen jährlichen Einflüsse durch unabhängige identisch verteilte Zufallsvariablen $Q_t > 0$ (Qualität von Jahr t) modellieren, d. h. man definiert die Verteilung von N durch

$$P[N = n | Q_t = q] = \frac{(\vartheta q)^n}{n!} \cdot e^{-\vartheta q} \quad (n \in \mathbb{N})$$

und Angabe der Verteilung von Q_t.

Wählt man für Q_t z. B. eine Gamma-Verteilung mit Parametern $\lambda, \nu > 0$, d. h. mit Dichte

$$f(x) = \begin{cases} \frac{\lambda^\nu}{\Gamma(\nu)} \cdot x^{\nu-1} \cdot e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0, \end{cases}$$

so gilt:

$$
\begin{aligned}
P[N=n] &= \int P[N=n|Q_t=q] \quad dQ_t(q) \\
&= \int_{\mathbb{R}} \frac{(\vartheta q)^n}{n!} \cdot e^{-\vartheta q} \cdot f(q)\, dq \\
&= \int_{\mathbb{R}_+} \frac{(\vartheta x)^n}{n!} \cdot e^{-\vartheta x} \cdot \frac{\lambda^\nu}{\Gamma(\nu)} \cdot x^{\nu-1} \cdot e^{-\lambda x}\, dx \\
&= \frac{\vartheta^n}{n!} \cdot \frac{\lambda^\nu}{\Gamma(\nu)} \cdot \int_{\mathbb{R}_+} x^{n+\nu-1} \cdot e^{-(\lambda+\vartheta)\cdot x}\, dx \\
&= \frac{\vartheta^n}{n!} \cdot \frac{\lambda^\nu}{\Gamma(\nu)} \cdot \frac{\Gamma(n+\nu)}{(\lambda+\vartheta)^{n+\nu}} \underbrace{\int_{\mathbb{R}_+} \frac{(\lambda+\vartheta)^{n+\nu}}{\Gamma(n+\nu)} \cdot x^{n+\nu-1} \cdot e^{-(\lambda+\vartheta)\cdot x} dx}_{=1}
\end{aligned}
$$

Mit

$$
\Gamma(k) = \int_0^\infty \underbrace{x^{k-1}}_{\substack{u(x) \\ u'(x)=(k-1)\cdot x^{k-2}}} \cdot \underbrace{e^{-x}}_{v'(x)} dx = (k-1)\cdot\Gamma(k-1)
$$

folgt

$$
\begin{aligned}
P[N=n] &= \frac{\vartheta^n}{n!} \cdot \frac{\lambda^\nu}{(\lambda+\vartheta)^{n+\nu}} \cdot (n+\nu-1)\cdot(n+\nu-2)\cdot \ldots \cdot \nu \\
&= \binom{n+\nu-1}{n} \cdot \left(\frac{\lambda}{\lambda+\vartheta}\right)^\nu \cdot \left(\frac{\vartheta}{\lambda+\vartheta}\right)^n \\
&= \binom{n+\nu-1}{n} \cdot p^\nu \cdot (1-p)^n \\
&\text{mit } p = \frac{\lambda}{\lambda+\vartheta}, \text{ d. h. } 1-p = \frac{\vartheta}{\lambda+\vartheta}.
\end{aligned}
$$

Also ist in diesem Fall N negativ binomialverteilt.

2.3 Modellierung der Schadenhöhen pro Schadenfall

Empirisch kann hier eine Reihe von Modellen als sinnvoll betrachtet werden. Im Folgenden behandeln wir von diesen die Lognormalverteilung:

Definition:
Ist X $N(\mu, \sigma)$-verteilt, so heißt die Verteilung von

$$Y = \exp(X) \quad \textbf{Lognormalverteilung}.$$

Hierbei ist $\log(Y)$ normalverteilt, was den Namen motiviert.

Wir leiten zunächst die Dichte der Lognormalverteilung her:
Für die Verteilungsfunktion F von Y gilt für $y > 0$:

$$\begin{aligned} F(y) &= P[Y \leq y] = P[\exp(X) \leq y] \\ &= P[X \leq \log(y)] \\ &= \int_{-\infty}^{\log(y)} f(x)\,dx, \end{aligned}$$

wobei f die Dichte von X ist. Durch Ableiten erhält man daraus die Dichte g von Y, d. h. für $y > 0$ gilt: $g(y) = P'(y) = f(\log(y)) \cdot \frac{1}{y}$ (sowie $g(y) = 0$ für $y \leq 0$).

Mit

$$f(x) = \frac{1}{\sqrt{2\alpha}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

folgt also für die Dichte g von Y:

$$g(y) = \begin{cases} \frac{1}{\sqrt{2\alpha\sigma}} \cdot \frac{1}{y} \cdot \exp\left(-\frac{(\log(y)-\mu)^2}{2\sigma^2}\right) & , y > 0 \\ 0 & , y \leq 0. \end{cases}$$

Anpassung der Lognormalverteilung an beobachteten Daten (also der Parameter μ und σ^2 an gegebene Schadenhöhen $y_1, \ldots, y_n$) kann mit Hilfe des Maximum-Likelihood-Prinzips erfolgen:

$$\begin{aligned} (\hat{\mu}, \hat{\sigma}^2) &= \text{argmax}_{(\mu,\sigma^2)} \quad L(\mu, \sigma^2) \\ &\text{mit} \\ L(\mu, \sigma^2) &= \prod_{i=1}^{n} g(y_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\alpha}\sigma} \cdot \frac{1}{y_i} \cdot e^{-\frac{(\log y_i - \mu)^2}{2\sigma^2}} \\ &= \left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\alpha}\, y_i}\right) \cdot \sigma^{-n} \cdot \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\log\, y_i - \mu)^2\right) \end{aligned}$$

Nullsetzen der partiellen Ableitungen der Loglikelihood-Funktion führt auf

$$\begin{aligned} 0 &\overset{(!)}{=} \frac{\partial}{\partial\mu} \log L(\mu, \sigma^2) = \frac{\partial}{\partial\mu} \left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\log(y_i) - \mu)^2\right] \\ &= -\frac{1}{2\sigma^2} \cdot \sum_{i=1}^{n} 2 \cdot (\log(y_i) - \mu) \cdot (-1) \\ &= \frac{1}{\sigma^2} \cdot \left(\sum_{i=1}^{n} \log(y_i) - n \cdot \mu\right) \\ \Rightarrow \hat{\mu} &= \frac{1}{n} \sum_{i=1}^{n} \log(y_i) \end{aligned}$$

sowie

$$\begin{aligned}
0 &= \frac{\partial}{\partial\sigma^2}\log\, L(\mu,\sigma^2)\\
&= 0 + \frac{\partial}{\partial\sigma^2}\left[-\frac{n}{2}\cdot\log\sigma^2\right] + \frac{\partial}{\partial\sigma^2}\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log(y_i)-\mu)^2\right]\\
&= -\frac{n}{2}\cdot\frac{1}{\sigma^2} + \frac{1}{2}\cdot\frac{1}{(\sigma^2)^2}\sum_{i=1}^{n}(\log(y_i)-\hat{\mu})^2\\
\Rightarrow \hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(\log(y_i)-\hat{\mu})^2 = \frac{1}{n}\sum_{i=1}^{n}(\log(y_i)-\hat{\mu})^2.
\end{aligned}$$

2.4 Gesamtschadenverteilung

Im kollektiven Modell modellieren wir den Gesamtschaden durch

$$S = \sum_{n=1}^{N} X_n,$$

wobei N die Schadenzahl ist und $X_1, X_2, \ldots$ die Schadenhöhen sind. Dabei setzen wir voraus, dass die Schadenhöhen unabhängig und identisch verteilt sind, und dass die Schadenzahl von den Schadenhöhen unabhängig ist.

Die letzte Annahme ist nicht immer realistisch: Z. B. führt in der Kfz-Haftpflicht-Versicherung ein Winter mit Glatteis zu vielen kleineren Schäden, d. h. dann ist N groß und gleichzeitig sind die Schadenhöhen klein.

