

Risikotheorie

Inhalt

Vorbemerkung	3
I Personenversicherungsmathematik	6
I. 1. Bewertung von Finanzströmen	6
I. 2. Lebensdauervertelungen und Sterbetafeln	10
I. 3. Lebensversicherungsprämien	21
I. 4. Deckungsrückstellungen	32
I. 5. Verbundene Leben	38
I. 6. Kostenrechnung und Überschussbeteiligung	41
I. 7. Die private Krankenversicherung	43
I. 8. Kopfschäden und Schadenprofile	46
I. 9. Alterungsrückstellungen und Prämienanpassungen	48
II Sachversicherungsmathematik	52
II. 1. Das kollektive Modell der Risikotheorie	52
II. 2. Die Panjer-Rekursion	71
II. 3. Die diskrete Fourier-Transformation	75
II. 4. Rückversicherung	82
II. 5. Schadenabwicklung und -reservierung	97
II. 6. Prämienprinzipien	102
III Credibility-Theorie	112
III. 1. Die Credibility-Prämie	119
III. 2. Die linearisierte Credibility-Prämie	122
III. 3. Die empirische Credibility-Prämie	125
III. 4. Modellerweiterungen	127
IV Ruintheorie	131
IV. 1. Zeitdiskrete Ruinprobleme	131
IV. 2. Zeitstetige Ruinprobleme	141
V Verallgemeinerte lineare Modelle	149
V. 1. Exponentialfamilien	149
V. 2. Lineare statistische Modelle	157
V. 3. Verallgemeinerte lineare Modelle	167
Verzeichnis der Definitionen, Sätze, Lemmata und Beispiele	184
Literatur	185
Danksagung	187

Vorbemerkung

Unter „Versicherung“ versteht man aus betriebswirtschaftlicher Sicht die *Deckung eines im einzelnen ungewissen, insgesamt geschätzten Mittelbedarfs auf der Grundlage des Risikoausgleichs im Kollektiv und in der Zeit*. Anders ausgedrückt bedeutet dies den Eintausch eines (zufallsbehafteten) Risikos des Versicherungsnehmers (VN) gegen eine (deterministische) Prämie an das Versicherungsunternehmen (VU), welches sich dadurch zur Übernahme der mit dem Risiko verbundenen Ansprüche oder Leistungen verpflichtet. Der VN erkauft sich damit einen gewissen finanziellen Schutz, da die in der Regel vergleichsweise kostengünstige Versicherungsprämie in ihrer Größenordnung bekannt und damit für den VN kalkulierbar ist, i. Allg. nicht jedoch die mögliche Belastung in einem Schadenfall (z.B. nicht gedeckte Hypothek für eine Immobilie nach einem Todesfall, Kosten einer aufwändigen Operation nach einem Unfall, Beschädigung oder Verlust eines Gebäudes durch einen Brandschaden usw.) Die mit solchen Risiken verbundenen finanziellen Gefahren kann das VU auf der anderen Seite nur durch Verteilung auf viele Versicherte (Ausgleich im Kollektiv) bzw. den Umstand der Nicht-Gleichzeitigkeit aller Schadenfälle (Ausgleich in der Zeit) minimieren. Eine gute Übersicht über die eher wirtschaftswissenschaftlichen Aspekte von Versicherung (Versicherungsbetriebslehre) gibt FARNY (2000).

Traditionell unterscheidet man die Versicherungsgeschäfte nach *Personenversicherung* (z.B. Lebens-, Kranken- und Pensionsversicherung), *Sachversicherung* (z.B. Feuerversicherung, Elementarschadenversicherung), *Vermögensversicherung* (z.B. Haftpflichtversicherung) und *Rückversicherung* (das ist die „Versicherung der Versicherer“). In jüngerer Zeit werden verstärkt auch alternative Versicherungskonzepte eingesetzt, die mit Instrumenten des Kapitalmarks arbeiten (Stichworte: *Index- und Fondsgebundene Lebensversicherung, Captives, Catastrophe Bonds* und *-Futures/Options, Securitisation, Alternative Risk Transfer*).

Naturgemäß fällt damit die Risikothorie und speziell die *Versicherungsmathematik* als Grundlage der Prämienkalkulation und der Risikomodellierung in den Bereich der Stochastik; zu ihren Themenschwerpunkten zählen insbesondere die Schätzung von Wahrscheinlichkeitsverteilungen und deren Parametern für die einzelnen Risiken (etwa auf der Basis von *Sterbe- und Invaliditätstafeln* im Bereich der Personenversicherungsmathematik oder von *Schadenfrequenz- und Schadenhöhenverteilungen* im Bereich der Sach- und Rückversicherung), die Modellierung von zeitabhängigen Risikoprozessen (insbesondere im Zusammenhang mit der Problematik von langfristigen Abwicklungen von Schadenfällen [*Loss Reserving*] und in der sog. *Ruintheorie*) oder die auf der zeitlichen Schadenerfahrung basierende *Prämiendifferenzierung* (*Credibility Theory* [Bayes-Verfahren], Bonus-Malus-Systeme [z.B. im Kfz-Haftpflichtbereich]).

Die Versicherungsmathematik zählt aufgrund ihrer großen wirtschaftlichen und sozialpolitischen Bedeutung mit zu den ältesten mathematischen Disziplinen, wie der folgende, teilweise ergänzte Auszug aus MILBRODT UND HELBIG (1999) zeigt (vgl. auch KOCH (1998) und KNOBLOCH UND VON DER SCHULENBURG (2000)):

Jahr	Ereignis
1308	Ältester bekannter Leibrentenvertrag, geschlossen zwischen dem Erzbischof von Köln und dem Kloster St. Denis bei Paris.
1370	Erste Seeversicherungsverträge in Genua (gemeinhin angesehen als Ursprung der Rückversicherung).
1583	<i>W. Gybbons</i> unterzeichnet in London den ersten bekannten Lebensversicherungsvertrag der Welt (einen Wettvertrag): Auszahlung von 400 £ bei Tod binnen eines Jahres (bei einer Einmalprämie von 30 £).
1585	<i>S. Stevin</i> stellt in der Schrift <i>Practique d'Arithmétique</i> eine Zinstafel sowie eine Tabelle von Endwerten von Zeitrenten in Abhängigkeit von der Laufzeit auf.
1590	Abschluss des sog. hamburgischen Seeversicherungsvertrags.
1591	Hamburger „Feuercontract“ zur Versicherung der städtischen Brauhäuser.
1662	<i>J. Graunt</i> verfasst auf Anregung von <i>W. Petty</i> die Schrift <i>Natural and political observations made upon the bills of mortality</i> mit einer Sterbetafel, die auf dem Londoner Todesregister beruht.
1669	<i>C.</i> und <i>L. Huygens</i> tauschen sich in einem Briefwechsel über Erwartungswert und Median der zukünftigen Lebensdauer unter Zugrundelegung von Graunts Sterbetafel aus, auch für verbundene Leben und Personengruppen, die beim letzten Tod erlöschen.
1670	Kampener „Kommunaltontine“, entsprechend einer Idee von <i>L. Tonti</i> , gestaltet als Rentenanleihe.
1671	<i>J. de Witt</i> verfasst die Schrift <i>Waerdye van Lyf-Renten naer Proportie van Los-Renten</i> (Prämienberechnung für Leibrenten, „Rechnungsgrundlagen erster Ordnung“) zum Zweck der Armee-Finanzierung im Niederländisch-Französischen Krieg.
1676	Gründung der Hamburger Feuerkasse, des ersten öffentlich-rechtlichen Versicherungsunternehmens der Welt.
1680/83	Zahlreiche Schriften von <i>G.W. Leibnitz</i> zu verschiedenen Problemen der Versicherungs- und Finanzmathematik, u.a. mit den Themen Öffentliche Assekuranzen (mit Bezug auf die kurz zuvor gegründete Hamburger Feuerkasse), verschiedene Arten der Zinsrechnung, Leibrenten, Pensionen, Lebensversicherungen (auch auf mehrere Leben), Bevölkerungsentwicklung.
1693	Der Astronom <i>E. Halley</i> verfasst die Schrift <i>An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and the funerals at the city of Breslaw; with an attempt to ascertain the price of annuitites upon lives</i> . Konstruktion einer Sterbetafel (Todesfälle von 1687 bis 1691 in Breslau), Darstellung von Leibrentenbarwerten.
1706	Gründung der <i>Amicable Society</i> , der ersten Lebensversicherungsgesellschaft der Welt, in London.
1725	<i>A. de Moivre</i> verfasst das erste Lehrbuch der Versicherungsmathematik mit dem Titel <i>Annuities upon Lives</i> . Sterbe-gesetz als Approximation von Halleys Sterbetafel, Rekursionsformeln für Leibrentenbarwerte.
1741	<i>J.P. Süßmilch</i> verfasst die Schrift <i>Die Göttliche Ordnung in den Veränderungen des menschlichen Geschlechts, aus der Geburt, dem Tode und der Fortpflanzung derselben</i> . Sterbetafel für Deutschland, mehr als 100 Jahre im Gebrauch.

- 1755 *J. Dodson* verfasst *The Mathematical Repository*. Lebensversicherung gegen laufende konstante Prämien, Einführung des Deckungskapitals.
- 1762 *Deed of Settlement* (Gründungsurkunde) der *Society for Equitable Assurances on Lives and Survivorships*. Erste Lebensversicherungsgesellschaft auf statistisch-mathematischer Basis. Wahl des auf Dodson zurückgehenden Begriffs des „Actuary“ (Aktuar) als Berufsbezeichnung des Versicherungsmathematikers.
- 1765 *D. Bernoulli* verfasst die Schrift *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir*. Zusammengesetzte Ausscheideordnung mit den Ausscheideursachen „Tod ohne vorherige Pockenerkrankung“ und „Ausscheiden durch Pockenerkrankung“.
- 1767/76 *L. Euler* verfasst die Schriften *Recherches générales sur la mortalité et la multiplication du genre humain* sowie *Sur les rentes viagères* und *Eclaircissements sur les établissements publics en faveur tant des veuves que des morts avec la description d'une nouvelle espèce de tontine aussi favorable au public qu'utile à l'état*. Erweiterung der Halleyschen Sterbetafelkonstruktion auf den Fall einer nichtstationären Bevölkerung. Jahresnettoprämien für Leibrenten (auch rekursiv), Bruttoprämien. Beschreibung einer „kontinuierlichen“ (zugangsoffenen) Tontinenversicherung.
- 1785/86 *N. Tetens* verfasst die Schrift *Einleitung zur Berechnung der Leibrenten und Anwartschaften, die vom Leben einer oder mehrerer Personen abhängen*. Erstes deutschsprachiges Lehrbuch der Lebensversicherungsmathematik (zweibändig); Einführung der Kommutationszahlen.
- 1792 Gründung der ersten Hagelversicherung in Neubrandenburg.
- 1820/25 *B. Gompertz* beschreibt das nach ihm benannte SterbeGesetz in den Texten *A sketch of an Analysis and Notation applicable to the Value of Life Contingencies* und *On the Nature of the Function Expressive of the Law of Human Mortality and on a new Method of Determining the Values of Life Contingencies*.
- 1845/51 *C.F. Gauß* erstellt ein Gutachten zur Prüfung der Professoren-Witwen- und Waisenkasse zu Göttingen.
- 1846 Gründung der Kölnischen Rückversicherungsgesellschaft.
- 1860/66 *W.M. Makeham* erweitert das Gompertz'sche SterbeGesetz in den Schriften *On the Law of Mortality* und *On the Principles to be observed in the Construction of Mortality Tables*.
- 1863 *A. Zillmer* entwickelt in *Beiträge zur Theorie der Prämienreserve bei Lebensversicherungsanstalten* eine Darstellung des Deckungskapitals unter Einschluss von Abschlusskosten.
- 1871/80 Erste Allgemeine Deutsche Sterbetafel (ADSt) für das gesamte Deutsche Reichsgebiet.
- 1898 Erste internationale Standardisierung versicherungsmathematischer Bezeichnungsweisen. Die Grundprinzipien dieser Notation gehen zurück auf *David Jones* (1843): *On the Value of Annuities and Reversionary Payments*.
- 1900 *L. Bachelier* leitet in der Schrift *Théorie de la Spéculation* eine Optionspreisformel unter Zugrundelegung einer Brown'schen Bewegung für die Aktienkursentwicklung her. Beginn der sog. *Stochastischen Finanzmathematik*.

I. Personenversicherungsmathematik

I. 1. Bewertung von Finanzströmen

In der Personenversicherungsmathematik spielen in der Regel Finanzströme eine wesentliche Rolle, z.B. bei der Pensionsversicherung durch regelmäßige Zahlungen des VU an den VN, aber auch in Form von Prämienzahlungen an das VU durch den VN z.B. in der Kapital-Lebensversicherung oder der Krankenversicherung. Hierdurch entsteht das Problem der finanztechnischen Bewertung solcher Zahlungsströme, entweder zum Ende der Laufzeit (Verzinsung) oder zum Anfang der Laufzeit (Diskontierung). Wir gehen im folgenden davon aus, dass Ein- und Auszahlungen K_j zu festen äquidistanten Zeitpunkten $j = 0, 1, \dots, n$ erfolgen (z.B. jährlich, monatlich usw.) und dass der Zinssatz i innerhalb der einzelnen Perioden konstant bleibt. Die Verzinsung erfolgt grundsätzlich zum Ende der jeweiligen Perioden. Die Größe $r = 1 + i$ heißt dabei *Zinsfaktor*, die Größe $v = \frac{1}{r} = \frac{1}{1+i}$ *Diskontfaktor*. Häufig wird auch noch die Größe $d = 1 - v$ betrachtet, die als *Diskont* (auf ein Kapital der Höhe 1) bezeichnet wird. Zur Herleitung der zugehörigen Bewertungsformeln betrachten wir folgende Graphik:

Zahlungszeitpunkt	0	1	2	3	...	n		
	----- ----- ----- ----- -----							
	K_0	rK_0	r^2K_0	r^3K_0	...	r^nK_0	Wertentwicklung von K_0	
		K_1	rK_1	r^2K_1	...	$r^{n-1}K_1$	Wertentwicklung von K_1	
			K_2	rK_2	...	$r^{n-2}K_2$	Wertentwicklung von K_2	
				K_3	...	$r^{n-3}K_3$	Wertentwicklung von K_3	
				:				
						K_n		
							$\sum_{j=0}^n r^{n-j} K_j$	Wertentwicklung gesamt

Der sog. *Endwert* W_n des durch die Zahlungen K_0, \dots, K_n gegebenen Finanzstroms ist also gegeben durch die Beziehung

$$W_n = \sum_{j=0}^n r^{n-j} K_j.$$

Durch Diskontierung mit dem Faktor v^n erhält man hieraus den sog. *Barwert* B_n des Finanzstroms:

$$B_n = v^n \sum_{j=0}^n r^{n-j} K_j = \sum_{j=0}^n v^j K_j.$$

In praktischen Anwendungen unterscheidet man noch die Fälle $K_0 = 0$ (*nachschüssige* Zahlungsweise) und $K_n = 0$ (*vorschüssige* Zahlungsweise).

Beispiel 1. Ein Sparer zahlt zum 1.1.2003 den Betrag von €1000,- auf ein Sparkonto ein, zum 1.1.2004 und 1.1.2005 jeweils den Betrag von €2000,- und hebt ohne weitere Zuzahlungen zum 31.12.2006 den Betrag von €3000,- ab. Der jährliche Zins betrage $i = 3\%$, also $i = 0,03$. Wie groß ist der Barwert des zum 31.12.2006 verbleibenden Kapitals?

Nach der obigen Bewertungsformel ergibt sich mit $n = 4$, $K_0 = 1000$, $K_1 = K_2 = 2000$, $K_3 = 0$ und $K_4 = -3000$ der Barwert (gerundet)

$$B_4 = \sum_{j=0}^4 v^j K_j = 1000 + v \cdot 2000 + v^2 \cdot 2000 - v^4 \cdot 3000 = 2161,48.$$

Der Sparer hätte also alternativ auch zum 1.1.2003 den Betrag von €2161,48 als Einmaleinlage einzahlen können, um zum 31.12.2006 auf denselben Endwert von €2432,76 zu kommen.

Barwerte spielen also insbesondere dann eine Rolle, wenn mehrere zukünftige Zahlungen durch eine Einmalzahlung zum Zeitpunkt 0 „abgelöst“ werden sollen.

Von besonderer Bedeutung in der Versicherungs- und Finanzmathematik sind wiederkehrende Zahlungen (sog. *Renten*) derselben festen Höhe. Obige Bewertungsformeln ergeben dann durch Anwendung der Formeln für die geometrische Reihe sofort folgende vereinfachte Ergebnisse, für die in der Versicherungsmathematik besondere Symbole verwendet werden:

$$a_{\overline{n}|} := v + v^2 + \dots + v^n = \sum_{k=1}^n v^k = v \frac{1-v^n}{1-v}$$

Barwert einer n -periodigen *nachschüssigen* Rente der Höhe 1

$$\ddot{a}_{\overline{n}|} := 1 + v + \dots + v^{n-1} = \sum_{k=0}^{n-1} v^k = \frac{1-v^n}{1-v}$$

Barwert einer n -periodigen *vorschüssigen* Rente der Höhe 1

$$s_{\overline{n}|} := r^{n-1} + r^{n-2} + \dots + 1 = \sum_{k=0}^{n-1} r^k = \frac{r^n - 1}{r - 1} = \frac{r^n - 1}{i}$$

Endwert einer n -periodigen *nachschüssigen* Rente der Höhe 1

$$\ddot{s}_{\overline{n}|} := r^n + r^{n-1} + \dots + r = \sum_{k=1}^n r^k = r \frac{r^n - 1}{r - 1} = r \frac{r^n - 1}{i}$$

Endwert einer n -periodigen *vorschüssigen* Rente der Höhe 1

Bar- und Endwerte von Rentenzahlungen der Höhe R sind hieraus entsprechend durch Multiplikation der entsprechenden Größen (für eine Rente der Höhe 1) mit R zu erhalten.

Hilfreich für einige Anwendungen ist die folgende Umrechnungstabelle für diese Größen.

	$\ddot{a}_{\overline{n} }$	$a_{\overline{n} }$	$\ddot{s}_{\overline{n} }$	$s_{\overline{n} }$	v^n	r^n
$\ddot{a}_{\overline{n} } =$		$r a_{\overline{n} }$	$\frac{\ddot{s}_{\overline{n} }}{1+d\ddot{s}_{\overline{n} }}$	$\frac{r s_{\overline{n} }}{1+i s_{\overline{n} }}$	$\frac{1-v^n}{d}$	$\frac{r^n-1}{r^n d}$
$a_{\overline{n} } =$	$v \ddot{a}_{\overline{n} }$		$\frac{v \ddot{s}_{\overline{n} }}{1+d\ddot{s}_{\overline{n} }}$	$\frac{s_{\overline{n} }}{1+i s_{\overline{n} }}$	$\frac{1-v^n}{i}$	$\frac{r^n-1}{r^n i}$
$\ddot{s}_{\overline{n} } =$	$\frac{\ddot{a}_{\overline{n} }}{1-d\ddot{a}_{\overline{n} }}$	$\frac{r a_{\overline{n} }}{1-i a_{\overline{n} }}$		$r s_{\overline{n} }$	$\frac{1-v^n}{v^n d}$	$\frac{r^n-1}{d}$
$s_{\overline{n} } =$	$\frac{v \ddot{a}_{\overline{n} }}{1-d\ddot{a}_{\overline{n} }}$	$\frac{a_{\overline{n} }}{1-i a_{\overline{n} }}$	$v \ddot{s}_{\overline{n} }$		$\frac{1-v^n}{v^n i}$	$\frac{r^n-1}{i}$
$v^n =$	$1-d\ddot{a}_{\overline{n} }$	$1-i a_{\overline{n} }$	$\frac{1}{1+d\ddot{s}_{\overline{n} }}$	$\frac{1}{1+i s_{\overline{n} }}$		$\frac{1}{r^n}$
$r^n =$	$\frac{1}{1-d\ddot{a}_{\overline{n} }}$	$\frac{1}{1-i a_{\overline{n} }}$	$1+d\ddot{s}_{\overline{n} }$	$1+i s_{\overline{n} }$	$\frac{1}{v^n}$	

Beispiel 2. Ein Studierender, der heute seinen 20. Geburtstag feiert, möchte durch eine Einmalzahlung auf einen Sparvertrag einen Beitrag zu seiner Alterssicherung leisten. Gedacht ist an eine 25-jährige monatliche vorschüssige Rente der Höhe €1000,-, beginnend mit dem vollendeten 65. Lebensjahr. Welchen Betrag muss er jetzt einzahlen, wenn das Kreditinstitut einen Jahreszins von 4% für die gesamte Laufzeit (also bis zum 90. Lebensjahr) garantiert?

Der Barwert der vorschüssigen Rente (vereinfachend angenommen: jährliche Höhe $R = 12000$) zum Abschluss des 65. Lebensjahres beträgt nach obigem

$$\ddot{a}_{\overline{25}|} \cdot R = 12000 \frac{1-v^{25}}{1-v} = 12000 \frac{1-1/1,04^{25}}{1-1/1,04} = 194963,56.$$

Für die notwendige Einmalzahlung muss dieser Betrag noch über 45 Jahre diskontiert werden, d.h. der entsprechende Barwert B ergibt sich zu

$$B = v^{45} \ddot{a}_{\overline{25}|} \cdot R = 33377,45.$$

(Dasselbe Ergebnis hätte man mit der zweiten allgemeinen Bewertungsformel oben erhalten unter Zugrundelegung von $K_0 = K_1 = \dots K_{44} = 0$, $K_{45} = K_{46} = \dots = K_{69} = 12000$.) Zur Finanzierung des nominalen Gesamtbetrags aller zukünftigen Rentenzahlungen in Höhe von $25 \cdot 12000 \text{ €} = 300000 \text{ €}$ ist also heute (nur) ein einmaliger finanzieller Aufwand von etwa 11% dieser Summe erforderlich.

