



TECHNISCHE  
UNIVERSITÄT  
WIEN

## DIPLOMARBEIT

# Anpassung von Leistungsprofilen in der Krankenversicherungsmathematik

ausgeführt in der

Forschungsgruppe für  
Finanz- und Versicherungsmathematik  
TU Wien

unter der Anleitung von

**Univ.Prof. Dipl.-Math. Dr.rer.nat. Thorsten Rheinländer**

durch

**Gregor Fleischmann**

Matrikelnummer: 01529195

Wien, am 17. Mai 2023

# Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich diese Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht.

Wien, am 17. Mai 2023

# Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich diese Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht.

Wien, am 17. Mai 2023

# Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich diese Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht.

Wien, am 17. Mai 2023

# Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich diese Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht.

Wien, am 17. Mai 2023

# Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich diese Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht.

Wien, am 17. Mai 2023

# Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich diese Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht.

Wien, am 17. Mai 2023

# Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich diese Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht.

Wien, am 17. Mai 2023

- **(Medizinische) Inflation** - Bei vielen Leistungen spielt die Entwicklung der Gehälter von medizinischem Personal oder auch die Kosten von einzelnen Medikamenten eine wesentliche Rolle.
- **Medizinischer Fortschritt** - Die Möglichkeit der Entwicklung von neuen medizinischen Behandlungsverfahren bzw. Medikamenten, welche mit entsprechend hohen Kosten verbunden sind, ist tendenziell schwer vorhersehbar. Man betrachte bspw. operative Eingriffe, welche vor einigen Jahren einen mehrtägigen Spitalsaufenthalt vorgesehen hätten und nun binnen eines Tages erledigt werden - dieser Sachverhalt hat für viele ambulante Tarife einen entsprechenden Mehraufwand nach sich gezogen. Gleichzeitig sorgen die verkürzten Krankenhausaufenthaltsdauern für eine Aufwandsverminderung bei stationären Tarifen.
- **Veränderung der Lebenserwartung** - Gemäß derzeitigen Prognosen wird erwartet, dass die Lebenserwartung auch in den nächsten Jahrzehnten ansteigen wird, für die Krankenversicherung ist hierbei die zugehörige Entwicklung der Morbidität entscheidend - bspw. gab es in den letzten Jahrzehnten auch in jungen Altersgruppen einen Anstieg in der Häufigkeit von chronischen Erkrankungen.

Zusätzlich zu diesen Umständen bei der Schätzung von zukünftigen Leistungsaufwendungen kommt noch eine wesentliche rechtliche Komponente: dem Versicherer ist es (bis auf wenige Ausnahmen) nicht erlaubt, den Versicherungsvertrag zu kündigen (§ 178i VersVG). Somit scheint hier ein Dilemma vorzuliegen: schätzt man die zukünftigen erwarteten Leistungen aus Vorsicht im Vergleich zum aktuellen Leistungsniveau sehr hoch ein, so wären die resultierenden Prämien naturgemäß ebenfalls hoch, also aus Sicht von potentiellen Versicherungsnehmern weniger attraktiv. Werden die Schätzwerte zu niedrig angesetzt, so besteht aufgrund der fehlenden Möglichkeit, Verträge zu kündigen, die Gefahr, in Zukunft hohe Verluste zu erleiden.

Um diese Problematik zu bewältigen existiert die bereits im ersten Abschnitt dieses Kapitels erwähnte Anpassungsklausel, d.h. der Versicherer besitzt unter gewissen Voraussetzungen die Möglichkeit der Veränderung bzw. Erhöhung der Prämie nach Vertragsabschluss. Somit erfolgt die Prämienberechnung bei Vertragsabschluss unter der Annahme, dass die zu diesem Zeitpunkt gültigen Annahmen (Rechnungsgrundlagen) auch in Zukunft gelten werden, also ist bspw. die erwartete Leistung für jenes zukünftige Jahr, in dem ein Versicherungsnehmer das Alter 80 erreicht, unabhängig davon, welches Alter er zum jetzigen Zeitpunkt besitzt. Kommt es in Zukunft zu einer Änderung der Annahmen - zumeist Steigerung der erwarteten Leistungen - so gelten diese nun gedanklich für immer und die Prämien der aktiven Versicherungsnehmer werden gemäß Äquivalenzprinzip und der Anrechnung der bisher angesparten Alterungsrückstellung neu bestimmt. In der Bundesrepublik Deutschland sind

die aktuariellen Mechanismen des Anpassungsprozesses weitgehend normiert - siehe insbesondere „Verordnung betreffend die Aufsicht über die Geschäftstätigkeit in der privaten Krankenversicherung“, (Krankenversicherungsaufsichtsverordnung - KVAV). Die in Österreich geübte actuarielle Praxis ist im Leitfaden der AVÖ [2] dargestellt.

### 2.1.4. Bewertungsunterschiede in der Bilanzierung

In der österreichischen Versicherungsbranche sind zumindest zwei Rechnungslegungsvorschriften relevant: die Vorschriften gemäß UGB (analog zu HGB in Deutschland) und die Vorschriften gemäß „International Financial Accounting Standards“ (IFRS). Sowohl UGB als auch IFRS verpflichten die Versicherer zur Erstellung von Bilanzen und entsprechenden Gewinn- und Verlustrechnungen, wobei die jeweils verwendeten Bewertungsregeln wesentlich voneinander abweichen. Zusätzlich sind Versicherer aufgrund gemäß Solvency II geltender aufsichtsrechtlicher Regelungen unter anderem dazu verpflichtet, eine sog. „Solvenzbilanz“ zu erstellen<sup>4</sup>. Die Bewertungsgrundsätze von Solvency II sind sehr ähnlich zu jenen von IFRS und unterscheiden sich somit ebenfalls wesentlich von den Bewertungsgrundsätzen von UGB, obwohl sowohl Solvency II als auch UGB dasselbe „Ziel“ verfolgen: den Gläubigerschutz, d.h. das Sicherstellen der dauerhaften Erfüllbarkeit der Verpflichtungen die durch den Abschluss von Versicherungsverträgen entstehen. Eine der - aus Sicht der Krankenversicherungsmathematik wohl wesentlichsten - Größen, die sowohl in der UGB-Bilanz als auch in Solvenzbilanz ausgewiesen wird, sind die versicherungstechnischen Rückstellungen (Deckungsrückstellung, Schadenrückstellung), d.h. die Bewertung aller zukünftigen Prämien und Leistungen anhand von entsprechenden Barwerten. In der UGB-Bilanz erfolgt die Bewertung der Rückstellung gemäß Vorsichtsprinzip, d.h. die für die Bestimmung des Barwerts (wie auch für die Prämienkalkulation) verwendeten Rechnungsgrundlagen beinhalten entsprechende Sicherheiten, bspw. werden zukünftige Leistungen höher angesetzt als erwartet. Gleichzeitig wird aber analog zur Prämienanpassung die Annahme getroffen, dass die zum Bilanzstichtag gültigen Rechnungsgrundlagen auch in Zukunft gelten werden. Zusätzlich werden bei der Bestimmung von Barwerten den jeweiligen Tarifen zugehörige Rechnungszinssätze, welche bei älteren Tarifen im Vergleich zur realistisch erzielbaren Rendite relativ hoch sein können, verwendet. Bei der Bewertung gemäß Solvency II gilt das Prinzip des „true and fair view“, d.h. zukünftige Zahlungen sind möglichst realistisch einzuschätzen - der Ausgleich für das Entfernen von Sicherheiten in den Rechnungsgrundlagen erfolgt bei Solvency II explizit durch die Bildung einer sog. „Risk Margin“. Für die Krankenversicherung bedeutet dies insbesondere die Berücksichtigung von zukünftigen Leistungs- und Prämienanpassungen, wobei die adäquate Schätzung

---

<sup>4</sup>siehe Richtlinie 2009/138/EG des Europäischen Parlaments und des Rates betreffend die Aufnahme und Ausübung der Versicherungs- und der Rückversicherungstätigkeit (Solvabilität II)

der jeweiligen zukünftigen Anpassungseffekte - wie bereits im vorherigen Abschnitt beschrieben - eine anspruchsvolle Aufgabe darstellt. Zusammengefasst existieren für jeden Tarif zwei Arten von Rechnungsgrundlagen, die je nach vorgesehener Bewertungsmethodik zur Anwendung kommen. Grundsätzlich gilt folgende Konvention: Rechnungsgrundlagen „erster Ordnung“ beinhalten entsprechende Sicherheiten und kommen sowohl bei der Prämienkalkulation, der Prämienanpassung, als auch bei der Bilanzierung (und Reservierung) gemäß UGB zur Anwendung. Rechnungsgrundlagen „zweiter Ordnung“ sind möglichst realistisch anzusetzen und bei Prognoseberechnungen werden i.d.R. zukünftige Anpassungen berücksichtigt. Eine in der Praxis gängige Vorgangsweise besteht darin, zunächst die Rechnungsgrundlagen zweiter Ordnung zu bestimmen, die Leistungsprofile der Rechnungsgrundlagen erster Ordnung entstehen dann durch Hinzugabe entsprechender Sicherheitszuschläge zu den Profilen zweiter Ordnung, bspw. durch prozentuelle Zuschlagsfaktoren. Da für Prämienanpassungen die Rechnungsgrundlagen erster Ordnung verwendet werden, kommen diese auch bei „realistischen“<sup>5</sup> Berechnungen - also bei IFRS bzw. Solvency II - zur Anwendung.

## 2.2. Modellierung und Berechnungsverfahren

Dieser Abschnitt behandelt die in der Praxis gängige mathematische Modellierung der in den obigen Abschnitten beschriebenen Sachverhalte.

### 2.2.1. Kopfschäden und Einsteigerprämien

Formeltechnisch werden die bereits erwähnten Kopfschäden mit  $K_x$  notiert, wobei  $x$  dem Alter des Versicherungsnehmers entspricht, d.h.  $K_{25}$  entspricht somit dem durchschnittlichen Leistungsaufwand für einen 25-jährigen Versicherungsnehmer der binnen eines Jahres - bzw. bis der Versicherungsnehmer das Alter 26 erreicht - anfällt. Der Vektor  $(K_{x_{\min}}, \dots, K_{x_{\max}})$  wird als Kopfschadenreihe bzw. (Leistungs-)Profil bezeichnet. Die Krankenversicherungsleistungen - sowie deren Leistungsbarwert - werden schließlich als Summe von Erlebensversicherungen, wobei die jeweilige Erlebenssumme dem Kopfschaden zum zugehörigen Laufzeitendalter des Versicherungsnehmers entspricht, modelliert. Prämien werden in der Krankenversicherung in der Regel in Form einer ewigen vorschüssigen - meist jährlichen - Leibrente modelliert. Endliche Prämienzahlungsdauern bzw. Einmalprämien sind angesichts frequenter Prämienanpassungen eher atypisch, etwaige spätere Prämienreduktion wird durch Abschluss von Optionsverträgen bewerkstelligt. Gemäß Äquivalenzprinzip lassen sich somit die

---

<sup>5</sup>Diese Stelle bezieht sich auf die Vorgaben zur sog. „Best-Estimate“ Berechnung.

Neuabschluss- bzw. Einsteigerjahresnettoprämien  $P_x$  berechnen:

$$\sum_{n=0}^{x_{\max}-x} {}_nE_x \cdot K_{x+n} \stackrel{!}{=} P_x \cdot \ddot{a}_x \Rightarrow P_x = \frac{1}{\ddot{a}_x} \sum_{n=0}^{x_{\max}-x} {}_nE_x \cdot K_{x+n}, \quad (2.1)$$

wobei  ${}_nE_x$  den Barwert einer normierten  $n$ -jährigen Erlebensversicherung und  $\ddot{a}_x$  den Barwert einer normierten ewigen Leibrente mit Startalter  $x$  bezeichnet.

### 2.2.2. Bestimmung von Leistungsprofilen

Wie anhand des vorherigen Abschnitts ersichtlich wird, handelt es sich bei Leistungsprofilen um Kalkulationsparameter. Da sich ein Leistungsprofil, wie bereits weiter oben ausgeführt wurde, aufgrund verschiedener Effekte im Laufe der Zeit verändert, ist es notwendig, die Kopfschäden regelmäßig neu zu berechnen. Um die Zeitspanne zwischen vollständigen Neuberechnungen von Leistungsprofilen zu maximieren, wird in der Praxis oftmals die sog. „Methode von Rusam“ angewandt. Hierbei wird ein Leistungsprofil mit einem altersunabhängigen Faktor  $G$ , welcher „Grundkopfschaden“ genannt wird, normiert, d.h. es gilt

$$K_x = G \cdot k_x, \quad (2.2)$$

wobei  $(k_{x_{\min}}, \dots, k_{x_{\max}})$  dem normierten Profil entspricht. Zusätzlich wird mit der Annahme gearbeitet, dass das normierte Profil sich im Lauf der Zeit nur sehr langsam verändert, während der Grundkopfschaden regelmäßig - oft sogar jährlich - angepasst wird. Der Grundkopfschaden entspricht i.d.R. dem Kopfschaden eines Wahlalters  $x_0$ , d.h.  $G := K_{x_0}$ . Für die Kalkulation des normierten Profils werden die Bestands- und Leistungsdaten mehrerer Jahre und - zwecks statistischer Signifikanz - oftmals mehrerer ähnlicher Tarife gemeinsam betrachtet. Zusätzlich werden verschiedene Leistungskategorien oftmals getrennt betrachtet, d.h. ein Tarif besitzt im Allgemeinen mehrere Profile. Dies hat unter anderem den Vorteil, dass zur Kalkulation eines normierten Profils die Daten von mehreren Tarifen herangezogen werden können, welche sich in der Gesamtheit ihrer versicherten Leistungen jeweils stark voneinander unterscheiden aber durchaus signifikante Überlappungen vorliegen. Ein in der Praxis gängiges Verfahren zur Ermittlung von normierten Profilen ist das „Gleichrichten von Kopfschäden“, welches an dieser Stelle für die Bestimmung eines Profils näher erläutert wird. Es sei  $T = \{t_1, \dots, t_n\}$  die Menge der für die Kalkulation betrachteten Jahre,  $S_x(t_i)$  die Gesamtleistungen für Personen mit Alter  $x$  im Jahr  $t_i$  und  $L_x(t_i)$  die zugehörige Anzahl an Versicherungsnehmern. Zunächst werden die „rohen“ Kopfschäden  $K_x^r(t_i)$  durch Berechnung der zugehörigen durchschnittlichen Leistungen, d.h. gemäß der Formel

$$K_x^r(t_i) = \frac{S_x(t_i)}{L_x(t_i)}, \quad (2.3)$$

für alle Betrachtungsjahre bestimmt. Im nächsten Schritt werden Skalierungsfaktoren  $f(t)$  ermittelt, um das Leistungsniveau der älteren Betrachtungsjahre auf jenes des Letztbetrachtungsjahres zu bringen. Die Faktoren  $f(t_i)$  bestimmen sich durch Berechnung des Quotienten von „künstlichen“ Gesamtleistungen  $S_x^*(t_i)$ , welche wiederum durch Multiplikation des zugehörigen jahres- und altersspezifischen rohen Kopfschadens mit dem gepoolten (summierten) altersspezifischen Versicherungsbestand berechnet werden, d.h.

$$S_x^*(t_i) = K_x^r(t_i) \cdot \sum_{j=1}^n L_x(t_j). \quad (2.4)$$

Für den Leistungs-Skalierungsfaktor des Jahres  $t_i$  gilt dann

$$f(t_i) = \frac{\sum_{x=x_{\min}}^{x_{\max}} S_x^*(t_n)}{\sum_{x=x_{\min}}^{x_{\max}} S_x^*(t_i)}. \quad (2.5)$$

Die zu bestimmenden „gleichgerichteten“ Kopfschäden  $K_x^g$  ermitteln sich dann durch Bildung einer Konvexkombination der skalierten rohen Kopfschäden. Der Konvexkombinationsparameter  $\alpha(t)$  ermittelt sich anhand der Gesamtbestandsverteilung im Verlauf der Betrachtungsjahre, d.h.

$$\alpha(t_i) = \frac{\sum_{x=x_{\min}}^{x_{\max}} L_x(t_i)}{\sum_{j=1}^n \sum_{x=x_{\min}}^{x_{\max}} L_x(t_j)}, \quad (2.6)$$

womit für  $K_x^g$  folgt, dass

$$K_x^g = \sum_{j=1}^n K_x^r(t_j) \cdot f(t_j) \cdot \alpha(t_j). \quad (2.7)$$

Da das gleichgerichtete Profil vor allem bei schwach besetzten Einzelaltern Sprungstellen aufweisen kann, wird an dieser Stelle i.d.R. eine Ausgleichskurve bestimmt. Gängige Approximationsfunktionen sind Polynome  $p(x)$ , welche gemäß der Methode der kleinsten Quadrate ermittelt werden, wobei bei der Parameterschätzung die Größe des Bestands für die jeweiligen Einzelalter berücksichtigt wird. Nach allfälliger Normierung erhält man somit das gewünschte Profil  $k_x := \frac{p(x)}{p(x_0)}$ . Zur vollständigen Bestimmung des Leistungsprofils fehlt somit nur noch der entsprechende Grundkopfschaden. Hierfür werden zunächst retrospektiv anhand des bestimmten Profils die Grundkopfschäden  $G(t_i)$  der Berechnungsjahre bestimmt, d.h.

$$G(t_i) = \frac{S_x(t_i)}{\sum_{x=x_{\min}}^{x_{\max}} L_x(t_i) \cdot k_x}.$$

Schlussendlich wird der Grundkopfschaden  $G$  des aktuellen Jahres  $t_{n+1}$  anhand einfacher exponentieller Regression der  $G(t_i)$  ermittelt. Vor der bereits betrachteten

Einsteigerprämienkalkulation werden die aus dem obigen Verfahren gewonnenen Leistungsprofile - gemäß dem Prinzip der aktuariellen Vorsicht - noch mit entsprechenden Sicherheiten versehen, d.h. die gewonnenen Rechnungsgrundlagen zweiter Ordnung werden in zugehörige Rechnungsgrundlagen erster Ordnung übergeführt, wobei die „Polsterung“ mit Sicherheiten bei Leistungsprofilen meist durch Zuschlag eines altersunabhängigen Prozentsatzes erfolgt.

### 2.2.3. Leistungs- und Prämienanpassung

Wie bereits mehrfach erwähnt wurde, kommt es während der Laufzeit von Krankenversicherungsverträgen zu regelmäßigen Annahmen- bzw. Rechnungsgrundlagenänderungen. Der gängigste Fall einer solchen Änderung ist die Erhöhung der (erwarteten) Leistungen, d.h. eine Steigerung der Kopfschäden - eine Verringerung der Kopfschäden ist natürlich ebenfalls möglich. Eine Möglichkeit der Feststellung des Vorliegens einer solchen Leistungssteigerung bzw. die Bemessung der Höhe ebendieser erfolgt durch retrospektive Gegenüberstellung der - um Sicherheiten bereinigten - erwarteten Leistungen und den tatsächlich eingetretenen Leistungen. Überschreitet die relative Abweichung beider Vergleichsgrößen einen kritischen Wert, so werden die hinterlegten Leistungsprofile neu kalkuliert. Die Neukalkulation erfolgt entweder analog zu den Ausführungen des vorherigen Abschnitts - Gleichrichtung, Kurvenanpassung, etc. - oder es wird lediglich ein neuer Grundkopfschaden bestimmt. Die ausführlichere Neukalkulation greift tendenziell dann, wenn zusätzlich zum Vorliegen einer großen relativen Abweichung auf der Ebene des gesamten Tarifbestands noch größere Abweichungen in einzelnen Altersgruppen vorliegen, also die „Form“ des verwendeten Profils nicht mehr stimmig ist.

Eine Erhöhung der Kopfschäden resultiert - unter der Annahme, dass Übergangswahrscheinlichkeiten und Rechnungszins unberührt bleiben - für alle aktiven Versicherungsverträge in einer entsprechenden Erhöhung der jeweiligen Leistungsbarwerte. Da die Alterungsrückstellung der Differenz von Leistungsbarwert und Prämienbarwert entspricht, ist die gemäß der bisher gültigen Rechnungsgrundlagen gebildete Reserve unmittelbar nach der Erhöhung der erwarteten Leistungen nicht mehr ausreichend. Um diesen „plötzlichen“ Mehrbedarf an Reserve auszugleichen bedarf es einer entsprechenden Steigerung des Prämienbarwerts, d.h. einer Prämienanpassung. Die Höhe der Prämie nach der Anpassung wird hierbei i.d.R. so bestimmt, dass durch die Anpassung **kein** Zuschuss zur Alterungsrückstellung erfolgt. Kalkulatorisch ist dies analog zu einem Vertragsneuabschluss, bei dem die neuen Prämien aus einer - zu bestimmenden - laufenden Prämie und einer Einmalprämie in der Höhe der bisher angesparten Reserve bestehen. Betrachtet man die Prämienanpassung aus Sicht eines Vertrags so erhält man die folgende Berechnung: sei  $P$  die bis

für alle Betrachtungsjahre bestimmt. Im nächsten Schritt werden Skalierungsfaktoren  $f(t)$  ermittelt, um das Leistungsniveau der älteren Betrachtungsjahre auf jenes des Letztbetrachtungsjahres zu bringen. Die Faktoren  $f(t_i)$  bestimmen sich durch Berechnung des Quotienten von „künstlichen“ Gesamtleistungen  $S_x^*(t_i)$ , welche wiederum durch Multiplikation des zugehörigen jahres- und altersspezifischen rohen Kopfschadens mit dem gepoolten (summierten) altersspezifischen Versicherungsbestand berechnet werden, d.h.

$$S_x^*(t_i) = K_x^r(t_i) \cdot \sum_{j=1}^n L_x(t_j). \quad (2.4)$$

Für den Leistungs-Skalierungsfaktor des Jahres  $t_i$  gilt dann

$$f(t_i) = \frac{\sum_{x=x_{\min}}^{x_{\max}} S_x^*(t_n)}{\sum_{x=x_{\min}}^{x_{\max}} S_x^*(t_i)}. \quad (2.5)$$

Die zu bestimmenden „gleichgerichteten“ Kopfschäden  $K_x^g$  ermitteln sich dann durch Bildung einer Konvexkombination der skalierten rohen Kopfschäden. Der Konvexkombinationsparameter  $\alpha(t)$  ermittelt sich anhand der Gesamtbestandsverteilung im Verlauf der Betrachtungsjahre, d.h.

$$\alpha(t_i) = \frac{\sum_{x=x_{\min}}^{x_{\max}} L_x(t_i)}{\sum_{j=1}^n \sum_{x=x_{\min}}^{x_{\max}} L_x(t_j)}, \quad (2.6)$$

womit für  $K_x^g$  folgt, dass

$$K_x^g = \sum_{j=1}^n K_x^r(t_j) \cdot f(t_j) \cdot \alpha(t_j). \quad (2.7)$$

Da das gleichgerichtete Profil vor allem bei schwach besetzten Einzelaltern Sprungstellen aufweisen kann, wird an dieser Stelle i.d.R. eine Ausgleichskurve bestimmt. Gängige Approximationsfunktionen sind Polynome  $p(x)$ , welche gemäß der Methode der kleinsten Quadrate ermittelt werden, wobei bei der Parameterschätzung die Größe des Bestands für die jeweiligen Einzelalter berücksichtigt wird. Nach allfälliger Normierung erhält man somit das gewünschte Profil  $k_x := \frac{p(x)}{p(x_0)}$ . Zur vollständigen Bestimmung des Leistungsprofils fehlt somit nur noch der entsprechende Grundkopfschaden. Hierfür werden zunächst retrospektiv anhand des bestimmten Profils die Grundkopfschäden  $G(t_i)$  der Berechnungsjahre bestimmt, d.h.

$$G(t_i) = \frac{S_x(t_i)}{\sum_{x=x_{\min}}^{x_{\max}} L_x(t_i) \cdot k_x}.$$

Schlussendlich wird der Grundkopfschaden  $G$  des aktuellen Jahres  $t_{n+1}$  anhand einfacher exponentieller Regression der  $G(t_i)$  ermittelt. Vor der bereits betrachteten

## 3. Erweiterung des traditionellen Modells

Die im vorherigen Kapitel betrachtete Modellierung der Krankenversicherungsleistungen anhand von deterministischen Leistungsprofilen, welche im Lauf der Zeit regelmäßig semi-manuell aktualisiert bzw. neu berechnet werden, ist sowohl simpel als auch vielseitig. Dennoch existieren seit längerer Zeit unter anderem in der Bepreisung aber auch im Meldewesen Entwicklungstrends, welche erahnen lassen, dass die derzeitige Form der Modellierung den zukünftigen aktuariellen Ansprüchen nicht unbedingt gerecht werden kann. Angesichts dieser Umstände beschäftigt sich diese Arbeit mit einer naheliegenden Erweiterung der derzeitig gängigen Methodik. In diesem Kapitel werden die Grundgedanken dieser Erweiterung näher erläutert. Die folgenden Kapitel widmen sich dann der allgemeinen Handhabung sowie konkreten Anwendungsfällen des erweiterten Modells.