Im kollektiven Modell lassen sich Erwartungswert und Varianz des Gesamtschadens leicht aus den entsprechenden Größen für Schadenhöhe und Schadenzahl bestimmen, denn es gilt:

Lemma 2.1. *Sind $X_1, X_2, \ldots$ unabhängig identisch verteilte reelle Zufallsvariablen, und ist N eine $\mathbb{N}_0$-wertige Zufallsvariable, die von $X_1, X_2, \ldots$ unabhängig ist, so gilt:*

a) $E\left[\sum_{i=1}^{N} X_i\right] = EN \cdot EX_1.$

b) $V\left(\sum_{i=1}^{N} X_i\right) = EN \cdot V(X_1) + V(N)\cdot(E(X_1))^2.$

Beweis: a) $0BdA$ gilt $X_1 \geq 0$ (sonst zerlegen wie die X_1 in Positiv- und Negativanteil). Durch Anwendung des Satzes von der monotonen Konvergenz erhalten wir dann:

$$
\begin{aligned}
E\left[\sum_{i=1}^{N} X_i\right] &= E\left(\sum_{k=0}^{\infty} 1_{[N=k]} \cdot \sum_{n=1}^{k} X_n\right) \\
&= \sum_{k=0}^{\infty} E\left[1_{[N=k]} \cdot \sum_{n=1}^{k} X_n\right] \\
&\overset{(\text{Unabhängigkeit})}{=} \sum_{k=0}^{\infty} P[N=k] \cdot E\left(\sum_{n=1}^{k} X_1\right) \\
&= EX_1 \cdot \sum_{k=0}^{\infty} k \cdot P[N=k] \\
&= EX_1 \cdot EN.
\end{aligned}
$$

b) Analog erhalten wir

$$
\begin{aligned}
E\left(\left(\sum_{i=1}^{N} X_i\right)^2\right) &= E\left(\sum_{k=0}^{\infty} 1_{[N=k]} \left(\sum_{n=1}^{k} X_n\right)^2\right) \\
&= \sum_{k=0}^{\infty} P[N=k] \cdot E\left(\left(\sum_{n=1}^{k} X_n\right)^2\right).
\end{aligned}
$$

Mit

$$
\begin{aligned}
E\left(\left(\sum_{n=1}^{k} X_n\right)^2\right) &= V\left(\sum_{n=1}^{k} X_n\right) + \left(E\left(\sum_{n=1}^{k} X_n\right)\right)^2 \\
&\overset{(\text{Unabhängigkeit})}{=} \sum_{n=1}^{k} V(X_n) + \left(E\left(\sum_{n=1}^{k} X_n\right)\right)^2 \\
&\overset{(\text{Identische Verteiltheit})}{=} k \cdot V(X_1) + (k \cdot EX_1)^2
\end{aligned}
$$

folgt

$$
\begin{aligned}
E\left(\left(\sum_{i=1}^{N} X_i\right)^2\right) &= \sum_{k=0}^{\infty} P[N=k](k \cdot V(X_1) + k^2 \cdot (EX_1)^2) \\
&= V(X_1) \cdot \sum_{k=0}^{\infty} k \cdot P[N=k] + (EX_1)^2 \cdot \sum_{k=0}^{\infty} k^2 \cdot P[N=k] \\
&= V(X_1) \cdot EN + (EX_1)^2 \cdot E(N^2).
\end{aligned}
$$

Damit erhalten wir

$$\begin{aligned}
V\left(\sum_{i=1}^{N} X_i\right) &= E\left(\left(\sum_{i=1}^{N} X_i\right)^2\right) - \left(E\left(\sum_{i=1}^{N} X_i\right)\right)^2 \\
&= V(X_1)\cdot EN + (EX_1)^2 \cdot E(N^2) - (EN \cdot EX_1)^2 \\
&= V(X_1)\cdot EN + (EX_1)^2 \cdot (E(N^2) - (EN)^2) \\
&= V(X_1)\cdot EN + (EX_1)^2 \cdot V(N). \quad \Box
\end{aligned}$$

Im Folgenden wollen wir nun die Verteilung des Gesamtschadens (z. B. beschrieben durch die Verteilungsfunktion) bestimmen, um damit wie in Abschnitt 1.3 den Schwankungszuschlag unter Verwendung der Illiquiditätswahrscheinlichkeit bestimmen zu können. Dazu approximieren wir zunächst die Verteilung P_{X_1} der Schadenhöhen durch eine diskrete Verteilung $P_{\underline{X}_1}$, die auf

$$\{k \cdot h \colon k = 0, 1, \ldots, K\}$$

konzentriert ist. $(f_k)_k = 0, 1, \ldots, K$ sei die Zähldichte von $P_{\underline{X}_1}$, d. h.

$$f_k = P[\underline{X}_1 = k \cdot h] \quad (k = 0, 1, \ldots, K).$$

Eine solche Approximation erhalten wir z. B. in dem wir setzen

$$f_k = P\left[\left(k - \frac{1}{2}\right) \cdot h < X_1 \leq \left(k + \frac{1}{2}\right) \cdot h\right]$$

für $k = 0, 1, \ldots, K$. Alternative Vorgehensweisen, die die Momente erhalten, finden sich in Mack (2002). Sodann approximieren wir den Gesamtschaden

$$S = \sum_{n=1}^{N} X_n$$

durch

$$\underline{S} = \sum_{n=1}^{N} \underline{X}_n.$$

$\underline{S}$ ist eine diskrete Zufallsvariable, deren Verteilung auf $\{k \cdot h \colon k \in \mathbb{N}_0\}$ konzentriert ist. Deren Zähldichte $(g_k)_{k\in\mathbb{N}_0}$ mit

$$g_k = P[\underline{S} = k \cdot h]$$

können wir mit Hilfe des folgenden Satzes bestimmen.

Satz 2.1. *(Formeln von Panjer).*
Sind $\underline{X}_1, \underline{X}_2, \ldots$ unabhängige identisch verteilte und auf

$$\{k \cdot h \colon k = 0, 1, \ldots, K\}$$

konzentrierte Zufallsvariablen mit Zähldichte

$$f_k = P[\underline{X}_1 = k \cdot h] \quad (k = 0, 1, \ldots, K),$$

und ist N eine von $\underline{X}_1, \underline{X}_2, \ldots$ unabhängige $\mathbb{N}_0$-wertige Zufallsvariable, für deren Zähldichte

$$p_n = P[N = n] \quad (n \in \mathbb{N}_0)$$

die Rekursion

$$n \cdot p_n = (n \cdot a + b) \cdot p_{n-1} \quad (n \in \mathbb{N}) \tag{2.1}$$

für Zahlen $a, b \in \mathbb{R}$ gilt, so lässt sich die Zähldichte

$$g_k = P[\underline{S} = k \cdot h] \quad (n \in \mathbb{N})$$

von

$$\underline{S} = \sum_{n=1}^{N} \underline{X}_n$$

rekursiv wie folgt bestimmen:

$$g_0 = \begin{cases} p_0 \cdot \exp(f_0 \cdot b) & \textit{für } a = 0 \\ \frac{p_0}{(1 - f_0 \cdot a)^{1 + \frac{b}{a}}} & \textit{für } a \neq 0 \end{cases} \tag{2.2}$$

und

$$g_k = \frac{1}{(1 - f_0 \cdot a)} \sum_{j=1}^{k} (a + b \cdot \frac{j}{k}) \cdot f_j \cdot g_{k-j} \textit{ für } k \geq 1. \tag{2.3}$$

Bemerkung: Voraussetzung (2.1) ist insbesondere für die Poisson-Verteilung erfüllt, da für $n \in \mathbb{N}$ gilt:

$$n \cdot \frac{\lambda^n}{n!} \cdot e^{-\lambda} = (n \cdot 0 + \lambda) \cdot \frac{\lambda^{n-1}}{(n-1)!} \cdot e^{-\lambda}.$$

Also gilt hier (2.1) mit $a = 0$ und $b = \lambda$.

Beweis von Satz 2.1:
Seien u, v und w die erzeugenden Funktionen von $N, \underline{X}_1/h$ und $\underline{S}/h$, d. h. für $z \in (-1, 1)$ gilt:

$$\begin{aligned} u(z) &= E(z^N) = \sum_{n=0}^{\infty} P[N = n] z^n = \sum_{n=0}^{\infty} p_n \cdot z^n, \\ v(z) &= E\left(z^{\underline{X}_1/h}\right) = \sum_{k=0}^{\infty} P[\underline{X}_1 = k \cdot h] \cdot z^k = \sum_{k=0}^{\infty} f_k \cdot z^k, \\ w(z) &= E\left(z^{\underline{S}/h}\right) = \sum_{k=0}^{\infty} P[\underline{S} = k \cdot h] \cdot z^k = \sum_{k=0}^{\infty} g_k \cdot z^k. \end{aligned}$$

Im ersten Schritt des Beweises zeigen wir:

$$w'(z) = a \cdot w'(z) \cdot v(z) + (a+b) \cdot v'(z) \cdot w(z) \text{ für } z \in (-1,1). \tag{2.4}$$

Dazu betrachten wir

$$\begin{aligned} w(z) &= E\left(z^{\sum_{n=1}^{N} \underline{X}_n/h}\right) \\ &= E\left(\sum_{k=0}^{\infty} I_{[N=k]} \cdot z^{\sum_{n=1}^{k} \underline{X}_n/h}\right). \end{aligned}$$

Anwendung des Satzes von der majorisierten Konvergenz (mit Majorante

$$\sum_{k=0}^{\infty} I_{[N=k]} \cdot 1,$$

die Erwartungswert 1 hat) liefert:

$$\begin{aligned} w(z) &= \sum_{n=0}^{\infty} E\left(I_{[N=k]} \cdot z^{\sum_{n=1}^{k} \underline{X}_n/h}\right) \\ &\overset{(\text{Unabhängigkeit})}{=} \sum_{n=0}^{\infty} P[N=k] \cdot \prod_{n=1}^{k} E\left(z^{\underline{X}_n/h}\right) \\ &= \sum_{n=0}^{\infty} P[N=k] \cdot v(z)^k \\ &= u(v(z)), \end{aligned}$$

d.h. wir haben gezeigt: $w(z) = u(v(z))$.