Natürlich spielt der garantierte Zinssatz hier eine wesentliche Rolle. Die nachfolgende Tabelle zeigt, welche Barwerte (Einmalzahlungen) sich bei unterschiedlichen Zinssätzen ergeben.

Zinssatz i	2%	3%	4%	5%	6%
Barwert B (in €)	98023,54	56914,20	33377,45	19764,45	11813,21

Hat man umgekehrt nur einen bestimmten Betrag K für die Einmalzahlung zur Verfügung, so kann man aus der Gleichung

$$K = R \cdot v^{45} \frac{1-v^{25}}{1-v} \quad \text{bzw.} \quad v^{70} - v^{45} - \frac{K}{R}v + \frac{K}{R} = 0$$

den Diskontfaktor v und damit den notwendigen Zinssatz i berechnen. Die folgende Tabelle zeigt, welchen Zinssatz man bei verschiedenen Werten von K (bei $R = 12000$) benötigt.

Einmalzahlung K (in €)	5000	10000	15000	20000	25000	30000
Zinssatz i (in %)	7,710	6,327	5,533	4,977	4,549	4,202

Beispiel 3 (Annuitätentilgung einer Hypothek). Die obigen Formeln können auch dazu verwendet werden, um Tilgungspläne für Hypotheken H mit konstanter Annuität A zu berechnen. Dazu setzen wir $K_0 = -H$ (Auszahlung der Hypothek H zur Zeit 0) und $K_1 = K_2 = \dots = K_n = A$ (Einzahlungen der Annuität A am Ende der jeweiligen Periode). Die Bedingung für die vollständige Tilgung der Hypothek einschließlich sämtlicher anfallender Zinszahlungen lautet dann

$$0 = W_n = -r^n H + A \sum_{k=0}^{n-1} r^k = -r^n H + \frac{r^n - 1}{r - 1} A, \text{ also } A = \frac{r^n (r - 1)}{r^n - 1} H = \frac{r^n}{s_{\overline{n}|}} H.$$

Die Annuität kann dabei je Periode zerlegt werden in einen Tilgungsanteil T_k und einen Zinsanteil Z_k , für die analog gilt:

$$T_k = r^{k-n-1} A = v^{n+1-k} A, \quad Z_k = A - T_k = (1 - v^{n+1-k}) A, \quad k = 1, \dots, n$$

mit der jeweiligen Restschuld H_k am Ende der Periode k , gegeben durch

$$H_k = \frac{r^n - r^k}{r^n - 1} H, \quad k = 1, \dots, n.$$

Die folgende Tabelle gibt beispielhaft einen Tilgungsplan wieder für die Situation $H = 100000$ € bei $n = 10$ Jahren Laufzeit und 4% Jahreszins, woraus sich nach obiger Formel für die Annuität der Wert $A = 12329,09$ € ergibt.

Ende Jahr	Vorschuld	Annuität	Zinsanteil	Tilgungsanteil	Restschuld
1	100000,00	12329,09	4000,00	8329,09	91670,91
2	91670,91	12329,09	3666,84	8662,26	83008,65
3	83008,65	12329,09	3320,35	9008,75	73999,90
4	73999,90	12329,09	2960,00	9369,10	64630,80
5	64630,80	12329,09	2585,23	9743,86	54886,94
6	54886,94	12329,09	2195,48	10133,62	44753,32
7	44753,32	12329,09	1790,13	10538,96	34214,36
8	34214,36	12329,09	1368,57	10960,52	23253,84
9	23253,84	12329,09	930,15	11398,94	11854,90
10	11854,90	12329,09	474,20	11854,90	0,00

I. 2. Lebensdauerverteilungen und Sterbetafeln

Bei *Leibrenten*, deren Zahlung (erst) mit dem Tode der begünstigten Person endet, lassen sich die obigen einfachen Rechnungen so nicht durchführen, da die gesamte Zahlungsdauer n hier die Zufallsvariable $K_x + 1$ ist, wobei K_x die nach unten abgerundete (sog. „gestutzte“) ganzzahlige restliche Lebensdauer für die begünstigte Person mit dem jetzigen Alter x ist.¹ Als „faire“ Prämie wird man hier den Erwartungswert des Barwerts ansehen (sog. *Äquivalenzprinzip* der Versicherung²), also etwa bei vorschüssiger Zahlungsweise und Auszahlungen der Höhe 1 die Größe

$$\ddot{a}_x := E \left(\sum_{k=0}^{K_x} v^k \right) = E \left(\frac{1 - v^{K_x+1}}{1 - v} \right) = \frac{1 - E(v^{K_x+1})}{1 - v} \quad \text{mit} \quad E(v^{K_x+1}) = \sum_{k=0}^{\infty} v^{k+1} P(K_x = k),$$

für deren Berechnung die Kenntnis der Wahrscheinlichkeitsverteilung von K_x notwendig ist. Der Ausdruck $E(v^{K_x+1})$ stimmt dabei gerade mit der *wahrscheinlichkeitserzeugenden Funktion* φ_{K_x+1} von K_x an der Stelle v überein; es gilt hier noch die Beziehung $\varphi_{K_x+1}(v) = v \cdot \varphi_{K_x}(v)$.

¹ Man beachte, dass die Anzahl der Zahlperioden tatsächlich $K_x + 1$ beträgt. Ist z.B. $K_x = 0$, so überlebt die betreffende Person das nächste Lebensjahr zwar nicht, erhält aber zu Beginn noch eine Rentenzahlung, d.h. die Anzahl der Zahlperioden ist hier gerade 1.

² Dies lässt sich durch das *Gesetz der Großen Zahlen* motivieren: verkauft ein VU in einem Portfolio eine große Zahl gleichartiger Leibrenten, d.h. an Personen mit demselben Alter x und gleichem Geschlecht, so stimmen die durchschnittlichen diskontierten zufälligen Kosten des VU aus dem Gesamtportfolio mit dem zugehörigen Erwartungswert, d.h. hier \ddot{a}_x , näherungsweise überein. Das Äquivalenzprinzip entspricht also einer Art Umlageverfahren, bei dem die diskontierten Kosten gleichmäßig auf die VN umgelegt werden.

Beispiel 4 (Fortsetzung von Beispiel 2). Wenn wir – zunächst fiktiv – annehmen, dass die restliche (gestutzte) Lebensdauerverteilung einer 65-jährigen männlichen Person beschrieben werden kann durch die abgeschnittene geometrische Verteilung

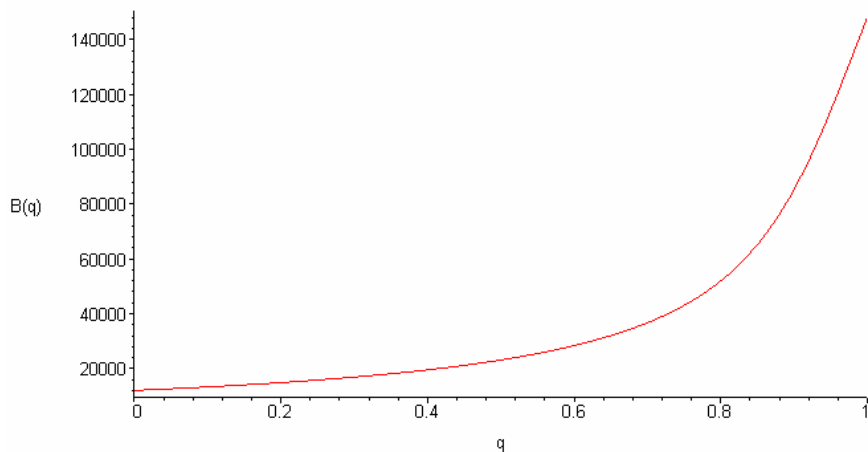
$$P(K_x = k) = q^k \frac{1-q}{1-q^{36}}, \quad k = 0, 1, \dots, 35 \quad \text{mit einem Parameter } q \in (0, 1),$$

d.h. das maximal erreichbare Alter („Endalter“) beträgt in unserer Beispielrechnung $\varpi^* = 100$, so folgt – wieder mit dem Rechnungszins $i = 4\%$ –

$$E(v^{K_x+1}) = \sum_{k=0}^{\infty} v^{k+1} P(K_x = k) = v \frac{1-q}{1-q^{36}} \sum_{k=0}^{35} (qv)^k = v \frac{1-q}{1-qv} \frac{1-(qv)^{36}}{1-q^{36}}, \quad \text{also}$$

$$\ddot{a}_x = \ddot{a}_x(q) = \frac{1 - E(v^{K_x+1})}{1-v} = \frac{1-v \frac{1-q}{1-qv} \frac{1-(qv)^{36}}{1-q^{36}}}{1-v}.$$

Die folgende Graphik zeigt den Verlauf des Barwerts $B(q) = 12000 \cdot \ddot{a}_x(q)$ mit $x = 65$, der für unser Beispiel relevant ist, im Bereich $q \in (0, 1)$.

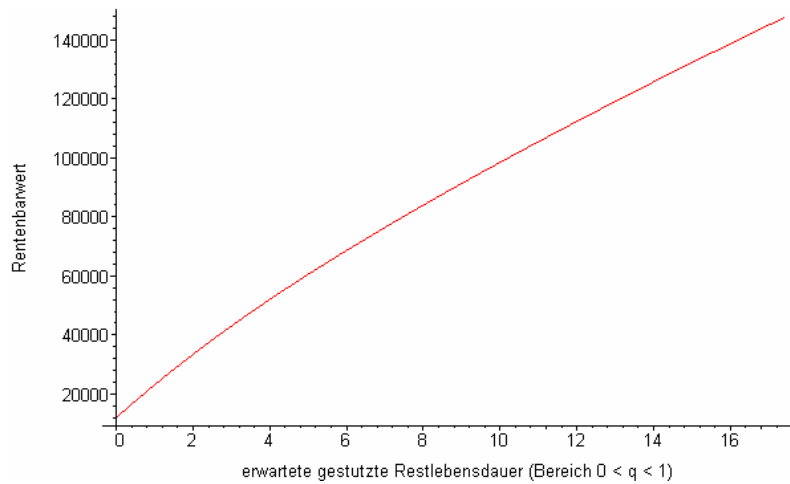


Für den Erwartungswert der restlichen gestutzten Lebensdauer erhalten wir in unserem Beispiel nach einiger Rechnung den Wert

$$E(K_x) = \frac{q - 36q^{36} + 35q^{37}}{(1-q)(1-q^{36})}.$$

Die folgende Graphik zeigt die Abhängigkeit des Rentenbarwerts von dieser erwarteten restlichen gestutzten Lebensdauer.

I Personenversicherungsmathematik



Bei einer beispielhaften erwarteten Restlebensdauer von 16 Jahren bei Männern im Alter 65 (d.h. $q = 0,98613$ in unserem Modell) ergibt sich damit der Rentenbarwert $B(q) = 138684,70 \text{ €}$ also ein etwas kleinerer Wert als bei der ersten deterministischen Rechnung über die ursprüngliche Rentenphase von 25 Jahren, obwohl in unserem Beispiel sogar ein Alter von 100 Jahren erreicht werden kann, die Rentenphase sich in diesem Fall also auf bis zu 35 Jahre verlängern würde.

Man beachte, dass der Grenzfall $q \rightarrow 1$ zu einer diskreten Gleichverteilung führt; dies entspricht genau dem Sterbegesetz von de Moivre (1724). Die erwartete Restlebensdauer ergibt sich hier zu $E(K_x) = 17,5$ Jahren mit einem zugehörigen Rentenbarwert von $B(q) = 148128,22 \text{ €}$, was zugleich dem maximal erreichbaren Wert in unserem Verteilungsmodell entspricht.

Für die weiteren Betrachtungen ist es hilfreich, zunächst eine etwas allgemeinere Darstellung von Lebensdauern und ihren Verteilungen zu betrachten. Dazu wollen wir mit T_x die kontinuierliche restliche Lebensdauer einer Person mit dem jetzigen (nicht notwendig ganzzahligen) Alter x bezeichnen. T_0 entspricht somit der gesamten Lebensdauer einer neugeborenen Person (Alter 0). Ferner wollen wir annehmen, dass die Verteilung der Lebensdauer T_0 eine Dichte f besitzt. Mit F werde die zugehörige Verteilungsfunktion bezeichnet. Da Lebensdauern nicht-negative Zufallsgrößen sind, können wir o.B.d.A. annehmen, dass $F(t) = f(t) = 0$ ist für $t \leq 0$. Mit ϖ wollen wir schließlich noch den rechten Endpunkt der Verteilung(sfunktion) bezeichnen, d.h. es ist

$$\varpi := \sup\{t \in \mathbb{R} \mid F(t) < 1\}.$$

Definition 1. Unter den obigen Annahmen heißt die Funktion

$$\mu(t) := \frac{f(t)}{1 - F(t)}, \quad 0 \leq t < \varpi$$

die *Sterbeintensität* zur Verteilung der Lebensdauer T_0 .

Wegen $P(s < T_0 < s + h) = \int_s^{s+h} f(t) dt \approx h \cdot f(s)$ für $s, h > 0$ und kleine Werte von h folgt hieraus sofort

$$P(s < T_0 < s + h | T_0 > s) = \frac{P(s < T_0 < s + h)}{P(T_0 > s)} \approx h \cdot \frac{f(s)}{1 - F(s)} = h \cdot \mu(s) \text{ für kleine Werte von } h,$$

d.h. für kleine h ist die Wahrscheinlichkeit, innerhalb der Zeitspanne h zu sterben, wenn man bis zum Zeitpunkt s überlebt hat, näherungsweise proportional zu h , mit Proportionalitätsfaktor $\mu(s)$. Eine alternative Darstellung der Sterbeintensität ist deshalb für stetige Dichten möglich vermöge der Grenzwertbeziehung

$$\mu(t) = \lim_{h \downarrow 0} \frac{1}{h} P(t < T_0 < t + h | T_0 > t), \quad 0 \leq t < \varpi.$$

Ähnlich wie in Definition 1 kann man auch eine Sterbeintensität zur Verteilung der allgemeineren Lebensdauer T_x definieren, wenn deren Dichte und Verteilungsfunktion bekannt sind.

Lemma 1. Unter den obigen Voraussetzungen gilt:

- a) Die Verteilungsfunktion F_x , Dichte f_x und rechter Endpunkt ϖ_x der Lebensdauer T_x sind gegeben durch

$$F_x(t) = \frac{F(x+t) - F(x)}{1 - F(x)}, \quad f_x(t) = \frac{f(x+t)}{1 - F(x)}, \quad x, t \geq 0, \quad t < \varpi_x = \varpi - x \text{ für } x < \varpi;$$

- b) die Sterbeintensität μ_x zur Verteilung der Lebensdauer T_x ist gegeben durch

$$\mu_x(t) := \frac{f_x(t)}{1 - F_x(t)} = \frac{f(x+t)}{1 - F(x+t)} = \mu(x+t), \quad 0 \leq t < \varpi_x = \varpi - x;$$

- c) $F_x(t) = 1 - \exp\left(-\int_0^t \mu_x(s) ds\right) = 1 - \exp\left(-\int_0^t \mu(x+s) ds\right), \quad 0 \leq t < \varpi_x = \varpi - x.$

Beweis: Es ist

$$F_x(t) = P(T_x \leq t) = P(T_0 \leq x+t | T_0 > x) = \frac{P(x < T_0 \leq x+t)}{P(T_0 > x)} = \frac{F(x+t) - F(x)}{1 - F(x)} = \int_0^t \frac{f(x+s)}{1 - F(x)} ds,$$

woraus Teil a) wegen

$$\varpi_x = \sup\{t \in \mathbb{R} | F_x(t) < 1\} = \sup\left\{t \in \mathbb{R} \left| \frac{F(x+t) - F(x)}{1 - F(x)} < 1 \right.\right\} = \sup\{t \in \mathbb{R} | F(x+t) < 1\} = \varpi - x$$

folgt. Teil b) ergibt sich unmittelbar aus a) unter Beachtung von

$$1 - F_x(t) = 1 - \frac{F(x+t) - F(x)}{1 - F(x)} = \frac{1 - F(x+t)}{1 - F(x)}.$$

Teil c) folgt aus der Gleichung

$$\int_0^t \mu_x(s) ds = \int_0^t \frac{f_x(s)}{1 - F_x(s)} ds = -\ln(1 - F_x(t)), \quad 0 \leq t < \varpi_x = \varpi - x.$$

Bemerkung: Eine alternative Darstellung der allgemeinen Sterbeintensität ist für stetige Dichten analog zu oben möglich vermöge der Grenzwertbeziehungen

$$\mu_x(t) = \lim_{h \downarrow 0} \frac{1}{h} P(t < T_x < t + h | T_x > t) = \frac{d}{dt} (-\ln(1 - F_x(t))), \quad 0 \leq t < \varpi_x = \varpi - x.$$

Bezeichnet wie oben T_x die kontinuierliche restliche Lebensdauer einer x -jährigen Person, so heißt in Anlehnung an die schon verwendete Begriffsbildung

$$K_x := \max \{k \in \mathbb{Z} | T_x \geq k\}$$

die *gestutzte* restliche Lebensdauer der x -jährigen Person, wobei hier sinnvollerweise x als ganzzahlig angenommen werden soll. Die gestutzte restliche Lebensdauer entspricht also gerade dem ganzzahligen Anteil der kontinuierlichen restliche Lebensdauer. Für ihre Verteilung gilt:

$$P(K_x = k) = P(k \leq T_x < k + 1) = F_x(k + 1) - F_x(k), \quad k = 0, \dots, \varpi^* - x, \quad x < \varpi^*$$

mit $\varpi^* := \max \{m \in \mathbb{Z} | m < \varpi\}$. (Falls ϖ ganzzahlig ist, gilt also speziell $\varpi^* = \varpi - 1$; ϖ^* heißt entsprechend *Endalter* der gestutzten Lebensdauer K_0 .) Ferner werden folgende Wahrscheinlichkeiten in aktuarieller Notation besonders gekennzeichnet:

$${}_t p_x := P(T_x > t), \quad {}_t q_x := 1 - {}_t p_x = P(T_x \leq t), \quad {}_{s|t} q_x := P(s < T_x \leq s + t), \quad s, t \geq 0$$

$$\text{sowie } p_x := {}_1 p_x, \quad q_x := {}_1 q_x, \quad t \geq 0.$$

${}_t p_x$ heißt die t -jährige Überlebenswahrscheinlichkeit der x -jährigen Person, ${}_t q_x$ die t -jährige Sterbewahrscheinlichkeit der x -jährigen Person. Weitere wichtige Symbole sind:

$$e_x^\circ := E(T_x) \quad \text{und} \quad e_x := E(K_x)$$

für die jeweilige erwartete Restlebensdauer (restliche [gestutzte] Lebenserwartung). Zwischen diesen Größen bestehen folgende Beziehungen:

Lemma 2. Unter den obigen Voraussetzungen gilt:

a) ${}_{s+t}p_x = {}_s p_x \cdot {}_t p_{x+s}$, ${}_{s|t}q_x = {}_s p_x \cdot {}_t q_{x+s}$, $s, t, x \geq 0$;

b) ${}_k p_x = \prod_{j=0}^{k-1} p_{x+j} = \prod_{j=0}^{k-1} (1 - q_{x+j})$, $x \geq 0$, $k \in \mathbb{N}$ mit ${}_0 p_x = 1$, ${}_0 q_x = 0$;

c) $P(K_x = k) = {}_k p_x q_{x+k}$ und $P(K_x > k) = {}_{k+1} p_x$, $x, k \geq 0$;

d) $e_x^\circ = \int_0^\infty t \cdot f_x(t) dt = \int_0^\infty 1 - F_x(t) dt = \int_0^\infty {}_t p_x dt$, $e_x = \sum_{k=0}^\infty k \cdot P(K_x = k) = \sum_{k=0}^\infty P(K_x > k) = \sum_{k=1}^\infty {}_k p_x$;

e) $\mu(x) = \frac{1 + e_x^\circ}{e_x^\circ}$, falls f stetig, und $q_x = \frac{1 + e_{x+1} - e_x}{1 + e_{x+1}}$, $x \geq 0$.

Beweis: Es ist

$$\begin{aligned} {}_{s+t}p_x &= P(T_x > s+t) = \frac{P(T_0 > x+s+t)}{P(T_0 > x)} = \frac{1 - F(x+s+t)}{1 - F(x)} = \frac{1 - F(x+s+t)}{1 - F(x+s)} \cdot \frac{1 - F(x+s)}{1 - F(x)} \\ &= \frac{P(T_0 > x+s+t)}{P(T_0 > x+s)} \cdot \frac{P(T_0 > x+s)}{P(T_0 > x)} = P(T_{x+s} > t) \cdot P(T_x > s) = {}_s p_x \cdot {}_t p_{x+s} \quad \text{und} \end{aligned}$$

$$\begin{aligned} \frac{{}_{s|t}q_x}{{}_s p_x} &= \frac{P(s < T_x \leq s+t)}{P(T_x > s)} = \frac{\frac{P(x+s < T_0 \leq x+s+t)}{P(T_0 > x)}}{\frac{P(T_0 > x+s)}{P(T_0 > x)}} = \frac{P(x+s < T_0 \leq x+s+t)}{P(T_0 > x+s)} \\ &= P(T_{x+s} \leq t) = {}_t q_{x+s}. \end{aligned}$$

Hieraus folgen a) und b) wegen

$${}_k p_x = p_x \cdot {}_{k-1} p_{x+1} = p_x \cdot p_{x+1} \cdot {}_{k-2} p_{x+2} = \dots = \prod_{j=0}^{k-1} p_{x+j}.$$

Ferner ist

$$P(K_x > k) = P(K_x \geq k+1) = P(T_x \geq k+1) = {}_{k+1} p_x$$

und somit

$$\begin{aligned} P(K_x = k) &= P(K_x \geq k) - P(K_x \geq k+1) = {}_k p_x - {}_{k+1} p_x = {}_k p_x - {}_k p_x p_{x+k} \\ &= {}_k p_x (1 - p_{x+k}) = {}_k p_x q_{x+k} \end{aligned}$$

für $x, k \geq 0$, woraus sich Teil c) ergibt.