### 3.1. Motivation

In diesem Abschnitt werden die teilweise oben bereits erwähnten Beweggründe für eine mathematisch komplexere Modellierung näher erläutert.

#### 3.1.1. Präzisere Modellierung

Wie bereits angemerkt handelt es sich bei Kopfschäden um Durchschnittswerte, bei deren Berechnung oftmals Versicherungsleistungen mit stark auseinandergelassenen Kostenaufwänden vermengt werden. Hinzukommt natürlich die - vor allem in jungen Altern wesentliche - Wahrscheinlichkeit, dass binnen einer Periode gar keine Leistungen in Anspruch genommen werden. Somit liegt je nach Leistungsart ein gewisses Ausmaß an Volatilität vor, welches in der derzeitigen Modellierung vernachlässigt wird. Insbesondere im Hinblick auf Solvency II, wo die Berücksichtigung der Wahrscheinlichkeitsverteilung von Schäden eine der Anforderungen für die Verwendung eines sog. „internen Modells“ - ein von diversen Versicherungsunternehmen angestrebtes Ziel - ist, besteht an dieser Stelle durchaus Verbesserungspotenzial.

### 3.1.2. Unternehmenssteuerung

Zwei insbesondere für Versicherungsunternehmen wichtige Kriterien sind Stabilität und Planbarkeit der Geschäftsentwicklung - bspw. Bilanzen und entsprechende Übergangsrechnungen. Demnach besitzt die Minimierung von Ergebnisvolatilität hohe Priorität. Zur Verringerung der oben beschriebenen Volatilität in den Leistungen eignet sich der Abschluss von - meist einjährigen - Rückversicherungsverträgen. Oftmals werden derartige Versicherungsverträge für Gesamtbestände eines oder mehrerer Tarife abgeschlossen und die Versicherungsleistung wird anhand der Gesamtleistungen des versicherten Bestands ermittelt. Zur adäquaten Bepreisung eines solchen Rückversicherungsvertrages ist die Kenntnis der Verteilung der Gesamtleistungen eines Gesamtbestands natürlich hilfreich. Ein weiterer, besonders für potenzielle Investoren, aber auch für Versicherungsnehmer relevanter Aspekt ist die Planbarkeit von Unternehmensgewinnen, welche bspw. in Form von Gewinnbeteiligungen oder Dividenden ausgeschüttet werden. Hier gilt abermals das Ziel der Volatilitätsminimierung.

### 3.1.3. Verminderung des Verwaltungsaufwands

Angesichts der langen Vertragslaufzeiten, der Unkündbarkeit seitens des Versicherungsunternehmens, sowie den Schwierigkeiten bei der Konvertierung von gesamten Tarifbeständen, wächst die Menge an zu verwaltenden Produkten bzw. Tarifen mit der Zeit äußerst schnell. Da die meisten Tarife zumindest alle zwei Jahre angepasst werden, wächst gleichermaßen der Arbeitsaufwand, der für diesen routinierten Prozess benötigt wird. Somit besteht Bedarf an einer Vereinheitlichung oder gar Automatisierung des Anpassungsprozesses.

### 3.1.4. Individuelle Produktgestaltung

In der derzeitigen Produktlandschaft gibt es oftmals mehrere Versionen desselben Tarifs, die sich untereinander nur im Vorhandensein bzw. der (fixen) Höhe eines Selbstbehalts unterscheiden. Die Profilberechnung für derartige Selbstbehaltstarife erfolgt oft umständlich anhand entsprechender „Kürzung“ der Rohleistungen des Vollkostentarifs und anschließender Gleichrichtung sowie Grundkopfschadenbestimmung. Demnach ist die Menge an verfügbaren Selbstbehalten oft stark eingeschränkt. Da die Leistungsverrechnung meist direkt zwischen dem Versicherer und der jeweiligen Instanz, welche die Rechnung ausstellt, erfolgt und etwaige Selbstbehaltsabrechnungen separat zwischen Versicherer und Versicherungsnehmer erfolgen, kennt der Versicherer im Regelfall die vollen Leistungskosten - mögliche Verzögerungen der Leistungsanspruchnahme aufgrund eines vorliegenden

periodenbezogenen Selbstbehalts (vgl. Sachversicherung) werden implizit ausgeschlossen. Somit wäre eine dynamischere Modellierung, welche für einen gegebenen Selbstbehalt ohne großen Aufwand die adäquate Prämie bestimmt, eine einfache Möglichkeit den Versicherungsnehmern mehr Gestaltungsspielraum zukommen zu lassen und gleichzeitig die verschiedenen Versionen eines Tarifs zu vereinheitlichen. Eine weitere Thematik, bei der durch dynamische Selbstbehalte verhältnismäßig viel Verwaltungsaufwand eingespart werden könnte, sind die sog. „Anpassungsablehner“. Hierbei handelt es sich um Versicherungsnehmer, welche der Erhöhung ihrer Prämien im Zuge einer Prämienanpassung widersprechen. Auch in einem solchen Szenario ist es dem Versicherer nicht erlaubt, den Versicherungsvertrag zu kündigen, aber die vertraglich vereinbarten Leistungen dürfen in einem der Prämienanpassung entsprechenden Ausmaß gekürzt werden, sodass kein Vorteil gegenüber den Versicherungsnehmern, die die Anpassung nicht ablehnen, entsteht. Eine einfache Möglichkeit, diese Leistungskürzung ohne großen Mehraufwand umzusetzen, ist die Festlegung eines Selbstbehalts, sodass die Prämie unter Verwendung der angepassten Leistungsprofile dieselbe Höhe besitzt, wie die bisherige Prämie. Ein solches Vorgehen lässt sich zusätzlich gut damit vereinbaren, dass das einmalige Ablehnen einer Prämienanpassung nicht zwangsweise auch das Ablehnen aller zukünftigen Prämienanpassungen impliziert.

Einen weiteren Aspekt, für dessen adäquate Bepreisung zumindest die kalkulatorische Trennung von Eintrittswahrscheinlichkeit und Leistungshöhe relevant ist, bildet die Prämienrückerstattung. Hierbei wird im Fall von längerer Leistungsfreiheit ein Teil der Prämie rückerstattet - man vergleiche dies mit Bonus-Malus-Systemen in der Kfz-Haftpflichtversicherung. Analog zu den Selbstbehalten wird hier - sofern überhaupt vertraglich vereinbart - derzeit zumeist ein fixes Auszahlungsmuster vorgegeben. Das Zulassen verschiedener Muster - sowohl Dauer der für Auszahlung benötigten Leistungsfreiheit als auch Höhe der Auszahlung (meist relativ zur Prämie) - würde dem Versicherungsnehmer ebenfalls mehr individuellen Gestaltungsspielraum bieten.

## 3.2. Stochastische Kopfschäden

Den obigen Ausführungen lässt sich entnehmen, dass zur Bewältigung vieler Problemstellungen, bei denen die derzeitige Modellierung unzureichend ist, wohl ein stochastischer Zugang benötigt wird. Angesichts dessen besteht der Kern der Modellerweiterung darin, die bekannten deterministischen Leistungsprofile durch entsprechende stochastische Gegenstücke zu ersetzen. Somit entsprechen die bekannten Kopfschäden nun nicht mehr Durchschnittswerten (Erwartungswerten),

sondern Zufallsvariablen, d.h.  $K_x \sim F_x$ , für eine Verteilungsfunktion  $F_x$ .

Wie bereits erwähnt, spielt die Wahrscheinlichkeit, dass keine Leistung in Anspruch genommen wird, eine gesonderte Rolle. Daher wird im Folgenden an diversen Stellen die Darstellung

$$K_x \stackrel{d}{=} B_x \cdot Y_x, \quad B_x \perp Y_x, \quad (3.1)$$

verwendet, wobei  $B_x$  Bernoulli-verteilt mit Parameter  $\theta_x$  ist und die Schadeneintrittswahrscheinlichkeit abbildet, also

$$\mathbb{P}[B_x = 1] = \mathbb{P}[K_x > 0] = \Theta_x. \quad (3.2)$$

Dementsprechend modelliert  $Y_x > 0$  die Schadenhöhe, d.h.

$$Y_x \stackrel{d}{=} K_x \mid \{K_x > 0\}. \quad (3.3)$$

Zusätzlich wird angenommen, dass  $Y_x$  eine Dichte- bzw. Wahrscheinlichkeitsfunktion  $f_{\Phi_x}(y)$  besitzt, welche durch den Parametervektor  $\Phi_x$  spezifiziert wird, d.h.

$$\mathbb{P}[Y_x = y] = f_{\Phi_x}(y). \quad (3.4)$$

Die Trennung von Schadeneintritt und Schadenhöhe wurde in Anlehnung an das in der Sachversicherung verwendete kollektive Modell gewählt. Ein ähnlicher Ansatz wird auch in „Zusatzversicherung für Krankenhauskosten mit wählbarem Selbstbehalt“ [4] betrachtet.

### 3.3. Profilberechnung mit Bayes'scher Inferenz

Analog zu Abschnitt 2.2.2 widmet sich dieser Abschnitt der Bestimmung von Leistungsprofilen, wobei der Fokus nun statt auf deterministischen Kopfschadenreihen, auf den Verteilungsparametern  $\theta_x$  und  $\phi_x$  mit  $x \in \{x_{\min}, \dots, x_{\max}\}$  liegt. Das Ziel besteht also darin, anhand von historischen Leistungs- und Bestandsdaten, sowie der Berücksichtigung von zeitlichen Trends, die Verteilung der Kopfschäden  $K_x$  zu schätzen. Für die Bestimmung einer solchen Schätzung gibt es natürlich mehrere Möglichkeiten. Beispielsweise könnte mit der Momentenmethode für jedes vergangene Jahr ein Punktschätzer für die Verteilungsparameter berechnet werden und schließlich anhand von Regression ein Endschätzwert extrapoliert werden. Dies würde jedoch voraussetzen, dass sich die jeweiligen Verteilungsparameter anhand der einzelnen Momente, bzw. anhand der bereits geschätzten Parameter bestimmen lassen, was bei komplexeren Verteilungen mit vielen zu schätzenden Parametern großen Aufwand bedeutet. Des Weiteren würde diese Methode für sich alleine lediglich einen

Schätzwert pro Parameter liefern, d.h. es ist nicht erkennbar, wie „passend“ bzw. unsicher die resultierende Schätzung ist - was bei geringen Bestandszahlen für einzelne (meist höhere) Alter durchaus relevant ist. Derartige Information könnte bspw. in die Festlegung des - gemäß des Prinzips der aktuariellen Vorsichts zu bildenden - Sicherheitszuschlags einfließen, oder sogar die Verwendung eines anderen Verteilungstyps veranlassen. Zusammengefasst wäre es also vorteilhaft, wenn das für die Parameterschätzung verwendete Verfahren möglichst flexibel, d.h. es bedarf keiner komplexen, vom gewählten Verteilungstyp abhängigen, Berechnungslogik, aber zugleich auch informativ im Bezug auf die Unsicherheit in den von ihm produzierten Schätzwerten ist. Eine Möglichkeit, um diesen Ansprüchen gerecht zu werden, bildet eine Modifikation des Verfahrens der Bayes'schen Inferenz, welches in den folgenden Teilabschnitten näher erläutert wird.

#### 3.3.1. Bayes'sche Inferenz allgemein

Die nachfolgende Darstellung folgt im Wesentlichen „The power prior: theory and applications“ [5]. Im Folgenden wird die Zufallsvariable  $Y$ , mit zugehöriger Dichte- bzw. Wahrscheinlichkeitsfunktion  $f_{\Theta}(y)$ , welche durch den (zu schätzenden) Parametervektor  $\Theta \in M \subset \mathbb{R}^d$  mit  $d \in \mathbb{N}$  spezifiziert wird, betrachtet. Zusätzlich ist eine Stichprobe  $\mathbf{y} = (y_1, \dots, y_n)$  gegeben, wobei die einzelnen  $y_i, 1 \leq i \leq n$  unabhängigen Realisationen von  $Y$  entsprechen - weiter unten werden hierfür die (bedingt) unabhängig identisch verteilten Zufallsvariablen  $Y_1, \dots, Y_n$  verwendet. Der Grundgedanke bei Bayes'scher Inferenz besteht - im Unterschied zur frequentistischen Statistik - in der Annahme, dass der zu schätzende Parameter(vektor)  $\Theta$  selbst ebenfalls eine Zufallsvariable ist, d.h.  $\Theta$  besitzt nun eine Dichtefunktion  $\pi(\theta)$  mit Träger  $U \subset M$ . Hinzukommt, dass  $Y$  nunmehr nur bedingt auf  $\Theta$  die Dichte  $f_{\Theta}$  besitzt, also gilt für alle  $\theta \in U$ , dass

$$Y \mid \{\Theta = \theta\} \sim F_{\theta}, \quad F_{\theta}(y) = \int_{-\infty}^y f_{\theta}(s) ds, \quad (3.5)$$

bzw. in vereinfachter Notation

$$Y \mid \Theta \sim F_{\Theta}. \quad (3.6)$$

Im Zuge des Verfahrens wird die Dichtefunktion  $\pi(\theta)$  bzw. die Verteilung von  $\Theta$  anhand der observierten Daten, also der Stichprobe  $\mathbf{y}$ , „aktualisiert“ - formal betrachtet wird die auf die Stichprobe bedingte Dichtefunktion  $\pi(\theta \mid \mathbf{y})$  bestimmt. Unter An-

wendung des - namensgebenden - Satzes von Bayes folgt

$$\begin{aligned}
 \pi(\theta | \mathbf{y}) &= \mathbb{P}[\Theta = \theta | Y_1 = y_1, \dots, Y_n = y_n] \\
 &= \frac{\mathbb{P}[\Theta = \theta, Y_1 = y_1, \dots, Y_n = y_n]}{\underbrace{\mathbb{P}[Y_1 = y_1, \dots, Y_n = y_n]}_{=:C(\mathbf{y})}} \\
 &= C(\mathbf{y})^{-1} \mathbb{P}[Y_1 = y_1, \dots, Y_n = y_n | \Theta = \theta] \cdot \mathbb{P}[\Theta = \theta] \\
 &= C(\mathbf{y})^{-1} \left( \prod_{i=1}^n \mathbb{P}[Y_i = y_i | \Theta = \theta] \right) \cdot \mathbb{P}[\Theta = \theta] \\
 &= C(\mathbf{y})^{-1} \underbrace{\left( \prod_{i=1}^n f_{\theta}(y_i) \right)}_{=: \mathcal{L}(\mathbf{y} | \theta)} \pi(\theta) \\
 &= C(\mathbf{y})^{-1} \mathcal{L}(\mathbf{y} | \theta) \pi(\theta)
 \end{aligned} \tag{3.7}$$

Im Allgemeinen wird die Ausgangsdichtefunktion  $\pi(\theta)$  als „Prior“ und die aktualisierte Dichtefunktion  $\pi(\theta | \mathbf{y})$  als „Posterior“ bezeichnet. Die Funktion  $\mathcal{L}(\mathbf{y} | \theta)$  wird „Likelihood“ genannt, während die - nicht von  $\theta$  abhängende - Normierungskonstante  $C(\mathbf{y})$ , also die unbedingte Verteilung von  $Y$ , als „Evidence“ bezeichnet wird. Gemäß des Satzes der totalen Wahrscheinlichkeit gilt

$$C(\mathbf{y}) = \int_U \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) d\theta. \tag{3.8}$$

Da die Bestimmung von  $C(\mathbf{y})$  - wie oben erkennbar - rechentechnisch meist aufwändig ist und statistische Methoden bzw. Schätzer welche auf Bayes'scher Inferenz basieren zumeist auch ohne korrekte Normierung auskommen, ist der Evidence im Vergleich zu den anderen Komponenten weniger bedeutsam. Dies spiegelt sich auch in der bekannten vereinfachten Formel

$$\pi(\theta | \mathbf{y}) \propto \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) \tag{3.9}$$

Posterior  $\propto$  Likelihood  $\times$  Prior,

wider. Zusammengefasst verbindet der Prozess der Bayes'schen Inferenz also eine Ausgangsschätzung in Form der Prior-Verteilung mit observierten Daten in Form der Likelihood zu einer aktualisierten Schätzung in Form der Posterior-Verteilung. Anhand der Posterior-Verteilung können dann sowohl Punkt-Schätzer für die Verteilungsparameter als auch Maße für die Unsicherheit dieser Schätzer ermittelt werden. Gängige Punktschätzer für  $\Theta$  sind bspw. der „Maximum-a-posteriori“-Schätzer  $\Theta^{\text{MAP}}$ , mit

$$\Theta^{\text{MAP}} = \arg \max_{\theta \in U} \pi(\theta | \mathbf{y}). \tag{3.10}$$

wendung des - namensgebenden - Satzes von Bayes folgt

$$\begin{aligned}
 \pi(\theta | \mathbf{y}) &= \mathbf{P}[\bullet = \theta | Y_1 = y_1, \dots, Y_n = y_n] \\
 &= \frac{\mathbf{P}[\bullet = \theta, Y_1 = y_1, \dots, Y_n = y_n]}{\underbrace{\mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n]}_{=:C(\mathbf{y})}} \\
 &= C(\mathbf{y})^{-1} \mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n | \bullet = \theta] \cdot \mathbf{P}[\bullet = \theta] \\
 &= C(\mathbf{y})^{-1} \left( \prod_{i=1}^n \mathbf{P}[Y_i = y_i | \bullet = \theta] \right) \cdot \mathbf{P}[\bullet = \theta] \\
 &= C(\mathbf{y})^{-1} \underbrace{\left( \prod_{i=1}^n f_{\theta}(y_i) \right)}_{=: \mathcal{L}(\mathbf{y} | \theta)} \pi(\theta) \\
 &= C(\mathbf{y})^{-1} \mathcal{L}(\mathbf{y} | \theta) \pi(\theta)
 \end{aligned} \tag{3.7}$$

Im Allgemeinen wird die Ausgangsdichtefunktion  $\pi(\theta)$  als „Prior“ und die aktualisierte Dichtefunktion  $\pi(\theta | \mathbf{y})$  als „Posterior“ bezeichnet. Die Funktion  $\mathcal{L}(\mathbf{y} | \theta)$  wird „Likelihood“ genannt, während die - nicht von  $\theta$  abhängende - Normierungskonstante  $C(\mathbf{y})$ , also die unbedingte Verteilung von  $Y$ , als „Evidence“ bezeichnet wird. Gemäß des Satzes der totalen Wahrscheinlichkeit gilt

$$C(\mathbf{y}) = \int_U \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) d\theta. \tag{3.8}$$

Da die Bestimmung von  $C(\mathbf{y})$  - wie oben erkennbar - rechtechnisch meist aufwändig ist und statistische Methoden bzw. Schätzer welche auf Bayes'scher Inferenz basieren zumeist auch ohne korrekte Normierung auskommen, ist der Evidence im Vergleich zu den anderen Komponenten weniger bedeutsam. Dies spiegelt sich auch in der bekannten vereinfachten Formel

$$\pi(\theta | \mathbf{y}) \propto \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) \tag{3.9}$$

Posterior  $\propto$  Likelihood  $\times$  Prior,

wieder. Zusammengefasst verbindet der Prozess der Bayes'schen Inferenz also eine Ausgangsschätzung in Form der Prior-Verteilung mit observierten Daten in Form der Likelihood zu einer aktualisierten Schätzung in Form der Posterior-Verteilung. Anhand der Posterior-Verteilung können dann sowohl Punkt-Schätzer für die Verteilungsparameter als auch Maße für die Unsicherheit dieser Schätzer ermittelt werden. Gängige Punktschätzer für  $\theta$  sind bspw. der „Maximum-a-posteriori“-Schätzer  $\theta^{\text{MAP}}$ , mit

$$\theta^{\text{MAP}} = \arg \max_{\theta \in U} \pi(\theta | \mathbf{y}). \tag{3.10}$$

Gemäß der Formel für die Varianz einer Beta-Verteilung gilt

$$\hat{\sigma} = \sqrt{\frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}} = \sqrt{\frac{112}{11132}} \approx 0.10. \quad (3.17)$$

Damit ergibt sich für den Variationskoeffizienten

$$\text{VarK}(\Theta) = \frac{\sqrt{\mathbb{V}[\Theta]}}{\mathbb{E}[\Theta]} = \frac{\hat{\sigma}}{\hat{\Theta}} \approx 0.16. \quad (3.18)$$

Die eben durchgeführte Berechnung kann direkt für die Schätzung von Schaden-  
 eintrittswahrscheinlichkeiten, also die Schätzung der Verteilungsparameter der  $B_x$ ,  
 herangezogen werden. Der oben bestimmte Variationskoeffizient fließt dann - zusam-  
 men mit einem entsprechenden Maß, welches aus der Schätzung der Parameter von  
 $Y_x$  resultiert - in die Wahl eines adäquaten Sicherheitszuschlags ein.