Ableiten liefert $w'(z) = u'(v(z)) \cdot v'(z)$.

Mit

$$\begin{aligned} u'(z) &= \sum_{n=1}^{\infty} n \cdot p_n \cdot z^{n-1} \\ &\overset{(2.1)}{=} \sum_{n=1}^{\infty} (n \cdot a + b) \cdot p_{n-1} \cdot z^{n-1} \\ &= \sum_{n=1}^{\infty} ((n-1) \cdot a + (a+b)) \cdot p_{n-1} \cdot z^{n-1} \\ &= a \cdot z \cdot \sum_{n=1}^{\infty} (n-1) \cdot p_{n-1} \cdot z^{n-2} + (a+b) \sum_{n=1}^{\infty} p_{n-1} \cdot z^{n-1}, \end{aligned}$$

also

$$u'(z) = a \cdot z \cdot u'(z) + (a+b) \cdot u(z) \tag{2.5}$$

folgt daraus

$$\begin{aligned} w'(z) &= (a \cdot v(z)u'(v(z)) + (a+b) \cdot u(v(z))) \cdot v'(z) \\ &= a \cdot v(z) \cdot u'(v(z)) \cdot v'(z) + (a+b) \cdot u(v(z)) \cdot v'(z) \\ &= a \cdot v(z) \cdot w'(z) + (a+b) \cdot w(z) \cdot v'(z), \end{aligned}$$

womit (2.4) gezeigt ist.

Im zweiten Schritt des Beweises zeigen wir (2.3) durch Koeffizientenvergleich in (2.4). Nach (2.4) gilt

$$w'(z) = a \cdot w'(z) \cdot v(z) + (a+b) \cdot v'(z) \cdot w(z).$$

Für die linke Seite erhalten wir:

$$w'(z) = \sum_{k=1}^{\infty} k \cdot g_k \cdot z^{k-1} = \sum_{k=0}^{\infty} (k+1) \cdot g_{k+1} \cdot z^k.$$

Die rechte Seite ist gleich

$$\begin{aligned} & a \cdot \left(\sum_{k=0}^{\infty} (k+1) \cdot g_{k+1} \cdot z^k \right) \cdot \left(\sum_{m=0}^{\infty} f_m \cdot z^m \right) \\ & + (a+b) \cdot \left(\sum_{m=0}^{\infty} (m+1) \cdot f_{m+1} \cdot z^m \right) \cdot \left(\sum_{k=0}^{\infty} g_k \cdot z^k \right) \\ = & \sum_{k=0}^{\infty} c_k \cdot z^k, \end{aligned}$$

wobei nach dem Multiplikationssatz für die Potenzreihen gilt:

$$c_k = a \cdot \sum_{m=0}^{k} (m+1) \cdot g_{m+1} \cdot f_{k-m} + (a+b) \cdot \sum_{m=0}^{k} (m+1) \cdot f_{m+1} \cdot g_{k-m}.$$

Daher führt ein Koeffizientenvergleich in (2.4) auf

$$(k+1) \cdot g_{k+1} = a \cdot \sum_{m=0}^{k} (m+1) \cdot g_{m+1} \cdot f_{k-m} + (a+b) \cdot \sum_{m=0}^{k} (m+1) \cdot f_{m+1} \cdot g_{k-m}$$

bzw.

$$
\begin{aligned}
&(k+1)\cdot g_{k+1} - a\cdot(k+1)\cdot g_{k+1}\cdot f_0\\
&\qquad = a\cdot\sum_{m=0}^{k-1}(m+1)\cdot g_{m+1}\cdot f_{k-m} + (a+b)\cdot\sum_{m=0}^{k}(m+1)\cdot f_{m+1}\cdot g_{k-m}\\
&\qquad = a\cdot\sum_{m=0}^{k} m\cdot g_m\cdot f_{k-m+1} + (a+b)\cdot\sum_{m=0}^{k}(k-m+1)\cdot f_{k-m+1}\cdot g_m,
\end{aligned}
$$

wobei in der ersten Summe der Summand für $m=0$ verschwindet, und in der zweiten Summe eine Indextransformation durchgeführt wurde, bei der m durch $k-m$ ersetzt wurde.

Dies impliziert

$$
\begin{aligned}
&(k+1)\cdot g_{k+1}\cdot(1-a\cdot f_0)\\
&\qquad = \sum_{m=0}^{k}\left(a\cdot m + (a+b)\cdot(k-m+1)\right)\cdot g_m\cdot f_{k-m+1}\\
&\qquad = \sum_{m=0}^{k}\left(a\cdot(k+1) + b\cdot(k-m+1)\right)\cdot g_m\cdot f_{k-m+1}\\
&\qquad = \sum_{j=1}^{k+1}\left(a\cdot(k+1) + b\cdot j\right)\cdot f_j\cdot g_{k+1-j}
\end{aligned}
$$

wobei in der letzten Summe $k-m+1$ durch j ersetzt wurde.

Insgesamt ist damit gezeigt:

$$
g_{k+1} = \frac{1}{1-a\cdot f_0}\sum_{j=1}^{k+1}\left(a + \frac{b}{k+1}\cdot j\right)\cdot f_j\cdot g_{k+1-j}
$$

für $k\in\mathbb{N}_0$, was (2.3) impliziert.

Im dritten (und letzten) Teil des Beweises zeigen wir (2.2). Nach (2.5) gilt

$$
u'(z) = a\cdot u'(z)\cdot z + (a+b)\cdot u(z),
$$

was impliziert

$$
\frac{u'(z)}{u(z)} = \frac{a+b}{1-a\cdot z}
$$

Daher gilt:

$$
\begin{aligned}
\log(u(z)) - \log(u(0)) \quad &= \quad \int_0^z \frac{d}{dt}\log(u(t))dt = \int_0^z \frac{u'(t)}{u(t)}dt\\
&\overset{\text{(s.o.)}}{=} \quad \int_0^z \frac{a+b}{1-a\cdot t} = \begin{cases} b\cdot z & \text{, falls } a=0\\ \frac{a+b}{-a}\cdot\log(1-a\cdot z) & \text{, falls } a\neq 0.\end{cases}
\end{aligned}
$$

Damit erhalten wir

$$
\begin{aligned}
g_0 &= w(0) \overset{\text{(Def.)}}{=} u(v(0)) = u(f_0) = \frac{u(f_0)}{u(0)} \cdot u(0) \\
&= \exp(\log u(f_0) - \log u(0)) \cdot p_0 \\
&\overset{\text{(s.o.)}}{=} \begin{cases} p_0 \cdot \exp(b \cdot f_0) & \text{, falls } a = 0 \\ p_0 \cdot \exp(-\frac{a+b}{a} \cdot \log(1 - a \cdot f_0)) & \text{, falls } a \neq 0 \end{cases} \\
&= \begin{cases} p_0 \cdot \exp(b \cdot f_0) & \text{, falls } a = 0 \\ p_0/(1 - a \cdot f_0)^{1+b/a} & \text{, falls } a \neq 0 \end{cases} . \quad \square
\end{aligned}
$$

3 Tarifkalkulation

3.1 Einführung

In diesem Kapitel soll für alle Risiken eines Versicherungszweiges (wie z. B. der Autohaftpflicht) der mittlere zukünftige Schaden pro Risiko geschätzt werden. Aus dieser Schätzung kann man unter Berücksichtigung des Schwankungszuschlags, der Betriebskosten und des gewünschten Gewinns wie in Abschnitt 1.3 erläutert die Prämie berechnen.

Im Folgenden werden wir zunächst die Gesamtheit der Risiken in Tarifklassen unterteilen, wobei innerhalb jeder Tarifklasse der mittlere zukünftige Schaden möglichst homogen sein soll, und anschließend pro Tarifklasse den mittleren zukünftigen Schaden schätzen.

3.2 Bildung von Tarifklassen

Im Folgenden bilden wir Tarifklassen, um jeweils Gruppen von Versicherten bei der Schätzung des mittleren Schadens zur Verfügung zu haben. Dazu gehen wir wie folgt vor:

1.) Auflistung aller Risikomerkmale, die einen potentiellen Einfluss auf den Schadenverlauf haben.
z. B. in der Kfz-Haftpflicht:
    - Fahrzeugmodell
    - Jahresfahrleistung
    - Fahrgebiet ...

2.) Bei jedem Risikomerkmal Ausprägungen mit ähnlichem Einfluss auf die Schadenerfahrung zu Ausprägungsklassen zusammenfassen.
Z. B. bei Jahresfahrleistung (in km) könnten wir Klassen bilden wie [0,10.000), [10.000, 15.000), [15.000, 20.000),...

3.) Diejenigen Risikomerkmale auswählen, die einen deutlichen Einfluss auf den Schadenverlauf haben.

Im Weiteren stellen wir exemplarisch jeweils ein Verfahren zur Durchführung von 2.) bzw. 3.) vor.