Teil d) folgt unmittelbar aus Eigenschaften des Erwartungswerts.³ Der erste Teil von e) folgt aus

$$e_x^\circ = \int_0^\infty \frac{1-F(x+t)}{1-F(x)} dt = \frac{1}{1-F(x)} \int_x^\infty 1-F(s) ds, \text{ also } e_x^\circ(1-F(x)) = \int_x^\infty 1-F(s) ds;$$

differenzieren ergibt: $-f(x)e_x^\circ + (1-F(x))e_x^{\circ\prime} = -(1-F(x))$, also $\mu(x) = \frac{f(x)}{1-F(x)} = \frac{1+e_x^{\circ\prime}}{e_x^\circ}$. Die restliche Beziehung ergibt sich aus

$$q_x e_{x+1} = e_{x+1} - p_x \sum_{k=1}^\infty {}_k p_{x+1} = e_{x+1} - \sum_{k=1}^\infty {}_{k+1} p_x = e_{x+1} + p_x - \sum_{k=1}^\infty {}_k p_x = 1 - q_x + e_{x+1} - e_x.$$

Beispiel 5. Es sei T_0 exponentialverteilt mit Parameter $\lambda > 0$, in Zeichen: $P^{T_0} = \mathcal{E}(\lambda)$. Dann ist $\varpi = \infty$, und es gilt:

$$f(t) = \lambda e^{-\lambda t}, F(t) = 1 - e^{-\lambda t}, \mu(t) = \frac{f(t)}{1-F(t)} \equiv \lambda, t \geq 0 \text{ sowie } \mu_x(t) = \mu(x+t) \equiv \lambda, x, t \geq 0,$$

d.h. es ist auch T_x exponentialverteilt mit (demselben) Parameter $\lambda > 0$; man sagt, die Exponentialverteilung sei *alterungsfrei* oder *gedächtnislos*. Die gestutzte (restliche) Lebensdauerverteilung ergibt sich zu

$$P(K_x = k) = F_x(k+1) - F_x(k) = e^{-\lambda k} - e^{-\lambda(k+1)} = e^{-\lambda k} (1 - e^{-\lambda}) = p q^k, k \in \mathbb{Z}^+ \text{ mit } p = 1 - e^{-\lambda},$$

d.h. K_x ist in diesem Fall *geometrisch* verteilt mit Parameter p , in Zeichen: $P^{K_x} = \mathcal{G}(p)$. Ferner ist

$$e_x^\circ = \frac{1}{\lambda}, e_x = \frac{q}{p} = \frac{e^{-\lambda}}{1 - e^{-\lambda}} = \frac{1}{e^\lambda - 1}.$$

(Achtung: p ist hier abweichend von der allgemeinen Notation die einjährige Sterbewahrscheinlichkeit!) Eine einfache Variante des obigen (für die Lebensversicherung unrealistischen) Modells erhält man durch das Abschneiden der Lebensdauern T_x bzw. der Verteilung an einem endlichen Wert $\varpi > 0$ vermöge der Definition

$$f^\varpi(t) := \frac{f(t)}{F(\varpi)}, 0 \leq t \leq \varpi.$$

(Die abgeschnittene Lebensdauer werde analog mit T_x^ϖ bezeichnet.) In diesem Fall ist ϖ natürlicherweise das Endalter, und es gilt

³ Mit Hilfe des Satzes von Fubini lässt sich etwa für eine nicht-negative Zufallsvariable Z mit Dichte h und Verteilungsfunktion H zeigen:

$$E(Z) = \int_0^\infty z \cdot h(z) dz = \int_0^\infty \int_0^z h(z) dw dz = \int_{0 \leq w \leq z} h(z) dw dz = \int_0^\infty \int_w^\infty h(z) dz dw = \int_0^\infty (1 - H(z)) dz.$$

$$F^{\overline{\omega}}(t) = \frac{F(t)}{F(\overline{\omega})}, \quad \mu^{\overline{\omega}}(t) = \frac{f^{\overline{\omega}}(t)}{1 - F^{\overline{\omega}}(t)} = \frac{f(t)}{F(\overline{\omega}) - F(t)}, \quad 0 \leq t < \overline{\omega} \text{ sowie allgemeiner}$$

$$f_x^{\overline{\omega}}(t) = \frac{f(x+t)}{F(\overline{\omega}) - F(x)}, \quad F_x^{\overline{\omega}}(t) = \frac{F(x+t) - F(x)}{F(\overline{\omega}) - F(x)},$$

$$\mu_x^{\overline{\omega}}(t) = \frac{f_x^{\overline{\omega}}(t)}{1 - F_x^{\overline{\omega}}(t)} = \frac{f(x+t)}{F(\overline{\omega}) - F(x+t)}, \quad 0 \leq t < \overline{\omega}_x = \overline{\omega} - x, \quad 0 \leq x < \overline{\omega}.$$

Für die restliche Lebenserwartung ergibt sich entsprechend

$$e_x^{\overline{\omega}} = \int_0^{\infty} 1 - F_x^{\overline{\omega}}(t) dt = \int_0^{\infty} \frac{F(\overline{\omega}) - F(x+t)}{F(\overline{\omega}) - F(x)} dt, \quad x < \overline{\omega}.$$

Im Fall von Exponentialverteilungen liefert dies speziell

$$f_x^{\overline{\omega}}(t) = \frac{f(x+t)}{F(\overline{\omega}) - F(x)} = \lambda \cdot \frac{e^{-\lambda t}}{1 - e^{-\lambda(\overline{\omega}-x)}}, \quad F_x^{\overline{\omega}}(t) = \frac{F(x+t) - F(x)}{F(\overline{\omega}) - F(x)} = \frac{1 - e^{-\lambda t}}{1 - e^{-\lambda(\overline{\omega}-x)}},$$

$$\mu_x^{\overline{\omega}}(t) = \frac{f_x^{\overline{\omega}}(t)}{1 - F_x^{\overline{\omega}}(t)} = \frac{\lambda}{1 - e^{-\lambda(\overline{\omega}-x-t)}}, \quad 0 \leq t < \overline{\omega}_x = \overline{\omega} - x,$$

$$e_x^{\overline{\omega}} = \frac{1}{\lambda} - \frac{\overline{\omega} - x}{e^{\lambda(\overline{\omega}-x)} - 1}, \quad 0 \leq x < \overline{\omega}.$$

Insbesondere nimmt also die Sterbeintensität $\mu_x^{\overline{\omega}}$ sowohl mit wachsendem Alter x als auch mit der Zeit t monoton zu.

Entsprechend lässt sich auch die geometrische Verteilung nach oben abschneiden, etwa so wie in Beispiel 4. Man beachte, dass nach obiger Bemerkung die abgeschnittene gestutzte Lebensdauer $K_x^{\overline{\omega}}$ hier rechnerisch nur den Endpunkt $\overline{\omega}^* = \overline{\omega} - 1$ erreicht, wenn $\overline{\omega}$ ganzzahlig ist. Es ergibt sich analog:

$$\begin{aligned} P(K_x^{\overline{\omega}} = k) &= F_x^{\overline{\omega}}(k+1) - F_x^{\overline{\omega}}(k) = \frac{F(x+k+1) - F(x+k)}{F(\overline{\omega}) - F(x)} \\ &= \frac{e^{-\lambda k} - e^{-\lambda(k+1)}}{1 - e^{-\lambda(\overline{\omega}-x)}} = \frac{pq^k}{1 - q^{\overline{\omega}-x}}, \quad k = 0, \dots, \overline{\omega}^* - x, \quad 0 \leq x \leq \overline{\omega}^* \text{ sowie} \end{aligned}$$

$${}_k p_x = P(K_x^{\overline{\omega}} \geq k) = \sum_{j=k}^{\overline{\omega}^*-x} \frac{pq^j}{1 - q^{\overline{\omega}-x}} = \frac{pq^k}{1 - q^{\overline{\omega}-x}} \sum_{j=0}^{\overline{\omega}^*-x-k} q^j = \frac{q^k - q^{\overline{\omega}-x}}{1 - q^{\overline{\omega}-x}}, \quad k = 0, \dots, \overline{\omega}^* - x, \quad 0 \leq x \leq \overline{\omega}^*$$

$$\text{und } e_x^{\overline{\omega}} = \sum_{k=1}^{\overline{\omega}^*-x} {}_k p_x = \sum_{k=1}^{\overline{\omega}^*-x} \frac{q^k - q^{\overline{\omega}-x}}{1 - q^{\overline{\omega}-x}} = \frac{q - (\overline{\omega} - x)q^{\overline{\omega}-x} + (\overline{\omega} - 1 - x)q^{\overline{\omega}+1-x}}{p(1 - q^{\overline{\omega}-x})}, \quad 0 \leq x \leq \overline{\omega}^*,$$

in Übereinstimmung mit Beispiel 4 (dort ist $\overline{\omega} - x = \overline{\omega} - 65 = 36$, also $\overline{\omega} = 101$, $\overline{\omega}^* = 100$).

Beispiel 6. Bekannte kontinuierliche Sterbegesetze sind:

1. Die stetige Gleichverteilung über dem Intervall $[0, \varpi]$ mit $\varpi > 0$ (de Moivre 1724) mit den Größen

$$f(t) = \begin{cases} \frac{1}{\varpi}, & 0 \leq t \leq \varpi \\ 0, & \text{sonst} \end{cases} \quad F(t) = \begin{cases} \frac{t}{\varpi}, & 0 \leq t \leq \varpi \\ 1, & t > \varpi \end{cases} \quad \mu(t) = \frac{1}{\varpi - t}, \quad 0 \leq t \leq \varpi$$

und der daraus folgenden diskreten Gleichverteilung über der Menge $\{0, 1, \dots, \varpi^* = \varpi - 1\}$ bei ganzzahligem ϖ für die gestutzte Lebensdauer K_0 ; die Verteilungen der Restlebensdauern T_x bzw. K_x sind ebenfalls Gleichverteilungen über dem Intervall $[0, \varpi - x]$ bzw. über der Menge $\{0, 1, \dots, \varpi^* - x = \varpi - x - 1\}$. Ebenso folgt die bei $0 < \tilde{\varpi} < \varpi$ abgeschnittene Lebensdauer $T_x^{\tilde{\varpi}}$ einer Gleichverteilung, und zwar über dem Intervall $[0, \tilde{\varpi} - x]$; die zugehörige abgeschnittene gestutzte Lebensdauer $K_x^{\tilde{\varpi}}$ ist entsprechend wieder gleichverteilt über der Menge $\{0, 1, \dots, \tilde{\varpi}^* - x = \tilde{\varpi} - x - 1\}$.

2. Die *Gompertz-Verteilung* (1824) mit der Sterbeintensität

$$\mu(t) = a e^{bt}, \quad t \geq 0 \quad (a, b > 0);$$

3. die *Makeham-Verteilung* (1860) mit der Sterbeintensität

$$\mu(t) = c + a e^{bt}, \quad t \geq 0 \quad (a, b, c > 0);$$

4. die *Weibull-Verteilung* (1939) mit der Sterbeintensität

$$\mu(t) = at^b, \quad t \geq 0 \quad (a, b > 0).$$

Für rechnerische Zwecke werden die einjährigen Überlebens- und Sterbewahrscheinlichkeiten üblicherweise in sogenannten *Sterbetafeln* zusammengefasst. Diese reichen nach Lemma 2 nämlich aus, um die Verteilung von K_x für alle relevanten Alter x vollständig zu beschreiben. Neben den Wahrscheinlichkeiten p_x und q_x enthalten solche Tafeln in der Regel auch die absoluten Größen l_x (Anzahl der Lebenden (*living*) mit Alter x) und d_x (Anzahl der im Alter x Gestorbenen (*dead*)). Standardmäßig beginnen diese Einträge im Alter $x = 0$ mit $l_0 = 100000$. Zwischen diesen Größen besteht ferner (bis auf Ungenauigkeiten durch Rundung) der Zusammenhang

$$q_x = \frac{d_x}{l_x}, \quad p_x = 1 - q_x = \frac{l_x - d_x}{l_x}, \quad x = 0, \dots, \varpi^*.$$

Wenn man von einem (theoretischen) Sterbegesetz bzw. der Verteilungsfunktion F der Gesamtlebensdauer ausgeht, kann man etwa

I Personenversicherungsmathematik

$$l_x = [100000 \cdot (1 - F(x))] \text{ mit } [z] := \max \{k \in \mathbb{Z} \mid k \leq z\} \text{ für } z \in \mathbb{R}$$

wählen. Die folgende Tabelle zeigt die Allgemeine Deutsche Sterbetafel aus den Jahren 1986/88 für die männliche Bevölkerung, die auf Auswertungen der damaligen Volkszählung beruht.

Bevölkerung

**Zum Aufsatz: „Allgemeine Sterbetafel“
Allgemeine Sterbetafel 1986/88 für die Bundesrepublik Deutschland
Männlich**

Vollendetes Alter	Überlebende im Alter x	Gestorbene im Alter x bis unter x + 1	Sterbe- wahrscheinlichkeit vom Alter x bis x + 1	Überlebens- wahrscheinlichkeit vom Alter x bis x + 1	Von den Überlebenden im Alter x		Durchschnittliche Lebenserwartung im Alter x in Jahren
					bis zum Alter x + 1 durchlebte	insgesamt noch zu durchlebende	
x	l_x	d_x	q_x	p_x	Jahre		e_x
					L_x	$e_x l_x$	
Wochen							
		während einer Woche	für eine Woche				
0	100 000	381	0,00380607	0,99619393	1 913	7 221 175	72,21
1	99 619	67	0,00067257	0,99932743	1 909	7 219 262	72,47
2	99 552	36	0,00036464	0,99963536	1 908	7 217 354	72,50
3	99 516	30 ¹⁾	0,00029945 ¹⁾	0,99970055 ¹⁾	2 571 ¹⁾	7 215 446	72,51
Monate							
		während eines Monats	für einen Monat				
0	100 000	514	0,00513709	0,99486291	8 300	7 221 175	72,21
1	99 486	86	0,00086295	0,99913705	8 287	7 212 875	72,50
2	99 400	75	0,00075282	0,99924718	8 280	7 204 588	72,48
3	99 326	61	0,00061732	0,99938268	8 275	7 196 308	72,45
4	99 264	48	0,00048815	0,99951185	8 270	7 188 034	72,41
5	99 216	35	0,00035623	0,99964377	8 267	7 179 764	72,37
6	99 180	27	0,00027482	0,99972518	8 264	7 171 497	72,31
7	99 153	23	0,00023541	0,99976459	8 262	7 163 233	72,24
8	99 130	19	0,00019441	0,99980559	8 260	7 154 971	72,18
9	99 111	14	0,00014581	0,99985419	8 259	7 146 711	72,11
10	99 096	13	0,00012645	0,99987355	8 257	7 138 453	72,04
11	99 084	9	0,00008997	0,99991003	8 257	7 130 195	71,98
Jahre							
		während eines Jahres	für ein Jahr				
0	100 000	925	0,00925273	0,99074727	99 237	7 221 175	72,21
1	99 075	70	0,00070529	0,99929471	99 040	7 121 939	71,88
2	99 005	49	0,00049307	0,99950693	98 980	7 022 899	70,93
3	98 956	35	0,00035530	0,99964470	98 938	6 923 918	69,97
4	98 921	30	0,00030589	0,99969411	98 906	6 824 980	68,99
5	98 891	29	0,00029097	0,99970903	98 876	6 726 074	68,02
6	98 862	27	0,00027641	0,99972359	98 848	6 627 198	67,03
7	98 835	25	0,00025775	0,99974225	98 822	6 528 350	66,05
8	98 809	23	0,00023576	0,99976424	98 797	6 429 528	65,07
9	98 786	21	0,00021668	0,99978332	98 775	6 330 731	64,09
10	98 764	20	0,00020463	0,99979537	98 754	6 231 956	63,10
11	98 744	20	0,00020218	0,99979782	98 734	6 133 201	62,11
12	98 724	21	0,00020926	0,99979074	98 714	6 034 467	61,12
13	98 704	23	0,00023283	0,99976717	98 692	5 935 753	60,14
14	98 681	29	0,00029344	0,99970656	98 668	5 837 061	59,15
15	98 652	40	0,00040200	0,99959800	98 632	5 738 395	58,17
16	98 612	55	0,00055634	0,99944366	98 584	5 639 764	57,19
17	98 557	74	0,00075125	0,99924875	98 520	5 541 179	56,22
18	98 483	94	0,00095645	0,99904355	98 438	5 442 659	55,26
19	98 389	105	0,00106913	0,99883087	98 336	5 344 223	54,32
20	98 284	108	0,00110194	0,99889806	98 229	5 245 887	53,37
21	98 175	107	0,00109139	0,99890861	98 122	5 147 657	52,43
22	98 068	104	0,00108503	0,99893497	98 016	5 049 536	51,49
23	97 964	102	0,00103813	0,99896187	97 913	4 951 520	50,54
24	97 862	99	0,00101645	0,99898355	97 812	4 853 607	49,60
25	97 763	98	0,00100325	0,99899675	97 714	4 755 795	48,65
26	97 664	98	0,00100130	0,99899870	97 618	4 658 081	47,69
27	97 567	99	0,00101285	0,99898715	97 517	4 560 465	46,74
28	97 468	101	0,00103836	0,99896184	97 417	4 462 948	45,79
29	97 367	105	0,00107643	0,99892357	97 314	4 365 531	44,84
30	97 262	109	0,00112194	0,99887806	97 207	4 268 217	43,88
31	97 153	114	0,00117247	0,99882753	97 096	4 171 009	42,93
32	97 039	119	0,00122951	0,99877049	96 979	4 073 914	41,98
33	96 920	126	0,00129498	0,99870502	96 857	3 976 934	41,03
34	96 794	133	0,00137346	0,99862654	96 728	3 880 078	40,09

¹⁾ In den übrigen Tagen des 1. Lebensmonats.