### 3.3.2. Zeitabhängigkeit und Power Prior

Betrachtet man erneut das obige Beispiel des Münzwurfs, so deutet die Höhe des be-  
 stimmten Variationskoeffizienten bzw. die Größe der Stichprobe auf eine verhältnis-  
 mäßig unsichere Schätzung hin. Diese Unsicherheit sollte sich verringern, wenn mehr  
 Daten, also mehr observierte Münzwürfe, in die Schätzung einfließen. Sei nun also  $\mathbf{z}$   
 eine zweite Stichprobe, für die die Münze  $m := 30$  Mal geworfen wurde und 14-mal  
 das Ereignis Kopf auftrat. Nun sollen die bisherigen Ergebnisse und die neue Stich-  
 probe zusammengeführt werden. Eine intuitive Lösung wäre die Vermengung beider  
 Stichproben und ein erneuter Beginn am Anfang, d.h. bei der uniform-verteilten Pri-  
 or. Aufgrund der Simplizität des verwendeten Modells und der bescheidenen Menge  
 an Daten, wäre dieser Zugang durchaus vertretbar, aber im Allgemeinen kommt im  
 Fall von mehreren Stichproben eine wesentliche - zum gleichen Endergebnis führende  
 - Eigenschaft der Bayes'schen Inferenz zur Geltung. Diese Eigenschaft ist die Verträ-  
 glichkeit bezüglich sequenzieller Anwendung des Verfahrens, d.h. die anhand der ersten  
 Stichprobe ermittelte Posterior-Verteilung  $\pi(\theta | \mathbf{y})$  kann als neuer Ausgangspunkt,  
 also als neue Prior-Verteilung verwendet werden. Für die neue Posterior-Verteilung  
 $\pi(\theta | \mathbf{y}, \mathbf{z})$  gilt dann

$$\pi(\theta | \mathbf{y}, \mathbf{z}) \propto \mathcal{L}(\mathbf{z} | \theta, \mathbf{y})\pi(\theta | \mathbf{y}). \quad (3.19)$$

wendung des - namensgebender - Satzes von Bayes folgt

$$\begin{aligned}
 \pi(\theta | \mathbf{y}) &= \mathbf{P}[\bullet = \theta | Y_1 = y_1, \dots, Y_n = y_n] \\
 &= \frac{\mathbf{P}[\bullet = \theta, Y_1 = y_1, \dots, Y_n = y_n]}{\underbrace{\mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n]}_{=: C(\mathbf{y})}} \\
 &= C(\mathbf{y})^{-1} \mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n | \bullet = \theta] \cdot \mathbf{P}[\bullet = \theta] \\
 &= C(\mathbf{y})^{-1} \left( \prod_{i=1}^n \mathbf{P}[Y_i = y_i | \bullet = \theta] \right) \cdot \mathbf{P}[\bullet = \theta] \\
 &= C(\mathbf{y})^{-1} \underbrace{\left( \prod_{i=1}^n f_{\theta}(y_i) \right)}_{=: \mathcal{L}(\mathbf{y} | \theta)} \pi(\theta) \\
 &= C(\mathbf{y})^{-1} \mathcal{L}(\mathbf{y} | \theta) \pi(\theta)
 \end{aligned} \tag{3.7}$$

Im Allgemeinen wird die Ausgangsdichtefunktion  $\pi(\theta)$  als „Prior“ und die aktualisierte Dichtefunktion  $\pi(\theta | \mathbf{y})$  als „Posterior“ bezeichnet. Die Funktion  $\mathcal{L}(\mathbf{y} | \theta)$  wird „Likelihood“ genannt, während die - nicht von  $\theta$  abhängende - Normierungskonstante  $C(\mathbf{y})$ , also die unbedingte Verteilung von  $Y$ , als „Evidence“ bezeichnet wird. Gemäß des Satzes der totalen Wahrscheinlichkeit gilt

$$C(\mathbf{y}) = \int_U \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) d\theta. \tag{3.8}$$

Da die Bestimmung von  $C(\mathbf{y})$  - wie oben erkennbar - rechentechnisch meist aufwändig ist und statistische Methoden bzw. Schätzer welche auf Bayes'scher Inferenz basieren zumeist auch ohne korrekte Normierung auskommen, ist der Evidence im Vergleich zu den anderen Komponenten weniger bedeutsam. Dies spiegelt sich auch in der bekannten vereinfachten Formel

$$\pi(\theta | \mathbf{y}) \propto \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) \tag{3.9}$$

Posterior  $\propto$  Likelihood  $\times$  Prior,

wieder. Zusammengefasst verbindet der Prozess der Bayes'schen Inferenz also eine Ausgangsschätzung in Form der Prior-Verteilung mit observierten Daten in Form der Likelihood zu einer aktualisierten Schätzung in Form der Posterior-Verteilung. Anhand der Posterior-Verteilung können dann sowohl Punkt-Schätzer für die Verteilungsparameter als auch Maße für die Unsicherheit dieser Schätzer ermittelt werden. Gängige Punktschätzer für  $\theta$  sind bspw. der „Maximum-a-posteriori“-Schätzer  $\theta^{\text{MAP}}$ , mit

$$\theta^{\text{MAP}} = \arg \max_{\theta \in U} \pi(\theta | \mathbf{y}). \tag{3.10}$$

Parametern  $(\alpha, \beta) = (A_n + 1, B_n + 1)$ . Setzt man in die bereits verwendete Formel für die Varianz der Beta-Verteilung ein und betrachtet den Limes für  $(A_n + B_n) \rightarrow \infty$ , so ist erkennbar, dass der Grenzwert der Posterior-Verteilung einer Dirac-Verteilung entspricht. Übernimmt man dieses Modell nun erneut direkt für die Schätzung von Schadeneintrittswahrscheinlichkeiten, welche sich im Lauf der Zeit durchaus verändern können, so entsteht das Problem, dass nahe und ferne Vergangenheit implizit in gleichem Maße gewichtet werden. Betrachtet man als Beispiel die zeitlich aufeinanderfolgend observierten Stichproben

$$\begin{aligned} \mathbf{y}_1 &= (0, 0, 0, 0, 0, 0, 0, 0, 1, 1), \\ \mathbf{y}_2 &= (0, 0, 0, 0, 0, 0, 0, 1, 1, 1), \\ \mathbf{y}_3 &= (0, 0, 0, 0, 0, 0, 1, 1, 1, 1), \\ \mathbf{y}_4 &= (0, 0, 0, 0, 0, 1, 1, 1, 1, 1), \\ \mathbf{y}_5 &= (0, 0, 0, 0, 1, 1, 1, 1, 1, 1), \end{aligned} \tag{3.23}$$

so ist ein wachsender Trend in der Schadeneintrittswahrscheinlichkeit erkennbar. Ausgehend von einer flachen Prior-Verteilung erhält man die Posterior-Verteilungen  $\pi_i := \pi(\theta \mid \mathbf{y}_1, \dots, \mathbf{y}_i)$  mit Posterior-Erwartungswerten  $\hat{\Theta}_i$  und Posterior-Standardabweichungen  $\hat{\sigma}_i$

$$\begin{aligned} \pi_1 &\sim \text{Beta}(3, 9), & \hat{\Theta}_1 &= 0.2500, & \hat{\sigma}_1 &\approx 0.1201, \\ \pi_2 &\sim \text{Beta}(6, 16), & \hat{\Theta}_2 &\approx 0.2727, & \hat{\sigma}_2 &\approx 0.0929, \\ \pi_3 &\sim \text{Beta}(10, 22), & \hat{\Theta}_3 &\approx 0.3125, & \hat{\sigma}_3 &\approx 0.0807, \\ \pi_4 &\sim \text{Beta}(15, 27), & \hat{\Theta}_4 &\approx 0.3571, & \hat{\sigma}_4 &\approx 0.0731, \\ \pi_5 &\sim \text{Beta}(21, 31), & \hat{\Theta}_5 &\approx 0.4038, & \hat{\sigma}_5 &\approx 0.0674. \end{aligned} \tag{3.24}$$

Somit ist der wachsende Trend zwar verzögert auch in den Punktschätzern (hier  $\mathbf{y}_i$ ) erkennbar, aber die - hier anhand  $\hat{\sigma}_i$  bemessene - Unsicherheit, welche intuitiv betrachtet aufgrund des Trends ebenfalls zunehmen sollte, sinkt sogar. Hieran lässt sich ableiten, dass für die adäquate Schätzung von Parametern, welche sich mit der Zeit verändern, eine Modifikation des Verfahrens benötigt wird. Eine Möglichkeit dies zu bewerkstelligen besteht in der Einführung einer weiteren Ebene von Hyperparametern, was im wesentlichen bedeuten würde, dass nunmehr nicht die Verteilung der zu schätzenden Parameter stationär ist, sondern dass die Verteilung der Veränderung der Parameter-Verteilung stationär ist. Bezogen auf das obige Beispiel wären die Beta-Parameter  $(\alpha, \beta)$  abermals Zufallsvariablen, welche nunmehr implizit eine stationäre Verteilung besitzen. Diese Vorgangsweise würde aber je nach (Basis-)Verteilungstyp schnell für enorme Zuwächse in der Modellkomplexität - und damit auch im Rechenaufwand - sorgen. Eine weitere Möglichkeit der Problembewältigung, welche im Zuge dieser Arbeit näher betrachtet wird, bildet die sog. „Power Prior“. Im Wesentlichen

wird hierbei im sequenziellen Inferenz-Verfahren ein weiterer Zwischenschritt eingebaut. Dieser Zwischenschritt entspricht der Diskontierung der bis zu diesem Schritt gesammelten Information, d.h. statt die zuvor gewonnene Posterior-Verteilung direkt als Prior-Verteilung für den nächsten Schritt zu verwenden, wird diese zuvor durch Anwendung einer geeigneten Wurzelfunktion, also Potenzieren mit Exponent  $p \in (0, 1)$ , „diskontiert“. Für aufeinanderfolgende Stichproben  $\mathbf{y}_1, \dots, \mathbf{y}_n$  bedeutet dies

$$\pi(\theta \mid \mathbf{y}_1, \dots, \mathbf{y}_{i+1}) \propto \mathcal{L}(\mathbf{y}_{i+1} \mid \theta) (\pi(\theta \mid \mathbf{y}_1, \dots, \mathbf{y}_i))^p. \quad (3.25)$$

Die Wahl von  $p$  steuert somit die „Anpassungswilligkeit“ der Posterior an neue Daten. Für  $p = 1$  entspricht die obige Formel der regulären Bayes’schen Inferenz, d.h. die Reihenfolge der Stichproben hat keinen Einfluss, für  $p = 0$  entspricht die in jedem Schritt verwendete Prior einer Uniformverteilung, womit zuvor observierte Daten nicht in die Schätzung einfließen. Zusätzlich muss  $p$  natürlich nicht als konstant angenommen werden, womit einzelne Stichproben auch unterschiedlich diskontiert werden können, sodass etwaigen Ausreißern weniger Gewicht zukommt.

### 3.3.3. Profilberechnung

Im Folgenden werden die Zerlegungsfaktoren der Kopfschäden -  $B_x$  und  $Y_x$  - im Kontext der Bayes’schen Inferenz betrachtet, d.h. die jeweiligen Parameter(vektoren)  $\theta_x$  bzw.  $\phi_x$  werden abermals als Zufallsvariablen aufgefasst und es gilt

$$B_x \mid \theta_x \sim B(\theta_x), \quad (3.26)$$

sowie

$$Y_x \mid \phi_x \sim F_{\phi_x}. \quad (3.27)$$

Die jeweiligen observierten Stichproben  $\mathbf{b}_{x,t} = (B_{x,t,1}, \dots, B_{x,t,n_t})$  bzw.  $\mathbf{y}_{x,t} = (Y_{x,t,1}, \dots, Y_{x,t,m_t})$ , d.h. der (Nicht-)Eintritt von Schäden, bzw. die im Eintrittsfall observierten Schadenhöhen, werden als bedingt auf den jeweiligen Parameter ( $\theta_x$  bzw.  $\phi_x$ ) paarweise unabhängige Zufallsvariablen betrachtet, d.h.

$$\mathbb{P}[B_{x,t,j}, B_{x,s,k} \mid \theta_x] = \mathbb{P}[B_{x,t,j} \mid \theta_x] \cdot \mathbb{P}[B_{x,s,k} \mid \theta_x], \quad (3.28)$$

sowie

$$\mathbb{P}[Y_{x,t,j}, Y_{x,s,k} \mid \phi_x] = \mathbb{P}[Y_{x,t,j} \mid \phi_x] \cdot \mathbb{P}[Y_{x,s,k} \mid \phi_x], \quad (3.29)$$

für  $t, s \in \{t_{\min}, \dots, t_{\max}\} =: T$ . Des Weiteren seien  $p_x(t), q_x(t) : T \rightarrow [0, 1]$  mit  $p_x(t_{\max}) = q_x(t_{\max}) = 1$  zeitabhängige Diskontierungsfaktoren für die Schätzung von  $\phi_x$  bzw.  $\theta_x$ . Da  $\theta_x \in (0, 1)$  kann hier erneut eine uniforme Initial-Prior  $\pi_x^B(\theta)$  herangezogen werden. Die (bedingte) Schadenhöhenverteilung  $F_{\phi_x}$  ist an dieser Stelle nicht näher spezifiziert, d.h. es kann nicht davon ausgegangen werden, dass die zulässige

wenigung des - namensgebenden - Satzes von Bayes folgt

$$\begin{aligned}
 \pi(\theta | \mathbf{y}) &= \mathbf{P}[\Theta = \theta | Y_1 = y_1, \dots, Y_n = y_n] \\
 &= \frac{\mathbf{P}[\Theta = \theta, Y_1 = y_1, \dots, Y_n = y_n]}{\underbrace{\mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n]}_{=: C(\mathbf{y})}} \\
 &= C(\mathbf{y})^{-1} \mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n | \Theta = \theta] \cdot \mathbf{P}[\Theta = \theta] \\
 &= C(\mathbf{y})^{-1} \left( \prod_{i=1}^n \mathbf{P}[Y_i = y_i | \Theta = \theta] \right) \cdot \mathbf{P}[\Theta = \theta] \\
 &= C(\mathbf{y})^{-1} \underbrace{\left( \prod_{i=1}^n f_{\theta}(y_i) \right)}_{=: \mathcal{L}(\mathbf{y} | \theta)} \pi(\theta) \\
 &= C(\mathbf{y})^{-1} \mathcal{L}(\mathbf{y} | \theta) \pi(\theta)
 \end{aligned} \tag{3.7}$$

Im Allgemeinen wird die Ausgangsdichtefunktion  $\pi(\theta)$  als „Prior“ und die aktualisierte Dichtefunktion  $\pi(\theta | \mathbf{y})$  als „Posterior“ bezeichnet. Die Funktion  $\mathcal{L}(\mathbf{y} | \theta)$  wird „Likelihood“ genannt, während die - nicht von  $\theta$  abhängende - Normierungskonstante  $C(\mathbf{y})$ , also die unbedingte Verteilung von  $Y$ , als „Evidence“ bezeichnet wird. Gemäß des Satzes der totalen Wahrscheinlichkeit gilt

$$C(\mathbf{y}) = \int_U \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) d\theta. \tag{3.8}$$

Da die Bestimmung von  $C(\mathbf{y})$  - wie oben erkennbar - rechtechnisch meist aufwändig ist und statistische Methoden bzw. Schätzer welche auf Bayes'scher Inferenz basieren zumeist auch ohne korrekte Normierung auskommen, ist der Evidence im Vergleich zu den anderen Komponenten weniger bedeutsam. Dies spiegelt sich auch in der bekannten vereinfachten Formel

$$\pi(\theta | \mathbf{y}) \propto \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) \tag{3.9}$$

Posterior  $\propto$  Likelihood  $\times$  Prior,

wider. Zusammengefasst verbindet der Prozess der Bayes'schen Inferenz also eine Ausgangsschätzung in Form der Prior-Verteilung mit observierten Daten in Form der Likelihood zu einer aktualisierten Schätzung in Form der Posterior-Verteilung. Anhand der Posterior-Verteilung können dann sowohl Punkt-Schätzer für die Verteilungsparameter als auch Maße für die Unsicherheit dieser Schätzer ermittelt werden. Gängige Punktschätzer für  $\Theta$  sind bsp.w. der „Maximum-a-posteriori“-Schätzer  $\Theta^{\text{MAP}}$ , mit

$$\Theta^{\text{MAP}} = \arg \max_{\theta \in U} \pi(\theta | \mathbf{y}). \tag{3.10}$$

wendung des - namensgebenden - Satzes von Bayes folgt

$$\begin{aligned}
 \pi(\theta | \mathbf{y}) &= \mathbf{P}[\bullet = \theta | Y_1 = y_1, \dots, Y_n = y_n] \\
 &= \frac{\mathbf{P}[\bullet = \theta, Y_1 = y_1, \dots, Y_n = y_n]}{\underbrace{\mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n]}_{=: C(\mathbf{y})}} \\
 &= C(\mathbf{y})^{-1} \mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n | \bullet = \theta] \cdot \mathbf{P}[\bullet = \theta] \\
 &= C(\mathbf{y})^{-1} \left( \prod_{i=1}^n \mathbf{P}[Y_i = y_i | \bullet = \theta] \right) \cdot \mathbf{P}[\bullet = \theta] \\
 &= C(\mathbf{y})^{-1} \underbrace{\left( \prod_{i=1}^n f_{\bullet}(y_i) \right)}_{=: \mathcal{L}(\mathbf{y} | \theta)} \pi(\theta) \\
 &= C(\mathbf{y})^{-1} \mathcal{L}(\mathbf{y} | \theta) \pi(\theta)
 \end{aligned} \tag{3.7}$$

Im Allgemeinen wird die Ausgangsdichtefunktion  $\pi(\theta)$  als „Prior“ und die aktualisierte Dichtefunktion  $\pi(\theta | \mathbf{y})$  als „Posterior“ bezeichnet. Die Funktion  $\mathcal{L}(\mathbf{y} | \theta)$  wird „Likelihood“ genannt, während die - nicht von  $\theta$  abhängende - Normierungskonstante  $C(\mathbf{y})$ , also die unbedingte Verteilung von  $Y$ , als „Evidence“ bezeichnet wird. Gemäß des Satzes der totalen Wahrscheinlichkeit gilt

$$C(\mathbf{y}) = \int_U \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) d\theta. \tag{3.8}$$

Da die Bestimmung von  $C(\mathbf{y})$  - wie oben erkennbar - rechentechnisch meist aufwändig ist und statistische Methoden bzw. Schätzer welche auf Bayes'scher Inferenz basieren zumeist auch ohne korrekte Normierung auskommen, ist der Evidence im Vergleich zu den anderen Komponenten weniger bedeutsam. Dies spiegelt sich auch in der bekannten vereinfachten Formel

$$\pi(\theta | \mathbf{y}) \propto \mathcal{L}(\mathbf{y} | \theta) \pi(\theta) \tag{3.9}$$

Posterior  $\propto$  Likelihood  $\times$  Prior,

wieder. Zusammengefasst verbindet der Prozess der Bayes'schen Inferenz also eine Ausgangsschätzung in Form der Prior-Verteilung mit observierten Daten in Form der Likelihood zu einer aktualisierten Schätzung in Form der Posterior-Verteilung. Anhand der Posterior-Verteilung können dann sowohl Punkt-Schätzer für die Verteilungsparameter als auch Maße für die Unsicherheit dieser Schätzer ermittelt werden. Gängige Punktschätzer für  $\theta$  sind bspw. der „Maximum-a-posteriori“-Schätzer  $\theta^{\text{MAP}}$ , mit

$$\theta^{\text{MAP}} = \arg \max_{\theta \in U} \pi(\theta | \mathbf{y}). \tag{3.10}$$

wendung des - namensgebenden - Satzes von Bayes folgt

$$\begin{aligned}
 \pi(\boldsymbol{\theta} \mid \mathbf{y}) &= \mathbf{P}[\boldsymbol{\theta} = \boldsymbol{\theta} \mid Y_1 = y_1, \dots, Y_n = y_n] \\
 &= \frac{\mathbf{P}[\boldsymbol{\theta} = \boldsymbol{\theta}, Y_1 = y_1, \dots, Y_n = y_n]}{\underbrace{\mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n]}_{=:C(\mathbf{y})}} \\
 &= C(\mathbf{y})^{-1} \mathbf{P}[Y_1 = y_1, \dots, Y_n = y_n \mid \boldsymbol{\theta} = \boldsymbol{\theta}] \cdot \mathbf{P}[\boldsymbol{\theta} = \boldsymbol{\theta}] \\
 &= C(\mathbf{y})^{-1} \left( \prod_{i=1}^n \mathbf{P}[Y_i = y_i \mid \boldsymbol{\theta} = \boldsymbol{\theta}] \right) \cdot \mathbf{P}[\boldsymbol{\theta} = \boldsymbol{\theta}] \\
 &= C(\mathbf{y})^{-1} \underbrace{\left( \prod_{i=1}^n f_{\boldsymbol{\theta}}(y_i) \right)}_{=: \mathcal{L}(\mathbf{y} \mid \boldsymbol{\theta})} \pi(\boldsymbol{\theta}) \\
 &= C(\mathbf{y})^{-1} \mathcal{L}(\mathbf{y} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta})
 \end{aligned} \tag{3.7}$$

Im Allgemeinen wird die Ausgangsdichtefunktion  $\pi(\boldsymbol{\theta})$  als „Prior“ und die aktualisierte Dichtefunktion  $\pi(\boldsymbol{\theta} \mid \mathbf{y})$  als „Posterior“ bezeichnet. Die Funktion  $\mathcal{L}(\mathbf{y} \mid \boldsymbol{\theta})$  wird „Likelihood“ genannt, während die - nicht von  $\boldsymbol{\theta}$  abhängende - Normierungskonstante  $C(\mathbf{y})$ , also die unbedingte Verteilung von  $Y$ , als „Evidence“ bezeichnet wird. Gemäß des Satzes der totalen Wahrscheinlichkeit gilt

$$C(\mathbf{y}) = \int_U \mathcal{L}(\mathbf{y} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{3.8}$$

Da die Bestimmung von  $C(\mathbf{y})$  - wie oben erkennbar - rechtechnisch meist aufwändig ist und statistische Methoden bzw. Schätzer welche auf Bayes'scher Inferenz basieren zumeist auch ohne korrekte Normierung auskommen, ist der Evidence im Vergleich zu den anderen Komponenten weniger bedeutsam. Dies spiegelt sich auch in der bekannten vereinfachten Formel

$$\pi(\boldsymbol{\theta} \mid \mathbf{y}) \propto \mathcal{L}(\mathbf{y} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) \tag{3.9}$$

Posterior  $\propto$  Likelihood  $\times$  Prior,

wieder. Zusammengefasst verbindet der Prozess der Bayes'schen Inferenz also eine Ausgangsschätzung in Form der Prior-Verteilung mit observierten Daten in Form der Likelihood zu einer aktualisierten Schätzung in Form der Posterior-Verteilung. Anhand der Posterior-Verteilung können dann sowohl Punkt-Schätzer für die Verteilungsparameter als auch Maße für die Unsicherheit dieser Schätzer ermittelt werden. Gängige Punktschätzer für  $\boldsymbol{\theta}$  sind bspw. der „Maximum-a-posteriori“-Schätzer  $\boldsymbol{\theta}^{\text{MAP}}$ , mit

$$\boldsymbol{\theta}^{\text{MAP}} = \arg \max_{\boldsymbol{\theta} \in U} \pi(\boldsymbol{\theta} \mid \mathbf{y}). \tag{3.10}$$

## 4. Prämienkalkulation und weitere Anwendungen

Dieses Kapitel schließt direkt an die Ausführungen des vorherigen Kapitels an, d.h. es wird angenommen, dass zunächst eine Gruppierung verschiedener (versicherter) Leistungsarten zu  $m \in \mathbb{N}$  Profilen stattgefunden hat und gemäß des beschriebenen Verfahrens für diese Profile Schätzungen der jeweiligen (Hyper-)Verteilungsparameter  $\theta_x^j$  bzw.  $\phi_x^j$  und damit auch der entsprechenden Kopfschäden  $K_x^j = B_x^j \cdot Y_x^j$ , wobei  $j \in \{1, \dots, m\}$ , bestimmt wurden. Ausgehend hiervon wird im Folgenden zunächst die Prämienberechnung unter Berücksichtigung von weiteren gängigen Vertragselementen betrachtet. In weiterer Folge wird sowohl das Risikoprofil selbst als auch zugehörige Steuerungsmöglichkeiten in Form des Abschlusses von Rückversicherungsverträgen auf der Ebene eines Portfolios von Versicherungsverträgen - bspw. alle zu einem Tarif versicherten Personen, oder ein Gewinnverband, etc. - näher betrachtet.