3.2.1 Ein agglomeratives Verfahren zur Bildung von Ausprägungsklassen

Wir betrachten ein nominal skaliertes Risikomerkmal mit vielen Ausprägungen (z. B. Postleitzahlengebiete oder Betriebsarten (Apotheken bis Zeitschriftenhandel)).

Zu Beginn ist jede Ausprägung in einer einzelnen Klasse enthalten. Dann werden solange Klassen miteinander verschmolzen, solange sich der erwartete Schaden nicht zu sehr unterscheidet.

Im Falle von log-normalverteilten Schadenhöhen können wir dazu einen Zweistichproben t-Test für die Gleichheit von Erwartungswerten verwenden: Sind nämlich $S_{1,1}, \ldots, S_{n_1,1}, S_{1,2}, \ldots, S_{n_2,2}$ die Schadenhöhen in Klasse 1 bzw. in Klasse 2, so setzen wir

$$X_1 = \log(S_{1,1}), \ldots, X_{n_1} = \log(S_{n_1,1})$$

als unabhängig $\mathcal{N}(\mu_X, \sigma^2)$-verteilt, und

$$Y_1 = \log(S_{1,2}), \ldots, Y_{n_2} = \log(S_{n_2,2})$$

als unabhängig $\mathcal{N}(\mu_Y, \sigma^2)$-verteilt voraus, wobei $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ unabhängig sind.

Der Zweistichproben t-Test lehnt

$$H_0 : \mu_X = \mu_Y$$

zum Niveau $\alpha \in (0, 1)$ ab, falls

$$\left| \sqrt{\tfrac{n_1 \cdot n_2}{n_1 + n_2}} \cdot \tfrac{\bar{X} - \bar{Y}}{S} \right| > t_{n_1+n_2-2;\,\alpha}$$

ist, wobei

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i, \quad \bar{Y} = \frac{1}{n_j} \sum_{j=1}^{n_2} Y_j$$

und

$$S^2 = \frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}{n_1 + n_2 - 2}$$

ist. Hierbei ist $t_{n_1+n_2-2;\,\alpha}$ das α-Fraktil der t-Verteilung mit $n_1 + n_2 - 2$ Freiheitsgraden.

Wir verschmelzen nun solange Klassen, solange die Hypothese der Erwartungswerte durch den t-Test zu einem vorgegebenen Niveau (z. B. α= 5%) nicht abgelehnt wird. Dabei kann zuerst noch mit dem F-Test überprüft werden, ob die Voraussetzung der Gleichheit der Varianzen in beiden Klassen überhaupt erfüllt ist.

3.2.2 Auswahl der Tarifmerkmale

Meist ist die Zahl der Risikomerkmale, die für sich alleine betrachtet einen Einfluss auf den Schadenverlauf haben, sehr groß. Aber es bestehen gegenseitige Abhängigkeiten zwischen den Risikomerkmalen, weswegen nicht alle gleichzeitig verwendet werden müssen. Z. B. besteht in der Kfz-Haftpflicht eine Abhängigkeit zwischen Geschlecht und Wagnisstärke, da Männer die stärker motorisierten PKWs fahren.

Wir gehen bei der Auswahl der Tarifmerkmale wieder schrittweise vor: Wir nehmen sukzessive das Risikomerkmal hinzu, das bei festgehaltenen Werten der anderen bereits ausgewählten Risikomerkmalen den deutlichsten Einfluss auf den Schadenverlauf hat. Anschließend streichen wir alle die Risikomerkmale, die nach Aufnahme des aktuellen Risikomerkmals keinen Einfluss auf den Schadenverlauf mehr haben.

Zur Entscheidung, ob ein Einfluss auf den Schadenverlauf vorliegt, kann bei log-normalverteilten Schadenhöhen die einfaktorielle Varianzanalyse (angewendet auf die logarithmierten Schadenhöhen) verwendet werden. Bei dieser sind unabhängige normalverteilte Zufallsvariablen

$$X_1^{(1)}, \ldots, X_{n_1}^{(1)}, X_1^{(2)}, \ldots, X_{n_2}^{(2)}, \ldots, X_1^{(k)}, \ldots, X_{n_k}^{(k)}$$

gegeben, wobei

$$X_1^{(j)}, \ldots, X_{n_j}^{(j)} \text{ identisch } \mathcal{N}(\mu_j, \sigma^2) - \text{verteilt}$$

sind ($j = 1, \ldots, k$).

Die Hypothese

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$$

wird dann abgelehnt, falls

$$\frac{SS_1^2}{SS_2^2} > F_{k-1,n-1;\alpha}$$

ist, wobei

$$SS_1^2 = \frac{1}{k-1} \sum_{j=1}^{k} n_j \cdot (\bar{X}^{(j)} - \bar{X})^2$$

mit

$$\bar{X} = \frac{1}{n} \sum_{j=1}^{k} \sum_{i=1}^{n_j} X_i^{(j)},$$

$$n = n_1 + \ldots + n_k,$$

$$\bar{X}^{(j)} = \frac{1}{n_j} \sum_{i=1}^{n_j} X_i^{(j)}$$

und

$$SS_2^2 = \frac{1}{n-k} \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_i^{(j)} - \bar{X}^{(j)})^2$$

Dabei muss das Verfahren modifiziert werden, sofern bereits mindestens ein Tarifmerkmal ausgewählt wurde, vgl. Mack (2002), Abschnitt 2.3.3.

3.3 Schätzung des mittleren zukünftigen Schadens pro Tarifklasse

Im Allgemeinen verwendet ein Tarif mehr als nur ein Risikomerkmal. Z. B. gilt beim deutschen PKW-Haftpflicht-Tarif:

Risikomerkmal	Anzahl Unterteilungen
Fahrzeugstärke	11
Tarifgruppe (Fahrgebiete und Berufsgruppen)	17
individueller Schadenverlauf	22

Damit gibt es hier insgesamt 11*17*22= 4114 Tarifklassen (oder Zellen).

Für eine isolierte Schätzung des mittleren zukünftigen Schadens für jede einzelne Tarifklasse liegen nun aber eventuell zu wenige Daten vor. Daher verwendet man zur Schätzung auch Werte von „Nachbarzellen“.

Üblich ist hierbei die Verwendung von sogenannten **Marginalfaktoren**: Die Schätzung für eine Tarifklasse ergibt sich dabei aus einer Grundprämie multipliziert mit Faktoren für jedes einzelne Risikomerkmal, dass bei der Bildung der Tarifklassen verwendet wurde.

Z. B. wird im Falle von zwei Risikomerkmalen die Schätzung wie folgt gebildet:

Schätzung		Ausprägung von Risikomerkmal 2			
		1	2	...	K
		y_1	y_2	$\cdots$	y_K
Ausprägung von Risikomerkmal 1					
1	x_1	$x_1 \cdot y_1$	$x_1 \cdot y_2$	...	$x_1 \cdot y_K$
2	x_2	$x_1 \cdot y_2$	$x_2 \cdot y_2$	...	$x_2 \cdot y_K$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
I	x_I	$x_I \cdot y_1$	$x_I \cdot y_2$	...	$x_I \cdot y_K$

bzw. Schätzung $(i, j) = b \cdot x_i \cdot y_j$ mit

$b =$ Grundprämie

$x_1, \ldots, x_I, y_1, \ldots, y_K$ Zu- bzw. Abschläge mit $x_1 = y_1 = 1$.

Dieser Ansatz bietet den großen Vorteil, dass man statt

$$I \cdot K$$

Werten nur

$$(I-1) + (K-1) + 1 = I + K - 1$$

Werten schätzen muss. Desweiteren kann der entstandene Tarif dem Kunden noch erklärt werden (z. B. indem man darauf hinweist, dass das Abstellen des

Autos bei Nacht in einer Garage den Betrag der KFZ-Versicherung um eine gewisse Prozentzahl senkt).

Offensichtlicher Nachteil ist jedoch, dass die angenommene multiplikative Struktur bei den mittleren Schadenhöhen keineswegs erfüllt sein muss, was zu schlechten Schätzungen führen kann.

Im Folgenden betrachten wir zwei Verfahren zur Bestimmung von Marginalfaktoren $x_1, \ldots, x_I$, $y_1, \ldots, y_K$ ausgehend von gegebenen

- Gesamtschäden $s_{i,k}$
- Volumen bzw. Gesamtversicherungssumme $v_{i,k}$
  pro Zelle (i, k) $\quad (i = 1, \ldots, I,\ k = 1, \ldots, K)$.

Beim Verfahren von Bailey/Simon werden die $x_1, \ldots, y_K$ durch Minimierung von

$$Q = \sum_{i=1}^{I} \sum_{k=1}^{K} (s_{i,k} - v_{i,k} \cdot x_i \cdot y_k)^2 / (v_{i,k} \cdot x_i \cdot y_k)$$

bestimmt.