I Personenversicherungsmathematik

Allgemeine Sterbetafel 1986/88 für die Bundesrepublik Deutschland

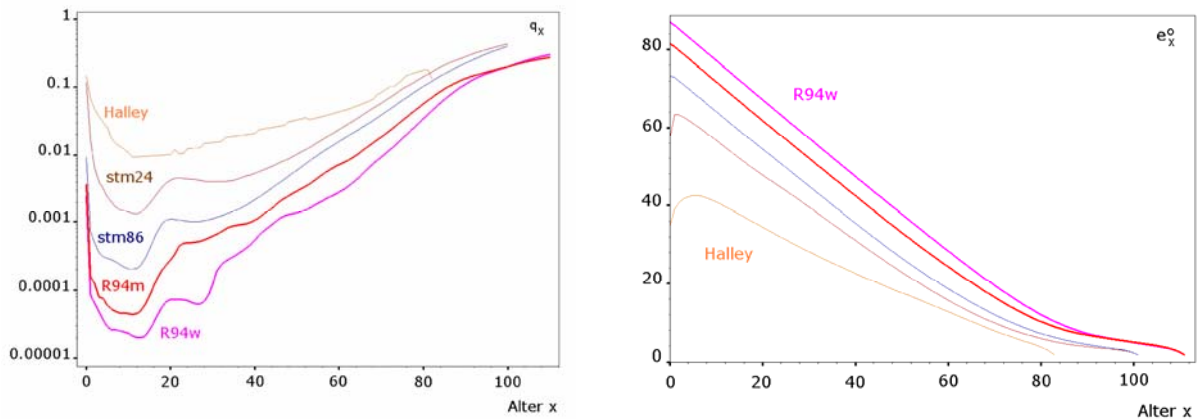
Männlich

Vollendetes Alter	Überlebende im Alter x	Gestorbene im Alter x bis unter x + 1	Sterbe- Überlebens- wahrscheinlichkeit vom Alter x bis x + 1		Von den Überlebenden im Alter x bis zum Alter x + 1 insgesamt noch zu durchlebende		Durchschnittliche Lebenserwartung im Alter x in Jahren
			q_x	p_x	Jahre		
x	l_x	d_x	q_x	p_x	L_x	$e_x l_x$	e_x
		während eines Jahres		für ein Jahr			
Jahre							
35	96 861	142	0,00146739	0,99853261	96 590	3 783 350	39,14
36	96 519	152	0,00157643	0,99842357	96 443	3 686 780	38,20
37	96 367	164	0,00170033	0,99829967	96 285	3 590 317	37,26
38	96 203	177	0,00184126	0,99815874	96 115	3 494 032	36,32
39	96 028	192	0,00200292	0,99799708	95 930	3 397 917	35,39
40	95 834	210	0,00218931	0,99781069	95 729	3 301 987	34,46
41	95 624	230	0,00240516	0,99759484	95 509	3 206 258	33,53
42	95 394	253	0,00265040	0,99734960	95 268	3 110 749	32,61
43	95 141	279	0,00292755	0,99707245	95 002	3 015 482	31,69
44	94 863	307	0,00323949	0,99676051	94 709	2 920 480	30,79
45	94 555	339	0,00358825	0,99641175	94 386	2 825 771	29,88
46	94 216	375	0,00397516	0,99602484	94 029	2 731 386	28,99
47	93 841	413	0,00440134	0,99559866	93 635	2 637 357	28,10
48	93 428	455	0,00487349	0,99512651	93 201	2 543 722	27,23
49	92 973	502	0,00540080	0,99460920	92 722	2 450 521	26,36
50	92 471	554	0,00599270	0,99400730	92 194	2 357 799	25,50
51	91 917	612	0,00665469	0,99334531	91 611	2 265 605	24,65
52	91 305	675	0,00739081	0,99260919	90 968	2 173 994	23,81
53	90 630	743	0,00820182	0,99179818	90 259	2 083 026	22,98
54	89 887	816	0,00908183	0,99091817	89 479	1 992 768	22,17
55	89 071	893	0,01002896	0,98997104	88 624	1 903 289	21,37
56	88 177	974	0,01104374	0,98895626	87 690	1 814 665	20,58
57	87 204	1 058	0,01212705	0,98787295	86 675	1 726 974	19,80
58	86 146	1 144	0,01328261	0,98671739	85 574	1 640 300	19,04
59	85 002	1 234	0,01452099	0,98547901	84 385	1 554 726	18,29
60	83 767	1 328	0,01585428	0,98414572	83 103	1 470 341	17,55
61	82 439	1 426	0,01729556	0,98270444	81 726	1 387 238	16,83
62	81 014	1 528	0,01885955	0,98114045	80 250	1 305 511	16,11
63	79 486	1 634	0,02056058	0,97943942	78 669	1 225 261	15,41
64	77 851	1 745	0,02241628	0,97756372	76 979	1 146 593	14,73
65	76 106	1 861	0,02445496	0,97554504	75 176	1 069 614	14,05
66	74 245	1 983	0,02671049	0,97328951	73 254	994 438	13,39
67	72 262	2 112	0,02922845	0,97077155	71 206	921 185	12,75
68	70 150	2 249	0,03205730	0,96794270	69 025	849 979	12,12
69	67 901	2 393	0,03523824	0,96476176	66 705	780 953	11,50
70	65 508	2 543	0,03881231	0,96118769	64 237	714 249	10,90
71	62 966	2 696	0,04281953	0,95718047	61 618	650 012	10,32
72	60 270	2 851	0,04729665	0,95270335	58 844	588 394	9,76
73	57 419	3 002	0,05227739	0,94772261	55 918	529 550	9,22
74	54 417	3 144	0,05778285	0,94221715	52 845	473 631	8,70
75	51 273	3 273	0,06382971	0,93617029	49 637	420 786	8,21
76	48 000	3 381	0,07042925	0,92957075	46 310	371 149	7,73
77	44 620	3 462	0,07759312	0,92240688	42 889	324 840	7,28
78	41 157	3 512	0,08533616	0,91466384	39 401	281 951	6,85
79	37 645	3 526	0,09367422	0,90632578	35 882	242 550	6,44
80	34 119	3 501	0,10262092	0,89737908	32 368	206 668	6,06
81	30 618	3 435	0,11218740	0,88781260	28 900	174 299	5,69
82	27 183	3 327	0,12237982	0,87762018	25 519	145 399	5,35
83	23 858	3 178	0,13320462	0,86679538	22 267	119 880	5,03
84	20 678	2 992	0,14466958	0,85533042	19 183	97 613	4,72
85	17 687	2 773	0,15678202	0,84321798	16 300	78 430	4,43
86	14 914	2 529	0,16954409	0,83045591	13 650	62 130	4,17
87	12 385	2 266	0,18295157	0,81704843	11 252	48 480	3,91
88	10 119	1 993	0,19699536	0,80300484	9 123	37 228	3,68
89	8 126	1 720	0,21166145	0,78833855	7 266	28 106	3,46
90	6 406	1 454	0,22693099	0,77306901	5 679	20 840	3,25
91	4 952	1 202	0,24277741	0,75722259	4 351	15 161	3,06
92	3 750	972	0,25916832	0,74083168	3 264	10 809	2,88
93	2 778	767	0,27606561	0,72393439	2 395	7 545	2,72
94	2 011	590	0,29342526	0,70657474	1 716	5 151	2,56
95	1 421	442	0,31119913	0,68880087	1 200	3 435	2,42
96	979	322	0,32933518	0,67066482	818	2 235	2,28
97	656	228	0,34777750	0,65222250	542	1 417	2,16
98	428	157	0,36646673	0,63353327	350	875	2,04
99	271	105	0,38534070	0,61465930	219	525	1,94
100	167	67	0,40433514	0,59566486	133	306	1,84

Quelle: Wirtschaft und Statistik 6 (1991)

Eine ausführliche Diskussion über die statistischen Aspekte der Erstellung und Glättung von Sterbetafeln findet man z.B. in GERBER (1997) oder KAKIES ET AL. (1985).

Die folgenden Graphiken zeigen die Entwicklung von Sterblichkeiten in den letzten 300 Jahren.



Erläuterung:

Halley: Sterbetafel von 1693, basierend auf Bevölkerungsdaten der Stadt Breslau

stm24: Sterbetafel der Jahre 1924/26 (männlich)⁴

stm86: die oben wiedergegebene Sterbetafel der Jahre 1986/88 (männlich)

R94m: Renten-Sterbetafel der Deutschen Aktuarvereinigung (DAV) von 1994 (männlich)⁵

R94w: Renten-Sterbetafel der Deutschen Aktuarvereinigung (DAV) von 1994 (weiblich)⁵

In allen Sterbetafeln fällt ein näherungsweise linearer Verlauf der einjährigen Sterbewahrscheinlichkeiten q_x in logarithmischer Skala im Altersbereich von etwa $30 \leq x \leq 80$ auf. Wegen

$$q_x = P(T_x \leq 1) = \int_0^1 \frac{f(x+t)}{1-F(x)} dt \approx \int_0^1 \frac{f(x)}{1-F(x)} dt = \frac{f(x)}{1-F(x)} = \mu(x)$$

ergibt sich also in diesem Bereich $\mu(x) \approx a e^{bx}$ mit $a, b > 0$, was den Gompertz-Makeham'schen Ansatz in gewisser Weise rechtfertigt.

I. 3. Lebensversicherungsprämien

Der am Anfang des vorigen Abschnitts betrachtete Leibrentenbarwert \ddot{a}_x lässt sich mit Hilfe von Lemma 2 nunmehr recht einfach darstellen, da hier gilt:

$$\begin{aligned} \ddot{a}_x &:= \frac{1 - E(v^{K_x+1})}{1-v} \quad \text{mit} \quad E(v^{K_x+1}) = \sum_{k=0}^{\infty} v^{k+1} P(K_x = k) = \sum_{k=0}^{\infty} v^{k+1} \cdot {}_k p_x q_{x+k} \\ &= \sum_{k=0}^{\infty} v^{k+1} q_{x+k} \prod_{j=0}^{k-1} (1 - q_{x+j}) \end{aligned}$$

⁴ vgl. TOSBERG (1957), S. 8 – 13.

⁵ vgl. MILBRODT UND HELBIG (1999).

(mit der üblichen Konvention, das „leere“ Produkt als 1 zu definieren) und die letzte Größe unter Verwendung von geeigneten Sterbetafeln leicht numerisch ausgewertet werden kann.

Eine alternative Darstellung für den Rentenbarwert erhalten wir, wenn wir die Indikatorfunktionen

$$I_{x,k} = \begin{cases} 1, & \text{falls die versicherte Person im Versicherungsjahr } k \text{ dem Kollektiv noch angehört} \\ 0, & \text{sonst} \end{cases}$$

der Kollektivzugehörigkeit betrachten. Dann ist nämlich $\{I_{x,k} = 1\} = \{K_x \geq k\}$ für $k \in \mathbb{Z}^+$ bzw.

$I_{x,k} = \mathbb{1}_{\{K_x \geq k\}}$ mit

$$\ddot{a}_x = E \left(\sum_{k=0}^{\infty} I_{x,k} v^k \right) = \sum_{k=0}^{\infty} P(I_{x,k} = 1) v^k = \sum_{k=0}^{\infty} P(K_x \geq k) v^k = \sum_{k=0}^{\infty} {}_k p_x v^k$$

(vgl. GERBER (1986), Kapitel 4). Hierbei bezeichnet wie in der Stochastik üblich $\mathbb{1}_M$ die *Indikatorfunktion* einer Menge M , d.h. es ist

$$\mathbb{1}_M(\omega) = \begin{cases} 1, & \text{falls } \omega \in M \\ 0, & \text{falls } \omega \notin M. \end{cases}$$

Beispiel 7 (Fortsetzung von Beispiel 4). Bei Zugrundelegung der Renten-Sterbetafel R94m der DAV (männlich) ergibt sich für die Situation in Beispiel 4 mit dem Alter $x = 65$ – wieder mit dem Rechnungszins $i = 4\%$ – der Wert $\ddot{a}_x = 158558,62 \text{ €}$ ⁶ bei Zugrundelegung der für Versicherungen mit Todesfallcharakter verwendeten Sterbetafel T94m der DAV (männlich) dagegen der kleinere Wert $\ddot{a}_x = 123204,11 \text{ €}$. Mit der abgekürzten Allgemeinen Sterbetafel von 1991 erhält man alternativ den Wert $\ddot{a}_x = 128620,54 \text{ €}$ also einen Wert, der näher an der entsprechenden Größe aus T94m liegt. Dies liegt daran, dass bei versicherungsmathematischen Berechnungen mit Rentencharakter aus Vorsichtsgründen meist eine Sterbetafel verwendet wird, die nicht die tatsächlichen, sondern geeignet angepasste – niedrigere – Sterbewahrscheinlichkeiten enthält, wodurch die Rentenbarwerte und damit auch eventuelle Prämienzahlungen an das VU höher ausfallen. Die Gründe für solche Anpassungen, ohne die ein VU nicht dauerhaft wirtschaften kann, werden wir später im Abschnitt *Prämienkalkulation* genauer untersuchen.

Wir wollen uns jetzt zunächst mit den sogenannten *Netto-Einmalprämien* für drei Grundtypen von Lebensversicherungsverträgen befassen. Diese Grundtypen sind die *Todesfallversicherung* (unbegrenzter oder begrenzter Dauer), die *Erlebensfallversicherung* sowie – als Kombination von beiden – die gemischte oder *Kapital-Lebensversicherung*.

Bei der Todesfallversicherung unbegrenzter Dauer (lebenslange Deckung) zahlt das VU die vereinbarte Versicherungssumme am Ende des Jahres, in dem der VN verstirbt. Bei der Todesfallversicherung begrenzter Dauer (temporäre Deckung) ist das VU nur innerhalb einer vereinbarten Frist zu einer solchen Zahlung verpflichtet. Stirbt der VN nach Ablauf dieser Frist, wird also nichts gezahlt.

⁶ Wir werden später sehen, wie solche Berechnungen mit Hilfe des Programms *leben.exe* oder des Maple-Worksheets *leben.mws* aus der Programmsammlung VMATH durchgeführt werden können.

Die Erlebensfallversicherung ist stets von begrenzter Dauer, da das VU die Versicherungssumme am Ende des Laufzeit -Vertragsjahres auszahlt, sofern der VN dieses erlebt. Bei vorherigem Ableben ist das VU wieder nicht zur Zahlung verpflichtet.

Die gemischte Versicherung vermeidet den Nachteil einer eventuellen Nicht-Zahlung dadurch, dass sie beide Vertragsarten (Todesfallversicherung begrenzter Dauer und Erlebensfallversicherung) in einer einzigen Police kombiniert, d.h. das VU zahlt die Versicherungssumme in jedem Fall aus, entweder bei Versterben vor Ablauf der Frist oder bei Erleben des Laufzeit –Vertragsjahres.

Unter der jeweiligen Netto-Einmalprämie wird der erwartete Barwert der Versicherungsleistung verstanden. Aufgrund des schon erwähnten *Äquivalenzprinzips* ist dies die „faire“ Prämie, d.h. auf lange Sicht und in einem großen Kollektiv von Versicherten mit gleichen Konditionen ist dies aufgrund des Gesetzes der Großen Zahlen der „Kostenanteil“, der entsteht, wenn die Versicherungsleistungen gleichmäßig auf alle VN umgelegt würden. Die aktuariellen Symbole hierfür sowie die entsprechenden Formeln lauten:

Symbol	Bedeutung
$A_x = E\left(v^{K_x+1}\right)$	Netto-Einmalprämie einer Todesfallversicherung der Höhe 1, unbegrenzte Deckung
$A_{x:n}^1 = E\left(v^{K_x+1} \mathbb{1}_{\{K_x < n\}}\right)$	Netto-Einmalprämie einer Todesfallversicherung der Höhe 1, n -jährige Deckung
$A_{x:n}^{\overline{1}} = E\left(v^n \mathbb{1}_{\{K_x \geq n\}}\right) = v^n P(K_x \geq n)$	Netto-Einmalprämie einer Erlebensfallversicherung der Höhe 1
$A_{x:n} = A_{x:n}^1 + A_{x:n}^{\overline{1}}$	Netto-Einmalprämie einer Kapital-Lebensversicherung der Höhe 1

Zur Erklärung der zugehörigen Formeln überlegen wir uns, dass bei einer Todesfallversicherung der Höhe 1 mit unbegrenzter Deckung die Versicherungssumme 1 am Ende des Todesjahres ausbezahlt wird, also $K_x + 1$ Jahre nach Vertragsabschluss. Zur Berechnung des (zufälligen) Barwerts der Versicherungsleistung muss der Betrag 1 also insgesamt $K_x + 1$ -mal diskontiert werden, das ergibt den Wert v^{K_x+1} . Der angegebene Ausdruck für A_x resultiert jetzt aus dem Äquivalenzprinzip als Erwartungswert dieses Barwerts.

Bei der entsprechenden Versicherung mit n -jähriger Deckung ist zu beachten, dass eine Auszahlung nur erfolgt, wenn der VN innerhalb der ersten n Jahre nach Vertragsabschluss verstirbt, also im Fall von $K_x < n$. (Bei $K_x \geq n$ wäre mit Wahrscheinlichkeit 1 $T_x > n$, d.h. der VN würde erst nach Ablauf der Deckungsdauer versterben.)

Bei der Erlebensfallsversicherung ist die Laufzeit n determiniert, d.h. es wird n -mal diskontiert, sofern der VN n Jahre nach Vertragsabschluss überlebt (d.h. im Fall von $K_x \geq n$.)

Alternativ lassen sich diese Einmalprämien natürlich auch mit Hilfe der jährlichen Sterbe- bzw. Überlebenswahrscheinlichkeiten nach Lemma 2 zusammenfassend wie folgt ausdrücken:

Lemma 3. Für die Netto-Einmalprämien der Todesfallversicherung, der Erlebensfallversicherung sowie der Kapital-Lebensversicherung gilt:

$$\begin{aligned}
 A_x &= E\left(v^{K_x+1}\right) = \sum_{k=0}^{\infty} v^{k+1} P(K_x = k) = \sum_{k=0}^{\infty} v^{k+1} \cdot {}_k p_x q_{x+k} \\
 A_{x:n|}^1 &= E\left(v^{K_x+1} \mathbb{1}_{\{K_x < n\}}\right) = \sum_{k=0}^{n-1} v^{k+1} P(K_x = k) = \sum_{k=0}^{n-1} v^{k+1} \cdot {}_k p_x q_{x+k} \\
 A_{x:n|}^{\frac{1}{}} &= E\left(v^n \mathbb{1}_{\{K_x \geq n\}}\right) = v^n P(K_x \geq n) = v^n \cdot {}_n p_x \\
 A_{x:n|} &= A_{x:n|}^1 + A_{x:n|}^{\frac{1}{}} = v^n \cdot {}_n p_x + \sum_{k=0}^{n-1} v^{k+1} \cdot {}_k p_x q_{x+k}.
 \end{aligned}$$

Hierbei ist zu beachten, dass im Falle eines endlichen Endalters ϖ die Summe in der Berechnung von A_x ebenfalls endlich ist, also $A_x = A_{x:n|}^1$ gilt mit $n = \varpi - x$.

Es ist interessant festzustellen, dass die Kenntnis aller Einmalprämien A_x für alle Alter x die Verteilungen der entsprechenden gestutzten Restlebensdauern K_x eindeutig bestimmt und damit auch die anderen Prämien durch A_x allein ausgedrückt werden können. Dies ergibt sich aus dem nachfolgenden

Lemma 4. Es gilt

$$p_x = \frac{1 - r A_x}{1 - A_{x+1}} \quad \text{bzw.} \quad q_x = \frac{r A_x - A_{x+1}}{1 - A_{x+1}} \quad \text{für } x = 0, \dots, \varpi^*$$

sowie

$$\begin{aligned}
 {}_n p_x &= \prod_{k=0}^{n-1} p_{x+k} = \prod_{k=0}^{n-1} \frac{1 - r A_{x+k}}{1 - A_{x+k+1}} \quad \text{und} \\
 A_{x:n|}^1 &= A_x - \frac{v^n A_{x+n}}{{}_n p_x} \quad \text{für } x = 0, \dots, \varpi^* - n, \quad n = 1, \dots, \varpi^*.
 \end{aligned}$$

Beweis: Es ist nach Lemma 2 a)

$$A_{x+1} = \sum_{k=0}^{\infty} v^{k+1} \cdot {}_k p_{x+1} q_{x+k+1} = \frac{1}{v p_x} \sum_{k=0}^{\infty} v^{k+2} \cdot {}_{k+1} p_x q_{x+(k+1)} = \frac{1}{v p_x} (A_x - v q_x) = \frac{A_x - v + v p_x}{v p_x},$$

woraus die ersten beiden Behauptungen folgen. Im zweiten Fall rechnet man analog

$$\sum_{k=n}^{\infty} v^{k+1} \cdot {}_k p_x q_{x+k} = \sum_{k=0}^{\infty} v^{n+1+k} \cdot {}_{n+k} p_x q_{x+n+k} = \frac{v^n}{{}_n p_x} \sum_{k=0}^{\infty} v^{k+1} \cdot {}_k p_{x+n} q_{x+n} = \frac{v^n}{{}_n p_x} A_{x+n},$$

also

$$A_{x:n}^1 = \sum_{k=0}^{n-1} v^{k+1} \cdot {}_k p_x q_{x+k} = A_x - \sum_{k=n}^{\infty} v^{k+1} \cdot {}_k p_x q_{x+k} = A_x - \frac{v^n}{{}_n p_x} A_{x+n}.$$

In der Versicherungspraxis sind Einmalprämien allerdings wegen der hohen finanziellen Belastung des VN unüblich. Statt dessen werden in der Regel *laufende Prämien* gleicher Höhe vereinbart, und zwar jährlich oder monatlich nach vorschüssiger Zahlweise, entweder bis zum Tod oder bis zum Ende der Vertragslaufzeit. Wir beschränken uns hier auf den Fall jährlicher Zahlungen, für die anderen Varianten – die u.a. unterjährige Sterbetafeln erfordern – verweisen wir auf GERBER (1986 und 1997).

Nach dem Äquivalenzprinzip ist die jährliche vorschüssige Prämie jeweils so zu bestimmen, dass die erwarteten Barwerte der Einzahlungen und der Versicherungsleistung (Auszahlung) gleich ausfallen. Für eine Todesfallversicherung unbegrenzter Deckung bedeutet das zunächst die Bestimmung des Leibrentenbarwerts \ddot{a}_x , der bereits zu Beginn dieses Abschnitts dargestellt wurde und mit Lemma 3 geschrieben werden kann als

$$\ddot{a}_x = E\left(\ddot{a}_{\overline{K_x+1}|}\right) = \frac{1 - E\left(v^{K_x+1}\right)}{1 - v} = \frac{1 - A_x}{d}.$$

Bezeichnen wir mit P_x die entsprechende jährliche Prämie für diese Versicherung, muss also die Gleichung

$$P_x \cdot \ddot{a}_x = A_x$$

gelten, woraus

$$P_x = \frac{A_x}{\ddot{a}_x} = \frac{d A_x}{1 - A_x}$$

folgt. Ähnlich lassen sich auch die übrigen jährlichen Prämien darstellen. Für den Barwert einer vorschüssigen Rente bis zum Tod, längstens n Jahre erhält man entsprechend:

$$\begin{aligned} \ddot{a}_{x:n} &:= E\left(\ddot{a}_{\overline{\min\{K_x+1, n\}}|}\right) = \frac{1 - E\left(v^{\min\{K_x+1, n\}}\right)}{1 - v} = \frac{1 - E\left(v^{K_x+1} \mathbb{1}_{\{K_x < n\}} + v^n \mathbb{1}_{\{K_x \geq n\}}\right)}{1 - v} \\ &= \frac{1 - \left(A_{x:n}^1 + A_{x:n}^1\right)}{1 - v} = \frac{1 - A_{x:n}^1}{d}. \end{aligned}$$

Dementsprechend bestimmen sich die jährlichen Prämien $P_{x:n}^1$ für die temporäre Todesfallversicherung und $P_{x:n}^{\overline{1}}$ für die Erlebensfallversicherung nach dem Äquivalenzprinzip aus den Gleichungen

$$P_{x:n}^1 \cdot \ddot{a}_{x:n} = A_{x:n}^1 \quad \text{und} \quad P_{x:n}^{\overline{1}} \cdot \ddot{a}_{x:n} = A_{x:n}^{\overline{1}}.$$

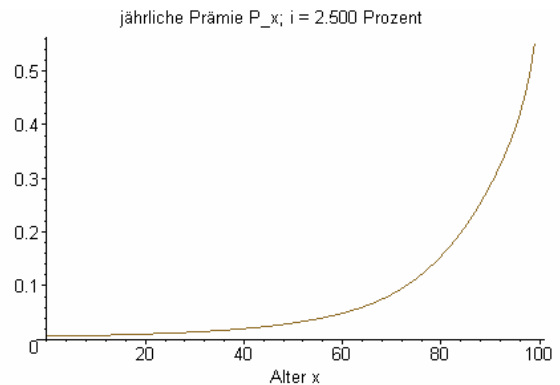
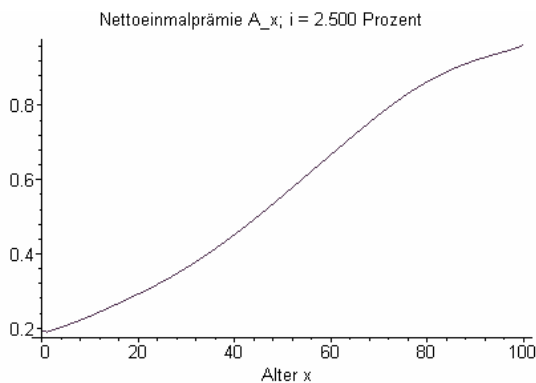
Die jährliche Prämie für die Kapital-Lebensversicherung ergibt sich hieraus wieder durch Addition.