### 4.1. Prämienberechnung

Analog zur traditionellen Modellierung bestimmen sich die (Einsteiger-)Prämien  $P_x$  durch Anwendung des Äquivalenzprinzips, d.h. zu Vertragsabschluss sind der Barwert der zukünftigen Leistungen und der Barwert der zukünftigen Prämien äquivalent. Im Wesentlichen entspricht dies weiterhin der Berechnung gemäß Formel (2.1), mit dem feinen Unterschied, dass die nun verwendeten Kopfschäden nicht deterministisch sind, was durch Bildung des Erwartungswerts berücksichtigt wird. Somit gilt

$$P_x = \frac{1}{\ddot{a}_x} \sum_{n=0}^{x_{\max}-x} {}_nE_x \cdot \mathbb{E}[K_{x+n}], \quad K_x = \sum_{j=1}^m K_x^j. \quad (4.1)$$

Da die ermittelten Kopfschadenverteilungen effektiv einer möglichst genauen Schätzung der (bis zur nächsten Anpassung) anfallenden Leistungen entsprechen, sowie der i.d.R. vorliegenden realen Leistungssteigerung (Inflation), würden Prämien, die gemäß der obigen Formel berechnet wurden, im Mittel zu einem Verlust für den Versicherer führen. Ein Ausgleich des Verlusts durch Gewinne bei der Kapitalveranlagung

(Realverzinsung höher als Rechnungszins) oder aufgrund „günstiger“ Abweichungen zwischen erwarteten und tatsächlichen Übergangs- bzw. Ausscheidewahrscheinlichkeiten (Sterblichkeiten, Storno, etc.) ist zwar möglich, aber i.d.R. unerwünscht (Leistungen als Hauptrisikotreiber). Somit bedarf es - wie auch schon bei der traditionellen Modellierung - der Verwendung von Sicherheitszuschlägen, oder allgemeiner, der Unterscheidung zwischen Rechnungsgrundlagen erster und zweiter Ordnung.

### 4.1.1. Sicherheitszuschläge

Wie schon in Abschnitt 2.1.4 angemerkt, ist es gängige Praxis, die Rechnungsgrundlagen erster Ordnung auf Basis der Rechnungsgrundlagen zweiter Ordnung zu bestimmen. Im Sinne der Verträglichkeit mit der traditionellen Modellierung eignet es sich, ausgehend von den stochastischen  $K_x$  erneut deterministische, mit Sicherheiten versehene,  $K_x^*$  zu bestimmen, welche dann in Formel (4.1) statt  $\mathbb{E}[K_{x+n}]$  verwendet werden - ein weiterer Grund für die Verwendung deterministischer Kopfschäden für die Rechnungsgrundlagen erster Ordnung, ist die Ausweispflicht von biometrischen Grundlagen (insbesondere Kopfschäden) im Anhang des Aktuarsberichts (siehe § 4 Abs. 14 AktBV). Zur Bestimmung der  $K_x^*$  lassen sich bekannte Prämienkalkulationsprinzipien (siehe [6]) heranziehen:

- **Nettoprämienprinzip** - analog zu Formel (4.1):

$$K_x^* := \mathbb{E}[K_x]. \quad (4.2)$$

- **Erwartungswertprinzip** - entspricht prozentuellem Zuschlag:

$$K_x^* := (1 + \alpha)\mathbb{E}[K_x]. \quad (4.3)$$

- **Varianzprinzip:**

$$K_x^* := \mathbb{E}[K_x] + \alpha\mathbb{V}[K_x]. \quad (4.4)$$

- **Standardabweichungsprinzip:**

$$K_x^* := \mathbb{E}[K_x] + \alpha\sqrt{\mathbb{V}[K_x]}. \quad (4.5)$$

- **Nullnutzenprinzip.** Sei  $u(x)$  eine Nutzenfunktion (konkav, wachsend) und  $w$  das Vermögen des Versicherers. Bestimme  $K_x^*$ , sodass gilt:

$$u(w) \stackrel{!}{=} \mathbb{E}[u(w + K_x^* - K_x)]. \quad (4.6)$$

- **Exponentialprinzip** - Spezialfall des Nullnutzenprinzips. Es gilt  $u_\alpha(x) = \frac{1}{\alpha}(1 - \exp(-\alpha x))$ , das Vermögen  $w$  kürzt sich im Zuge der Umformungen, schlussendlich gilt:

$$K_x^* := \frac{1}{\alpha} \log \mathbb{E}[\exp(\alpha K_x)]. \quad (4.7)$$

Bei der Wahl des zu verwendenden Prinzips spielen mehrere Faktoren, wie bspw. die Risikoaversion des Versicherers (Wahl von  $\alpha$ ), oder gewünschte Eigenschaften wie Cash-Invarianz, eine Rolle. Zusätzlich zu den obigen Sicherheiten, welche aufgrund der Verteilung der  $K_x^j$  zustande kommen, können optional auch zusätzliche Sicherheiten aufgrund von Schätzunsicherheiten in den Verteilungsparametern ebendieser miteinkalkuliert werden, bspw. können im Fall, dass der Posterior-Erwartungswert als Punktschätzer für  $\theta_x^j$  bzw.  $\phi_x^j$  verwendet wurde, die Posterior-Variationskoeffizienten  $\gamma(\theta_x^j)$  bzw.  $\gamma(\phi_x^j)$  zusätzlich in Form einer vorgeschobenen Anwendung des Erwartungswertprinzips in Form von  $\alpha_x^j := g(\gamma(\theta_x^j), \gamma(\phi_x^j))$ , für eine entsprechende Funktion  $g$ , verwendet werden, d.h. es würde

$$K_x := \sum_{j=1}^m (1 + \alpha_x^j) K_x^j, \quad (4.8)$$

gelten.

#### 4.1.2. Selbstbehalte

Wie bereits in Abschnitt 3.1.4 angemerkt, spielen Selbstbehalte in der Krankenversicherung eine nicht unwesentliche Rolle. Im Regelfall werden in der Krankenversicherung jährliche Selbstbehalte vereinbart, aber es existieren ebenfalls leistungsfallbezogene Selbstbehalte. Letztere werden an dieser Stelle nicht näher behandelt, aber sie lassen sich bspw. anhand eines Compound-Poisson Ansatzes für die Kopfschäden und einer leichten Modifikation der folgenden Ausführungen ebenfalls abbilden.

Da nun die Verteilung der  $K_x^j$  bestimmt wurde, lässt sich die Berechnung der Prämie unter der Berücksichtigung von Selbstbehalten verhältnismäßig einfach bewerkstelligen. Je nachdem, ob ein Selbstbehalt für alle versicherten Leistungen, oder verschiedene Selbstbehalte für verschiedene Leistungsarten (Profile), vorgesehen sind, ergibt sich eine entsprechende Modifikation von Formel (4.8). Im Fall, dass ein „Gesamt-Selbstbehalt“  $S \in \mathbb{R}^+$  vorliegt, gilt

$$K_x := \left( \left( \sum_{j=1}^m (1 + \alpha_x^j) K_x^j \right) - S \right)^+. \quad (4.9)$$

(Realverzinsung höher als Rechnungszins) oder aufgrund „günstiger“ Abweichungen zwischen erwarteten und tatsächlichen Übergangs- bzw. Ausscheidewahrscheinlichkeiten (Sterblichkeiten, Storno, etc.) ist zwar möglich, aber i.d.R. unerwünscht (Leistungen als Hauptrisikotreiber). Somit bedarf es - wie auch schon bei der traditionellen Modellierung - der Verwendung von Sicherheitszuschlägen, oder allgemeiner, der Unterscheidung zwischen Rechnungsgrundlagen erster und zweiter Ordnung.

#### 4.1.1. Sicherheitszuschläge

Wie schon in Abschnitt 2.1.4 angemerkt, ist es gängige Praxis, die Rechnungsgrundlagen erster Ordnung auf Basis der Rechnungsgrundlagen zweiter Ordnung zu bestimmen. Im Sinne der Verträglichkeit mit der traditionellen Modellierung eignet es sich, ausgehend von den stochastischen  $K_x$  erneut deterministische, mit Sicherheiten versehene,  $K_x^*$  zu bestimmen, welche dann in Formel (4.1) statt  $\mathbf{E}[K_{x+n}]$  verwendet werden - ein weiterer Grund für die Verwendung deterministischer Kopfschäden für die Rechnungsgrundlagen erster Ordnung, ist die Ausweispflicht von biometrischen Grundlagen (insbesondere Kopfschäden) im Anhang des Aktuarsberichts (siehe § 4 Abs. 14 AktBVG). Zur Bestimmung der  $K_x^*$  lassen sich bekannte Prämienkalkulationsprinzipien (siehe [6]) heranziehen:

- **Nettoprämienprinzip** - analog zu Formel (4.1):

$$K_x^* := \mathbf{E}[K_x]. \quad (4.2)$$

- **Erwartungswertprinzip** - entspricht prozentuellem Zuschlag:

$$K_x^* := (1 + \alpha)\mathbf{E}[K_x]. \quad (4.3)$$

- **Varianzprinzip**:

$$K_x^* := \mathbf{E}[K_x] + \alpha\mathbf{V}[K_x]. \quad (4.4)$$

- **Standardabweichungsprinzip**:

$$K_x^* := \mathbf{E}[K_x] + \alpha\sqrt{\mathbf{V}[K_x]}. \quad (4.5)$$

- **Nullnutzenprinzip**. Sei  $u(x)$  eine Nutzenfunktion (konkav, wachsend) und  $w$  das Vermögen des Versicherers. Bestimme  $K_x^*$ , sodass gilt:

$$u(w) \stackrel{!}{=} \mathbf{E}[u(w + K_x^* - K_x)]. \quad (4.6)$$

(Realverzinsung höher als Rechnungszins) oder aufgrund „günstiger“ Abweichungen zwischen erwarteten und tatsächlichen Übergangs- bzw. Ausscheidewahrscheinlichkeiten (Sterblichkeiten, Storno, etc.) ist zwar möglich, aber i.d.R. unerwünscht (Leistungen als Hauptrisikotreiber). Somit bedarf es - wie auch schon bei der traditionellen Modellierung - der Verwendung von Sicherheitszuschlägen, oder allgemeiner, der Unterscheidung zwischen Rechnungsgrundlagen erster und zweiter Ordnung.

#### 4.1.1. Sicherheitszuschläge

Wie schon in Abschnitt 2.1.4 angemerkt, ist es gängige Praxis, die Rechnungsgrundlagen erster Ordnung auf Basis der Rechnungsgrundlagen zweiter Ordnung zu bestimmen. Im Sinne der Verträglichkeit mit der traditionellen Modellierung eignet es sich, ausgehend von den stochastischen  $K_x$  erneut deterministische, mit Sicherheiten versehene,  $K_x^*$  zu bestimmen, welche dann in Formel (4.1) statt  $E[K_{x+n}]$  verwendet werden - ein weiterer Grund für die Verwendung deterministischer Kopfschäden für die Rechnungsgrundlagen erster Ordnung, ist die Ausweispflicht von biometrischen Grundlagen (insbesondere Kopfschäden) im Anhang des Aktuarsberichts (siehe § 4 Abs. 14 AktBV). Zur Bestimmung der  $K_x^*$  lassen sich bekannte Prämienkalkulationsprinzipien (siehe [6]) heranziehen:

- **Nettoprämienprinzip** - analog zu Formel (4.1):

$$K_x^* := F[K_x]. \quad (4.2)$$

- **Erwartungswertprinzip** - entspricht prozentuellem Zuschlag:

$$K_x^* := (1 + \alpha)E[K_x]. \quad (4.3)$$

- **Varianzprinzip:**

$$K_x^* := E[K_x] + \alpha V[K_x]. \quad (4.4)$$

- **Standardabweichungsprinzip:**

$$K_x^* := E[K_x] + \alpha \sqrt{V[K_x]}. \quad (4.5)$$

- **Nullnutzenprinzip.** Sei  $u(x)$  eine Nutzenfunktion (konkav, wachsend) und  $w$  das Vermögen des Versicherers. Bestimme  $K_x^*$ , sodass gilt:

$$u(w) \stackrel{!}{=} E[u(w + \bar{K}_x^* - \bar{K}_x)]. \quad (4.6)$$

Für alle  $i \in \{0, \dots, H + 1\}$  gilt

$$V_i(t) = a_i^{\text{Pre}}(t) + v \sum_{j=0}^{H+1} p_{i,j}(t, t+1) \left( a_{i,j}^{\text{Post}}(t) + V_j(t+1) \right), \quad (4.15)$$

mit der Startbedingung  $V_i(t) = 0$  für  $t \geq x_{\max} - x + 1$ . Die Prämie  $P_x$  ergibt sich dann durch die Forderung

$$V_0(0) \stackrel{!}{=} 0. \quad (4.16)$$

## 4.2. Analysen auf Portfolioebene

Aufgrund des Wesens der Krankenversicherung - (lebens-)lange Vertragslaufzeiten, Prämienanpassungen, etc. - ist die einzelvertragliche Betrachtung wesentlicher Maße bzw. Kennzahlen wie bspw. der Profitabilität nicht sonderlich sinnvoll. Zusätzlich gibt es Steuerungsmechanismen, wie etwa das Abschließen von Rückversicherungsverträgen, aber auch Beteiligungsmechanismen, wie bspw. (unternehmens-)erfolgsabhängige Prämienrückerstattung, oder auch die Prämienanpassung selbst, für die eine granularere Betrachtungsebene benötigt wird (Planungsrechnung). Der folgende Abschnitt widmet sich einigen Analysen, die auf einer solchen Ebene, d.h. für ein Portfolio bzw. eine Gruppe von Versicherungsverträgen, Anwendung finden und für die erwähnten Vertrags-Mechanismen von Bedeutung sind.

Im Folgenden wird ein Portfolio von  $n$  Versicherungsverträgen betrachtet. Der Vektor  $X(t) := (x_1(t), \dots, x_n(t))$  beschreibt das Alter der Versicherungsnehmer zu Beginn der Periode (des Jahres)  $t \geq 0$ . Des Weiteren beschreibt der Zufallsvektor  $O(t) = (o_1(t), \dots, o_n(t))$  die Ausscheideordnung, d.h. es gilt  $o_i(t) = \mathbb{1}_{\{T_{x_i(0)} > t\}}$ , wobei  $T_x$  die gestutzte Restlebenszeit eines  $x$ -jährigen beschreibt. OBdA gelte, dass für alle Verträge dieselben Rechnungsgrundlagen (Profile, Selbstbehalte, etc.) verwendet werden, d.h. es gibt für jedes Alter  $x$  einen stochastischen Kopfschaden  $K_x$ , welcher die realistisch „erwartbaren“ Leistungen repräsentiert (zweite Ordnung) und einen (mit Sicherheiten versehenen) deterministischen Kopfschaden  $K_x^*$  für Prämienkalkulation und Reservierung (erste Ordnung).

### 4.2.1. Leistungsverteilung eines Portfolios

Eine für viele Analysen relevante bzw. maßgebliche Größe ist die Verteilung der Versicherungsleistungen für eine Betrachtungsperiode (bspw. Planungsrechnung). Hierzu betrachte man den Prozess  $K(t) := \sum_{i=1}^n o_i(t) \cdot K_{x_i(t)}$ , welcher die Leistungen für das

Gesamtportfolio in der Periode  $t$  abbildet. Die Verteilungsfunktion  $F_{K,t}(y)$  von  $K(t)$  bestimmt sich aufgrund der Annahme, dass die einzelnen versicherten Risiken unabhängig sind, durch Faltung der Verteilungsfunktionen der einzelnen Summanden  $o_i(t) \cdot K_{x_i(t)}$ . Seien

$$G_{m,t}(z) := \mathbb{P} \left[ \sum_{i=m}^n o_i(t) \cdot K_{x_i(t)} \leq z \right], \quad (4.17)$$

und

$$H_{m,t}(z) := \mathbb{P} \left[ o_m(t) \cdot K_{x_m(t)} \leq z \right], \quad (4.18)$$

wobei  $1 \leq m \leq n$ , dann gilt

$$\begin{aligned} F_{K,t}(y) &= \mathbb{P}[K(t) \leq y] = G_{1,t}(y) \\ &= \mathbb{P} \left[ o_1(t)K_{x_1(t)} + \sum_{i=2}^n o_i(t) \cdot K_{x_i(t)} \leq y \right] \\ &= \int_0^y H_{1,t}(y-z) dG_{2,t}(z) \\ &= (H_{1,t} * G_{2,t})(y) \\ &\vdots \\ &= (H_{1,t} * \dots * H_{n,t})(y) \end{aligned} \quad (4.19)$$

Die Komplexität des obigen Terms legt nahe, dass die Berechnung der obigen Verteilungsfunktion in der Praxis näherungsweise anhand eines geeigneten Approximationsverfahrens (bspw. Monte-Carlo Simulation) erfolgt. Mit  $F_{K,t}$  lassen sich u.a. Konfidenzintervalle bestimmen, welche - nach Festlegung einer passenden Zerlegung des zu analysierenden Bestandes in Teilportfolios (bspw. Altersgruppen) - zur Validierung der Passform von Leistungsprofilen herangezogen werden können.

#### 4.2.2. Portfolio-Rückversicherung

Wie in anderen Sparten, kommt es auch in der Krankenversicherung zum Abschluss von Mit- und Rückversicherungsverträgen - wenngleich letztere bemessen am Prämienvolumen meist eine Nebenrolle spielen. Im Kontext der Mitversicherung ist der Abschluss von Quoten-Versicherungsverträgen üblich, während bei der Rückversicherung tendenziell Summenexzedenten-Versicherungsverträge gezeichnet werden. Bei Mitversicherungsverträgen handelt es sich erwarteterweise um langfristige Vertragsverhältnisse, bei denen alle beteiligten Unternehmen jeweils einen vereinbarten Teil der Prämie erhalten, sowie der Leistungen erbringen. Somit haben auch alle Vertragspartner eine entsprechende Alterungsrückstellung zu bilden, welche aber aufgrund der Linearität des Vertragsanteils relativ einfach bestimmbar ist. Ausgehend von Formel

(2.8) und dem Prämien- und Leistungsanteil  $\alpha$  gilt

$${}_mV_x^\alpha = \left( \sum_{n=0}^{x_{\max}-x-m} {}_nE_{x+m} \cdot \alpha \cdot K_{x+m+n}^* \right) - \alpha P \cdot \ddot{a}_{x+m} = \alpha \cdot {}_mV_x. \quad (4.20)$$

Etwaige Zahlungen aufgrund von Prämienrückerstattungen lassen sich in der obigen Formel bei Bedarf ebenfalls leicht ergänzen (analog zu Kopfschäden). Im Gegensatz hierzu sind Rückversicherungsverträge meist kurzfristige Verträge, deren Laufzeit nur wenige Jahre beträgt und die in Form von Einmal- oder jährlichen Prämien abgegolten werden. Die Ausprägungen der Leistungsversprechen ist sehr vielseitig, so können bspw. alle über einem vereinbarten Betrag (Priorität) anfallenden Leistungen zur Gänze vom Rückversicherer übernommen werden, oder aber nur ein Teil und dann nur bis zu einer vereinbarten Obergrenze (Plafond). Seien  $S_0(t), \dots, S_m(t) \in \mathbb{R}_+ \cup \{\infty\}$  die Grenzen von „Deckungsintervallen“ für die Periode  $t$  und  $\alpha_1(t), \dots, \alpha_m(t)$  die zugehörigen Leistungsanteile. Die erwarteten Rückversicherungsleistungen für die Periode  $t$  bestimmen sich dann folgendermaßen

$$K^{\text{RV}}(t) = \sum_{i=1}^m \alpha_i \cdot \left( S_i(t) - S_{i-1}(t) - (K(t) - S_i(t))^- \right)^+. \quad (4.21)$$

Der Eigenbehalt des Erstversicherers ist dann durch

$$K^{\text{EV}}(t) = K(t) - K^{\text{RV}}(t), \quad (4.22)$$

gegeben. Anhand der bereits angeführten Prämienkalkulationsprinzipien - siehe Formeln (4.2) bis (4.7) - lässt sich dann eine Rückversicherungsprämie  $P^{\text{RV}}$  bestimmen<sup>2</sup>.

### 4.2.3. Verteilung des erwarteten Gewinns

Wie auch in anderen Versicherungssparten, entstehen Gewinne in der Krankenversicherung im Wesentlichen dadurch, dass die Annahmen, welche für die Kalkulation der Prämien verwendet werden, entsprechende Sicherheiten beinhalten. Dies betrifft im Allgemeinen nicht nur Sicherheiten in den erwarteten Leistungen, sondern auch in der erwarteten Ausscheideordnung (Sterblichkeiten und Stornoannahmen) sowie der erzielbaren Verzinsung. Der - versicherungstechnische - Gewinn gliedert sich in folgende Komponenten:

- Abgegrenzte (Netto-)Prämien - die Prämien der Versicherungsnehmer werden im Folgenden mit  $P(t) = \sum_{i=1}^n o_i(t) \cdot P_i(t)$  bezeichnet

<sup>2</sup>bei mehrjährigen Verträgen sind (erwartete) Bestandsentwicklung sowie Barwertierung zu berücksichtigen

- Tatsächliche Leistungen des Erstversicherers - diese werden durch  $K^{\text{EV}}(t)$  modelliert.
- Veränderung der Alterungsrückstellung - diese besteht aus Zuführung bzw. Entnahme der Sparprämie, Verzinsung und Vererbung. Für einen  $(x+m)$ -jährigen Versicherungsnehmer mit Jahresnettoprämie  $P$  lässt sich die Veränderung der Rückstellung im Fall des Überlebens des nächsten Jahres, d.h. bis zum Erreichen des Alters  $x+m+1$ , folgendermaßen zerlegen:

$${}_{m+1}V_x - {}_mV_x = \underbrace{(P - K_{x+m}^*)}_{\text{Sparprämie}} + i \underbrace{({}_mV_x + P - K_{x+m}^*)}_{\text{Verzinsung}} + \underbrace{{}_1q_{x+m} \cdot {}_{m+1}V_x}_{\text{Vererbung}}. \quad (4.23)$$

Intuitiv lässt sich die obige Formel folgendermaßen beschreiben: zu Beginn der Periode wird die bereits angesparte Rückstellung  ${}_mV_x$  um die Sparprämie erhöht bzw. verringert. Anschließend wird der nun vorliegende Betrag gemäß des Rechnungszinssatzes verzinst (Zinserfordernis). Die angesparte Rückstellung jener Versicherungsnehmer, die binnen der Periode ausscheiden, wird zur Finanzierung der Rückstellung für die verbleibenden Versicherungsnehmer verwendet. Durch Subtrahieren des Vererbungsterms erhält man dann die zu Beginn der Periode erwartete Veränderung der Deckungsrückstellung:

$${}_1p_{x+m} \cdot {}_{m+1}V_x - {}_mV_x = (P - K_{x+m}^*) + i({}_mV_x + P - K_{x+m}^*). \quad (4.24)$$

- Kapitalerträge des technischen Geschäfts - sei  $r(t)$  die in Periode  $t$  erzielte Verzinsung und  $m_i(t)$  die Vertragsdauern der einzelnen Versicherungsverträge zu Beginn der Periode, dann gilt für die Kapitalerträge  $G(t)$ , dass

$$G(t) = r(t) \sum_{i=1}^n o_i(t) \cdot ({}_{m_i(t)}V_{x_i(0)} + P_i(t) - K_{x_i(t)}^*), \quad (4.25)$$

d.h. es wird ausgehend vom Verzinsungsterm der Deckungsrückstellung lediglich der Rechnungszinssatz durch  $r(t)$  ersetzt. In der Praxis ist es üblich, dass eine sog. Überdeckung vorliegt, d.h. dass der (Buch-)Wert der den Rückstellungen gewidmeten Vermögenswerten (Deckungsstock) diese übersteigt und somit auch im Fall, dass  $r(t) \leq i$ , ein das Zinserfordernis übersteigender Kapitalertrag entstehen kann (größerer Ausgangswert).