Hierbei gilt

$$\begin{aligned} Q &= \sum_{i=1}^{I} \sum_{k=1}^{K} v_{i,k}^2 \cdot (\frac{s_{i,k}}{v_{i,k}} - x_i \cdot y_k)^2 / (v_{i,k} \cdot x_i \cdot y_k) \\ &= \sum_{i=1}^{I} \sum_{k=1}^{K} v_{i,k} \cdot (\frac{s_{i,k}}{v_{i,k}} - x_i \cdot y_k)^2 / (x_i \cdot y_k) \\ &= \sum_{i=1}^{I} \sum_{k=1}^{K} v_{i,k} \cdot (z_{i,k} - x_i \cdot y_k)^2 / (x_i \cdot y_k) \end{aligned}$$

mit $z_{i,k} = s_{i,k} / v_{i,k}$.

Nullsetzen der partiellen Abteilungen von Q nach x_{i_0} liefert

$$
\begin{aligned}
0 &\overset{(!)}{=} \frac{\partial}{\partial_{x_{i_0}}} \sum_{i=1}^{I}\sum_{k=1}^{K} v_{i,k}\cdot (z_{i,k}-x_i\cdot y_k)^2/(x_i\cdot y_k) \\
&= \frac{\partial}{\partial_{x_{i_0}}} \sum_{k=1}^{K} v_{i_0,k}\cdot (z_{i_0,k}-x_{i_0,k}\cdot y_k)^2/(x_{i_0}\cdot y_k) \\
&= \sum_{k=1}^{K} v_{i_0,k}\cdot \left(2(z_{i_0,k}-x_{i_0}\cdot y_k)\cdot(-y_k)/(x_{i_0}\cdot y_k) + (z_{i_0,k}-x_{i_0}\cdot y_k)^2\cdot \frac{-1}{x_{i_0}^2\cdot y_k}\right) \\
&= \sum_{k=1}^{K} v_{i_0,k}\cdot \left(-2\frac{z_{i_0,k}}{x_{i_0}} + 2y_k - \frac{z_{i_0,k}^2}{x_{i_0}^2\cdot y_k} + 2\frac{z_{i_0,k}}{x_{i_0}} - y_k\right) \\
&= \sum_{k=1}^{K} v_{i_0,k}\cdot y_k - \frac{1}{x_{i_0}^2}\cdot \sum_{k=1}^{K} v_{i_0,k}\cdot \frac{z_{i_0,k}^2}{y_k}.
\end{aligned}
$$

Daraus folgt

$$
x_{i_0} = \sqrt{\frac{\sum_{k=1}^{K}(v_{i_0,k}\cdot z_{i_0,k}^2)/y_k}{\sum_{k=1}^{K} v_{i_0,k}\cdot y_k}} \tag{3.1}
$$

Analog führt

$$
0 \overset{(!)}{=} \frac{\partial}{\partial y_{k_0}} Q
$$

auf

$$
y_{k_0} = \sqrt{\frac{\sum_{i=1}^{I}(v_{i,k_0}\cdot z_{i,k_0}^2)/x_i}{\sum_{i=1}^{I} v_{i,k_0}\cdot x_i}} \tag{3.2}
$$

(Symmetrie beachten!).

Die Gleichungen (3.1) und (3.2) stellen Fixpunktgleichungen dar. Man kann diese näherungsweise lösen, indem man mit $y_k = 1\,(k = 1,\dots,K)$ startet, und dann sukzessive in (3.1) bzw. (3.2) einsetzt.

Bemerkung:
Wegen des Quadrats in der Definition von Q hat das obige Verfahren den Nachteil, dass es empfindlich auf Ausreißer reagiert.

Beim **Marginalverfahren** versucht man, $x_1,\dots,y_K$ so zu bestimmen, dass gilt:

$$
\sum_{k=1}^{K} v_{i,k}\cdot x_i\cdot y_k = \sum_{k=1}^{K} s_{i,k} \quad (i = 1,\dots,I) \tag{3.3}
$$

$$
\sum_{i=1}^{I} v_{i,k}\cdot x_i\cdot y_k = \sum_{i=1}^{I} s_{i,k} \quad (k = 1,\dots,K) \tag{3.4}
$$

Dazu löst man (3.3) bzw. (3.4) wieder nach x_i bzw. y_k auf und setzt sukzessive in die entstehende Fixpunktgleichung ein.

Nach Konstruktion stimmt hier für jede Ausprägungsklasse jedes Tarifmerkmals die Summe der Bedarfsprämien mit dem dort beobachteten Gesamtschaden überein.

Weiter ist das obige Verfahren nicht so empfindlich gegen Ausreißer wie das Verfahren von Bailey/Simon, da oben kein Quadarat auftaucht.

4 Schadenreservierung bei lang andauernder Schadenabwicklung

4.1 Einführung

Zwischen Eintritt eines versicherten Schadenfalles und dessen endgültiger Regulierung kann unter Umständen eine lange Zeit verstreichen, da:

1. Schäden evtl. erst lange nach ihrer Untersuchung bemerkt werden, z. B. Fehler eines Architekten, Notars oder eines industriellen Produktes. Man spricht hierbei von IBNR-Schäden (incurred but not reported).
2. es unter Umständen lange dauert, bis die endgültige Schadenhöhe feststeht, z. B. wenn sie vom Ausgang eines Gerichtsprozesses oder vom Erfolg einer ärztlichen Behandlung abhängt. Für diese Schäden ist eine IBNER-Reserve nötig (incurred but not enough reserved).

Zur Begleichung der daraus resultierenden Kosten ist eine **Spätschadenreserve** erforderlich. Diese dient insbesondere zur:

- externen Rechnungslegung
- Prämienkalkulation (denn die Spätschäden müssen von den Prämien im Jahr des Anfalls des Schadens bezahlt werden).

Idee zur Berechnung einer Spätschadenreserve (oder Schadenreserve):
Übertrage Erfahrungen aus früheren Anfalljahren (d. h. Schadeneintrittsjahren) auf spätere Anfalljahre.

4.2 Das Abwicklungsdreieck

Die beobachteten Daten werden in sogenannten Abwicklungsdreiecken dargestellt.

Beispiel:
Unmittelbar nach Ende des Jahres 2008 liegt das folgenden Abwicklungsdreieck vor, das die für Schäden aus den Anfalljahren 2005-2008 in den einzelnen Abwicklungsjahren geleisteten Zahlungen (in 1000 €) enthält:

Anfalljahr	Abwicklungsjahr			
	2005	2006	2007	2008
2005	1001	854	568	347
2006		1113	671	422
2007			1265	1168
2008				1725

Wir gehen jetzt zu einer Darstellung mit relativen Abwicklungsjahren, d. h. mit Angabe der Verzögerungen in Bezug auf die Anfalljahre, über.

Dafür erhalten wir:

Anfalljahr	Abwicklungsjahr k			
	0	1	2	3
2005	1001	854	568	347
2006	1113	671	422	
2007	1265	1168		
2008	1725			

Diese Darstellung hat den Vorteil, dass Daten mit gleicher Verzögerung in Bezug auf die Anfalljahre jetzt untereinander stehen.

Anschließend ist es üblich, auch zu relativen Anfalljahren überzugehen, d. h. man schreibt:

Anfalljahr	Abwicklungsjahr k			
i	0	1	2	3
0	1001	854	568	347
1	1113	671	422	
2	1265	1168		
3	1725			

Ziel im Folgenden:
Vorhersage der Werte für die „leeren Felder“, d. h. im Beispiel oben für die Felder (i, k) mit $i + k > 3$.

Bemerkungen:

a) Die Tabelle ist in realen Anwendungen natürlich viel größer.

b) Im Folgenden gehen wir davon aus, dass die eingetragenen Schäden (z. B. unter Verwendung eines geeigneten Preisindex) inflationsbereinigt sind. Falls das nicht der Fall ist, wird die Inflation sonst in die Zukunft fortgeschrieben, ohne dass Änderungen bei der Inflation erkannt werden.

c) Als Schäden können in das Abwicklungsdreieck sowohl bezahlte Schäden (deren Höhe genau bekannt ist) als auch angefallene Schäden (die geschätzt, aber dafür auch sehr aktuell sind) eingetragen werden.

d) Das auftretende Schätzproblem ist schwierig, da die Zahl der Schäden mit langer Abwicklungsdauer nur ein kleiner Teil der Gesamtzahl aller Schäden ist. Folglich stehen zur Schätzung nur wenige Daten zur Verfügung.

4.3 Das Chain-Ladder-Verfahren

Gegeben ist Abwicklungsdreieck

$$(S_{i,k})_{0 \le i \le I, \quad 0 \le k \le I}$$

mit Anfalljahr i und Abwicklungsjahr k. Bei diesem sollen ausgehend von beobachteten Werten von

$$(S_{i,k})_{0\leq k\leq I-i, \quad 0\leq i\leq I}$$

die Werte von

$$(S_{i,k})_{I-i<k\leq I, \quad 0\leq i\leq I}$$

vorhergesagt werden.

Damit kann dann insbesondere der für die Prämienkalkulation benötigte Endschaden

$$C_{i,I} = S_{i,0} + S_{i,1} + \ldots + S_{i,I},$$

von dem bisher nur

$$C_{i,I-i} = S_{i,0} + S_{i,1} + \ldots + S_{i,I-i}$$

bekannt ist, geschätzt werden.