I Personenversicherungsmathematik

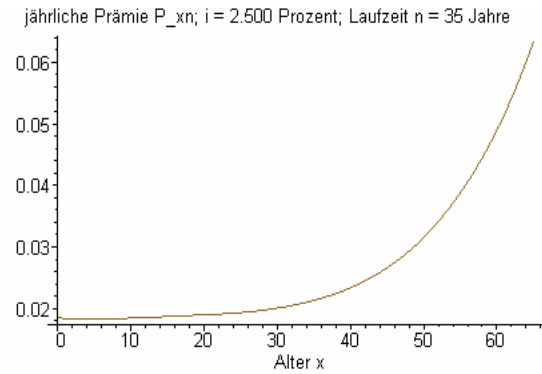
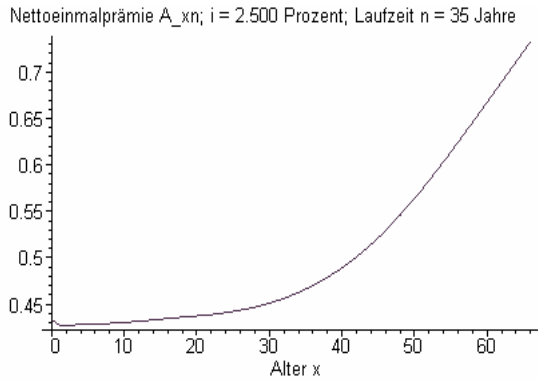
Alle relevanten Größen sind exemplarisch noch einmal in der folgenden Tabelle zusammengefasst:

Symbol	Bedeutung
$\ddot{a}_x = \frac{1 - A_x}{1 - v} = \frac{1 - A_x}{d}$	erwarteter Rentenbarwert der Höhe 1 bei lebenslanger Zahlung
$\ddot{a}_{x:n} = \frac{1 - A_{x:n}}{1 - v} = \frac{1 - A_{x:n}}{d}$	erwarteter Rentenbarwert der Höhe 1 bei maximal n Jahre dauernder Zahlung
$P_x = \frac{A_x}{\ddot{a}_x} = \frac{d A_x}{1 - A_x}$	jährliche Netto-Prämie einer Todesfallversicherung der Höhe 1, unbegrenzte Deckung
$P_{x:n}^1 = \frac{A_{x:n}^1}{\ddot{a}_{x:n}} = \frac{d A_{x:n}^1}{1 - A_{x:n}^1}$	jährliche Netto-Prämie einer Todesfallversicherung der Höhe 1, n -jährige Deckung
$P_{x:n}^{\overline{1}} = \frac{A_{x:n}^{\overline{1}}}{\ddot{a}_{x:n}} = \frac{d A_{x:n}^{\overline{1}}}{1 - A_{x:n}^{\overline{1}}}$	jährliche Netto-Prämie einer Erlebensfallversicherung der Höhe 1
$P_{x:n} = P_{x:n}^1 + P_{x:n}^{\overline{1}} = \frac{A_{x:n}}{\ddot{a}_{x:n}} = \frac{d A_{x:n}}{1 - A_{x:n}}$	jährliche Netto-Prämie einer Kapital-Lebensversicherung der Höhe 1

Die folgenden Graphiken zeigen den Verlauf einiger Prämien für die Versicherungssumme 1 in Abhängigkeit vom Alter x unter Zugrundelegung der Sterbetafel T94m der DAV (erstellt mit dem Maple-Worksheet `leben.mws`).



I Personenversicherungsmathematik



Mit Hilfe des Programms `leben.exe` oder des Maple-Worksheets `leben.mws` lassen sich damit auch Leibrentenbarwerte wie in Beispiel 7 berechnen, weil natürlich auch die umgekehrten Beziehungen

$$\ddot{a}_x = \frac{A_x}{P_x} \quad \text{und} \quad \ddot{a}_{:xn} = \frac{A_{:xn}^1}{P_{:xn}^1} = \frac{A_{:xn}}{P_{:xn}}$$

gelten. Für die Rechnungen in Beispiel 7 erhält man etwa mit der Rententafel DAV R94m

$$\ddot{a}_x = \frac{A_x}{P_x} = \frac{0,491796}{0,03722} = 13,21225 \quad (\text{für die Rentenhöhe } 1),$$

was nach Multiplikation mit 12000 € zu dem angegebenen Wert führt.

Ähnlich wie oben lassen sich auch aus den jährlichen Netto-Prämien die Verteilungen der gestutzten Restlebensdauern K_x eindeutig rekonstruieren:

Lemma 5. Es gilt

$$p_x = (1 - r P_x) \frac{d + P_{x+1}}{d + P_x} \quad \text{bzw.} \quad q_x = \frac{r P_x (1 + P_{x+1}) - P_{x+1}}{d + P_x} \quad \text{für } x = 0, \dots, \varpi^*$$

sowie

$${}_n p_x = \prod_{k=0}^{n-1} p_{x+k} = \prod_{k=0}^{n-1} (1 - r P_{x+k}) \frac{d + P_{x+k+1}}{d + P_{x+k}} \quad \text{für } x = 0, \dots, \varpi^* - n, \quad n = 1, \dots, \varpi^*.$$

Beweis: Mit Lemma 4 folgt wegen $P_x = \frac{d A_x}{1 - A_x}$, also $A_x = \frac{P_x}{d + P_x}$ sofort

$$p_x = \frac{1 - r A_x}{1 - A_{x+1}} = \frac{1 - r \frac{P_x}{d + P_x}}{1 - \frac{P_{x+1}}{d + P_{x+1}}} = \frac{d + P_x - r P_x}{d} \cdot \frac{d + P_{x+1}}{d + P_x} = \left(1 - \frac{i}{1 - v} P_x\right) \frac{d + P_{x+1}}{d + P_x} = (1 - r P_x) \frac{d + P_{x+1}}{d + P_x}$$

und

$$q_x = 1 - p_x = 1 - (1 - r P_x) \frac{d + P_{x+1}}{d + P_x} = \frac{d + P_x - (1 - r P_x)(d + P_{x+1})}{d + P_x}$$

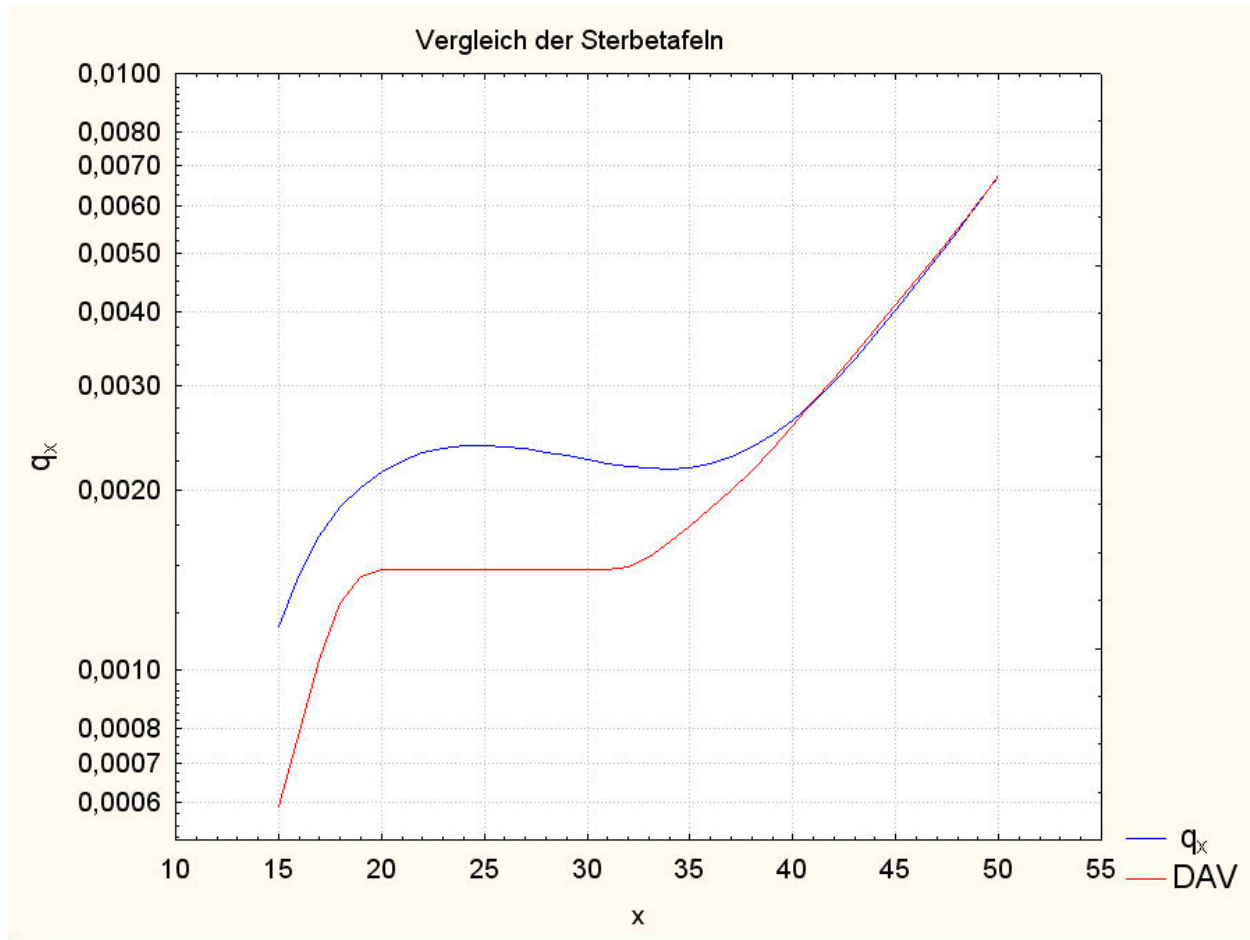
$$= \frac{d + P_x - d - P_{x+1} + i P_x + r P_x P_{x+1}}{d + P_x} = \frac{r P_x (1 + P_{x+1}) - P_{x+1}}{d + P_x}$$

wie behauptet.

Die nachfolgende Tabelle enthält jährliche Netto-Prämien P_x für eine Todesfallversicherung unbegrenzter Deckung für einen männlichen VN in Anlehnung an einen realen Tarif sowie die daraus nach Lemma 5 abgeleiteten einjährigen Sterbewahrscheinlichkeiten q_x , zum Vergleich die entsprechenden Sterbewahrscheinlichkeiten der DAV-Tafel T94m (Rechnungszins 2%). (Der Unterschied in jüngeren Jahren stammt aus einer zusätzlichen Deckung für eine Unfallversicherung, insofern geben die q_x strenggenommen dort nicht exakt die Sterbewahrscheinlichkeiten wieder.)

Alter x	P_x	q_x	q_x (DAV)
15	0,009715	0,001181081	0,000593
16	0,009973	0,001452549	0,000792
17	0,010234	0,001682537	0,001040
18	0,010498	0,001873983	0,001298
19	0,010766	0,002029812	0,001437
20	0,011041	0,002152951	0,001476
21	0,011324	0,002246343	0,001476
22	0,011615	0,002312964	0,001476
23	0,011916	0,002355827	0,001476
24	0,012229	0,002377996	0,001476
25	0,012554	0,002382591	0,001476
26	0,012894	0,002372790	0,001476
27	0,013249	0,002351833	0,001476
28	0,013621	0,002323023	0,001476
29	0,014011	0,002289720	0,001476
30	0,014420	0,002255345	0,001476
31	0,014850	0,002223367	0,001476
32	0,015302	0,002197298	0,001489

Alter x	P_x	q_x	q_x (DAV)
33	0,015778	0,002180689	0,001551
34	0,016278	0,002177114	0,001641
35	0,016804	0,002190161	0,001747
36	0,017358	0,002223423	0,001869
37	0,017941	0,002280482	0,002007
38	0,018554	0,002364897	0,002167
39	0,019198	0,002480192	0,002354
40	0,019875	0,002629843	0,002569
41	0,020586	0,002817265	0,002832
42	0,021332	0,003045798	0,003087
43	0,022115	0,003318700	0,003387
44	0,022936	0,003639135	0,003726
45	0,023797	0,004010162	0,004100
46	0,024699	0,004434732	0,004522
47	0,025642	0,004915674	0,004983
48	0,026629	0,005455700	0,005508
49	0,027661	0,006057392	0,006049
50	0,028738	0,006723205	0,006751



Grundsätzlich lassen sich alle Versicherungen auch mit zeitabhängigen Versicherungssummen gestalten. Wir wollen das hier exemplarisch nur für die Todesfallversicherung temporärer Dauer behandeln, da solche Policen z.B. bei der Restschuldversicherung für eine Hypothek (vgl. Beispiel 3) eine Rolle spielen. Für eine Darstellung in voller Allgemeinheit verweisen wir auf GERBER (1986 und 1997).

Wenn wir vereinbaren, dass am Ende des Jahres $k = 0, \dots, n-1$ nach Abschluss der Versicherung die Versicherungssumme c_{k+1} durch das VU ausbezahlt wird, wenn der Versicherte im Alter $x+k$ stirbt, d.h. wenn $K_x = k$ gilt, erhalten wir als erwarteten Barwert der Versicherungsleistung in Analogie zum vorherigen Fall die Größe

$$A_{x:n|}^1(\mathbf{c}) = E\left(c_{K_x+1}v^{K_x+1}\right) = \sum_{k=0}^{n-1} c_{k+1}v^{k+1}P(K_x = k) = \sum_{k=0}^{n-1} c_{k+1}v^{k+1} \cdot {}_k p_x q_{x+k}.$$

Eine besonders einfache Form nimmt dieser Ausdruck für den Fall einer geometrischen Progression an, d.h. wenn

$$c_k = \beta^k \text{ gilt für } k = 1, \dots, n \text{ mit } \beta > 0.$$

Wir erhalten dann

$$A_{x:n|}^1(\beta) = \sum_{k=0}^{n-1} \beta^{k+1}v^{k+1} \cdot {}_k p_x q_{x+k} = \sum_{k=0}^{n-1} (\beta v)^{k+1} \cdot {}_k p_x q_{x+k},$$

also den Barwert einer modifizierten temporären Todesfallversicherung mit der Versicherungssumme 1, aber einem anderen Rechnungszins i_β , der sich aus der Gleichung

$$v_\beta = \frac{1}{1+i_\beta} = \beta v = \frac{\beta}{1+i} \text{ mit der Lösung } i_\beta = \frac{1+i-\beta}{\beta}$$

ergibt. Theoretisch kann dieser „Zins“ rechnerisch sogar negativ ausfallen. Jährliche (konstante) Prämien für solche Versicherungen lassen sich analog dem obigen Vorgehen bestimmen, etwa bei der temporären Todesfallversicherung als

$$P_{x:n}^1(\mathbf{c}) = \frac{A_{x:n}^1(\mathbf{c})}{\ddot{a}_{x:n}} = \frac{d A_{x:n}^1(\mathbf{c})}{1 - A_{x:n}^1}.$$

Beispiel 8 (Ergänzung zu Beispiel 3). Wir wollen die jährliche (konstante) Prämie für eine *Restschuldversicherung* für einen 20-jährigen Mann bestimmen, die im Falle des vorzeitigen Todes zum Jahresende die Restschuld der Hypothek nebst den bis dahin anfallenden Zinsen ablöst. Da bei der Annuitätentilgung die Zahlungen des Schuldners stets zum Jahresende erfolgen und dies bei vorzeitigem Tod im Jahr des Versterbens nicht mehr möglich ist, beläuft sich die *zu versichernde Restschuld* HV_k für das Todesjahr auf den Betrag

$$HV_k = r \frac{r^n - r^k}{r^n - 1} H = \frac{r^{n+1} - r^{k+1}}{r^n - 1} H, \quad k = 0, \dots, n-1.$$

Hierbei bezieht sich $r = 1 + i_H$ auf den Hypothekenzins i_H . Die jährliche (vorschüssige) Prämie P für die entsprechende Restschuldversicherung bestimmt sich nach obiger Rechnung also zu

$$P = H \frac{r^{n+1} A_{x:n}^1 - A_{x:n}^1(\beta)}{r^n - 1} \cdot \frac{d_v}{1 - A_{x:n}^1} \quad \text{mit} \quad v = \frac{1}{1+i_v}, \quad d_v = 1 - \frac{1}{1+i_v} = \frac{i_v}{1+i_v},$$

wobei i_v den Rechnungszins für die Versicherung bezeichnet, der üblicherweise vom Hypothekenzins i_H verschieden ist. Der ersatzweise Zins i_β bestimmt sich hier also zu

$$i_\beta = \frac{i_v - i_H}{1 + i_H},$$

der in der Praxis fast immer negativ ist. Wählen wir als Rechnungszins für die Versicherung $i_v = 2,5\%$, so ergeben sich unter den Konditionen von Beispiel 3 unter Verwendung der DAV-Sterbetafel T94m folgende Größen (angegeben ist das jeweilige Ergebnis nach `leben.exe`):

I Personenversicherungsmathematik

```
File mit Sterbewahrscheinlichkeiten [ohne .QX:'q' für quit]? dau94t_n
Zinssatz i (in %)? 2.5
Alter x? 20
Laufzeit n? 10 [ maximal omega-x = 81 ]
a_x = .291574 [Todesfall, lebenslange Deckung]
P_x = .010039
a1_xn= .012836 [Todesfall, temporäre Deckung]
P1_xn= .00144
a_xn1= .769744 [Erlebensfall]
P_xn1= .08635
a_xn = .782581 [gemischte Versicherung]
P_xn = .08779
Optionen:
's': Speichern der Ergebnisse in Datei [File].[Nummer]
'c': neu, 'x': neues Alter, 'n': neue Laufzeit, 'i': neuer Zinssatz
Graphik:
F1: Sterbewahrscheinlichkeiten, F2: restliche Lebenserwartung
F3: Deckungskapital gemischte, F4: temporäre Todesfallversicherung
F5: Tabellen, F6: Kurzanleitung
```

```
File mit Sterbewahrscheinlichkeiten [ohne .QX:'q' für quit]? dau94t_n
Zinssatz i (in %)? -1.44231
Alter x? 20
Laufzeit n? 10 [ maximal omega-x = 81 ]
a_x = 2.185182 [Todesfall, lebenslange Deckung]
P_x = .026982
a1_xn= .015893 [Todesfall, temporäre Deckung]
P1_xn= .001498
a_xn1= 1.139409 [Erlebensfall]
P_xn1= .107367
a_xn = 1.155302 [gemischte Versicherung]
P_xn = .108864
Optionen:
's': Speichern der Ergebnisse in Datei [File].[Nummer]
'c': neu, 'x': neues Alter, 'n': neue Laufzeit, 'i': neuer Zinssatz
Graphik:
F1: Sterbewahrscheinlichkeiten, F2: restliche Lebenserwartung
F3: Deckungskapital gemischte, F4: temporäre Todesfallversicherung
F5: Tabellen, F6: Kurzanleitung
```

$$i_\beta = \frac{i_V - i_H}{1 + i_H} = -\frac{3}{208} = -0,0144231, \quad A_{xn}^1 = 0,012836, \quad A_{xn} = 0,782581, \quad A_{xn}^1(\beta) = 0,015893,$$

woraus $P = 90,34 \text{ €}$ folgt im Gegensatz zu der einfachen Hypothekenversicherung (über die volle Summe inklusive Anfangszinsen) in Höhe von $100000 r \cdot P_{xn}^1 = 149,76 \text{ €}$ Für einen 40-jährigen Mann bei 20-jähriger Laufzeit hätte sich analog ergeben:

```
File mit Sterbewahrscheinlichkeiten [ohne .QX:'q' für quit]? dau94t_n
Zinssatz i (in %)? 2.5
Alter x? 40
Laufzeit n? 20 [ maximal omega-x = 61 ]
a_x = .45197 [Todesfall, lebenslange Deckung]
P_x = .020115
a1_xn= .10195 [Todesfall, temporäre Deckung]
P1_xn= .006656
a_xn1= .524486 [Erlebensfall]
P_xn1= .034244
a_xn = .626435 [gemischte Versicherung]
P_xn = .04098
Optionen:
's': Speichern der Ergebnisse in Datei [File].[Nummer]
'c': neu, 'x': neues Alter, 'n': neue Laufzeit, 'i': neuer Zinssatz
Graphik:
F1: Sterbewahrscheinlichkeiten, F2: restliche Lebenserwartung
F3: Deckungskapital gemischte, F4: temporäre Todesfallversicherung
F5: Tabellen, F6: Kurzanleitung
```

```
File mit Sterbewahrscheinlichkeiten [ohne .QX:'q' für quit]? dau94t_n
Zinssatz i (in %)? -1.44231
Alter x? 40
Laufzeit n? 20 [ maximal omega-x = 61 ]
a_x = 1.659045 [Todesfall, lebenslange Deckung]
P_x = .036839
a1_xn= .17119 [Todesfall, temporäre Deckung]
P1_xn= .007819
a_xn1= 1.149212 [Erlebensfall]
P_xn1= .052489
a_xn = 1.320402 [gemischte Versicherung]
P_xn = .060308
Optionen:
's': Speichern der Ergebnisse in Datei [File].[Nummer]
'c': neu, 'x': neues Alter, 'n': neue Laufzeit, 'i': neuer Zinssatz
Graphik:
F1: Sterbewahrscheinlichkeiten, F2: restliche Lebenserwartung
F3: Deckungskapital gemischte, F4: temporäre Todesfallversicherung
F5: Tabellen, F6: Kurzanleitung
```

also

$$A_{xn}^1 = 0,10195, \quad A_{xn} = 0,626435, \quad A_{xn}^1(\beta) = 0,17119, \quad P = 562,40 \text{ €}, \quad 100000 r \cdot P_{xn}^1 = 692,22 \text{ €}$$

I. 4. Deckungsrückstellungen

Werden bei einer Lebensversicherung jährliche Prämienzahlungen vereinbart, besteht in der Regel nur zu Beginn der Versicherung eine Äquivalenz im Sinne der Gleichheit der erwarteten Barwerte der künftigen Versicherungsleistung und der künftigen Prämienzahlungen. In späteren Jahren verschiebt sich dieses Gleichgewicht, und zwar zu Gunsten des VU, wenn die Sterblichkeit des VN mit der Zeit zunimmt. Die Differenz der genannten erwarteten Barwerte heißt das *Deckungskapital* (Deckungsrückstellung) des Versicherungsvertrags und wird üblicherweise mit dem Symbol V bezeichnet. Das Deckungskapital „gehört“ eigentlich dem VN, das VU muss diesen Betrag also für die Finanzierung der zukünftigen Versicherungsleistungen reservieren. Die folgende Tabelle enthält eine Übersicht über die konkreten Bezeichnungen für die geläufigen Versicherungsarten und ihre formelmäßigen Zusammenhänge.