Sei nun  $E(t)$  die Zufallsvariable, die den (versicherungstechnischen) Gewinn in Periode  $t$  beschreibt. OBdA gelte zusätzlich, dass das Kostenergebnis stets Null ist, d.h. Kosteneinnahmen aus der Differenz zwischen Netto- und Bruttoprämien und tatsächlich anfallende Kosten (Verwaltung, Provisionen, etc.) gleichen einander aus - i.d.R. werden auch hier Gewinne erzielt, da Kostenmargen ebenfalls Sicherheiten

(Realverzinsung höher als Rechnungszins) oder aufgrund „günstiger“ Abweichungen zwischen erwarteten und tatsächlichen Übergangs- bzw. Ausscheidewahrscheinlichkeiten (Sterblichkeiten, Storno, etc.) ist zwar möglich, aber i.d.R. unerwünscht (Leistungen als Hauptrisikotreiber). Somit bedarf es - wie auch schon bei der traditionellen Modellierung - der Verwendung von Sicherheitszuschlägen, oder allgemeiner, der Unterscheidung zwischen Rechnungsgrundlagen erster und zweiter Ordnung.

### 4.1.1. Sicherheitszuschläge

Wie schon in Abschnitt 2.1.4 angemerkt, ist es gängige Praxis, die Rechnungsgrundlagen erster Ordnung auf Basis der Rechnungsgrundlagen zweiter Ordnung zu bestimmen. Im Sinne der Verträglichkeit mit der traditionellen Modellierung eignet es sich, ausgehend von den stochastischen  $K_x$  erneut deterministische, mit Sicherheiten versehene,  $K_x^*$  zu bestimmen, welche dann in Formel (4.1) statt  $\mathbf{E}[K_{x+n}]$  verwendet werden - ein weiterer Grund für die Verwendung deterministischer Kopfschäden für die Rechnungsgrundlagen erster Ordnung, ist die Ausweispflicht von biometrischen Grundlagen (insbesondere Kopfschäden) im Anhang des Aktuarsberichts (siehe § 4 Abs. 14 AktBV). Zur Bestimmung der  $K_x^*$  lassen sich bekannte Prämienkalkulationsprinzipien (siehe [6]) heranziehen:

- **Nettoprämienprinzip** - analog zu Formel (4.1):

$$K_x^* := \mathbf{E}[K_x]. \quad (4.2)$$

- **Erwartungswertprinzip** - entspricht prozentuellem Zuschlag:

$$K_x^* := (1 + \alpha)\mathbf{E}[K_x]. \quad (4.3)$$

- **Varianzprinzip:**

$$K_x^* := \mathbf{E}[K_x] + \alpha\mathbf{V}[K_x]. \quad (4.4)$$

- **Standardabweichungsprinzip:**

$$K_x^* := \mathbf{E}[K_x] + \alpha\sqrt{\mathbf{V}[K_x]}. \quad (4.5)$$

- **Nullnutzenprinzip.** Sei  $u(x)$  eine Nutzenfunktion (konkav, wachsend) und  $w$  das Vermögen des Versicherers. Bestimme  $K_x^*$ , sodass gilt:

$$u(w) \stackrel{!}{=} \mathbf{E}[u(w + K_x^* - K_x)]. \quad (4.6)$$

# 5. Implementierung des Inferenz Algorithmus

Wie bereits in Abschnitt 3.3.1 angemerkt wurde, ist die exakte Bestimmung einer Posterior-Verteilung in den meisten Fällen rechentechnisch anspruchsvoll - insbesondere da sich die resultierende Verteilung oft nicht (einfach) parametrisieren lässt. Demnach wird die Posterior-Verteilung in der Praxis näherungsweise anhand von entsprechenden Algorithmen bestimmt. Eine Klasse dieser Algorithmen bilden sog. „Sampling-Algorithmen“, also Verfahren, bei denen Stichproben einer gewünschten, oft nur indirekt spezifizierten, Verteilung generiert werden. Eine im Zuge von Bayes'scher Inferenz oft zur Anwendung kommende Subklasse bilden die sog. „Markov-Chain-Monte-Carlo-Verfahren“ („MCMC-Verfahren“). Eines dieser Verfahren ist auch ein wesentlicher Teil der konkreten Implementierung der theoretischen Ausführungen der vorangegangenen Kapitel, welche im nächsten Kapitel behandelt wird. Dieses Kapitel widmet sich sowohl der mathematischen Theorie, auf die sich dieses Verfahren stützt, als auch der Implementierung eines allgemeinen Bayes'schen Modells in der Programmiersprache **R**.

## 5.1. MCMC allgemein

Wie die Bezeichnung dieser Art von Verfahren vermuten lässt, spielen hier Markov Ketten eine wesentliche Rolle. Sei nun  $K(A | x) : \mathcal{B}(U) \times U \rightarrow [0, 1]$  ein Markov-Kern, d.h. für alle  $x \in U$  entspricht  $K(\cdot | x)$  einem Wahrscheinlichkeitsmaß und für alle  $A \in \mathcal{B}(U)$  ist  $K(A | \cdot)$  eine messbare Funktion. Sei nun  $\mathbf{X} := (X_n)_{n \geq 0}$  eine Folge von Zufallsvariablen, sodass für alle  $n \geq 0$  gilt, dass

$$\begin{aligned}
 \mathbb{P}[X_{n+1} \in A | X_0 = x_0, \dots, X_n = x_n] &= \mathbb{P}[X_{n+1} \in A | X_n = x_n] \\
 &= K(A | x_n).
 \end{aligned}
 \tag{5.1}$$

Dann erfüllt  $\mathbf{X}$  die sog. Markov-Eigenschaft und wird als Markov-Kette mit Übergangskern  $K$  bezeichnet, d.h. für beliebiges  $n$  lässt sich die Verteilung von  $X_n$  durch mehrmaliges Anwenden des Übergangskerns auf die Initialverteilung  $X_0$  bestimmen,

was sich vereinfacht folgendermaßen darstellen lässt

$$X_{n+1} \sim K(X_n, X_{n+1}). \quad (5.2)$$

Markov-Ketten lassen sich anhand ihrer Eigenschaften abermals in verschiedene Subtypen gliedern. Der für MCMC relevante Subtyp sind sog. „ergodische“ Markov-Ketten, welche eine eindeutige stationäre Verteilung  $\pi$  besitzen, d.h. es gilt

$$X_n \sim \pi \Rightarrow X_{n+1} \sim \pi. \quad (5.3)$$

Des Weiteren konvergiert die Verteilung von  $X_n$  für  $n \rightarrow \infty$  unabhängig von der Verteilung von  $X_0$  gegen  $\pi$ . Hier setzt dann schlussendlich auch der „Monte-Carlo-Teil“ des Verfahrens ein, d.h. ausgehend von einem beliebigen Startwert  $x_0$  und des entsprechend konstruierten Markov-Kerns  $K$ , dessen induzierte stationäre Verteilung der „Zielverteilung“ entspricht, wird dann eine Folge von Stichproben  $(x_n)_{n \geq 1}$  generiert, wobei die jeweilige Sample-Verteilung für  $x_{n+1}$  durch  $K(x_n, \cdot)$  gegeben ist. Für genügend großes  $m \in \mathbb{N}$  entsprechen Vektoren der Form  $(x_{k_1}, \dots, x_{k_2})$ , mit  $m \leq k_1 \leq k_2$  dann (approximativ) einer Stichprobe der Zielverteilung.

## 5.2. Algorithmen

In diesem Abschnitt werden zwei Algorithmen betrachtet, die sich auf den obigen Grundgedanken aufsetzen. Der erste Algorithmus ist verhältnismäßig simpel, dafür konvergiert die Verteilung der erzeugten Stichproben - insbesondere bei mehrdimensionalen Verteilungen - entsprechend langsam. Beim zweiten Algorithmus handelt es sich um eine Erweiterung des ersten Algorithmus, welche die Konvergenz entsprechend beschleunigt. Dieser Algorithmus kommt letztendlich auch bei der Implementierung zur Anwendung - wenngleich auch nur im Hintergrund.

### 5.2.1. Metropolis-Hastings

Das Metropolis-Hastings Verfahren ist eines der ersten MCMC-Verfahren. Hierbei wird die Übergangsdichte  $T(x | y)$  der Markov-Kette in zwei Komponenten (bzw. Algorithmus-Schritte) zerlegt: eine Vorschlagsdichte  $\rho(x | y)$  und eine Akzeptanz-Wahrscheinlichkeit  $\alpha(x | y)$ , wobei gilt, dass

$$T(x | y) = \rho(x | y)\alpha(x | y). \quad (5.4)$$

was sich vereinfacht folgendermaßen darstellen lässt

$$X_{n+1} \sim K(X_n, X_{n+1}). \quad (5.2)$$

Markov-Ketten lassen sich anhand ihrer Eigenschaften abermals in verschiedene Subtypen gliedern. Der für MCMC relevante Subtyp sind sog. „ergodische“ Markov-Ketten, welche eine eindeutige stationäre Verteilung  $\pi$  besitzen, d.h. es gilt

$$X_n \sim \pi \Rightarrow X_{n+1} \sim \pi. \quad (5.3)$$

Des Weiteren konvergiert die Verteilung von  $X_n$  für  $n \rightarrow \infty$  unabhängig von der Verteilung von  $X_0$  gegen  $\pi$ . Hier setzt dann schlussendlich auch der „Monte-Carlo-Feil“ des Verfahrens ein, d.h. ausgehend von einem beliebigen Startwert  $x_0$  und des entsprechend konstruierten Markov-Kerns  $K$ , dessen induzierte stationäre Verteilung der „Zielverteilung“ entspricht, wird dann eine Folge von Stichproben  $(x_n)_{n \geq 1}$  generiert, wobei die jeweilige Sample-Verteilung für  $x_{n+1}$  durch  $K(x_n, \cdot)$  gegeben ist. Für genügend großes  $m \in \mathbb{N}$  entsprechen Vektoren der Form  $(x_{k_1}, \dots, x_{k_2})$ , mit  $m \leq k_1 \leq k_2$  dann (approximativ) einer Stichprobe der Zielverteilung.

## 5.2. Algorithmen

In diesem Abschnitt werden zwei Algorithmen betrachtet, die sich auf den obigen Grundgedanken aufsetzen. Der erste Algorithmus ist verhältnismäßig simpel, dafür konvergiert die Verteilung der erzeugten Stichproben - insbesondere bei mehrdimensionalen Verteilungen - entsprechend langsam. Beim zweiten Algorithmus handelt es sich um eine Erweiterung des ersten Algorithmus, welche die Konvergenz entsprechend beschleunigt. Dieser Algorithmus kommt letztendlich auch bei der Implementierung zur Anwendung - wenngleich auch nur im Hintergrund.

### 5.2.1. Metropolis-Hastings

Das Metropolis-Hastings Verfahren ist eines der ersten MCMC-Verfahren. Hierbei wird die Übergangsdichte  $T(x | y)$  der Markov-Kette in zwei Komponenten (bzw. Algorithmus-Schritte) zerlegt: eine Vorschlagsdichte  $\rho(x | y)$  und eine Akzeptanz-Wahrscheinlichkeit  $\alpha(x | y)$ , wobei gilt, dass

$$T(x | y) = \rho(x | y)\alpha(x | y). \quad (5.4)$$

### 5.2.2. Hamiltonian Monte Carlo

Ein Nachteil des Metropolis Hastings Algorithmus besteht darin, dass die erzeugte Markov-Kette bei mehrdimensionalen Verteilungen - im Kontext der Bayes'schen Inferenz also bei Posterior-Verteilungen mit mehr als einem Parameter - oft sehr lange in einer kleinen Teilregion des Trägers der Zielverteilung verweilt. Dies liegt daran, dass das Volumen von Regionen mit geringen Wahrscheinlichkeiten in Relation zum Volumen von Regionen mit hohen Wahrscheinlichkeiten (insbesondere denen des „typical sets“) beim Erhöhen der Dimension disproportional stark ansteigt, d.h. die Menge an Kandidaten mit geringen Akzeptanzwahrscheinlichkeiten wächst schneller als jene mit hohen Akzeptanzwahrscheinlichkeiten, womit die Kette oft lange im selben Punkt verweilt. Demnach muss eine Stichprobe, um repräsentativ für die Zielverteilung zu sein, entsprechend groß gewählt werden, was wiederum einen hohen Rechenaufwand impliziert.

In diesem Abschnitt wird das sog. Hamiltonian Monte Carlo (HMC) Verfahren betrachtet. Hierbei handelt es sich um eine Erweiterung des Metropolis Hastings Algorithmus, welche ein wesentlich praktikableres Konvergenzverhalten besitzt. Die im Folgenden präsentierten Inhalte wurden aus „A conceptual introduction to Hamiltonian Monte Carlo“ [8] selektiert und zusammengefasst.

Der Grundgedanke von HMC besteht darin, an jedem Punkt  $\theta$  des Trägers von  $\pi(\theta)$  eine „Richtung“ vorzugeben, entlang der sich die Markov-Kette orientiert<sup>3</sup>. Dies erfolgt durch Zerlegung des Parameterraums in unterschiedliche „Energieniveaus“ und anschließender Bestimmung eines Vektorfelds, das die Markov-Kette für einen Startwert  $\theta$  entlang einer Trajektorie mit konstantem Energieniveau zum nächsten Wert leitet. Die für die Konstruktion dieses Vektorfelds benötigte (geometrische) Information ist nicht zur Gänze in der Struktur der Zielverteilung  $\pi(\theta)$  enthalten. Demnach findet eine Erweiterung des  $d$ -dimensionalen Parameterraums zum  $2d$ -dimensionalen Phasenraum statt, d.h. neben den Elementen  $\theta$  existieren nun auch zugehörige (duale) „Impulse“  $\psi$ . Zusätzlich zur Erweiterung der Parameter, wird auch die Zielverteilung  $\pi(\theta)$  auf den Phasenraum ausgeweitet. Dies geschieht durch Verwendung einer bedingten Dichte, d.h. für die erweiterte Dichte  $\pi(\theta, \psi)$  gilt

$$\pi(\theta, \psi) = \pi(\psi \mid \theta) \cdot \pi(\theta). \quad (5.8)$$

Somit gewinnt man durch marginalisieren der Verteilung auf dem Phasenraum erneut die Zielverteilung, d.h. Trajektorien, die die Verteilung auf dem Phasenraums

<sup>3</sup>Vergleichbar mit der Bestimmung von stabilen Orbits für verschiedene Distanzen zu einem Gravitationszentrum

was sich vereinfacht folgendermaßen darstellen lässt

$$X_{n+1} \sim K(X_n, X_{n+1}). \quad (5.2)$$

Markov-Ketten lassen sich anhand ihrer Eigenschaften abermals in verschiedene Subtypen gliedern. Der für MCMC relevante Subtyp sind sog. „ergodische“ Markov-Ketten, welche eine eindeutige stationäre Verteilung  $\pi$  besitzen, d.h. es gilt

$$X_n \sim \pi \Rightarrow X_{n+1} \sim \pi. \quad (5.3)$$

Des Weiteren konvergiert die Verteilung von  $X_n$  für  $n \rightarrow \infty$  unabhängig von der Verteilung von  $X_0$  gegen  $\pi$ . Hier setzt dann schlussendlich auch der „Monte-Carlo-Teil“ des Verfahrens ein, d.h. ausgehend von einem beliebigen Startwert  $x_0$  und des entsprechend konstruierten Markov-Kerns  $K$ , dessen induzierte stationäre Verteilung der „Zielverteilung“ entspricht, wird dann eine Folge von Stichproben  $(x_n)_{n \geq 1}$  generiert, wobei die jeweilige Sample-Verteilung für  $x_{i+1}$  durch  $K(x_i, \cdot)$  gegeben ist. Für genügend großes  $m \in \mathbb{N}$  entsprechen Vektoren der Form  $(x_{k_1}, \dots, x_{k_2})$ , mit  $m \leq k_1 \leq k_2$  dann (approximativ) einer Stichprobe der Zielverteilung.

## 5.2. Algorithmen

In diesem Abschnitt werden zwei Algorithmen betrachtet, die sich auf den obigen Grundgedanken aufsetzen. Der erste Algorithmus ist verhältnismäßig simpel, dafür konvergiert die Verteilung der erzeugten Stichproben - insbesondere bei mehrdimensionalen Verteilungen - entsprechend langsam. Beim zweiten Algorithmus handelt es sich um eine Erweiterung des ersten Algorithmus, welche die Konvergenz entsprechend beschleunigt. Dieser Algorithmus kommt letztendlich auch bei der Implementierung zur Anwendung - wenngleich auch nur im Hintergrund.

### 5.2.1. Metropolis-Hastings

Das Metropolis-Hastings Verfahren ist eines der ersten MCMC-Verfahren. Hierbei wird die Übergangsdichte  $T(x | y)$  der Markov-Kette in zwei Komponenten (bzw. Algorithmus-Schritte) zerlegt: eine Vorschlagsdichte  $\rho(x | y)$  und eine Akzeptanz-Wahrscheinlichkeit  $\alpha(x | y)$ , wobei gilt, dass

$$T(x | y) = \rho(x | y)\alpha(x | y). \quad (5.4)$$

Es sei noch zusätzlich angemerkt, dass die Dichte der Zielverteilung  $\pi(\theta)$  ausschließlich in Form ihrer (numerisch approximierten) logarithmischen Ableitung in das Verfahren einfließt, womit auch hier die Notwendigkeit der Normierung von  $\pi(\theta)$  entfällt.

### 5.3. Modellierung mit Stan

In diesem Abschnitt wird anhand einfacher Beispiele illustriert, wie die theoretischen Ausführungen bezüglich Bayes'scher Inferenz und der Power Prior programmier-technisch implementiert werden können. Hierzu wird die Programmiersprache R verwendet, wobei für den wesentlichsten Teil, also das Erzeugen von Approximationen - in Form von Stichproben - der Posterior-Verteilung, auf „Stan“ zurückgegriffen wird. Stan ist eine Plattform für statistische Modellierung mit besonderem Fokus auf Bayes'scher Inferenz und MCMC-Verfahren [9].

Zunächst wird in Anlehnung an Abschnitt 3.3.1 erneut das Beispiel eines Münzwurfs betrachtet. Sei  $\theta \in (0, 1)$  der zu schätzende Bernoulli-Parameter bzw. die Wahrscheinlichkeit des Ereignisses „Kopf“ und  $\mathbf{y} = (y_1, \dots, y_n) \in \{0, 1\}^n$  eine Stichprobe von Münzwürfen. Des Weiteren wird zunächst eine flache Prior Verteilung verwendet, d.h.  $\pi(\theta) = \mathbb{1}_{(0,1)}(\theta)$  - also eine Beta-Verteilung mit Parametern  $(\alpha, \beta) = (1, 1)$ . Diese Zusammenhänge werden nun in folgendem Stan-Modell - Dateityp „stan“ - eingepflegt:

```
data {
  int n; // sample size
  int<lower=0, upper=1> y[n]; // sample data
  real<lower=0> a; // alpha parameter of Beta prior
  real<lower=0> b; // beta parameter of Beta prior
}

parameters {
  real<lower=0, upper=1> theta; // Bernoulli parameter
}

model {
  y ~ binomial(1, theta); // set sample distribution

  theta ~ beta(a, b); // set prior distribution
}
```

Die Stan-Datei gliedert sich in folgende Blöcke:

- Die Input-Daten, welche im Wesentlichen jenen Werten bzw. Parametern entsprechen, die für die Berechnung von Likelihood und Prior benötigt werden.
- Die zu schätzenden Modell-Parameter.
- Die Modellspezifikationen, welche beschreiben, wie Input-Daten und Modell-Parameter zusammenhängen.