Im Folgenden definieren wir Schadenstände $C_{i,k}$ durch

$$C_{i,k} = S_{i,0} + S_{i,1} + \ldots + S_{i,k},$$

d. h. $C_{i,k}$ beschreibt die für Schäden aus dem Anfalljahr i bis zum Abwicklungsjahr k insgesamt geleisteten Zahlungen.

Fassen wir $(C_{i,k})_{0\leq i\leq I, \quad 0\leq k\leq I}$ als Zufallsvariablen auf, so liegen dem Chain-Ladder-Verfahren die folgenden Annahmen zugrunde:

(CL_1) Es gibt Abwicklungsfaktoren $f_0, \ldots, f_{I-1} \in \mathbb{R}$ mit

$$E(C_{i,k+1}|C_{i,1}, \ldots, C_{i,k}) = C_{i,k} \cdot f_k$$

für alle $0 \leq i \leq I, \quad 0 \leq k \leq I-1$.

(CL_2) Die Zufallsvektoren $(C_{i,1}, \ldots, C_{i,I}) \quad (0 \leq i \leq I)$ sind unabhängig.

Voraussetzung (CL_1) impliziert für $C_{i,k} > 0$:

$$\begin{aligned}
E\left(\frac{C_{i,k+1}}{C_{i,k}}\right) &= E\left(E\left(\frac{C_{i,k+1}}{C_{i,k}}|C_{i,1}, \ldots, C_{i,k}\right)\right) \\
&= E\left(\frac{1}{C_{i,k}}E(C_{i,k+1}|C_{i,1}, \ldots, C_{i,k})\right) \\
&\overset{(CL_1)}{=} E\left(\frac{1}{C_{i,k}} \cdot C_{i,k} \cdot f_k\right) \\
&= f_k,
\end{aligned}$$

d. h. unabhängig vom Anfalljahr i ändert sich der Schadenstand nach dem k-ten Abwicklungsjahr im Mittel prozentual um einen nur von k abhängenden Faktor.

Wir schätzen nun die Faktoren

$$f_k = E\left(\frac{C_{i,k+1}}{C_{i,k}}\right)$$

durch ein gewichtetes Mittel der beobachteten prozentualen Änderungen

$$\frac{C_{j,k+1}}{C_{j,k}} \quad (j = 0, 1, \dots, I-k-1).$$

Genauer verwenden wir

$$\hat{f}_k = \frac{\sum_{j=0}^{I-k-1} \frac{C_{j,k+1}}{C_{j,k}} \cdot C_{j,k}}{\sum_{j=0}^{I-k-1} C_{j,k}} = \frac{\sum_{j=0}^{I-k-1} C_{j,k+1}}{\sum_{j=0}^{I-k-1} C_{j,k}}.$$

Sodann schätzen wir den Endschadenstand

$$C_{i,I} = C_{i,I-i} \cdot \frac{C_{i,I-i+1}}{C_{i,I-1}} \cdot \frac{C_{i,I-i+2}}{C_{i,I-i+1}} \cdot \ldots \cdot \frac{C_{i,I}}{C_{i,I-1}}$$

durch

$$\hat{C}_{i,I} = C_{i,I-i} \cdot \hat{f}_{I-i} \cdot \hat{f}_{I-i+1} \cdot \ldots \cdot \hat{f}_{I-1}.$$

Beispiel: Für das Abwicklungsdreieck

Anfalljahr	Abwicklungsjahr			
	0	1	2	3
0	1001	854	568	347
1	1113	671	422	
2	1265	1168		
3	1725			

berechnen wir zuerst die **Schadenstände**

Anfalljahr	Abwicklungsjahr			
	0	1	2	3
0	1001	1855	2423	2770
1	1113	1784	2206	
2	1265	2433		
3	1725			

und schätzen damit die Faktoren für die Steigerungen der Schadenstände durch:

$$\hat{f}_0 = \frac{1855 + 1784 + 2433}{1001 + 1113 + 1265} \approx 1,797$$

$$\hat{f}_1 = \frac{2433 + 2206}{1855 + 1784} \approx 1,272 \quad , \hat{f}_3 = \frac{2770}{2423} \approx 1,143$$

Damit können wir die fehlenden Werte vorhersagen gemäß

Anfalljahr	Abwicklungsjahr			
	0	1	2	3
0	1001	1855	2423	2770
1	1113	1784	2206	$2206 \cdot \hat{f}_2 \approx \underline{2524,5}$
2	1265	2433	$2433 \cdot \hat{f}_1 \approx \underline{3094,8}$	$3094{,}8 \cdot \hat{f}_2 \approx \underline{3537,3}$
3	1725	$1725 \cdot \hat{f}_0 \approx \underline{3099,8}$	$\underline{3943}$	$\underline{4506{,}8}$

Wir zeigen im Folgenden:

Satz 4.1. *Unter den Annahmen* (CL_1) *und* (CL_2) *sind die Schätzer* $\hat{f}_k$ *und* $\hat{C}_{i,I}$ *erwartungstreu, d. h. es gilt*

$$E(\hat{f}_k) = f_k \text{ und } E(\hat{C}_{i,I}) = E(C_{i,I}),$$

Beweis: Es gilt

$$\begin{aligned}
& E(\hat{f}_k \,|\, C_{l,k} \quad (l = 0, \dots, I-k-1)) \\
= \; & \frac{\sum_{j=0}^{I-k-1} E(C_{j,k+1}) | \, C_{l,k} \, (l = 0, \dots, I-k-1))}{\sum_{j=0}^{I-k-1} C_{j,k}} \\
\overset{(CL_2)}{=} \; & \frac{\sum_{j=0}^{I-k-1} E(C_{j,k+1} | \, C_{j,k})}{\sum_{j=0}^{I-k-1} C_{j,k}} \\
= \; & \frac{\sum_{j=0}^{I-k-1} E(E(C_{j,k+1} | \, C_{j,1}, \dots, C_{j,k}) | C_{j,k})}{\sum_{j=0}^{I-k-1} C_{j,k}} \\
\overset{(CL_1)}{=} \; & \frac{\sum_{j=0}^{I-k-1} E(C_{j,k} \cdot f_k | C_{j,k})}{\sum_{j=0}^{I-k-1} C_{j,k}} \\
= \; & \frac{\sum_{j=0}^{I-k-1} C_{j,k} \cdot f_k}{\sum_{j=0}^{I-k-1} C_{j,k}} = f_k.
\end{aligned}$$

Daraus folgt die erste Behauptung, da

$$E(\hat{f}_k) = E(E\hat{f}_k | C_{l,k} \, (l = 0, \dots, I-k-1)) \overset{\text{s.o}}{=} E(f_k) = f_k.$$

Zum Beweis der zweiten Behauptung beachten wir

$$\begin{aligned}
E(\hat{C}_{i,I}) \quad & = \quad E(C_{i,I-i} \cdot \hat{f}_{I-i} \cdot \hat{f}_{I-i+1} \cdot \ldots \cdot \hat{f}_{I-1}) \\
& = \quad E(C_{i,I-1} E(\hat{f}_{I-i} \cdot \ldots \cdot \hat{f}_{I-1} | C_{j,k} \, (k \le I-1, j \le I-k-1))) \\
& = \quad E(C_{i,I-1} E(\hat{f}_{I-i} \cdot \ldots \cdot \hat{f}_{I-2} \cdot E(\hat{f}_{I-1} | C_{j,k} \, (k \le I-1, j \le I-k-1))) \\
& \overset{\text{analog oben}}{=} \quad E(C_{i,I-1} E(\hat{f}_{I-i} \cdot \ldots \cdot \hat{f}_{I-2} \cdot f_{I-1})) \\
& = \quad E(C_{i,I-1} \cdot E(E(\hat{f}_{I-i} \cdot \ldots \cdot \hat{f}_{I-2} | C_{j,k} \, (k \le I-2, j \le I-k-1))) \cdot f_{I-1}) \\
& \overset{\text{analog}}{=} \quad \ldots = E(C_{i,I-1} \cdot f_{I-i} \cdot f_{I-i+1} \cdot \ldots \cdot f_{I-1}) = E(C_{i,I}) \quad \square
\end{aligned}$$

4.4 Das Cape-Cod-Verfahren

Beim Chain-Ladder-Verfahren wird der Endschadenstand

$$C_{i,I}$$

durch mulitplikative Fortschreibung

$$C_{i,I-i} \cdot \hat{f}_{I-i} \cdot \hat{f}_{I-i+1} \cdot \ldots \cdot \hat{f}_{I-1}$$

des zuletzt beobachteten Schadenstandes $C_{i,I-i}$ des Anfalljahres i geschätzt. Dies führt immer dann zu Problemen, wenn dieser beobachtete Schadenstand durch Großschäden oder ausgebliebene Schadenlasten stark verfälscht ist.