Symbol	Bedeutung
${}_kV_x = A_{x+k} - P_x \ddot{a}_{x+k}, k = 0, 1, \dots$	Deckungskapital zu Beginn des Jahres k der Versicherung, Todesfallversicherung mit unbegrenzter Deckung
${}_kV_{x:n}^1 = A_{x+k;n-k}^1 - P_{x:n}^1 \ddot{a}_{x+k;n-k}, k = 0, \dots, n-1$	Deckungskapital zu Beginn des Jahres k der Versicherung, Todesfallversicherung mit n -jähriger Deckung
${}_kV_{x:n}^{\overline{1}} = A_{x+k;n-k}^{\overline{1}} - P_{x:n}^{\overline{1}} \ddot{a}_{x+k;n-k}, k = 0, \dots, n-1$	Deckungskapital zu Beginn des Jahres k der Versicherung, Erlebensfallversicherung mit n -jähriger Dauer
${}_kV_{x:n} = A_{x+k;n-k} - P_{x:n} \ddot{a}_{x+k;n-k}, k = 0, \dots, n-1$	Deckungskapital zu Beginn des Jahres k einer Kapital-Lebensversicherung der Höhe 1

Aufgrund der Definition der jährlichen Prämien ist hier stets ${}_0V = 0$ für alle Vertragsarten; entsprechend wird ${}_nV_{x:n}^1 = 0$ und ${}_nV_{x:n}^{\overline{1}} = {}_nV_{x:n} = 1$ gesetzt, da entweder am Ende der Versicherung weder Leistungen noch Zahlungen fällig sind (temporäre Todesfallversicherung) oder aber die Versicherungssumme 1 ausbezahlt wird (Erlebensfall- bzw. Kapital-Lebensversicherung).

Es ist auch möglich, aufgrund der Definition der jährlichen Prämien die Deckungskapitalien ohne die Leibrentenbarwerte darzustellen. Es ergibt sich dann:

$$\begin{aligned}
 {}_kV_x &= A_{x+k} \left(1 - \frac{P_x}{P_{x+k}} \right), \quad k = 0, 1, \dots, \\
 {}_kV_{x:n|}^1 &= A_{x+k;n-k}^1 \left(1 - \frac{P_{x:n|}^1}{P_{x+k;n-k}^1} \right), \quad k = 0, \dots, n-1 \\
 {}_kV_{x:n|}^{\overline{1}} &= A_{x+k;n-k}^{\overline{1}} \left(1 - \frac{P_{x:n|}^{\overline{1}}}{P_{x+k;n-k}^{\overline{1}}} \right), \quad k = 0, \dots, n-1 \\
 {}_kV_{x:n|} &= A_{x+k;n-k} \left(1 - \frac{P_{x:n|}}{P_{x+k;n-k}} \right) = {}_kV_{x:n|}^1 + {}_kV_{x:n|}^{\overline{1}}, \quad k = 0, \dots, n-1.
 \end{aligned}$$

Beispiel 9 (Ergänzung zu Beispiel 5). Wenn wir annehmen, dass die Lebensdauer einer Person einer Exponentialverteilung $P^{T_0} = \mathcal{E}(\lambda)$ folgt, d.h. die gestutzte Lebensdauer K_x einer x -jährigen Person – unabhängig vom Alter – geometrisch verteilt ist mit Parameter $p = 1 - e^{-\lambda}$, so folgt

$$A_x = E(v^{K_x+1}) = vp \sum_{k=0}^{\infty} (vq)^k = \frac{vp}{1-vq}, \quad \ddot{a}_x = \frac{1-A_x}{1-v} = \frac{1}{1-vq}, \quad P_x = \frac{A_x}{\ddot{a}_x} = vp$$

und damit

$${}_kV_x = A_{x+k} - P_x \ddot{a}_{x+k} = 0, \quad k = 0, 1, \dots,$$

d.h. der Verlauf der Versicherung ist zu jedem Zeitpunkt ausgeglichen. Ähnlich folgt

$$A_{x:n|}^1 = vp \sum_{k=0}^{n-1} (vq)^k = \frac{vp(1-(vq)^n)}{1-vq}, \quad A_{x:n|}^{\overline{1}} = (vq)^n, \quad A_{x:n|} = A_{x:n|}^1 + A_{x:n|}^{\overline{1}} = \frac{vp + d(vq)^n}{1-vq}$$

und somit

$$\begin{aligned}
 P_{x:n|}^1 &= \frac{A_{x:n|}^1}{\ddot{a}_{x:n|}^1} = \frac{dA_{x:n|}^1}{1-A_{x:n|}^1} = vp = P_x, \\
 P_{x:n|}^{\overline{1}} &= \frac{A_{x:n|}^{\overline{1}}}{\ddot{a}_{x:n|}^{\overline{1}}} = \frac{dA_{x:n|}^{\overline{1}}}{1-A_{x:n|}^{\overline{1}}} = \frac{(1-vq)(vq)^n}{1-(vq)^n}, \\
 P_{x:n|} &= \frac{dA_{x:n|}}{1-A_{x:n|}} = \frac{vp + d(vq)^n}{1-(vq)^n}.
 \end{aligned}$$

Man beachte, dass erwartungsgemäß die Prämien für die temporäre und die unbegrenzte Todesfallversicherung hier identisch sind. Für die übrigen Deckungskapitalien erhält man:

$$\begin{aligned}
 {}_kV_{x:n}^1 &= A_{x+k;n-k}^1 \left(1 - \frac{P_{x:n}^1}{P_{x+k;n-k}^1} \right) = 0, \quad k = 0, \dots, n-1 \\
 {}_kV_{x:n}^{\frac{1}{2}} &= A_{x+k;n-k}^{\frac{1}{2}} \left(1 - \frac{P_{x:n}^{\frac{1}{2}}}{P_{x+k;n-k}^{\frac{1}{2}}} \right) = (vq)^{n-k} \frac{1 - (vq)^k}{1 - (vq)^n}, \quad k = 0, \dots, n-1 \\
 {}_kV_{x:n} &= {}_kV_{x:n}^1 + {}_kV_{x:n}^{\frac{1}{2}} = {}_kV_{x:n}^{\frac{1}{2}} = (vq)^{n-k} \frac{1 - (vq)^k}{1 - (vq)^n}, \quad k = 0, \dots, n-1.
 \end{aligned}$$

Die Bedeutung von Deckungskapitalien kann auch noch auf andere Weise veranschaulicht werden. Wir werden dies hier explizit jedoch nur für eine Todesfallversicherung darstellen, für den allgemeinen Fall verweisen wir auf GERBER (1986 oder 1997).

Lemma 6. Die jährliche (konstante) Prämie P_x einer Todesfallversicherung unbegrenzter Dauer kann durch

$$P_x = {}_kP_x^s + {}_kP_x^r, \quad k = 0, 1, \dots$$

für jedes Versicherungsjahr k zerlegt werden in eine *Sparprämie* ${}_kP_x^s$ und eine *Risikoprämie* ${}_kP_x^r$ mit den Darstellungen

$${}_kP_x^s = v \cdot {}_{k+1}V_x - {}_kV_x \quad \text{und} \quad {}_kP_x^r = (1 - {}_{k+1}V_x) \cdot vq_{x+k}.$$

Beweis. Aus dem Beweis zu Lemma 4 ergibt sich durch Ersetzung von x durch $x+k$ sofort auch die Beziehung

$$A_{x+k+1} = \frac{A_{x+k} - vq_{x+k}}{vp_{x+k}} \quad \text{bzw.} \quad vp_{x+k} A_{x+k+1} - A_{x+k} + vq_{x+k} = 0.$$

Ähnlich ergibt sich

$$\ddot{a}_{x+k+1} = \frac{1 - A_{x+k+1}}{d} = \frac{v - A_{x+k}}{dvp_{x+k}}$$

und somit

$$\ddot{a}_{x+k} - vp_{x+k} \ddot{a}_{x+k+1} = \frac{1 - A_{x+k}}{d} - \frac{v - A_{x+k}}{d} = \frac{1 - v}{d} = 1.$$

Insgesamt erhält man nun

$$\begin{aligned}
 {}_k P_x^s + {}_k P_x^r &= v \cdot {}_{k+1} V_x - {}_k V_x + (1 - {}_{k+1} V_x) \cdot v q_{x+k} = v(1 - q_{x+k}) \cdot {}_{k+1} V_x - {}_k V_x + v q_{x+k} \\
 &= v p_{x+k} \cdot {}_{k+1} V_x - {}_k V_x + v q_{x+k} = v p_{x+k} \cdot (A_{x+k+1} - P_x \ddot{a}_{x+k+1}) - (A_{x+k} - P_x \ddot{a}_{x+k}) + v q_{x+k} \\
 &= v p_{x+k} A_{x+k+1} - A_{x+k} + v q_{x+k} - P_x (v p_{x+k} \ddot{a}_{x+k+1} - \ddot{a}_{x+k}) = P_x,
 \end{aligned}$$

wie behauptet.

Die Sparprämie ${}_k P_x^s$ und die Risikoprämie ${}_k P_x^r$ haben folgende versicherungstechnische Interpretation:

Wenn der VN das Versicherungsjahr k überlebt, wird die Versicherung fortgesetzt. Zur Finanzierung des Fehlbedarfs zwischen der künftigen erwarteten Versicherungsleistung und den künftigen erwarteten (konstanten) Prämieinnahmen muss also das Deckungskapital ${}_{k+1} V_x$ zu Beginn des Jahres $k+1$ (d.h. zum Ende des Jahres k) reserviert werden; sein Barwert zu Beginn des Jahres k beträgt dann $v \cdot {}_{k+1} V_x$. Für den Fall des Versterbens des VN innerhalb des Jahres k muss zu Beginn des Jahres k der Betrag v (gewichtet mit der einjährigen Sterbewahrscheinlichkeit q_{x+k}) reserviert werden; der Unterschiedsbetrag ist also gerade die Risikoprämie ${}_k P_x^r = (1 - {}_{k+1} V_x) \cdot v q_{x+k}$. Die Sparprämie ${}_k P_x^s$ stellt dann die Differenz aus der konstanten Prämie P_x und der Risikoprämie ${}_k P_x^r$ dar. Ferner gilt hier

$$\sum_{j=0}^{k-1} r^{k-j} {}_j P_x^s = \sum_{j=0}^{k-1} (r^{k-j-1} \cdot {}_{j+1} V_x - r^{k-j} \cdot {}_j V_x) = {}_k V_x - r^k {}_0 V_x = {}_k V_x,$$

d.h. die akkumulierten und verzinsten Sparprämien ergeben genau die entsprechenden Deckungskapitalien. Dies zeigt noch einmal aus anderer Sicht, warum dem VN die Deckungsrückstellungen „gehören“, da sie diejenigen Anteile aus der konstanten Prämienzahlung darstellen, die nicht zur Abdeckung des aktuellen (auf das Prämienjahr bezogenen) Risikos gebraucht werden. Sie können also – zumindest zu Beginn der Versicherung – als eine „Vorauszahlung“ auf das spätere erhöhte Sterblichkeitsrisiko betrachtet werden.

Würde man die Versicherung statt mit konstanten jährlichen Prämien mit einer variablen, dem sich ändernden Sterblichkeitsrisiko genau angemessenen Risikoprämie ${}_k P_x^r$ als jährliche Prämienzahlung vereinbaren, so wären gerade alle künftigen Deckungskapitalien Null:

$${}_k V_x^r := A_{x+k} - E \left(\sum_{j=0}^{K_x} v^j \cdot {}_{k+j} P_x^r \right) = A_{x+k} - E \left(\sum_{j=0}^{\infty} v^j \cdot {}_{k+j} P_x^r \cdot \mathbb{1}_{\{K_x \geq j\}} \right) = A_{x+k} - \sum_{j=0}^{\infty} v^j \cdot {}_{k+j} P_x^r \cdot {}_j p_{x+k} = 0$$

(wie im Fall der geometrischen Verteilung in Beispiel 8). Betrachtet man zwei dieser Gleichungen für die Jahre k und $k+1$, so ergibt sich aus der zweiten Gleichung

$$v p_{x+k} A_{x+k+1} = v p_{x+k} \sum_{j=0}^{\infty} v^j \cdot {}_{k+1+j} P_x^r \cdot {}_j p_{x+k+1} = \sum_{j=0}^{\infty} v^{j+1} \cdot {}_{k+1+j} P_x^r \cdot {}_{j+1} p_{x+k} = A_{x+k} - {}_k P_x^r,$$

also unter Beachtung der vorigen Rechnungen (vgl. den Beweis zu Lemma 6)

I Personenversicherungsmathematik

$${}_k P_x^r = A_{x+k} - v p_{x+k} A_{x+k+1} = v q_{x+k},$$

d.h. gerade die (übliche) Prämie für eine einjährige Todesfallversicherung im Alter $x+k$. Dies deckt sich mit der Formel ${}_k P_x^r = (1 - {}_{k+1}V_x) \cdot v q_{x+k}$ für den hier betrachteten Fall, dass die Deckungskapitalien sämtlich verschwinden. Bei jährlich konstanter Prämie ergibt sich die Differenz zwischen dieser „reinen“ Risikoprämie und der anteiligen Risikoprämie in Höhe von $v \cdot {}_{k+1}V_x \cdot q_{x+k}$ also deswegen, weil die im Laufe der Zeit verzinslich akkumulierten Sparanteile (also die Deckungskapitalien) zur Abdeckung des aktuellen jährlichen Risikos herangezogen werden können.

Analoge Zerlegungen der Deckungskapitalien in Spar- und Risikoprämie sind auch für die anderen Vertragsarten möglich, vgl. GERBER (1986 oder 1997).

Die Deckungsrückstellungen spielen bei zwei Aspekten des Versicherungsvertrags eine besondere Rolle, zum einen bei vorzeitiger Kündigung durch den VN, zum anderen bei der Umwandlung einer bestehenden Versicherung in eine künftig beitragsfreie Versicherung (mit reduzierter Versicherungssumme), falls der VN die Prämienzahlungen nicht mehr aufbringen kann oder will.

Bei der vorzeitigen Kündigung zum Anfang des Versicherungsjahres k hat nach den obigen Ausführungen der VN nur Anspruch auf das aktuelle Deckungskapital ${}_k V$, weil dieses den nicht für die jährlichen Risiken verbrauchten verzinslich akkumulierten Sparanteilen („Vorauszahlungen“) entspricht. Man spricht in diesem Zusammenhang auch vom *Rückkaufswert* der Versicherung. In der Praxis wird allerdings nicht das volle rechnerische Deckungskapital ausgezahlt, sondern eine um Stornokosten und eventuelle Provisionsanteile für den Vertragsabschluss reduzierte Summe. Deshalb sind Rückkaufswerte vor allem bei Kapital-Lebensversicherungen in der Regel erst ab dem zweiten Versicherungsjahr positiv.

Bei der Umwandlung in eine beitragsfreie Versicherung wird das bis zum Stichtag erreichte Deckungskapital ${}_k V$ als Einmalprämie für die neue Versicherung (mit Restlaufzeit $n-k$ Jahre) verwendet. Dementsprechend lässt sich die reduzierte neue Versicherungssumme VS_k berechnen als ${}_k V / A$, wobei A die übliche Einmalprämie für die entsprechende Versicherungsart ist. Für eine Kapital-Lebensversicherung ergibt sich so beispielsweise

$$VS_k = \frac{{}_k V_{x:n}}{A_{x+k;n-k}} = 1 - \frac{P_{x:n}}{P_{x+k;n-k}}, \quad k = 1, \dots, n-1.$$

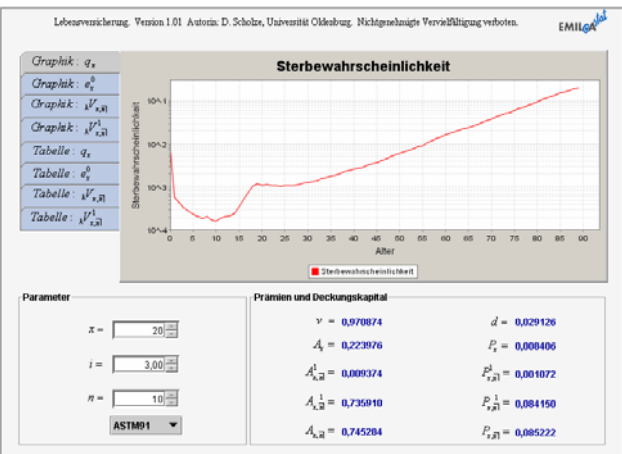
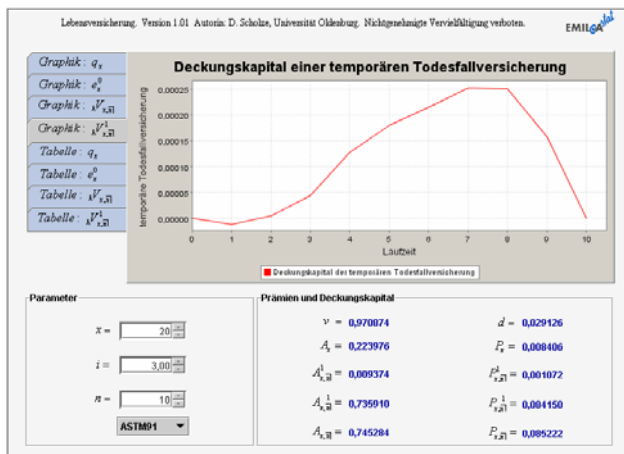
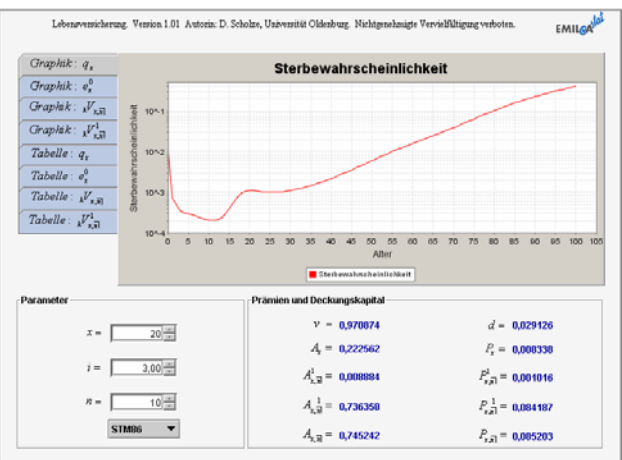
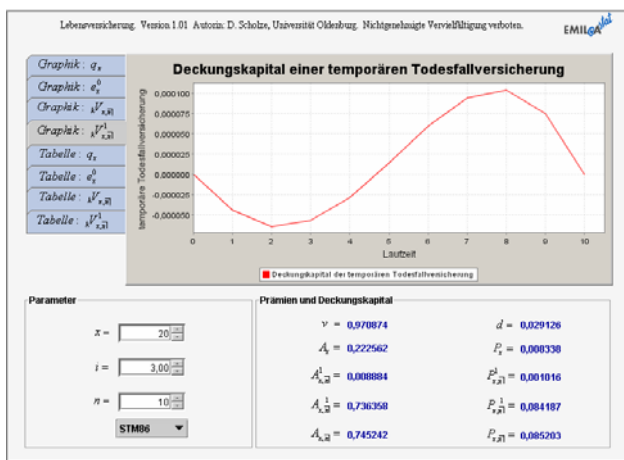
Beispiel 10. Ein 30jähriger Mann schließt eine Kapital-Lebensversicherung mit 35jähriger Laufzeit über die Versicherungssumme 100000 € ab. Bei Kündigung / Umwandlung in eine beitragsfreie Versicherung nach 20 Jahren ergeben sich bei 3% Rechnungszins je nach Sterbetafel folgende Rückkaufswerte / reduzierte Versicherungssummen:

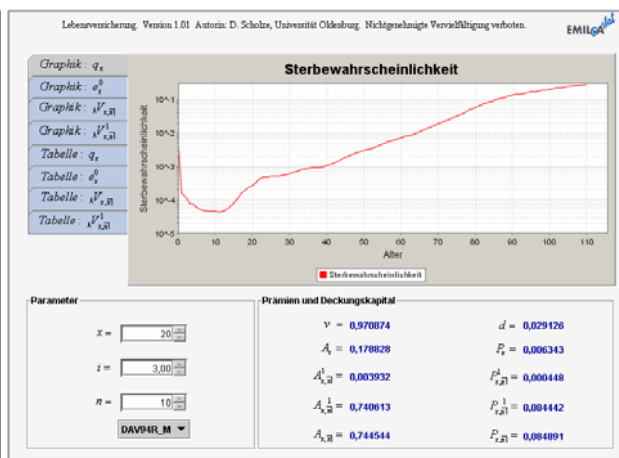
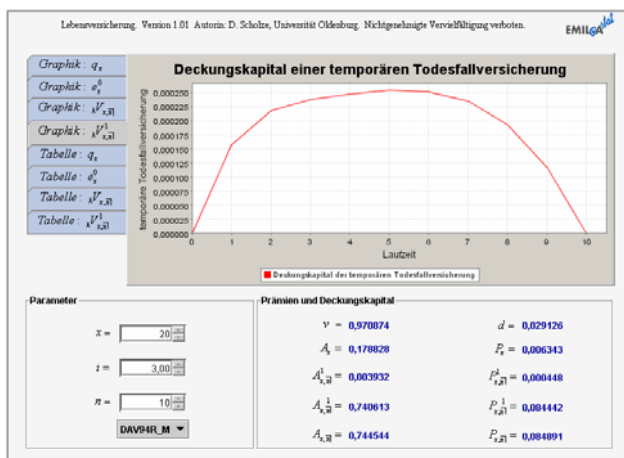
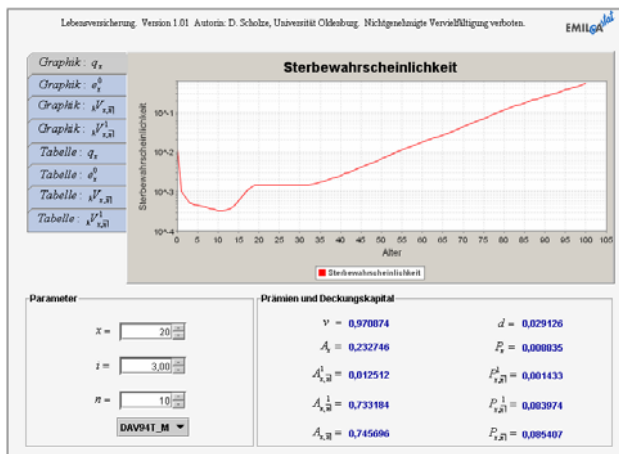
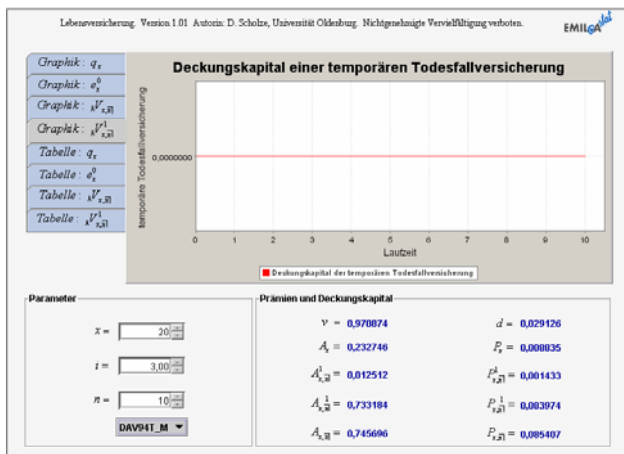
Sterbetafel	stm86	astm91	dav94t_m	dav94r_m
${}_k V_{x:n}$	45412 €	45200 €	45454 €	44798 €
$A_{x+k;n-k}$	0,663096	0,662676	0,665494	0,651967
VS_k	68485 €	68208 €	68301 €	68712 €

I Personenversicherungsmathematik

Die obigen Ausführungen machen aus verschiedener Sicht deutlich, dass es für ein VU unvorteilhaft ist, wenn negative Deckungskapitalien entstehen, etwa in den Altersstufen, in denen die Sterblichkeit rückläufig ist. Bei vorzeitiger Kündigung der Versicherung durch den VN würden dann Ansprüche seitens des VU an den VN entstehen, d.h. der VN wäre bei Kündigung ggf. nachschusspflichtig. Auch Umwandlungen in eine beitragsfreie Versicherung wären nicht möglich. Negative Deckungskapitalien wirken sich auch unvorteilhaft auf die Solvabilität des Unternehmens aus, weil das VU zur Absicherung seiner Risiken dann z.B. Fremdkapital aufnehmen müsste. Aus diesem Grund werden Sterbetafeln manchmal „geglättet“ (z.B. die DAV-Tafel T94m), d.h. in Altersstufen mit abnehmender jährlicher Sterbewahrscheinlichkeit werden die Sterblichkeiten solange als konstant angenommen, bis sie wieder steigen.

Die folgenden Graphiken zeigen diesen Effekt bei einer temporären Todesfallversicherung mit Abschlussalter 20 und 10jähriger Laufzeit und verschiedenen Sterbetafeln.





Man beachte, dass die durch Glättung bedingte Konstanz der Sterblichkeit in der DAV-Tafel T94m zu verschwindenden Deckungskapitalien führt, da hier lokal (d.h. bedingt bei Bezug auf die Altersspanne von 20 bis 30 Jahren) die Situation einer exponentialverteilten Lebensdauer und damit einer geometrisch verteilten gestutzten Lebensdauer wie in Beispiel 9 vorliegt.

I. 5. Verbundene Leben

Es ist manchmal wünschenswert, nicht nur eine einzelne Person, sondern eine ganze Personengemeinschaft zu versichern, z.B. ein Ehepaar oder eine andere Lebensgemeinschaft. Wir beschränken uns hier der Einfachheit halber auf Gemeinschaften mit zwei Personen; der allgemeine Fall wird z.B. in GERBER (1986 oder 1997) behandelt.

Bei Personen-Paaren sind grundsätzlich zwei Optionen bezüglich der Auszahlung der Versicherungssumme im Todesfall denkbar:

1. Zahlung beim Tod der *ersten* Person
2. Zahlung beim Tod der *zweiten* Person.

Die erste Variante ist sinnvoll, wenn der oder die jeweils Überlebende noch von der Versicherungssumme profitieren möchte.

Für die folgenden Ausführungen bezeichnen wir das Alter der ersten Person mit x , das der zweiten mit y sowie mit $T^{[1]}$ die (gesamte) Lebensdauer der ersten und mit $T^{[2]}$ die (gesamte) Lebensdauer

der zweiten Person der Lebensgemeinschaft. Entsprechend sind $T_x^{[1]}$ und $T_y^{[2]}$ die restlichen Lebensdauern der Personen, analog $K_x^{[1]}$ und $K_y^{[2]}$ die zugehörigen gestutzten Rest-Lebensdauern. Mit $F^{[k]}$, $f^{[k]}$, $\mu^{[k]}$ usw. werden die entsprechenden Größen "Verteilungsfunktion", "Dichte", "Sterbeintensität" usw. für die jeweilige Person k bezeichnet.

Man kann gemäß den beiden obigen Varianten der Lebensgemeinschaft jeweils eine neue, gemeinsame Rest-Lebensdauer zuordnen, und zwar

$$T_{xy}^{\min} := \min \{T_x^{[1]}, T_y^{[2]}\} \text{ für die erste Variante}$$

$$T_{xy}^{\max} := \max \{T_x^{[1]}, T_y^{[2]}\} \text{ für die zweite Variante,}$$

entsprechend für die gestutzten Rest-Lebensdauern. Einige Eigenschaften dieser „synthetischen“ Rest-Lebensdauern zeigt das folgende Resultat auf.

Lemma 7. Unter der Annahme der stochastischen Unabhängigkeit der Lebensdauern gilt für $t \geq 0$:

- a) $F_{xy}^{\min}(t) := P(T_{xy}^{\min} \leq t) = 1 - (1 - F_x^{[1]}(t))(1 - F_y^{[2]}(t))$, $F_{xy}^{\max}(t) := P(T_{xy}^{\max} \leq t) = F_x^{[1]}(t)F_y^{[2]}(t)$
- b) $\mu_{xy}^{\min}(t) = \mu_x^{[1]}(t) + \mu_y^{[2]}(t)$, $\mu_{xy}^{\max}(t) = \mu_x^{[1]}(t) \frac{F_y^{[2]}(t) - F_x^{[1]}(t)F_y^{[2]}(t)}{1 - F_x^{[1]}(t)F_y^{[2]}(t)} + \mu_y^{[2]}(t) \frac{F_x^{[1]}(t) - F_x^{[1]}(t)F_y^{[2]}(t)}{1 - F_x^{[1]}(t)F_y^{[2]}(t)}$
- c) ${}_t p_{xy}^{\min} = {}_t p_x^{[1]} \cdot {}_t p_y^{[2]} = \prod_{j=0}^{t-1} (1 - q_{x+j}^{[1]})(1 - q_{y+j}^{[2]})$, $q_{xy}^{\min} = 1 - {}_1 p_{xy}^{\min} = q_x^{[1]} + q_y^{[2]} - q_x^{[1]}q_y^{[2]}$.

Beweis: Es ist

$$F_{xy}^{\min}(t) = P(T_{xy}^{\min} \leq t) = 1 - P(T_{xy}^{\min} > t) = 1 - P(\{T_x^{[1]} > t\} \cap \{T_y^{[2]} > t\}) = 1 - (1 - F_x^{[1]}(t))(1 - F_y^{[2]}(t))$$

und

$$F_{xy}^{\max}(t) = P(T_{xy}^{\max} \leq t) = P(\{T_x^{[1]} \leq t\} \cap \{T_y^{[2]} \leq t\}) = F_x^{[1]}(t)F_y^{[2]}(t).$$

Dies zeigt a). Für b) kann man etwa so argumentieren:

$$F_{xy}^{\min}(t) = 1 - (1 - F_x^{[1]}(t))(1 - F_y^{[2]}(t)) = 1 - \exp\left(-\int_0^t \mu_x^{[1]}(s) ds\right) \exp\left(-\int_0^t \mu_y^{[2]}(s) ds\right)$$

$$= 1 - \exp\left(-\int_0^t \{\mu_x^{[1]}(s) + \mu_y^{[2]}(s)\} ds\right),$$

woraus der erste Teil der Behauptung folgt. Den zweiten Teil zeigen wir unter der vereinfachenden Annahme, dass (stetige) Dichten existieren. Dann gilt

I Personenversicherungsmathematik

$$\begin{aligned}\mu_{xy}^{\max}(t) &= \frac{f_{xy}^{\max}(t)}{1 - F_{xy}^{\max}(t)} = \frac{f_x^{[1]}(t)F_y^{[2]}(t) + f_y^{[2]}(t)F_x^{[1]}(t)}{1 - F_x^{[1]}(t)F_y^{[2]}(t)} \\ &= \mu_x^{[1]}(t) \frac{(1 - F_x^{[1]}(t))F_y^{[2]}(t)}{1 - F_x^{[1]}(t)F_y^{[2]}(t)} + \mu_y^{[2]}(t) \frac{(1 - F_y^{[2]}(t))F_x^{[1]}(t)}{1 - F_x^{[1]}(t)F_y^{[2]}(t)},\end{aligned}$$

woraus der zweite Teil der Behauptung folgt. Teil c) ergibt sich aus

$${}_tP_{xy}^{\min} = P(T_{xy}^{\min} \geq t) = P(\{T_x^{[1]} \geq t\} \cap \{T_y^{[2]} \geq t\}) = {}_tP_x^{[1]} \cdot {}_tP_y^{[2]} = \prod_{j=0}^{t-1} (1 - q_{x+j}^{[1]})(1 - q_{y+j}^{[2]})$$

und

$$q_{xy}^{\min} = 1 - {}_1P_{xy}^{\min} = 1 - (1 - q_x^{[1]})(1 - q_y^{[2]}) = 1 - (1 - q_x^{[1]} - q_y^{[2]} + q_x^{[1]}q_y^{[2]}) = q_x^{[1]} + q_y^{[2]} - q_x^{[1]}q_y^{[2]}.$$

Bei Versicherungen auf den ersten Todesfall (Variante 1 oben) lassen sich somit Prämien leicht aus gegebenen Sterbetafeln berechnen.

Beispiel 11. Es soll eine 20-jährige Todesfallversicherung für ein Ehepaar ($x = 35$ (Mann), $y = 32$ (Frau)) über die Versicherungssumme 100000 € abgeschlossen werden. Die Versicherungssumme wird fällig, sobald einer von beiden innerhalb der Vertragslaufzeit verstirbt.

Nach Lemma 7 c) reicht es, eine „gemeinsame“ Sterbetafel für das Ehepaar zu erstellen, wobei der Eintrag zum Jahr k durch den Ausdruck

$$q_{x+k \ y+k}^{\min} = q_{x+k}^{[1]} + q_{y+k}^{[2]} - q_{x+k}^{[1]}q_{y+k}^{[2]}, \quad k = 0, 1, \dots$$

gegeben ist. Die folgende Tabelle zeigt das entsprechende Ergebnis bei Zugrundelegung der DAV-Sterbetafeln T94m und T94w:

x	$q_x^{[1]}$	y	$q_y^{[2]}$	Alter neu	q_{xy}^{\min}
35	0,001747	32	0,000783	0	0,00252863
36	0,001869	33	0,000833	1	0,00270044
37	0,002007	34	0,000897	2	0,00290220
38	0,002167	35	0,000971	3	0,00313590
39	0,002354	36	0,001057	4	0,00340851
40	0,002569	37	0,001156	5	0,00372203
41	0,002832	38	0,001267	6	0,00409541
42	0,003087	39	0,001390	7	0,00447271
43	0,003387	40	0,001524	8	0,00490584
44	0,003726	41	0,001672	9	0,00539177
45	0,004100	42	0,001812	10	0,00590457
46	0,004522	43	0,001964	11	0,00647712
47	0,004983	44	0,002126	12	0,00709841
48	0,005508	45	0,002295	13	0,00779036
49	0,006049	46	0,002480	14	0,00851400
50	0,006751	47	0,002676	15	0,00940893
51	0,007485	48	0,002902	16	0,01036528
52	0,008302	49	0,003151	17	0,01142684
53	0,009215	50	0,003425	18	0,01260844
54	0,010195	51	0,003728	19	0,01388499
55	0,011236	52	0,004066	20	0,01525631

Mit dem Programm `leben.exe` oder dem Maple-Worksheet `leben.mws` ergibt sich aus der entsprechenden Tabelle rechts bei 3% Rechnungszins eine jährliche Prämie in Höhe von 571,20 € zahlbar bis zum ersten Todesfall, längstens 20 Jahre.

Hätte jeder von beiden eine separate Todesfallversicherung über diese Summe abgeschlossen, hätte sich folgendes ergeben:

Jährliche Prämie Mann: 406,20 € jährliche Prämie Frau: 170,00 €

Für gewisse Sterblichkeitsmodelle lassen sich bei Versicherungen auf den ersten Tod elegante Vereinfachungen vornehmen, z.B. im Modell von Gompertz-Makeham mit gleichen Parametern a und b (aber ggf. unterschiedlichen Parametern $c^{[1]}$ und $c^{[2]}$) für beide Personen (Beispiel 6). Hier gilt:

$$\mu_{xy}^{\min}(t) = \mu_x(t) + \mu_y(t) = \mu(x+t) + \mu(y+t) = c^{[1]} + a e^{b(x+t)} + c^{[2]} + a e^{b(y+t)} = (c^{[1]} + c^{[2]}) + a e^{b(z+t)}$$

mit $z = \frac{1}{b} \ln(e^{bx} + e^{by})$ für $t \geq 0$, d.h. die „gemeinsame“ Restlebensdauer der Gemeinschaft ist

wieder nach einem Gompertz-Makeham-Modell verteilt, mit dem neuen Parameter $c = c^{[1]} + c^{[2]}$ und einem fiktiven Anfangs-„Alter“ z , das sich durch eine nicht-lineare Mittelung der Alter x und y ergibt.

Ähnliche Überlegungen können natürlich auch zur Prämienfindung für verbundene Leben bei Erlebensfall- und Kapital-Lebensversicherungsformen angestellt werden.

I. 6. Kostenrechnung und Überschussbeteiligung

Im Lebensversicherungsgeschäft werden typischerweise die folgenden drei Kostenarten berücksichtigt (für technische Einzelheiten vgl. WOLFSBORF (1997)):

- Abschlusskosten (α -Kosten):

Dies beinhaltet diejenigen Kosten des VU, die mit dem Neuabschluss von Versicherungsverträgen entstehen (Außendienst, Maklerprovision, ärztliche Untersuchungskosten, Werbekosten usw.). Sie werden i.d.R. in Promille der Versicherungssumme oder der Beitragssumme angeben.

- Inkassokosten (β -Kosten):

Dies beinhaltet diejenigen Kosten des VU, die durch die technische Abwicklung der Beitragszahlung entstehen (Bankkosten, früher auch Kosten durch persönliches Inkasso durch Versicherungsagenten).

- Verwaltungskosten (γ -Kosten):

Dies beinhaltet diejenigen Kosten des VU, die durch die allgemeine Verwaltung entstehen (Personalkosten, Gebäudemiete, EDV-Kosten, Steuern usw.).

I Personenversicherungsmathematik

Berücksichtigt man diese Kostenarten angemessen in der (jährlichen) Prämie, so gelangt man zur sog. *adäquaten Prämie* P^a , gegeben durch

$$P^a = P + P^\alpha + P^\beta + P^\gamma.$$

Beispielsweise ergibt sich für eine Kapital-Lebensversicherung die adäquate Prämie aus der Beziehung

$$P_{x:n}^a \ddot{a}_{x:n} = A_{x:n} + \alpha + \beta P_{x:n}^a \ddot{a}_{x:n} + \gamma \ddot{a}_{x:n},$$

woraus mit $A_{x:n} + d \ddot{a}_{x:n} = 1$ folgt:

$$P_{x:n}^a = \frac{A_{x:n} + \alpha}{\ddot{a}_{x:n}} + \beta P_{x:n}^a + \gamma, \text{ also } P_{x:n}^a = \frac{A_{x:n} + \alpha (A_{x:n} + d \ddot{a}_{x:n})}{(1 - \beta) \ddot{a}_{x:n}} + \frac{\gamma}{1 - \beta} = \frac{1 + \alpha}{1 - \beta} P_{x:n} + \frac{\alpha d + \gamma}{1 - \beta}.$$

Die adäquate Prämie ergibt sich also direkt durch eine positiv-lineare Transformation aus der Netto-Prämie.

Beispiel 12: Wählen wir als Kostensätze $\alpha = 3,5\%$, $\beta = 3\%$ und $\gamma = 0,5\%$, so ergibt sich etwa bei 2,75% Rechnungszins die adäquate Prämie für die Versicherungssumme 1 zu

$$P_{x:n}^a = \frac{1 + \alpha}{1 - \beta} P_{x:n} + \frac{\alpha d + \gamma}{1 - \beta} = 1,067 P_{x:n} + 0,00612,$$

d.h. die adäquate Prämie ist hier etwa um 7% größer als die Netto-Prämie.

Eine entsprechende Modifikation ergibt sich auch für die *adäquate Deckungsrückstellung*. Bei der auf Zillmer (1863) zurückgehenden Berücksichtigung der Abschlußkosten erhalten wir

$${}_k V^a = {}_k V + {}_k V^\alpha.$$

Für eine Kapital-Lebensversicherung gilt etwa

$${}_k V_{x:n}^\alpha = -P^\alpha \ddot{a}_{x+k:n-k} = -\alpha \frac{\ddot{a}_{x+k:n-k}}{\ddot{a}_{x:n}} = -\alpha (1 - {}_k V_{x:n})$$

und somit

$${}_k V_{x:n}^a = (1 + \alpha) {}_k V_{x:n} - \alpha, \quad k = 0, 1, \dots, n.$$

Das "gezillmerte" Deckungskapital wird also wegen ${}_0 V_{x:n} = 0$ in den ersten Jahren regelmäßig negativ sein und erst zum Ende der Vertragslaufzeit den Wert

$${}_n V_{x:n}^a = (1 + \alpha) {}_n V_{x:n} - \alpha = (1 + \alpha) - \alpha = 1$$

erreichen. Aus diesem Grund sind die tatsächlichen Rückkaufswerte einer Kapital-Lebensversicherung auch niedriger als die ungezillmerten Deckungsrückstellungen; sie sind insbesondere für das erste Jahr der Versicherung fast immer null.

Die Zillmerung insbesondere im Zusammenhang mit den Rückkaufswerten einer Lebensversicherung war wegen teilweise unklarer Formulierungen in den Versicherungsverträgen auch Gegenstand rechtlicher Entscheidungen (etwa BGH-Urteil von 2005). Demnach ist die Zillmerung als Methodik nicht zu beanstanden, wenn eine ausreichende Transparenz der Berechnung für den VN gegeben ist.