Im nächsten Schritt wird dieses Modell in R verwendet, um ausgehend von der - zuvor generierten - Stichprobe  $y$  von Münzwürfen eine Stichprobe der Posterior-Verteilung von  $\theta$  zu erzeugen:

```
# Stan library
library(rstan)

# set seed of random number generator
set.seed(100)

# random bernoulli sample, size 10000, theta = 0.6
Samples.v <- rbinom(1e4, 1, 0.6)

# model file
model <- stan_model('Bernoulli_Beta.stan')

# posterior MCMC samples
fit <- sampling(
  model,
  data = list( # model-input
    n = length(Samples.v), # number of samples
    y = Samples.v, # sample data
    a = 1, # prior parameters
    b = 1
  ),
  iter = 3e3, # length of each markov chain
  warmup = 5e2, # burn in samples per chain
  chains = 4, # number of chains
  control = list(max_treedepth = 10), # max integration steps
  algorithm = 'HMC' # Hamiltonian Monte Carlo
)
```

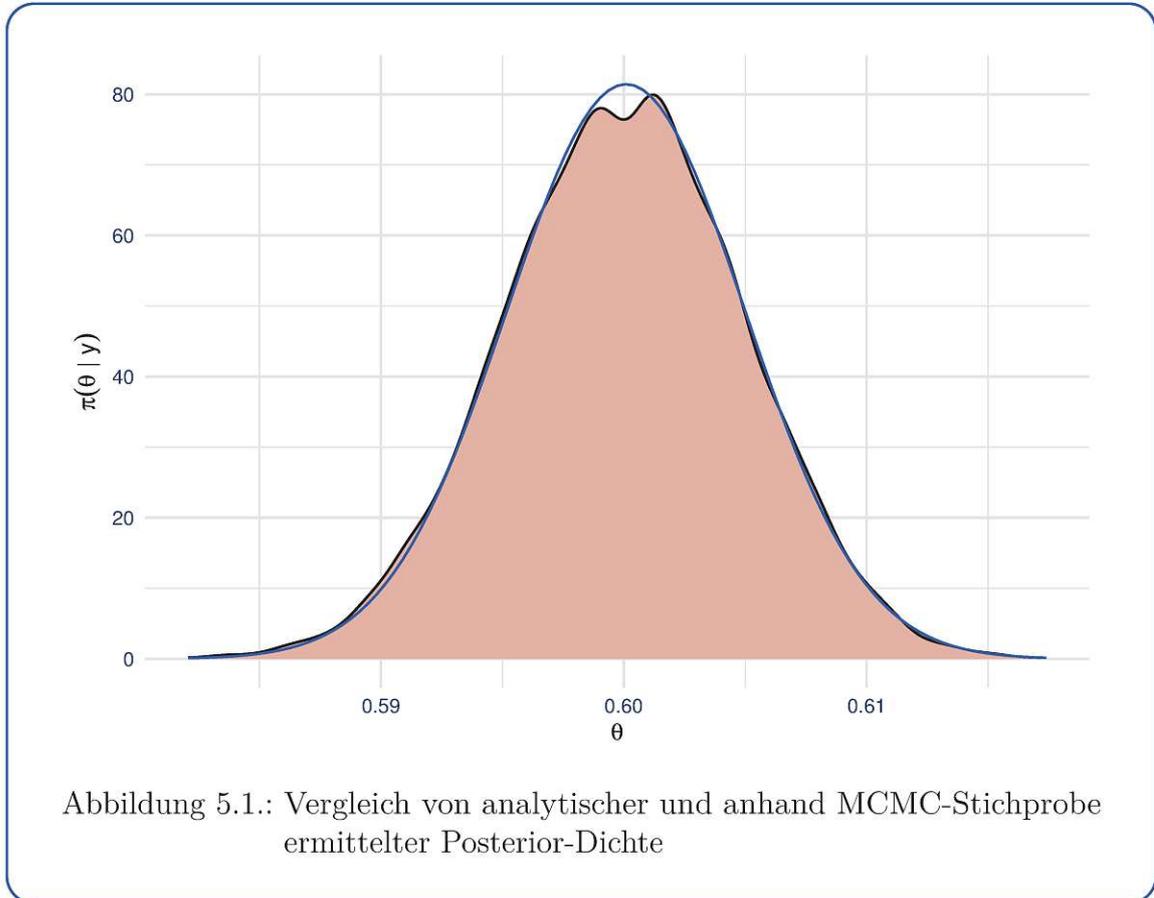
Im obigen Code wurden durch Simulation von Markov-Ketten vier Stichproben der Posterior-Verteilung mit jeweils 3000 Elementen erzeugt. Da die Verteilungen der

Ketten sich zunächst an die stationäre Verteilung annähern müssen, werden die ersten 500 Elemente der einzelnen Stichproben verworfen („Burn-In Periode“). Schlussendlich werden die restlichen Elemente zu einer großen Stichprobe mit 10000 Elementen vermengt. Die Beschränkung der Integrationsschritte, also der Anzahl an Schritten entlang der durch die Hamilton’schen Gleichungen induzierten Bahnen, dient in erster Linie der Verkürzung der Laufzeit. Bei der Erstellung der Input-Bernoulli-Stichprobe wurde eine Erfolgswahrscheinlichkeit von 0.6 verwendet. Die empirische Verteilung der MCMC-Stichprobe liefert folgende Ergebnisse:

##	mean	sd	2.5%	25%	50%	75%	97.5%
##	0.60001	0.00494	0.59026	0.59666	0.60004	0.60338	0.60958

Wie ebenfalls bereits in Abschnitt 3.3.1 ausgeführt wurde, lässt sich die Posterior-Verteilung für dieses Beispiel sogar analytisch bestimmen. Die erzeugte Münzwurf-Stichprobe enthält 6001 „Kopf“ und 3999 „Zahl“ Ereignisse. Demnach entspricht die Posterior einer Beta-Verteilung mit Parametern  $(\alpha, \beta) = (6002, 4000)$ .

Ein Vergleich der zugehörigen Dichte und einem anhand der MCMC-Stichprobe erstellten Histogramm liefert folgendes Bild:



Aufgrund der flachen Prior Verteilung, sowie der für dieses einfache Beispiel vergleichsweise großen Stichprobe sind diese Ergebnisse wohl nicht allzu überraschend, wengleich in der Nähe des Modus vergleichsweise größere Abweichungen auftreten. Dies ist im Kontext von HMC wenig überraschend, da das Verfahren darauf abzielt, möglichst gute Schätzer für Erwartungswerte von Funktionen der zugrundeliegenden Zufallsgröße (hier der Bernoulli-Parameter) zu liefern, wie bspw. anhand des obigen Stichprobenmittelwerts erkennbar ist.

Im nächsten Schritt wird das eben betrachtete Beispiel um zwei Komponenten erweitert:

- Der zu schätzende Bernoulli-Parameter  $\theta$  ist nun nicht mehr konstant, d.h. es werden nun mehrere Perioden mit zugehörigen Bernoulli-Stichproben betrachtet, wobei der für die Erstellung der Stichproben verwendete Parameter für jede Periode unterschiedlich ist.

- Die Daten der verschiedenen Perioden werden im Zuge der Schätzung durch Anwendung der Power Prior unterschiedlich gewichtet.

Um diese Aspekte in Stan zu berücksichtigen, muss die Modelldatei entsprechend adaptiert werden:

```
functions {
  real foo_lpdf(real theta, matrix PD){ // Bernoulli Power Prior
    int m = rows(PD);

    real sum_ = sum(
      ( PD[,1] * log(theta) +
        (1 - PD[,1]) * log(1 - theta) ) .* PD[,2]
    );

    return(sum_);
  }
}

data {
  int n; // sample size
  int<lower=0, upper=1> y[n]; // sample data
  int m; // prior data size
  matrix[m,2] PD; // prior data + PP discounting factors
}

parameters {
  real<lower=0, upper=1> theta; // Bernoulli parameter
}

model {
  y ~ binomial(1, theta); // set sample distribution

  theta ~ foo(PD); // set prior distribution
}
```

Betrachtet man zunächst die Änderungen im Input-Block, so ist erkennbar, dass nun auch die Stichproben  $\mathbf{y}_t = (y_{t,1}, \dots, y_{t,d_t})$ ,  $t \in \{t_{\min}, \dots, t_{\max}\}$ , vergangener Perioden

sowie die zugehörigen Diskontierungsfaktoren  $g(t)$  (Gewichtung) in Form der  $m \times 2$ -Matrix „PD“ an das Modell übergeben werden. Zusätzlich gibt es nun einen weiteren „Funktionen“-Block, in dem unter anderem eigene (Log-)Dichtefunktionen - erkennbar an Suffix „\_lpdf“ - definiert werden können. Durch Anwendung von Formel (3.32) erhält man für die (finale) Posterior Verteilung, dass

$$\pi(\theta \mid \mathbf{y}_{t_{\min}}, \dots, \mathbf{y}_{t_{\max}}) \propto \prod_{t=t_{\min}}^{t_{\max}} \mathcal{L}(\mathbf{y}_t \mid \theta)^{g(t)}. \quad (5.13)$$

Herausheben der Likelihood (hier  $g(t_{\max}) = 1$ ), sowie Einsetzen für die Bernoulli-Likelihood ergibt

$$\prod_{t=t_{\min}}^{t_{\max}} \mathcal{L}(\mathbf{y}_t \mid \theta)^{g(t)} = \mathcal{L}(\mathbf{y}_{t_{\max}} \mid \theta) \cdot \left( \prod_{t=t_{\min}}^{t_{\max}-1} \prod_{s=1}^{d_t} (\theta^{y_{t,s}} \cdot (1-\theta)^{1-y_{t,s}})^{g(t)} \right). \quad (5.14)$$

Separiert man nun die Power Prior und logarithmiert, so ergibt sich

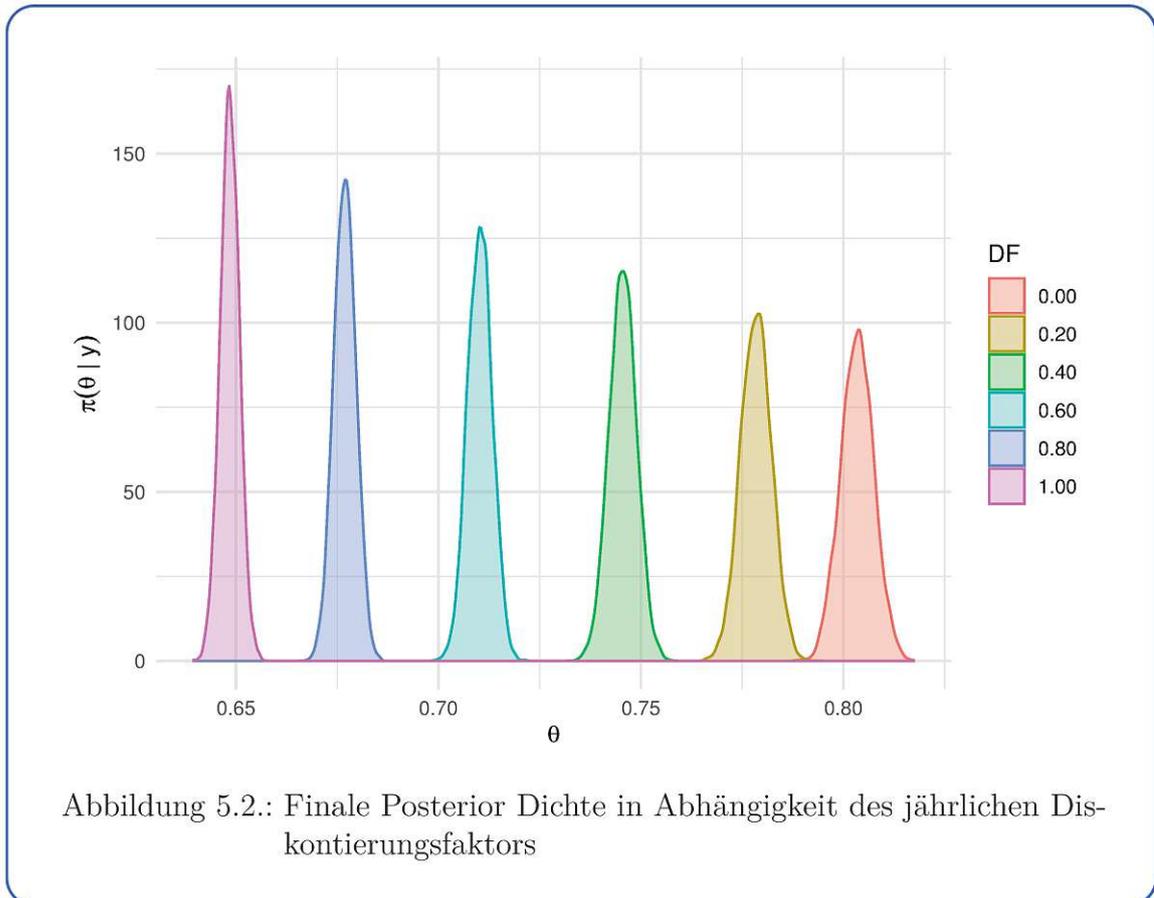
$$\log \left( \prod_{t=t_{\min}}^{t_{\max}-1} \prod_{s=1}^{d_t} (\theta^{y_{t,s}} \cdot (1-\theta)^{1-y_{t,s}})^{g(t)} \right) = \sum_{t=t_{\min}}^{t_{\max}-1} \sum_{s=1}^{d_t} g(t) (y_{t,s} \log \theta + (1-y_{t,s}) \log(1-\theta)). \quad (5.15)$$

Jede Zeile der Input-Matrix beinhaltet das Paar  $(y_{t,s}, g(t))$ . Die Berechnung der obigen Summe lässt sich somit vektorisieren, wie anhand der Syntax erkennbar ist. Im nächsten Schritt werden die Input-Stichproben erstellt.

```
# multi period Bernoulli samples
samples.n <- 1e4 # sample size 10000 for each period
theta.v <- 5:8 / 10 # theta changing from 0.5 to 0.8

# combine within data frame
# TIME_DELTA = 0 --> new data
DATA.df <- data.frame(
  TIME_DELTA = rep(
    length(theta.v):1 - 1,
    each = samples.n
  ),
  SAMPLE = rbinom(
    samples.n * length(theta.v),
    1,
    rep(theta.v, each = samples.n)
  )
)
```

Es werden vier Perioden betrachtet. Die Erfolgswahrscheinlichkeit entwickelt sich in gleichmäßigen Inkrementen von 0.5 zu 0.8. Abschließend werden für  $g(t) := v^{t_{\max}-t}$  und  $v \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$  MCMC-Stichproben der (finalen) Posterior Verteilung erstellt. Betrachtet man die Ergebnisse erneut in Form von Histogrammen, so ergibt sich folgendes Bild:



Wie erwartet hat die Wahl des Diskontierungsfaktors maßgeblichen Einfluss auf die Position und die Höhe der Spitze der Verteilung.

## 6. Beispielberechnungen

Dieses Kapitel widmet sich der programmatischen Implementierung und Analyse der in den vorangegangenen Kapiteln erarbeiteten theoretischen Ausführungen. Hierzu wird die historische Entwicklung eines fiktiven Versicherungsbestands betrachtet, d.h. die Bestands- und Leistungsdaten mehrerer vergangener Perioden werden für die Schätzung der Leistungsverteilungsparameter für die aktuelle bzw. nächste Periode herangezogen. Hierfür wird erneut die Programmiersprache **R** - mit entsprechender Unterstützung durch Stan - verwendet. Der Grund für das Betrachten simulierter Daten liegt darin, dass somit die Genauigkeit der Schätzungsergebnisse überprüft werden kann, insbesondere im Bezug auf die Größe der Stichprobe - hohe Alter sind im Regelfall schwach besetzt.

### 6.1. Bestands- und Leistungssimulation

Wie oben erwähnt, werden die Daten eines fiktiven Versicherungsbestands betrachtet, d.h. sie müssen zunächst simuliert werden. Neben der Alterung oder dem Ausscheiden von Versicherungsnehmern aufgrund von Storno oder Tod, sind bei der Simulation natürlich auch Neuzugänge und insbesondere die zeitliche Veränderung der Leistungen zu berücksichtigen. In diesem Abschnitt erfolgt die Simulation dieser Daten, wobei die verwendeten Simulationsparameter (Leistungsverteilungsparameter) nach der auf Basis der simulierten Daten erstellten Parameterschätzung ebenfalls für die Validierung der Ergebnisse herangezogen werden. Das Ergebnis der Simulation ist eine Tabelle deren Zeilen die Leistungen für ein Leistungsprofil eines aktiven Versicherungsnehmers, sowie dessen (Vertrags-)Alter für diese Periode festhalten. Die Tabelle besitzt folgende Felder:

- **Period** - gibt die Periode an.
- **Profile** - gibt das betrachtete Leistungsprofil an.
- **x** - gibt das Alter des Versicherungsnehmers in der Betrachtungsperiode an.
- **Kx** - gibt die Leistungen an den Versicherungsnehmer für diese Periode und das betrachtete Leistungsprofil an. Sollten in der Periode keine Leistungsfälle eintreten, ist dieser Wert gleich Null.

Die Simulation erfolgt in zwei Schritten: zunächst wird der Versicherungsbestand und dessen Veränderung von Periode zu Periode erzeugt und anschließend werden die Leistungsparameter festgelegt, mit dem Bestand verknüpft und schlussendlich werden die Leistungen der einzelnen Versicherungsnehmer simuliert. Im Folgenden wird der zur Erstellung der Daten verwendete Code stückweise erläutert.

Zunächst werden einige Parameter festgelegt bzw. eingelesen. Diese umfassen die Anzahl an zu simulierenden Perioden, Storno- und Sterbewahrscheinlichkeiten - siehe A.1 - sowie Start- und Endalter für Kinder und Erwachsene. Die Simulation des Bestands greift auf das gesamte zulässige Altersband zurück, d.h. Alter 0 bis 110, während die Leistungssimulation und spätere Parameterschätzung auf Erwachsene beschränkt ist - diese Einschränkung ist natürlich nicht zwingend, sondern als Anlehnung an die oftmals unterschiedliche (vertragliche, kalkulatorische, etc.) Behandlung von Kindern und Erwachsenen gedacht.

```
# libraries
library(openxlsx)
library(tidyverse)
library(gsubfn)

# parameters
{
  # number of periods (i.e. years) to be simulated
  periods.n <- 8

  # mortality and lapse rates for ages x = 0,...,113
  probs.df <- read.xlsx('mortality and lapse rates.xlsx')

  # add survival probabilities
  probs.df$px <- 1 - probs.df$qx - probs.df$wx

  # age limits
  {
    x_min_child.n <- 0
    x_max_child.n <- 17
    x_min_adult.n <- 18
    x_max_adult.n <- 110
  }
}
```

Im nächsten Schritt wird die Struktur des Initialbestands, also des Bestands für

Periode 0, spezifiziert. Zunächst wird die Gesamtzahl an Verträgen bzw. Versicherungsnehmern - hier synonym, im Allgemeinen ist dies selbstverständlich nicht der Fall - vorgegeben. Die Altersverteilung innerhalb des Portfolios wird indirekt spezifiziert.

```
# initial portfolio parameters
# using (scaled & shifted) beta distribution
# as portfolio age distribution

Portfolio_Age_Distr <- list(
  TotalIP = 1e4, # initial total number of insured persons
  alpha = 2, # alpha parameter of beta distribution
  beta = 3 # beta parameter of beta distribution
)
```

Für die Altersstruktur wird eine Beta-Verteilung mit Parametern  $(\alpha, \beta) = (2, 3)$  herangezogen, entsprechend auf das Altersintervall  $[0, 110]$  umgelegt und diskretisiert, d.h. die Wahrscheinlichkeit für das Alter  $x$  entspricht dann der anhand der transformierten Verteilungsfunktion gewonnenen Wahrscheinlichkeit für das Intervall  $(x, x + 1]$ . Schlussendlich wird der Initialbestand durch Erzeugung einer Stichprobe dieser Verteilung gewonnen.

```
sample_beta <- function(
  n,
  x_min,
  x_max,
  alpha,
  beta
){
  # draws n samples from the set {x_min, ... , x_max}
  # probabilities for each element are given by
  # the beta distribution using the chosen parameters
  # returns drawn samples as vector

  # map set to unit interval
  denom.n <- x_max - x_min + 1
  num.v <- x_min:x_max - x_min

  # determine probabilities using the cdf
  sample_distr.v <- pbeta(
    (num.v + 1) / denom.n,
```

## 6. Beispielberechnungen

```
    shape1 = alpha,
    shape2 = beta
  ) - pbeta(
    (num.v    ) / denom.n,
    shape1 = alpha,
    shape2 = beta
  )

  # draw samples
  out.v <- sample(
    x = x_min:x_max,
    size = n,
    prob = sample_distr.v,
    replace = T
  )

  return(out.v)
}
```

Für das weitere Verarbeiten, ist es günstiger, wenn der Output der obigen Funktion - also die Stichprobe - als Tabelle, statt als Vektor angeliefert wird. Demnach wird zusätzlich eine Wrapper-Funktion definiert, welche dann in weiterer Folge zur Anwendung kommt.

```
sample_beta_df <- function(
  n,
  x_min,
  x_max,
  alpha,
  beta
){
  # wrapper for sample_beta
  # returns a single column data frame
  # containing the drawn samples

  out.df <- data.frame(
    x = sample_beta(
      n = n,
      x_min = x_min,
      x_max = x_max,
      alpha = alpha,
      beta = beta
    )
  )
}
```

## 6. Beispielberechnungen

```
)  
) %>%  
  as_tibble()  
  
return(out.df)  
}
```

Nun wird der Initialbestand erstellt.

```
# initial portfolio  
Lx_Base.df <- sample_beta_df(  
  n = Portfolio_Age_Distr$TotalIP,  
  x_min = x_min_child.n,  
  x_max = x_max_adult.n,  
  alpha = Portfolio_Age_Distr$alpha,  
  beta = Portfolio_Age_Distr$beta  
) %>%  
  mutate(  
    Period = 0  
) %>%  
  select(  
    Period,  
    x  
)
```

Im nächsten Schritt werden die Parameter für die Simulation des Neugeschäfts festgelegt. Auch hier erfolgt die Ermittlung des Bestands anhand der Festlegung einer Gesamtanzahl und anschließender Erzeugung einer Stichprobe gemäß einer spezifizierten Beta-Verteilung. Die für die jeweiligen Perioden verwendeten Parameter werden durch Wahl eines zufälligen relativen Anteils, bzw. einer zufälligen Translation der Parameter des Initialbestands ermittelt.

```
# number of insured persons for each period is  
# determined as (bounded) random fraction of initial total  
# random offsets to beta parameters for each period  
  
NB_Perc <- list(  
  TotalIP = c(0.01, 0.03), # 1 to 3 percent of initial total  
  alpha = 0.50 * c(-1, 1), # offset in interval [-0.5, 0.5]
```

## 6. Beispielberechnungen

```
beta = 0.20 * c(-1, 1) # offset in interval [-0.2, 0.2]
)
```

Die Parameter werden in einer Tabelle erfasst.

```
# totals and beta parameters
# for new business of each period
NB_Age_Dist <- data.frame(
  Period = 1:(periods.n - 1),

  TotalIP = round(
    runif(
      n = periods.n - 1,
      min = NB_Perc$TotalIP[1],
      max = NB_Perc$TotalIP[2]
    ) * Portfolio_Age_Distr$TotalIP,
    0),

  alpha = Portfolio_Age_Distr$alpha + runif(
    n = periods.n - 1,
    min = NB_Perc$alpha[1],
    max = NB_Perc$alpha[2]
  ),

  beta = Portfolio_Age_Distr$beta + runif(
    n = periods.n - 1,
    min = NB_Perc$beta[1],
    max = NB_Perc$beta[2]
  )

) %>%
as_tibble()
```

Anschließend werden die Neugeschäftsportfolios für jede Periode erstellt.

```
# draw new business portfolio for each period
Lx_NB.df <- NB_Age_Dist %>%
  nest(DAT = !Period) %>%
  mutate(
    DAT = map(DAT, ~sample_beta_df(
      n = .$TotalIP,
```

## 6. Beispielberechnungen

```
x_min = x_min_child.n,  
x_max = x_max_adult.n,  
alpha = .$alpha,  
beta = .$beta  
))  
) %>%  
unnest(DAT)
```

Im wesentlichsten Schritt des ersten Teils erfolgt nun Periode für Periode die Bestimmung des Ist-Bestands durch Fortschreibung des Vorperiodenbestands gemäß Überlebenswahrscheinlichkeiten und anschließender Vermengung mit dem entsprechenden Neugeschäftsportfolio.

```
# portfolio progression (aging, lapsing, deaths)  
  
# initialise  
Lx.df <- Lx_Base.df  
  
# process  
for(p in 1:(periods.n - 1)){  
  # survivors of previous period  
  Lx_Surv.df <- Lx.df %>%  
    filter(  
      Period == p - 1  
    ) %>%  
    left_join(  
      probs.df %>%  
        select(  
          x,  
          px  
        ),  
      by = c('x')  
    ) %>%  
    mutate(  
      # random draw to determine survival  
      SURVIVAL_IND = mapply(  
        rbernoulli,  
        n = 1,  
        p = px  
      )  
    ) %>%  
    filter(  

```

```

    # only contracts that survived
    SURVIVAL_IND == 1
  ) %>%
  mutate(
    Period = p, # new period
    x = x + 1 # new age
  ) %>%
  select(
    Period,
    x
  )

# append data
Lx.df <- Lx.df %>%
  union_all(
    Lx_Surv.df
  ) %>%
  union_all(
    # add new business
    Lx_NB.df %>%
      filter(
        Period == p
      ) %>%
      select(
        Period,
        x
      )
  )

rm(Lx_Surv.df)
}

```

Nun folgt der zweite Teil, d.h. die Simulation der Leistungen. Im Folgenden werden zwei Leistungsprofile betrachtet. Die Einzelschadenverteilungen werden erwartungsgemäß in Schadeneintrittswahrscheinlichkeit und Schadenhöhenverteilung getrennt. Erstere wird durch eine Bernoulli-Verteilung reflektiert, für die zweite werden hier jeweils Gamma-Verteilungen verwendet, d.h. die Kopfschadenverteilung eines Profils lässt sich durch Tripel der Form  $(\theta_x, \alpha_x, \beta_x)$  spezifizieren, wobei  $\theta_x$  der Schadeneintrittswahrscheinlichkeit und  $(\alpha_x, \beta_x)$  den Gamma-Parametern entspricht. Zusätzlich werden für jede Periode und beide Profile Inflationsparameter ermittelt, welche in weiterer Folge auf  $(\alpha_x, \beta_x)$  Einfluss nehmen.