Die Idee beim Cape-Cod-Verfahren ist es, diesen Effekt zu vermeiden, indem der zuletzt beobachtete Schadenstand in einen Cape-Cod-Schadenstand und einen Ausreißereffekt zerlegt wird, und nur ersterer fortgeschrieben wird.

Genauer nehmen wir an, dass die Prämien $\pi_0, \pi_1, \ldots, \pi_I$ für die Anfalljahre bekannt sind. Weiter nehmen wir an, dass für die Schadenstände gilt:

(CC 1) Es gibt Abwicklungsfaktoren $g_0, g_1, \ldots, g_I$ mit $g_I = 1$ und eine globale Endschadenquote $\kappa > 0$ so, dass gilt:
$E[C_{i,k}] = \kappa \cdot \pi_i \cdot g_k$.

Wegen $g_I = 1$ gilt dann

$$\frac{E[C_{i,k}]}{E[C_{i,I}]} = \frac{\kappa \cdot \pi_i \cdot g_k}{\kappa \cdot \pi_i \cdot g_I} = g_k,$$

d. h. g_k beschreibt gerade den prozentualen Anteil des Schadenstandes $C_{i,k}$ (nach k Abwicklungsjahren) am Endschaden $C_{i,I}$, der als unabhängig vom Anfalljahr i angestrebt wird.

Beim Cape-Cod-Verfahren wird g_k mit Hilfe der Schätzer $\hat{f}_k$ des Chain-Ladder-Verfahrens für

$$f_k = E\left(\frac{C_{i,k+1}}{C_{i,k}}\right)$$

geschätzt durch

$$\hat{g}_k = \frac{1}{\hat{f}_k} \cdot \frac{1}{\hat{f}_{k+1}} \cdot \ldots \cdot \frac{1}{\hat{f}_{I-1}}.$$

Zur Schätzung der Endschadenquote

$$\kappa = \frac{E[C_{i,k}]}{\pi_i \cdot g_k}$$

beachten wir $g_I = 1$, was

$$\kappa = \frac{E[C_{i,I}]}{\pi_i} \quad (i = 0, \ldots, I)$$

impliziert. Wir schätzen nun κ, indem wir $C_{i,I}$ durch den Chain-Ladder-Schätzer $\hat{C}_{i,I}$ ersetzt und über die resultierenden Schätzer ein gewichtetes Mittel bilden:

$$\hat{\kappa} = \sum_{i=0}^{I} \frac{\hat{g}_{I-i} \cdot \pi_i}{\sum_{j=0}^{I} \hat{g}_{I-j} \cdot \pi_j} \cdot \frac{\hat{C}_{i,I}}{\pi_i} = \frac{\sum_{i=0}^{I} C_{i,I-i}}{\sum_{j=0}^{I} \hat{g}_{I-j} \cdot \pi_j}.$$

Abschließend zerlegen wir nun die zuletzt beobachteten Endschadenstände

$$C_{i,I-i}$$

in den Cape-Cod Schadenstand

$$\hat{T}_{i,I-i} = \hat{g}_{I-1} \cdot \pi_i \cdot \hat{\kappa}$$

(vgl. (CC 1)) und den Ausreißereffekt

$$\hat{X}_i = C_{i,I-1} - \hat{g}_{I-1} \cdot \pi_i \cdot \hat{\kappa},$$

lassen den Ausreißereffekt bei der Schätzung von $C_{i,j}$ mit $j > I-i$ unverändert, und schreiben hier den Cape-Cod Schadenstand gemäß unserem Modell fort, d. h. wir schätzen $C_{i,j}$ durch

$$\begin{aligned} \hat{C}_{i,j} &= \hat{X}_i + \hat{g}_j \cdot \frac{\hat{T}_{i,I-j}}{\hat{g}_{I-i}} \\ &= \hat{X}_i + \hat{g}_j \cdot \pi_i \cdot \hat{\kappa}. \end{aligned}$$

Beispiel:
Wir betrachten nochmals das Beispiel von oben, diesmal aber zusätzlich mit Prämien pro Anfalljahr.

Wir schätzen die Quoten durch

$$\hat{g}_3 = 1,\ \hat{g}_2 = \frac{1}{\hat{f}_2} = \frac{1}{1,143} \approx 0,875,\ \hat{g}_1 = \frac{1}{\hat{f}_1 \cdot \hat{f}_2} \approx 0,688, \text{etc.}$$

Abschließend schätzen wir die Endschadenquote durch

$$\hat{\kappa} = \frac{2770 + 2206 + 2433 + 1725}{1 \cdot 3000 + 0,875 \cdot 3000 + 0,688 \cdot 3000 + 0,383 \cdot 5000} \approx 0,951$$

Für die Schadenstände haben wir: (Cape-Cod Schadenstand)

Anfalljahr	Verdiente Prämie	Abwicklungsjahr			
		0	1	2	3
0	3000	1001	1855	2423	2770
1	3000	1113	1784	2206	
				(2496)	(2853)
2	3000	1265	2433		
			(1962)	(2497)	(2854)
3	5000	1725			
		(1821)	(3272)	(4162)	(4758)
Chain-Ladder	Faktoren $\hat{f}_k$	1,797	1,272	1,143	
Quoten	$\hat{g}_k$	0,383	0,688	0,875	1

Damit erhalten wir für die Außreißereffekte:

$$\begin{aligned}
\hat{X}_1 &= 2206 - 2496 = -290, \\
\hat{X}_2 &= 2433 - 1962 = 471, \\
\hat{X}_3 &= 1725 - 1821 = -96.
\end{aligned}$$

Die geschätzten Endschadenstände sind damit:

$$\begin{aligned}
\hat{C}_{1,3} &= 2853 + \hat{X}_1 = 2563, \\
\hat{C}_{2,3} &= 2854 + 471 = 3325, \\
\hat{C}_{3,3} &= 4758 - 96 = 4662.
\end{aligned}$$

5 Risikoteilung

5.1 Einführung

Üblich: Versicherungsunternehmen ...

1.) übernimmt Risiken nicht ganz, sondern nur zum Teil,

2.) gibt Teile übernommener Risiken wieder ab.

z. B. zu 1.):
Nichtproportionale Risikoteilung (nur das ist mathematisch interessant!) der Schadenvariable X:

$$X = \min\{X, a\} + \max\{X - a, 0\}$$

für ein $a > 0$.

Hierbei: $-\min\{X, a\} \ldots$ sog. **Erstrisiko** z. b. Deckungssumme bei Haftpflicht mit a groß

- bei KFZ-Haftpflicht:
  z. B. 50 Mio. Euro für Personen-, Sach- und Personenschäden, jedoch max. 8 Mio. Euro je geschädigter Person

- bei Privat-Haftpflicht
  z. B. 50 Mio. Euro, max. 50.000 für Vermögensschäden und max. 8 Mio. Euro pro geschädigter Person.

- $\max\{X - a, 0\} \ldots$ sog. **Zweitrisiko**
  z. B. Selbstbeteiligung bei KFZ-Vollkasko oder privater Krankenversicherung, hier a klein.

Vorteil bei Begrenzung des Risikoschutzes auf das Zweitrisiko:

- Verringerung des Verwaltungs- und Regulierungsaufwandes durch Ausschluss von Kleinschäden
- Beeinflussung des moralischen Risikos (z. B. bei KFZ-Vollkasko).

z. B. zu 2.):
Schadenexzedenten-Rückversicherung übernimmt gegen festen Betrag den zufälligen Teil

$$\min\{\max\{X - a, 0\}, b\} = \begin{cases} 0, & \text{falls } X \leq a \\ X - a, & \text{falls } a < X \leq a + b \\ b, & \text{falls } X > a + b \end{cases}$$

(sog. „Layer“) des Schadens, wobei $a, b \in \mathbb{R}_+$.
Z. B. falls Illiquiditätswahrscheinlichkeit $P[S > E(S) + c]$ zu groß ist. Hier verringert das Versicherungsunternehmen gegen Zahlung eines festen Betrags sein Zufallsrisiko.

Frage:
Welche Prämie soll man für Erstrisiko, Zweitrisiko oder „Layer“verlangen?

Zur Beantwortung dieser Frage müssen wir bestimmen, wie sich die Risikoteilung auf

- „Nettoprämie“, also den erwarteten Schaden
- Schwankungszuschlag

auswirkt.