## 6. Beispielberechnungen

```
# inflation for each period and profile
# age independent inflation
# affects both gamma parameters
infl.df <- data.frame(
  Period = 1:(periods.n - 1),
  P1 = runif(periods.n - 1, min = 0, max = 0.05),
  P2 = runif(periods.n - 1, min = 0, max = 0.05)
) %>%
gather(
  Profile,
  Infl,
  !Period
)
```

Die Tripel werden pro Profil nur für Anfangs- und Endalter explizit vorgegeben, die restlichen Werte werden durch Interpolation ermittelt.

```
# initial claim distribution parameters
# parameter vectors of the form (theta, alpha, beta_inv)
# are defined for both minimum and maximum (adult) age
# values for other ages are derived via interpolation

init_claim_params <- list(
  P1 = list(
    x_min = c(0.20, 40, 25),
    x_max = c(0.80, 50, 30)
  ),

  P2 = list(
    x_min = c(0.15, 40, 30),
    x_max = c(0.85, 50, 40)
  )
)
```

Im nächsten Schritt werden die für die Interpolation verwendeten Funktionen spezifiziert.

```
# functions used for interpolation
inter_funcs <- list(
  P1 = list(
```

## 6. Beispielberechnungen

```
theta = ~log(1 + (.x)^2),
alpha = ~exp((.x)^(1/6)),
beta_inv = ~log(1 + (.x)^4)
),

P2 = list(
  theta = ~(0.1 + .x)^(1/2),
  alpha = ~(.x)^(1/4),
  beta_inv = ~log(1 + .x)
)
)
```

Die jeweilige Interpolationsfunktion wird dann mit einer Affin-Transformation verkettet, sodass die vorher festgelegten Randwerte erreicht werden.

```
interpolate_profiles <- function(
  inter_funcs,
  x_min,
  x_max,
  edge_values
){

  # given interpolation function f(x),
  # create affine transform
  # g(x) = k * f(x) + d
  # to hit edge values

  out <- data.frame(
    x = x_min:x_max
  )

  for(i in 1:length(inter_funcs)){

    f <- inter_funcs[[i]]

    if(is_formula(f)){
      f <- match.funfn(f)

      y_min <- edge_values$x_min[i]
      y_max <- edge_values$x_max[i]

      k_ <- (y_max - y_min) / (f(x_max) - f(x_min))
```

## 6. Beispielberechnungen

```
d_ <- y_min - f(x_min) * k_  
  
g <- function(x){  
  out <- k_ * f(x) + d_  
  
  return(out)  
}  
  
param_name <- names(inter_funcs)[i]  
  
out <- out %>%  
  mutate(  
    !!param_name := g(x)  
  )  
}  
  
return(out)  
}
```

Im Folgenden werden die jeweiligen Verteilungsparameter für jede Periode ermittelt und in einer Tabelle gespeichert. Diese Werte werden später für die Validierung der Ergebnisse der Parameterschätzung herangezogen.

```
# determine claims parameters for each period  
Kx_Param.df <- data.frame()  
  
for(i in 1:length(init_claim_params)){  
  
  profile_name <- names(init_claim_params)[i]  
  
  # initial period  
  Kx_Param.df <- Kx_Param.df %>%  
    union_all(  
      interpolate_profiles(  
        inter_funcs = inter_funcs[[i]],  
        x_min = x_min_adult.n,  
        x_max = x_max_adult.n,  
        edge_values = init_claim_params[[i]]  
      ) %>%  
      mutate(  
        Period = 0,  

```

```

        Profile = profile_name
      ) %>%
      select(
        c(Period, Profile),
        !c(Period, Profile)
      )
    )

# later periods
for(j in 1:(periods.n - 1)){
  Kx_Param.df <- Kx_Param.df %>%
  union_all(
    Kx_Param.df %>%
    filter(
      Profile == profile_name,
      Period == j - 1
    ) %>%
    mutate(
      Period = j
    ) %>%
    left_join(
      infl.df,
      by = c(
        'Profile',
        'Period'
      )
    ) %>%
    mutate(
      across(c(alpha, beta_inv), ~ .x * sqrt(1 + Infl))
    ) %>%
    select(
      all_of(names(Kx_Param.df))
    )
  )
}

rm(profile_name)
}

# obtain beta (rate) from beta_inv (scale)
Kx_Param.df <- Kx_Param.df %>%
  mutate(

```

## 6. Beispielberechnungen

```
    beta = 1 / beta_inv
  ) %>%
  select(
    !beta_inv
  )

# save parameters
write_csv2(Kx_Param.df, 'SIM_Params.csv')
```

Schlussendlich werden gemäß der Verteilungsparameter für jeden Versicherungsnehmer Leistungen pro Periode und Profil simuliert.

```
simulate_claim <- function(n, theta, alpha, beta){
  # create n claim samples for
  # given claim parameters (theta, alpha, beta)
  round(
    as.numeric(rbernoulli(n, theta)) *
    rgamma(n, shape = alpha, rate = beta),
    2
  )
}
```

Die Ergebnisse werden dann in Form der zu Beginn dieses Abschnitts beschriebenen Tabelle abgespeichert.

```
# joining simulation parameters to
# contract portfolio and then draw sample
Final.df <- Lx.df %>%
  filter(
    x_min_adult.n <= x & x <= x_max_adult.n
  ) %>%
  left_join(
    Kx_Param.df,
    by = c(
      'x',
      'Period'
    )
  ) %>%
  mutate(
    Kx = mapply(
      simulate_claim,
```

```
n = 1,
  theta = theta,
  alpha = alpha,
  beta = beta
)
) %>%
select(
  Period,
  Profile,
  x,
  Kx
) %>%
arrange(across())

# save output
Final.df %>% write_csv2('Kx_SIM.csv')
```

### 6.2. Parameterschätzung

Ausgehend von den soeben erstellten „Rohdaten“ werden nun die Verteilungsparameter der Kopfschäden, d.h. für jedes Profil die Tripel  $(\theta_x, \alpha_x, \beta_x)$ , geschätzt.

Zunächst werden einige Parameter, wie bspw. der jährliche Power Prior Diskontierungsfaktor, festgelegt. Um auszugleichen, dass die Stichprobengröße für Einzelalter - insbesondere bei hohen Altern - verhältnismäßig gering sein kann, werden die Stichproben benachbarter Einzelalter mehrfach verwendet, d.h. um die Verteilungsparameter für das Alter 25 zu schätzen, werden bspw. die Rohdaten (observierte Einzelschäden) der Alter 23 bis 27 herangezogen. Für Alter 26 wird das Intervall entsprechend verschoben, in höheren Altern wird das Intervall zusätzlich verlängert. Somit entsteht - insbesondere bei „Randaltern“ - ein glättender Effekt, welcher beim späteren Vergleich mit den für die Simulation verwendeten Parametern zu berücksichtigen ist.

```
# libraries
library(rstan)
library(tidyverse)
library(purrr)
library(furrr)
```

## 6. Beispielberechnungen

```
# open (previously simulated) contract and claims data
Kx.df <- read_csv2('Kx_SIM.csv')

# parameters
{
  # age limits
  x_min <- 18
  x_max <- 110

  # age bands
  # to combat small sample sizes for single ages
  # intervals of neighboring ages will be used
  age_bands.df <- data.frame(
    x = x_min:x_max
  ) %>%
  mutate(
    x_lb = case_when(
      x < 65 ~ x - 2,
      x < 75 ~ x - 3,
      x < 85 ~ x - 4,
      TRUE ~ pmin(85, x-5)
    ),
    x_lb = pmax(x_lb, x_min),
    x_ub = case_when(
      x < 65 ~ x_lb + 4,
      x < 75 ~ x_lb + 6,
      x < 85 ~ x_lb + 8,
      TRUE ~ x_lb + x_max - 85
    ),
    x_ub = pmin(x_ub, x_max)
  ) %>%
  group_by(
    x_lb,
    x_ub
  ) %>%
  mutate(
    # avoid sampling the same data multiple times
    Band_ID = cur_group_id()
  ) %>%
  ungroup()

  max_ID <- max(age_bands.df$Band_ID)
  count.v <- rep(1, 3)
```

## 6. Beispielberechnungen

```
# annual (periodical) discounting factor
v.n <- 0.70

# number of past periods to be used for estimation of
# parameters for current period
n_periods_used <- 4

# number of MCMC samples to be drawn for each age
n_samples <- 3e3

}
```

Als Nächstes folgen Hilfsfunktionen zur Auffächerung der Rohdaten gemäß der gewählten Altersbänder sowie zur Erstellung der MCMC-Stichproben.

```
# utility functions
{
gather_data <- function(data.df){
  out <- age_bands.df %>%
    distinct(
      Band_ID,
      x_lb,
      x_ub
    ) %>%
    left_join(
      data.df %>%
        as.data.frame(),
      by = character(0)
    ) %>%
    filter(
      x_lb <= x,
      x <= x_ub
    ) %>%
    select(
      -c(
        x,
        x_lb,
        x_ub
      )
    )

  return(out)
}
```

```

create_fit <- function(HISTORIC.df, NEW.df){

  cat(
    ifelse(count.v[1] == 1, 'Occurence', 'Claim Size'),
    ' - ',
    'Profile ',
    count.v[2],
    ' - ',
    'Band_ID ',
    count.v[3],
    '\n'
  )

  if(count.v[3] == max_ID){
    if(count.v[2] == 2){
      count.v[1] <<- count.v[1] + 1
    }

    count.v[2] <<- count.v[2] %% 2 + 1
  }

  count.v[3] <<- count.v[3] %% max_ID + 1

  Samples.v <- NEW.df %>%
    pull(Kx) %>%
    as.vector()

  HData.m <- HISTORIC.df %>%
    mutate(
      Disc = v.n ^ Disc
    ) %>%
    as.matrix()

  out <- sampling(
    model,
    list(
      n = length(Samples.v),
      m = nrow(HData.m),
      y = as.numeric(Samples.v),
      PD = as.matrix(HData.m)
    ),

```

## 6. Beispielberechnungen

```
    iter = n_samples + 500,
    warmup = 500,
    chains = 1,
    cores = getOption("mc.cores", 1L),
    refresh = 0
  )

  return(out)
}
}
```

Die letzte bzw. aktuelle Periode, d.h. die „Stichprobe“, also jener Teil welcher im Inferenz-Verfahren die Likelihood bestimmt, wird separiert.

```
# separation of current period
{
  Kx.df <- Kx.df %>%
  select(
    Period,
    Profile,
    x,
    Kx
  ) %>%
  as_tibble()

  Kx_new.df <- Kx.df %>%
  filter(Period == max(Period)) %>%
  select(-Period)

  Kx_hist.df <- Kx.df %>%
  filter(Period < max(Period))
}
```

Im nächsten Schritt werden die Rohdaten für die Schätzung der Schadeneintrittswahrscheinlichkeit modifiziert, d.h. falls ein Schaden eingetreten ist, d.h. ein Leistungsbetrag größer als Null vorliegt, so entspricht dies einem Erfolg-Ereignis der Bernoulli Zufallsvariable.

## 6. Beispielberechnungen

```
# claim occurrence probability

# historic data
Kx_Occur_hist.df <- Kx_hist.df %>%
  filter(
    Period > max(Period) - n_periods_used
  ) %>%
  mutate(
    Disc = max(Period) - Period + 1,
    Kx = ifelse(Kx > 0, 1, 0),
  ) %>%
  select(
    -Period
  ) %>%
  nest(SUBTAB = c(x, Kx)) %>%
  mutate(SUBTAB = future_map(SUBTAB, gather_data)) %>%
  unnest(SUBTAB) %>%
  nest(HISTORIC = c(Kx, Disc))

# new data
Kx_Occur_new.df <- Kx_new.df %>%
  mutate(
    Kx = ifelse(Kx > 0, 1, 0)
  ) %>%
  nest(SUBTAB = c(x, Kx)) %>%
  mutate(SUBTAB = future_map(SUBTAB, gather_data)) %>%
  unnest(SUBTAB) %>%
  nest(NEW = Kx)
```

Nun erfolgt die Parameterschätzung. Für dieses Beispiel wurde für alle Parameter der Posterior-Erwartungswert als Schätzer herangezogen.

```
# occurrence parameter estimation
model <- stan_model('Bernoulli_PP.stan')

Occur_Est.df <- Kx_Occur_new.df %>%
  full_join(
    Kx_Occur_hist.df,
    by = c(
      'Profile',
      'Band_ID'
    )
  )
```

## 6. Beispielberechnungen

```
)
) %>%
mutate(
  Fit = map2(HISTORIC, NEW, create_fit),
  Post_Mean_theta = map(
    Fit,
    function(x){as.numeric(get_posterior_mean(x, 'theta'))}
  ),
  Post_Mean_theta = as.numeric(Post_Mean_theta)
)

Occur_Est_Summary.df <- Occur_Est.df %>%
select(
  Profile,
  Band_ID,
  Post_Mean_theta
)
```

Derselbe Prozess erfolgt nun zur Ermittlung der Schadenhöhenparameter, wobei schadenfreie Einträge zunächst aussortiert werden.

```
# claim size estimation

# historic data
Kx_Claim_hist.df <- Kx_hist.df %>%
  filter(
    Period > max(Period) - n_periods_used,
    Kx > 0
  ) %>%
  mutate(
    Disc = max(Period) - Period + 1
  ) %>%
  select(
    -Period
  ) %>%
  nest(SUBTAB = c(x, Kx)) %>%
  mutate(SUBTAB = future_map(SUBTAB, gather_data)) %>%
  unnest(SUBTAB) %>%
  nest(HISTORIC = c(Kx, Disc))

# new data
Kx_Claim_new.df <- Kx_new.df %>%
  filter(
```

## 6. Beispielberechnungen

```
  Kx > 0
) %>%
nest(SUBTAB = c(x, Kx)) %>%
mutate(SUBTAB = future_map(SUBTAB, gather_data)) %>%
unnest(SUBTAB) %>%
nest(NEW = Kx)
```

Die Schätzergebnisse werden schlussendlich in einer Tabelle zusammengefasst.

```
# claim size parameter estimation
model <- stan_model('Gamma_PP.stan')

Claim_Est.df <- Kx_Claim_new.df %>%
  full_join(
    Kx_Claim_hist.df,
    by = c(
      'Profile',
      'Band_ID'
    )
  ) %>%
  mutate(
    Fit = map2(HISTORIC, NEW, create_fit),
    Post_Mean_alpha = map(
      Fit,
      function(x){get_posterior_mean(x, 'alpha')}
    ),
    Post_Mean_beta = map(
      Fit,
      function(x){get_posterior_mean(x, 'beta')}
    ),
    across(
      c(Post_Mean_alpha, Post_Mean_beta),
      ~as.numeric(.x)
    )
  )

Claim_Est_Summary.df <- Claim_Est.df %>%
  select(
    Profile,
    Band_ID,
    Post_Mean_alpha,
    Post_Mean_beta
```

```

)

Param_Est_Summary.df <- Claim_Est_Summary.df %>%
  full_join(
    Occur_Est_Summary.df,
    by = c(
      'Profile',
      'Band_ID'
    )
  ) %>%
  full_join(
    age_bands.df,
    by = c('Band_ID')
  ) %>%
  select(
    !Band_ID
  ) %>%
  select(
    c(x, x_lb, x_ub),
    !c(x, x_lb, x_ub)
  )

# save results
Param_Est_Summary.df %>%
write_csv2(
  paste0(
    'EST_Params_',
    formatC(v.n * 1e3, width = 4, flag = '0'),
    '.csv'
  )
)

```

### 6.3. Ergebnisanalyse

Im Folgenden werden die anhand des im vorherigen Abschnitt beschriebenen Prozesses ermittelten Schätzungen für die Verteilungsparameter der Kopfschäden näher analysiert. Zunächst werden die - durch Simulation gewonnenen - für die Schätzung verwendeten Ausgangsdaten ein wenig näher betrachtet. Von den acht simulierten Leistungsjahren (Periode 0 bis 7) wurden die letzten vier Jahre für die Parameterbestimmung herangezogen. In der folgenden Grafik ist die Altersverteilung und Größe des Versicherungsbestandes erkennbar.

## 6. Beispielberechnungen

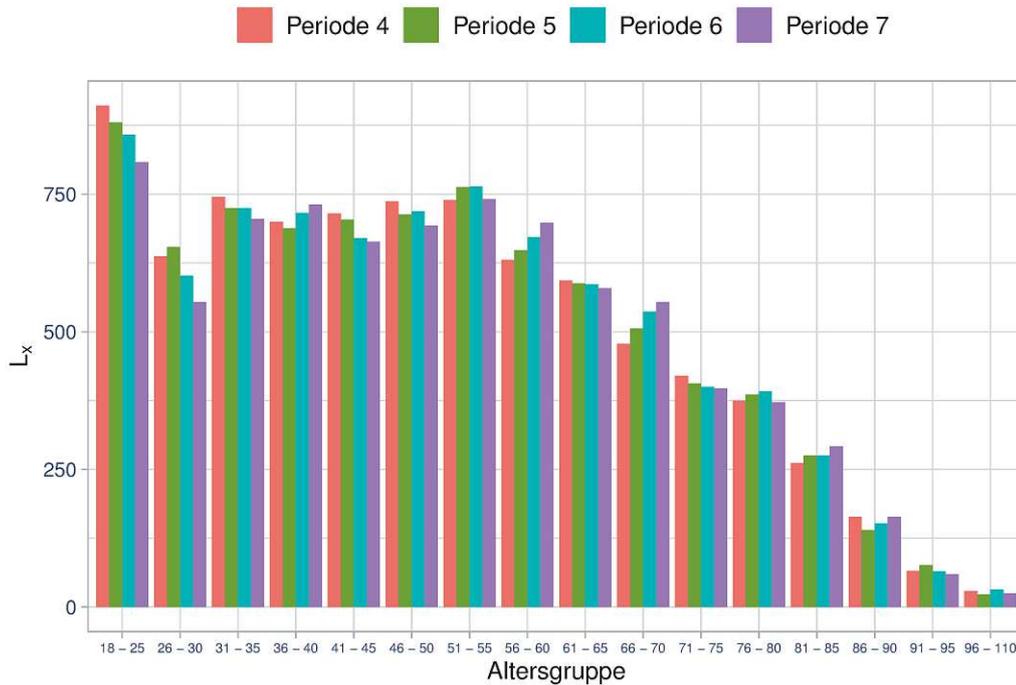


Abbildung 6.1.: Anzahl versicherter Personen pro Altersgruppe und Periode

In der Entwicklung des Gesamtbestandes ist ein leichter Rückgang von 8202 Personen in Periode 4 auf 8037 Personen in Periode 7 zu verzeichnen, d.h. die Neugesäftssegmente können das Ausscheiden aufgrund Storno und Tod nicht ausgleichen. Die nächste Grafik zeigt die Anzahl der eingetretenen Leistungsfälle.

## 6. Beispielberechnungen

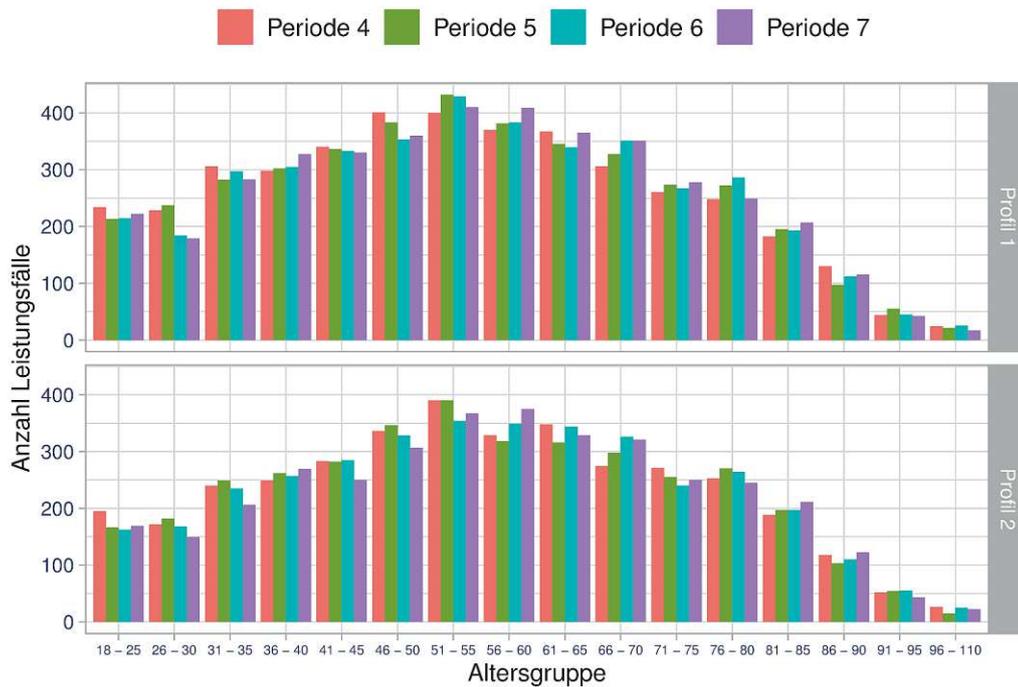


Abbildung 6.2.: Anzahl der Leistungsfälle pro Altersgruppe, Periode und Leistungsprofil

Die Anzahl der Leistungsfälle korrespondiert bei der Schätzung der Schadenhäufigkeit, also der  $\theta_x$ , zur Anzahl an „Erfolg“-Ereignissen (Bernoulli-Zufallsereignis) - die Gesamtanzahl an Ereignissen entspricht natürlich der Größe des Bestands - während sie bei der Schätzung der Schadenhöhenparameter, d.h. den Tupeln  $(\alpha_x, \beta_x)$ , die Anzahl an Gesamt ereignissen vorgibt. Als Nächstes wird die Entwicklung der Gesamtleistungen betrachtet.

## 6. Beispielberechnungen

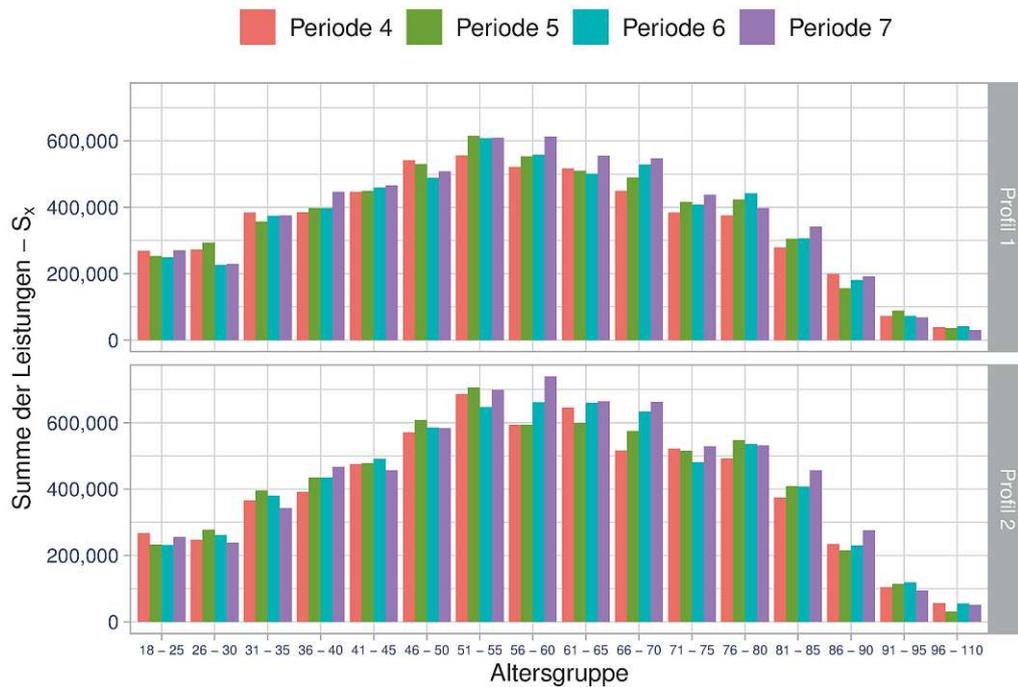


Abbildung 6.3.: Summe der Gesamtleistungen pro Altersgruppe, Periode und Leistungsprofil

In mehreren Altersgruppen, aber insbesondere in der Altersgruppe der 56- bis 60-jährigen sind hier durchaus stärkere Schwankungen zu erkennen. Ein Teil dieser lässt sich - wie anhand eines erneuten Blicks auf 6.3 erkennbar - anhand der Bestandsentwicklung (überwiegend Alterung) erklären, aber auch die auf die Schadenhöhenverteilungsparameter wirkende (simulierte) Inflation spielt eine tragende Rolle.

Nun folgen die Schätzergebnisse. Hier werden zunächst die Erwartungswerte der Kopfschäden  $\mathbb{E}[K_x] = \frac{\theta_x \alpha_x}{\beta}$  betrachtet.

## 6. Beispielberechnungen

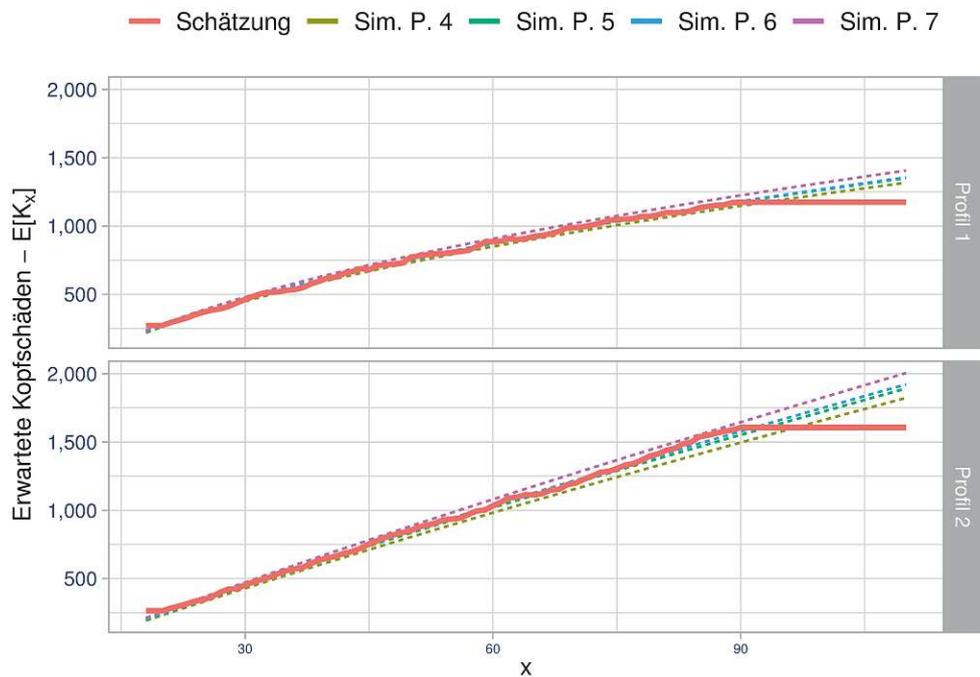


Abbildung 6.4.: Erwartete Kopfschäden gemäß Schätzungs- und Simulationsparametern pro Einzelalter

Die Randwerte sind auf die gewählte Gruppierung der Rohdaten zurückzuführen, bspw. werden für die Schätzung der Parameter  $(\theta_{18}, \alpha_{18}, \beta_{18})$  die Werte der Alter 18 bis 22 herangezogen, womit eine entsprechende Verzerrung nach oben entsteht. Abseits der Ränder liegen die Schätzwerte stets unter den anhand der Simulationsparameter der letzten Periode errechneten Werte, was wenig überraschen sollte, da die geringeren Erwartungswerte vergangener Perioden die Schätzung entsprechend „nach unten“ ziehen - wenngleich dieser Effekt durch die Wahl des Diskontierungsparameters beeinflussbar ist.

## 6. Beispielberechnungen

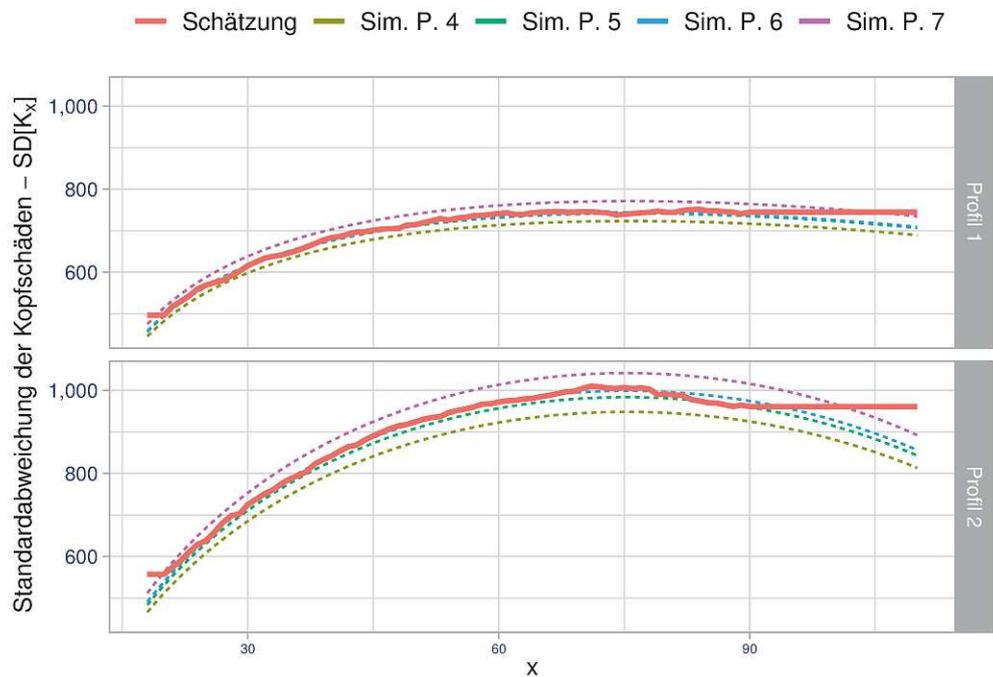


Abbildung 6.5.: Standardabweichung der Kopfschäden gemäß Schätzungs- und Simulationsparametern pro Einzelalter

Die Abweichungen bei den Schätzern der Standardabweichungen lassen sich gleichermaßen erklären. Die nächste Grafik zeigt ausgehend vom Letztperiodenbestand anhand Monte-Carlo Simulation ermittelte Dichtefunktionen der Gesamtportfolioleistungen, d.h. ausgehend vom selben Bestand wurden lediglich die Verteilungsparameter der Kopfschäden mit den Simulationsparametern der für die Schätzung betrachteten Perioden, sowie den Schätzparametern selbst, herangezogen, um Stichproben (Größe 10.000) zu erzeugen.

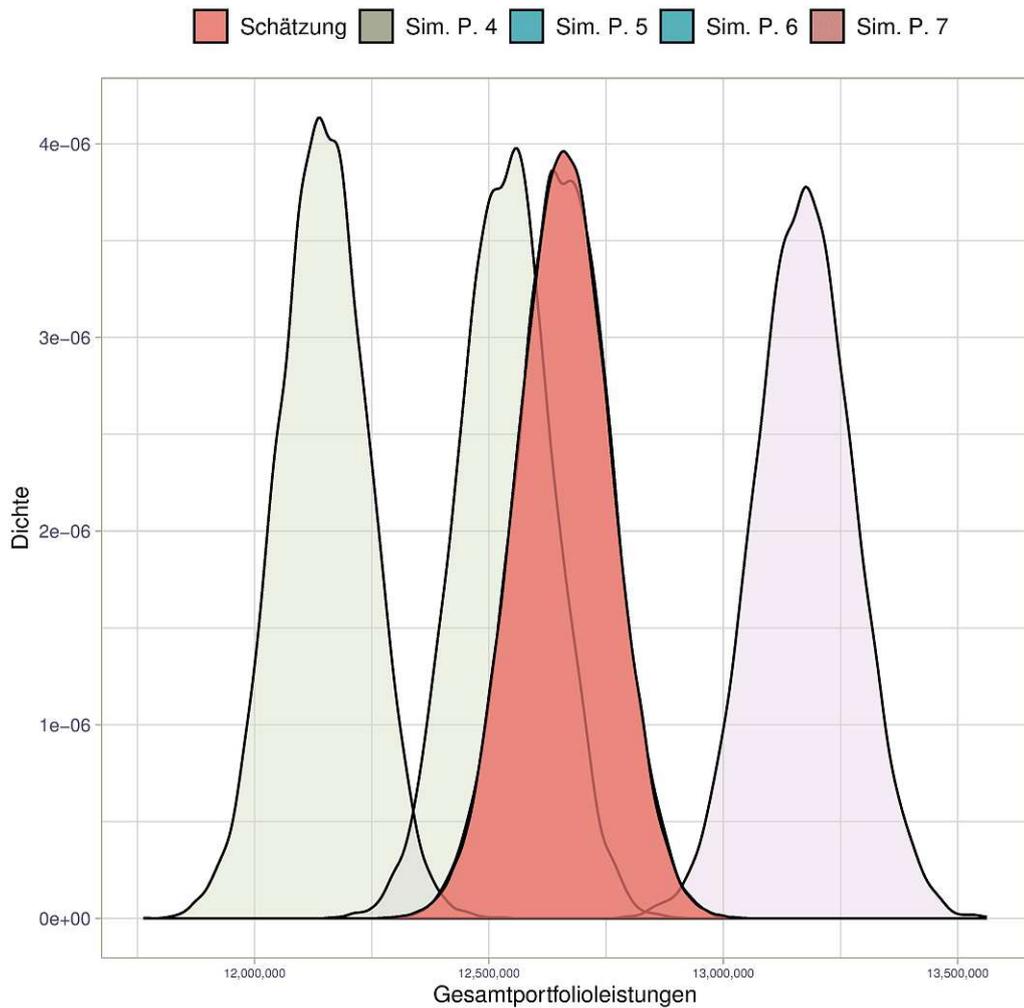


Abbildung 6.6.: Wahrscheinlichkeitsdichte der Gesamtportfolioleistungen

Wie auch schon in Abbildungen (6.4) und (6.5) ersichtlich, liegt das Schätzergebnis - abseits der Ränder - sehr nahe bei den Simulationsparametern der vorletzten Periode. Gepaart mit der im Vergleich zu den Vorperioden höheren (simulierten) Inflation von Periode 6 auf Periode 7, entstehen somit auf Gesamtportfolioebene größere Abweichungen zwischen Schätzung und „Echtniveau“ (Leistungen der letzten Periode). Zum Ausgleich dieser werden die geschätzten Kopfschäden um entsprechende Sicherheiten ergänzt. Die folgende Grafik zeigt eine Gegenüberstellung der observierten (simulierten) Leistungen der Letztperiode und den gemäß Schätzungsparametern erwarteten Leistungen mit Hinzugabe von Sicherheiten durch Anwendung des Standardabweichungsprinzips mit Parameter 5% bzw. 10% .

## 6. Beispielberechnungen

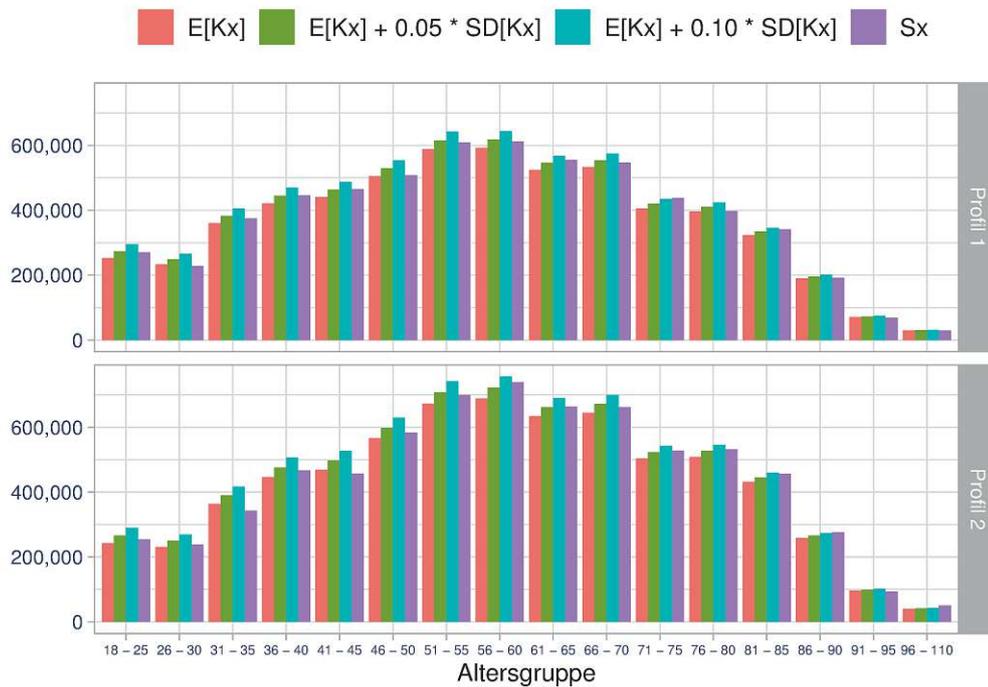


Abbildung 6.7.: Gegenüberstellung der tatsächlichen Leistungen und der erwarteten Leistungen zuzüglich Sicherheiten pro Altersgruppe

Die Grafik gibt zu erkennen, dass hier Sicherheiten in den Kopfschäden in der Höhe von 10% der Standardabweichung ebendieser bereits mehr als ausreichend - abseits geringer systematischer Effekte für hohe Alter - sind, um die vorliegende Unterschätzung auszugleichen.

## 7. Zusammenfassung

Die traditionelle Modellierung von Leistungsprofilen in der Krankenversicherung ist zwar grundsätzlich praktikabel, aber ihre Verfahren und deren Ergebnisse sind für viele aktuariell relevante bzw. interessante Prozesse (Pricing, Prämienanpassung) und Analysen (Bestimmung verteilungsbezogener KPIs, internes Modell) oft nur mit zusätzlichem manuellen Aufwand - inklusive entsprechendem aktuariellen Judgement - anwendbar oder schlichtweg unzureichend.

Im Zuge dieser Arbeit wird eine - mit der bisherigen Form kompatible - Erweiterung der traditionellen Modellierung betrachtet, welche viele der obigen Einschränkungen abmildert bzw. beseitigt. Hierbei werden Leistungsprofile durch Zufallsvariablen abgebildet und es wird ein auf Bayes'scher Inferenz basierendes adaptives Verfahren zur Schätzung von Verteilungsparametern vorgestellt. Des Weiteren werden einige Anwendungsmöglichkeiten (Analysen) vorgestellt, welche vom stochastischen Zugang profitieren bzw. erst durch diesen ermöglicht werden. Schlussendlich wird das Verfahren zur Veranschaulichung auf die Bestands- und Leistungshistorie eines simulierten Versicherungsbestands angewandt.

Das vorgestellte Verfahren ist weitgehend automatisierbar und ermöglicht durch Berechnung von Konfidenzintervallen oder anderer statistischer Größen zusätzlich das Erstellen genauerer Validierungs- und Entscheidungsprozesse (Backtesting, statistische Signifikanz, etc.). Nichtsdestotrotz hat die Wahl diverser Modellparameter - Diskontierung (Adaptivität) und „Verteilungstyp“ - in diversen Konstellationen (Bestandsgröße, Art der Leistungen) signifikanten Einfluss auf die Ergebnisse des Modells. In solchen Situationen bedarf es auch weiterhin an aktuariellem Judgement, wenngleich dieses nun aufgrund des Modelloutputs besser begründbar ist.

## A. Simulationsparameter

x	qx	wx	x	qx	wx	x	qx	wx
0	0.00354995	0.0500	38	0.0007843	0.0332	76	0.030344	0.0164
1	0.0002301	0.0496	39	0.00085145	0.0327	77	0.03421525	0.0159
2	0.00018355	0.0491	40	0.00093225	0.0323	78	0.0387107	0.0155
3	0.00014235	0.0487	41	0.00103165	0.0319	79	0.04390295	0.0150
4	0.00010925	0.0482	42	0.00115035	0.0314	80	0.04986515	0.0146
5	0.0000857	0.0478	43	0.0012888	0.0310	81	0.0566697	0.0142
6	0.00007275	0.0473	44	0.001447	0.0305	82	0.0643873	0.0137
7	0.0000689	0.0469	45	0.0016255	0.0301	83	0.0730854	0.0133
8	0.0000695	0.0465	46	0.0018234	0.0296	84	0.08282685	0.0128
9	0.00007075	0.0460	47	0.0020403	0.0292	85	0.09366805	0.0124
10	0.0000732	0.0456	48	0.00227695	0.0288	86	0.10565645	0.0119
11	0.00008075	0.0451	49	0.0025355	0.0283	87	0.1188287	0.0115
12	0.00009535	0.0447	50	0.0028169	0.0279	88	0.13320825	0.0111
13	0.0001193	0.0442	51	0.00312255	0.0274	89	0.14880285	0.0106
14	0.0001573	0.0438	52	0.0034556	0.0270	90	0.1656034	0.0102
15	0.0002087	0.0434	53	0.0038204	0.0265	91	0.1835818	0.0097
16	0.00027005	0.0429	54	0.00422185	0.0261	92	0.20269	0.0093
17	0.00033705	0.0425	55	0.004667	0.0257	93	0.2228596	0.0088
18	0.00040395	0.0420	56	0.00516095	0.0252	94	0.2440016	0.0084
19	0.0004622	0.0416	57	0.00570405	0.0248	95	0.26600695	0.0080
20	0.0005037	0.0412	58	0.00629635	0.0243	96	0.28874835	0.0075
21	0.0005239	0.0407	59	0.0069359	0.0239	97	0.31208235	0.0071
22	0.0005218	0.0403	60	0.0076167	0.0235	98	0.3358518	0.0066
23	0.0005064	0.0398	61	0.0083286	0.0230	99	0.3598896	0.0062
24	0.00049	0.0394	62	0.00906315	0.0226	100	0.3854483	0.0058
25	0.00047935	0.0389	63	0.00981825	0.0221	101	0.41038495	0.0053
26	0.0004735	0.0385	64	0.01059465	0.0217	102	0.4352779	0.0049
27	0.00047015	0.0381	65	0.0113973	0.0212	103	0.45996615	0.0044
28	0.00047	0.0376	66	0.01223935	0.0208	104	0.484293	0.0040
29	0.00047465	0.0372	67	0.0131415	0.0204	105	0.50810915	0.0035
30	0.0004847	0.0367	68	0.01413325	0.0199	106	0.53127535	0.0031
31	0.00050115	0.0363	69	0.01524805	0.0195	107	0.5537525	0.0027
32	0.00052415	0.0358	70	0.01652495	0.0190	108	0.5752946	0.0022
33	0.00055425	0.0354	71	0.01800775	0.0186	109	0.5957871	0.0018
34	0.00059125	0.0350	72	0.01974485	0.0181	110	0.6151349	0.0013
35	0.00063215	0.0345	73	0.021789	0.0177	111	0.633812	0.0009
36	0.0006768	0.0341	74	0.02419635	0.0173	112	0.65159495	0.0004
37	0.00072685	0.0336	75	0.02702675	0.0168	113	1	-

Tabelle A.1.: Sterbe- und Stornowahrscheinlichkeiten

## 7. Zusammenfassung

Die traditionelle Modellierung von Leistungsprofilen in der Krankenversicherung ist zwar grundsätzlich praktikabel, aber ihre Verfahren und deren Ergebnisse sind für viele aktuariell relevante bzw. interessante Prozesse (Pricing, Prämienanpassung) und Analysen (Bestimmung verteilungsbezogener KPIs, internes Modell) oft nur mit zusätzlichem manuellen Aufwand - inklusive entsprechendem aktuariellen Judgement - anwendbar oder schlichtweg unzureichend.

Im Zuge dieser Arbeit wird eine - mit der bisherigen Form kompatible - Erweiterung der traditionellen Modellierung betrachtet, welche viele der obigen Einschränkungen abmildert bzw. beseitigt. Hierbei werden Leistungsprofile durch Zufallsvariablen abgebildet und es wird ein auf Bayes'scher Inferenz basierendes adaptives Verfahren zur Schätzung von Verteilungsparametern vorgestellt. Des Weiteren werden einige Anwendungsmöglichkeiten (Analysen) vorgestellt, welche vom stochastischen Zugang profitieren bzw. erst durch diesen ermöglicht werden. Schlussendlich wird das Verfahren zur Veranschaulichung auf die Bestands- und Leistungshistorie eines simulierten Versicherungsbestands angewandt.

Das vorgestellte Verfahren ist weitgehend automatisierbar und ermöglicht durch Berechnung von Konfidenzintervallen oder anderer statistischer Größen zusätzlich das Erstellen genauerer Validierungs- und Entscheidungsprozesse (Backtesting, statistische Signifikanz, etc.). Nichtsdestotrotz hat die Wahl diverser Modellparameter - Diskontierung (Adaptivität) und „Verteilungstyp“ - in diversen Konstellationen (Bestandsgröße, Art der Leistungen) signifikanten Einfluss auf die Ergebnisse des Modells. In solchen Situationen bedarf es auch weiterhin an aktuariellem Judgement, wiewohl dieses nun aufgrund des Modelloutputs besser begründbar ist.