5.2 Einfluss der Risikoteilung auf die Schadenvariablen

Betrachtet wird die nichtproportionale Risikoteilung

$$X = \min\{X, a\} + \max\{X - a, 0\}$$

des Schadens X in Erstrisiko $\min\{X, a\}$ und Zweitrisiko $\max\{X - a, 0\}$. Wir stellen den Gesamtschaden im kollektiven Modell dar gemäß

$$S = \sum_{n=1}^{N} X_n$$

mit $X_1, X_2, \ldots$ u.i.v. und N unabhängig von $X_1, X_2, \ldots$. Für die Anzahl Schäden $\underline{N}$ beim Erstrisiko gilt dann

$$\underline{N} = N,$$

während für die entsprechenden Anzahl $\underline{\underline{N}}$ beim Zweitrisiko gilt:

$$\underline{\underline{N}} = \sum_{m=1}^{N} I_{\{X_m > a\}}.$$

Hierbei

$$\begin{aligned}
P[\underline{\underline{N}} = k] &= P[\sum_{m=1}^{N} I_{\{X_m > a\}} = k] \\
&= \sum_{n=k}^{\infty} P[\sum_{m=1}^{N} I_{\{X_m > a\}} = k, \quad N = n] \\
& \text{(da für } \underline{\underline{N}} = k \text{ der Wert von } N \text{ mindestens } k \text{ sein muss)} \\
&= \sum_{n=k}^{\infty} P[\sum_{m=1}^{n} I_{\{X_m > a\}} = k, \quad N = n] \\
&\overset{\text{Unabhängigkeit}}{=} \sum_{n=k}^{\infty} P[\sum_{m=1}^{n} I_{\{X_m > a\}} = k] \cdot P[N = n] \\
&= \sum_{n=k}^{\infty} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \cdot P[N = n],
\end{aligned}$$

mit

$$p = P[X > a],$$

wobei wir benutzt haben, dass

$$\sum_{m=1}^{n} I_{\{X_m > a\}}$$

$b(n, p)$-verteilt ist.

Ist nun speziell N $\pi(\lambda)$-verteilt, so folgt:

$$\begin{aligned}
P[\underline{\underline{N}} = k] &= \sum_{n=k}^{\infty} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \cdot \frac{\lambda^n}{n!} e^{-\lambda} \\
&= \frac{p^k \lambda^k}{k!} \left(\sum_{n=k}^{\infty} \frac{1}{(n-k)!} \cdot ((1-p) \cdot \lambda)^{n-k} \right) \cdot e^{-\lambda} \\
&= \frac{(p \cdot \lambda)^k}{k!} \cdot e^{(1-p) \cdot \lambda} \cdot e^{-\lambda} \\
&= \frac{(p \cdot \lambda)^k}{k!} \cdot e^{-p \cdot \lambda},
\end{aligned}$$

d. h. $\underline{\underline{N}}$ ist $\pi(\lambda P)$-verteilt.

Insbesondere verringert sich bei Übergang zum Zweitrisiko die erwartete Schadenzahl von λ auf

$$E\underline{\underline{N}} = \lambda \cdot p = \lambda \cdot P[X > a].$$

Für die Schadenhöhe X des Erstrisikos $\min\{X, a\}$ gilt:

$$E\min\{X,a\} = \int_0^\infty P\{\min\{X,a\} > t\}\,dt,$$

da für die nichtnegative Zufallsvariante Z die Beziehung

$$\begin{aligned} EZ &= \int_\Omega \int_0^\infty I_{\{t<Z\}} dt\, dP \\ &\overset{\text{Fubini}}{=} \int_0^\infty \int_\Omega I_{\{t<Z\}} dP\, dt \\ &= \int_0^\infty P[Z > t]\, dt \end{aligned}$$

gilt. Es folgt

$$\begin{aligned} E\min\{X,a\} &= \int_0^a P\{\min\{X,a\} > t\}\,dt \\ &= \int_0^a P\{X > t\}\,dt, \end{aligned}$$

was für $P[X > a] > 0$ kleiner als

$$EX = \int_0^\infty P[X > t]\,dt$$

ist.
Analog für Zweitrisiko (s.u.): $E\{\max\{X - a, 0\}\} = \int_a^\infty P\{X > t\}\,dt$.

Für den Gesamtschaden

$$S = \sum_{n=1}^N X_n = \sum_{n=1}^N \min\{X_{n,a}\} + \sum_{n=1}^N \max\{X_n - a, 0\} = \underline{S} + \underline{\underline{S}}$$

gilt im kollektiven Modell:

$$E\underline{S} = EN \cdot E\min\{X_1, a\} \overset{\text{s. o.}}{=} EN \cdot \int_0^a P[X > t]\,dt$$

und

$$\begin{aligned} E\underline{\underline{S}} = EN \cdot E\max\{X_1 - a, 0\} &= EN \cdot \int_0^\infty P\{\max\{X_1 - a, 0\} > t\}\,dt \\ &= EN \cdot \int_0^\infty P\{X_1 - a > t\}\,dt \\ &= EN \cdot \int_0^\infty P\{X_1 > t + a\}\,dt \\ &= EN \cdot \int_a^\infty P\{X > s\}\,ds. \end{aligned}$$

5.3 Der Entlastungseffekt

Der sogenannte Entlastungseffekt

$$r(a) = \frac{E(\underline{S})}{E(S)}$$

beschreibt den Anteil des Erstschadens am Gesamtschaden und gibt damit den Teil der Schadenerwartung an, um den der Versicherer beim Übergang zum Zweitrisiko (d. h. Einführung eines Selbstbehalts) entlastet wird.

Hierbei ist

$$1 - r(a) = 1 - \frac{E(\underline{S})}{E(S)} = \frac{E(\underline{\underline{S}})}{E(S)}$$

die Entlastung beim Übergang zum Erstrisiko (also z. B. Einführung einer Deckungssumme).
Im kollektiven Modell gilt:

$$\begin{aligned} r(a) &= \frac{E\left(\sum_{n=1}^{N} \min\{X_n, a\}\right)}{E\left(\sum_{n=1}^{N} X_1\right)} \overset{\text{Unabhängigkeit}}{=} \frac{EN \cdot E\min\{X_1, a\}}{EN \cdot EX_1} \\ &= \frac{\int_0^a P\{X_1 > t\}\, dt}{\int_0^\infty P\{X_1 > t\}\, dt}. \end{aligned}$$

Man sieht, dass der Entlastungseffekt nur von der Verteilung der Schadenhöhen abhängt.

Man sieht:

1.) $r(a) \in [0, 1]$, wobei $r(0) = 0$ und $r(a) \to 1\,(a \to \infty)$.

2.) r ist monoton wachsend in a, d. h. je größer der Selbstbehalt, umso größer der Entlastungseffekt.

3.) $r'(a) = \frac{P\{X_1 > a\}}{EX_1}$, und sofern $a \to P\{X_1 > a\}$ differenzierbar ist, folgt, da diese Abbildung monoton fallend ist:

$$r''(a) \leq 0.$$

Nach 3.) ist $a \to r(a)$ konkav. Ist dann v die maximale Schadenhöhe (also $r(v) = 1$), so gilt

$$\begin{aligned} r(c \cdot v) &= r(c \cdot v + (1-c) \cdot 0) \geq c \cdot r(v) + (1-c) \cdot r(0) \\ &= c \cdot 1 + (1-c) \cdot 0 = c \end{aligned}$$

für $c \in (0,1)$.

Also gilt für die Entlastung bei Einführung eines Selbstbehalts von $c \cdot v$, dass diese den Versicherer mindestens soviel entlastet wie eine Selbstbeteiligung von $100 \cdot c\%$, die zu einer Entlastung um den Faktor c beim Schaden führt.

5.4 Prämienkalkulation bei Risikoteilung

Wie wir bereits in Kapitel 1 gesehen haben, enthält die Prämie als Bestandteile:

- Schadenerwartung $E(S)$
- den Schwankungszuschlag s_0
- einen Fixkostenzuschlag k_0 und einen Proportionalkostenanteil p

Genauer berechnet sie sich zu:

$$b = \frac{E(S) + s_0 + k_0}{1-p}.$$

Wir haben im vorigen Abschnitt gesehen, dass die Änderung von $E(S)$ zu $E(\underline{S})$ beim Erstrisiko durch den Entlastungseffekt

$$r(a) = \frac{E(\underline{S})}{E(S)} = \frac{\int_0^a P\{X > t\} dt}{E(S)}$$

beschrieben wird. In Abhängigkeit der Deckungssumme a kann dieser Effekt ausgehend von der Verteilungsfunktion der Schadenhöhe berechnet werden.

Der Schwankungszuschlag s_0 kann im kollektiven Modell

$$\underline{S} = \sum_{n=1}^{N} \min\{X_n, a\}$$

(bei Einführung einer Deckungssumme a) ausgehende von der mit Hilfe der Formel von Panjer geschätzten Verteilungsfunktion von $\underline{S}$ bestimmt werden.

Da hierbei die Varianz der Schadenhöhen kleiner wird, wird auch der Schwankungszuschlag bei Einführung einer nichttrivialen Deckungssumme kleiner.

Die beiden Kostenzuschläge bleiben, sofern die Provisionen für die Vertreter nicht verändert werden, gleich.

Probleme bei Risikoteilung:

a) Antiselektion
Die Risiken, die z. B. eine Selbstbeteiligung akzeptieren, haben eventuell eine andere Schadenhöhenverteilung, so dass der Versicherer Prämien mit und ohne Selbstbeteiligung dann für die falschen Verteilungen berechnet.

b) Verlust an Information
Versicherer erfährt nichts mehr über Schäden unterhalb von a, sofern er nur das Zweitrisiko $\max\{X - a, 0\}$ übernimmt.

c) Inflation
Ändert sich durch Inflation die Schadenhöhe, der Selbstbehalt aber nicht, so ändert sich der Entlastungseffekt.