Ein weiterer wichtiger Aspekt von Lebensversicherungen ist die – gesetzlich sogar vorgeschriebene – *Überschussbeteiligung* des VN. Eine (zumindest in früheren Jahren mögliche) Anlage der Prämien zu regelmäßig höheren Zinsen als der Rechnungszins erlaubt es nämlich, höhere Erträge zu erwirtschaften, die bis auf geringfügige Abschläge vollständig an den VN weiterzugeben sind. Der der Sparprämie entsprechende Teil dieses Mehrertrags (Gewinns) kann dargestellt werden als

$${}_{k+1}G^s = ({}_kV + {}_kP^s)(\hat{i} - i),$$

wobei ${}_kV$ das entsprechende Deckungskapital, ${}_kP^s$ die Sparprämie und \hat{i} den tatsächlichen Zins bezeichnet. Für die Todesfallversicherung unbegrenzter Dauer erhält man so (vgl. Lemma 6):

$${}_{k+1}G_x^s = ({}_kV_x + v \cdot {}_{k+1}V_x - {}_kV_x)(\hat{i} - i) = v \cdot {}_{k+1}V_x(\hat{i} - i).$$

Aktuell sind bei der Lebensversicherung die Möglichkeiten zu einer Überschussbeteiligung wegen des dauerhaft niedrigen Zinsniveaus allerdings stark eingeschränkt. Gebräuchliche Arten der Überschussbeteiligung sind z.B. eine entsprechende Erhöhung der Versicherungssumme bei Fälligkeit oder eine Verkürzung der Vertragslaufzeit bei gleichbleibender Versicherungssumme. Für weitere Einzelheiten sei hier auf GERBER (1986 und 1997) oder WOLFSDORF (1997) verwiesen.

Abschließende Bemerkung: Nach einem Urteil des EuGH vom März 2011 müssen in der Europäischen Union ab dem 21.12.2012 alle Lebens- und Rentenversicherungsverträge für Männer und Frauen in gleicher Weise tarifiert werden (sog. *Unisex-Tarife*), um eine (vermeintliche) geschlechtsspezifische Diskriminierung zu vermeiden. Zur Umgehung sogenannter *Antisielektionseffekte* wird dies aber vermutlich nicht zu einer Kostensenkung für die VU führen, sondern im Gegenteil zu einer Verteuerung.

I. 7. Die private Krankenversicherung

Die Möglichkeit, sich gegen die wirtschaftlichen Folgen einer Krankheit abzusichern, wurde bereits im Mittelalter durch zahlreiche, vorwiegend berufsständisch und regional abgegrenzte Selbsthilfeorganisationen geboten. Grundlage für deren Funktionsfähigkeit waren einerseits starke persönliche und ideelle Bindungen an die Gemeinschaft (später auch gesetzliche Zwangsmaßnahmen), andererseits eine im Wesentlichen gleichartige Gefährdung der einzelnen Mitglieder. Einen deutlichen Einschnitt in das frühe "Krankenversicherungswesen" Deutschlands brachte 1883 das "Gesetz betreffend die Krankenversicherung der Arbeiter". In gewisser Weise begann damit die Trennung zwischen gesetzlicher Krankenversicherung und privater Vorsorge für die wirtschaftlichen Folgen von Krankheitsfällen. Ein Wendepunkt für die auf Privatinitiative basierende "Krankenversicherung" Deutschlands kam in den Jahren 1924/25; ein großer Nachholbedarf an Behandlungsbedürftigkeit und das Streben nach wirtschaftlicher Sicherung erzeugten einen beispiellosen Boom. Die

Jahre 1924/25 können als der eigentliche Beginn der privaten Krankenversicherung (PKV) bezeichnet werden. Die starke Nachfrage nach privatem Versicherungsschutz stellte die damalige PKV (besonders vor dem Hintergrund einer inzwischen modernen Wirtschaftsordnung) zwangsläufig vor eine Reihe neuer Probleme:

- Die ursprüngliche Struktur, gekennzeichnet durch persönliche und ideelle Bindung einerseits, durch gleichartige Gefährdung der Mitglieder andererseits änderte sich; sie wurde weitgehend von der Form des auf dem Markt anbietenden Unternehmens verdrängt.
- Die Ausdehnung der Geschäftsgebiete vieler Unternehmen und zahlreiche Neugründungen erzeugten einen harten Konkurrenzkampf.
- Ohne Erfahrungen, ohne gesicherte Rechnungsgrundlagen und ohne der Versicherungsform adäquate Modelle musste (zudem noch unter dem Druck des Konkurrenzkampfes) die Krankenkostenversicherung im Sinne einer Schadenversicherung entwickelt werden.
- Vertragsrechtlich ergaben sich erhebliche Unklarheiten, da das Versicherungsvertragsgesetz (VVG) von 1908 die sich neu entwickelnde Form der PKV noch nicht berücksichtigen konnte.

Die Anforderungen an die für das deutsche Versicherungsaufsichtsgesetz (VAG) neue Institution des Verantwortlichen Aktuars werden in §11a VAG niedergelegt. Sie gelten nach §12 Abs. 2 VAG und §12 Abs. 3 VAG für die Krankenversicherung in gleichem Maße. Wesentliche Aufgabe des Aktuars ist es dabei, darüber zu wachen, dass bei der Berechnung von Prämien und Alterungsrückstellungen eine dauerhafte Erfüllbarkeit der Versicherungsverträge gewährleistet wird. Der Aktuar wird grundsätzlich vom Unternehmen bestellt und hat mit diesem einen Arbeitsvertrag. Er ist weisungsgebunden und unterliegt der arbeitsvertraglichen Treuepflicht.

§12 VAG enthält die speziellen Vorschriften für die substitutive PKV. Geregelt werden die Kalkulationsvorschrift nach Art der Lebensversicherung und die Verfahren zur Beitragsentlastung älterer Versicherte in §12a VAG. Besonders hinzuweisen ist auf die Funktion des Treuhänders im Rahmen der Prämienanpassungsklausel nach §12b VAG. Dem Treuhänder obliegt die Prüfung, "ob die Berechnung der Prämien mit den dafür bestehenden Rechtsvorschriften in Einklang steht". Eine Tarifanpassung erfolgt nur mit Zustimmung des vom Unternehmen unabhängigen Treuhänders. Letzterer wird von der Aufsichtsbehörde bestellt und ist ansonsten nur ihr verpflichtet (nach SCHMID-GROTJOHANN (1995), S. 10f).

Ausgelöst durch Beitragssteigerungen insbesondere für ältere Versicherte wird die Beitragsgestaltung der deutschen PKV seit einigen Jahren in der Tages- und Wirtschaftspresse sowie im politischen Raum verstärkt diskutiert. Im Mittelpunkt steht dabei die Frage nach der Bezahlbarkeit des Versicherungsschutzes im Alter. Diese öffentliche Diskussion mündete schließlich u.a. in die oben bereits erwähnte Neufassung der betreffenden Teile des VAG.

Parallel dazu fand und findet eine öffentliche Diskussion des Einflusses der Kostenentwicklung im Gesundheitswesen und der demographischen Entwicklung auf die Leistungs- und Beitragsperspektiven im Bereich der gesetzlichen Krankenversicherung (GKV) statt, die einerseits allgemeine Aufmerksamkeit für diese Kostenproblematik geschaffen hat und andererseits im Gesundheitsreformgesetz (GRG) und im Gesundheitsstrukturgesetz (GSG) von 1989 bzw. 1993 auch Regelungen zur Folge hatte, die versicherten Personen der PKV den Zugang zur GKV der Rentner erschweren. Das Beitragsproblem der älteren Versicherten in der PKV hat mehrere zusammenwirkende Ursachen:

- Im Zeitablauf zunehmende Krankheitskosten: Seit mehreren Jahrzehnten liegt die Kostensteigerung im Gesundheitswesen über der allgemeinen Preissteigerung.
- Das "mathematische Altenproblem": Auch falls die Kopfschäden altersunabhängig steigen, ist die bei Beitragsanpassung resultierende Preissteigerung höher als die Kostensteigerung. Dieser

wohlbekannte Effekt, auf den vielfach hingewiesen wird, beruht darauf, dass die durch die Kostendynamik erforderlichen Mehrleistungen der PKV zum *erreichten* Alter und nicht zum *ursprünglichen* Eintrittsalter der versicherten Personen neu versichert werden müssen.

- Im Zeitablauf steiler werdende Profile: Die Zunahme der Kopfschäden mit der Zeit erfolgt nicht gleichmäßig über die Alter, die Kostensteigerungen wachsen etwa ab dem Alter 50.

Charakteristisch für das System der GKV ist das Solidaritätsprinzip, das letztlich im Sozialstaatlichkeitsprinzip des Grundgesetzes begründet liegt. Grundgedanke des Solidaritätsprinzips ist, dass neben einem versicherungstechnischen Risikoausgleich auch ein sozialer Ausgleich zum Tragen kommt. Diesem sozialen Ausgleich entspricht das sogenannte Umlageverfahren in der GKV. Die Beiträge für grundsätzlich gleiche Leistungen werden proportional zum Bruttoarbeitseinkommen erhoben, so dass zwischen dem Wert des übernommenen Versicherungsschutzes und dem Beitrag des gesetzlich Versicherten keine erkennbare Relation besteht.

Im Gegensatz dazu postuliert das der PKV eigene individuelle Äquivalenzprinzip (wenn auch nur theoretisch) wie in der Lebensversicherung die Gleichheit von zu erwartender Gesamtnettoprämieinnahme und dem Erwartungswert der zugehörigen Versicherungsleistungen im Rahmen der vereinbarten Risikoübernahme über die gesamte Vertragsdauer hinweg, bezogen auf den Zeitpunkt des Vertragsabschlusses. Es handelt sich damit um die Gleichheit zweier Barwerte, deren Berechnung die Verwendung eines Rechnungszinses impliziert (nach SCHMID-GROTJOHANN (1995), S. 18f). Dabei ist zu berücksichtigen, dass ein Krankenversicherungsvertrag in der Regel auf lebenslange Dauer abgeschlossen wird.

Bezeichnet man wie in der Lebensversicherung mit ${}_k p_x$ die Wahrscheinlichkeit dafür, dass eine versicherte Person mit dem bei Versicherungsvertragsabschluss gegebenen Alter x (sogenanntes Eintrittsalter) im k -ten Jahr dem Kollektiv noch angehört, sowie mit $K_{x,k}$ die von einem "typischen" Versicherten des Eintrittsalters x im Versicherungsjahr k verursachten Kosten (dies sind also Zufallsvariablen im Sinne der mathematischen Stochastik) und $P_{x,k}$ die von den versicherten Personen desselben Eintrittsalters im Versicherungsjahr k vorschüssig gezahlte, deterministische Jahres-Prämie, so drückt sich das Äquivalenzprinzip in Formeln also folgendermaßen aus:

$$\sum_{k=0}^{\infty} {}_k p_x P_{x,k} v^k = \sum_{k=0}^{\infty} {}_k p_x E(K_{x,k}) v^k.$$

Dabei wird angenommen, dass die verursachten Kosten in jedem Jahr rechnungstechnisch dem jeweiligen Jahresbeginn zugerechnet werden. $E(K_{x,k})$ bezeichnet hierbei den Erwartungswert der "typischen" Kosten im Versicherungsjahr k . Eine Erklärung dieser Gleichung kann man wie folgt geben. Wir bezeichnen wieder mit

$$I_{x,k} = \begin{cases} 1, & \text{falls die versicherte Person im Versicherungsjahr } k \text{ dem Kollektiv noch angehört} \\ 0, & \text{sonst} \end{cases}$$

die (übliche) Indikatorfunktion der Kollektivzugehörigkeit. Da die versicherte Person dem Kollektiv zu einem zufälligen Zeitpunkt dauerhaft nicht mehr angehört (z.B. aufgrund von Tod oder Storno), bilden diese Indikatorfunktionen ebenfalls Zufallsvariablen im Sinne der mathematischen Stochastik, auch wenn die gezahlten Prämien selbst nicht-stochastisch sind. Es ergeben sich damit folgende formale Darstellungen des (zufälligen) Barwerts BP_x der Gesamtnettoprämieinnahme und des (zufälligen) Barwerts BV_x der zugehörigen Versicherungsleistungen:

$$BP_x = \sum_{k=0}^{\infty} I_{x,k} P_{x,k} v^k$$

$$BV_x = \sum_{k=0}^{\infty} I_{x,k} K_{x,k} v^k.$$

Aufgrund der offenkundigen Beziehung ${}_k p_x = E(I_{x,k})$ folgt somit für die zugehörigen Erwartungswerte:

$$E(BP_x) = \sum_{k=0}^{\infty} {}_k p_x P_{x,k} v^k$$

$$E(BV_x) = \sum_{k=0}^{\infty} {}_k p_x E(K_{x,k}) v^k.$$

Die Gleichsetzung der beiden Erwartungswerte entspricht gerade dem in der obigen Gleichung formalisierten Äquivalenzprinzip. Die Verwendung von Erwartungswerten gründet sich dabei wie in der Lebensversicherung auf das Gesetz der großen Zahlen. Diese Herangehensweise impliziert, dass man es für sozial geboten hält, alle Versicherten mit denselben Kollektivmerkmalen (Alter, ggf. Geschlecht, Gesundheitsstatus usw.) auch finanziell in gleicher Weise zu belasten, unabhängig davon, ob die versicherte Person die vertraglich zugesagten Leistungen auch tatsächlich in Anspruch nimmt oder nicht.

Die Verbleibewahrscheinlichkeiten ${}_k p_x$ werden dabei – ähnlich wie in der Lebensversicherung – empirisch aus rohen Todeswahrscheinlichkeiten sowie Stornowahrscheinlichkeiten durch relativ aufwendige mathematisch-statistische Ausgleichsverfahren ermittelt.

Ein Vergleich der Größen BP_x und BV_x zeigt, dass der VN bei Anwendung des Äquivalenzprinzips einen für ihn "fairen" Risikotransfer vornimmt, indem er die ihm unbekanntem zufälligen zukünftigen Kosten $K_{x,k}$ gegen die ihm bekannten Prämien $P_{x,k}$ in geeigneter Weise "eintauscht". Das VU trägt dabei im wesentlichen zwei elementare Risiken: zum einen die Zufälligkeit der insgesamt eingenommenen Prämien (bei frühzeitigem Ausscheiden eines Versicherten aus dem Kollektiv entsteht eine Kostenunterdeckung), zum anderen die Zufälligkeit der nicht genau vorhersehbaren Kosten (was ebenfalls zu einer Kostenunterdeckung führen kann, wenn das bisherige Beitragsaufkommen nicht zur Regulierung der Kosten ausreicht).

I. 8. Kopfschäden und Schadenprofile

Ausgangspunkt der Betrachtung sind die jährlich in einem bestimmten Tarif mit der Anzahl L_x im Kollektiv der x -jährigen Versicherten anfallenden gesamten Versicherungsleistungen S_x , wobei L_x und S_x wieder als Zufallsvariablen aufgefasst werden.

Unter dem Kopfschaden KS_x versteht man nun den Erwartungswert der relativen Versicherungsleistung

$$KS_x = E\left(\frac{S_x}{L_x}\right).$$

Wiederum nach dem Gesetz der großen Zahlen gilt näherungsweise

$$KS_x \approx \frac{S_x}{L_x},$$

wenn L_x "genügend groß" ist. Die relative Versicherungsleistung $\frac{S_x}{L_x}$ kann also in ausreichend großen Kollektiven in erster Näherung als Schätzgröße für die Kopfschäden KS_x verwendet werden.

Die in der Regel mit dem Alter steigenden Kopfschäden werden in sogenannten *Kopfschadenreihen* alters- und geschlechtsabhängig dargestellt. Wählt man ein bestimmtes Alter x_0 (meist 28 oder 43) oder auch eine bestimmte Altersgruppe (meist 41 bis 45) und setzt dazu die Kopfschäden der anderen Alter bzw. Altersgruppen in Beziehung, so erhält man die sogenannten *normierten Kopfschäden* κ_x , auch *Profile* genannt:

$$\kappa_x := \frac{KS_x}{KS_{x_0}}.$$

Sie erweisen sich empirisch als weitgehend tarifunabhängig. Als Grundkopfschaden G bezeichnet man den Kopfschaden des ausgewählten Alters x_0 . Es gilt damit

$$KS_x = G \kappa_x$$

(nach SCHMID-GROTJOHANN (1995), S. 72f). Setzt man die im vorigen Abschnitt bestimmten Größen in die Formel zur Bestimmung des erwarteten Leistungsbarwerts ein, so ergibt sich

$$A_x = E(BV_x) = \sum_{k=0}^{\infty} {}_k p_x E(K_{x,k}) v^k = G \sum_{k=0}^{\infty} {}_k p_x \kappa_{x+k} v^k.$$

Wird auf der Beitragsseite eine konstante Prämie $P_{x,k} = P_x$ für alle Versicherungsjahre k gefordert, so erhält man entsprechend

$$E(BP_x) = P_x \sum_{k=0}^{\infty} {}_k p_x v^k = P_x \ddot{a}_x,$$

wobei \ddot{a}_x wieder den vorschüssigen Leibrentenbarwert bezeichnet. Aus dem Äquivalenzprinzip ergibt sich somit die folgende Formel für die Jahresnettoprämie

$$P_x = \frac{A_x}{\ddot{a}_x} = G \frac{\sum_{k=0}^{\infty} {}_k p_x \kappa_{x+k} v^k}{\sum_{k=0}^{\infty} {}_k p_x v^k}.$$

Die Summationen sind hier de facto natürlich nicht bis ∞ zu erstrecken, da die Lebensdauern der Versicherten eine biologisch bedingte natürliche Obergrenze besitzen; sie enden entsprechend mit dem aus der Lebensversicherung bekannten Endalter $\varpi - x$.

I. 9. Alterungsrückstellungen und Prämienanpassungen

Die Prämienkalkulation nach dem Äquivalenzprinzip geht stillschweigend davon aus, dass sich die verwendeten Rechnungsgrundlagen (Sterblichkeit, Stornowahrscheinlichkeiten, Kopfschadenreihen) für die nach Alter und Geschlecht differenzierten Kollektive über die Zeit nicht ändern. Es ist aber eine empirisch belegbare Tatsache, dass alle genannten Rechnungsgrundlagen in der Vergangenheit einer mehr oder minder starken Veränderung ausgesetzt waren und es auch in Zukunft sein werden. So werden etwa regelmäßig den neuen Verhältnissen angepasste Sterbetafeln oder andere Ausscheideordnungen entwickelt. Eine besonders starke Veränderung erfuhr in der Vergangenheit die u.a. durch medizinischen Fortschritt bedingte Kostenstruktur in der PKV, mit der Konsequenz, dass in kurzer Folge z.T. erhebliche Beitragserhöhungen durchgeführt wurden. Auf die entsprechenden sozialen Auswirkungen wurde bereits weiter oben hingewiesen.

Für die versicherungstechnische Abwicklung solcher Beitragsanpassungen existieren allerdings einschlägige Berechnungsmodalitäten, die ebenfalls auf dem Äquivalenzprinzip beruhen. Der in diesem Zusammenhang wesentliche Begriff ist der der *Alterungsrückstellung*.

In der Alterungsrückstellung werden ähnlich der Deckungsrückstellung der Lebensversicherung nach dem Versicherungsbeginn zunächst die Teile der Jahresnettoprämie P_x verzinslich angesammelt, die nicht unmittelbar zur Abdeckung der rechnungsmäßigen Kopfschäden KS_x und eines gewissen Teils der Verwaltungskosten benötigt werden. Mit zunehmenden Alter des Versicherten und steigenden Kopfschäden nehmen wie in der Lebensversicherung diese (zumeist) positiven "Sparanteile" ab und werden negativ in der Phase, in der die rechnungsmäßigen Kopfschäden die Jahresnettoprämie übersteigen. Damit wird die Alterungsrückstellung nach und nach wieder abgebaut und ist am rechnungsmäßigen Endzeitpunkt der Versicherung, in der Regel zum Endalter $\omega = 100$, voll aufgebraucht. (In der Sterbetafel PKV95 beträgt das Endalter abweichend $\omega = 103$.) Man kann somit die Alterungsrückstellung beschreiben als ausgleichendes Element zwischen den mit dem Alter steigenden Kopfschäden KS_x und der im Zeitablauf konstant konzipierten Nettoprämie P_x . Mit anderen Worten soll wie in der Lebensversicherung die Alterungsrückstellung die Gültigkeit des individuellen Äquivalenzprinzips zu jedem Zeitpunkt des im Prinzip lebenslang angelegten Vertrages gewährleisten. Die Berechnung der Alterungsrückstellung ist damit prinzipiell identisch mit der Berechnung der Deckungsrückstellung in der Lebensversicherung.

Nach dem Äquivalenzprinzip muss zu jedem Betrachtungszeitpunkt $x + m$ Äquivalenz zwischen dem Erwartungsbarwert zukünftiger Aufwendungen, also dem Leistungsbarwert A_{x+m} einerseits, sowie der Summe aus dem Barwert $P_x \ddot{a}_{x+m}$ künftiger Nettoprämien und der bis zum Betrachtungszeitpunkt $x + m$ angesammelten Alterungsrückstellung ${}_m V_x$ andererseits bestehen, in Formeln:

$${}_m V_x = A_{x+m} - P_x \ddot{a}_{x+m} = (P_{x+m} - P_x) \ddot{a}_{x+m}, \quad k = 0, 1, \dots$$

Mit Bildung der Alterungsrückstellung an sich kommt der Versicherer also bereits der Verpflichtung nach, eine Beitragserhöhung oder eine Leistungsminderung aufgrund des Älterwerdens des Versicherungsnehmers auszuschließen. Ein Rückkaufswert bzw. ein individueller Anspruch des

Versicherungsnehmers auf "seine" Alterungsrückstellung bei Storno (etwa beim Wechsel des Versicherers) ist dagegen – anders als in der Lebensversicherung – nicht gegeben, da hier keine vorher definierte Versicherungsleistung (etwa garantierte Auszahlung eines Betrags anlässlich bestimmter Ereignisse wie Tod oder Invalidität bzw. Erleben eines bestimmten Zeitpunktes) vorliegt. Vielmehr orientiert sich die auf dem Äquivalenzprinzip beruhende Alterungsrückstellung an durch das Kollektiv definierten Kopfschäden, die nicht einzelnen Versicherten als geldwerte Leistung zuzurechnen sind. Ein Anspruch auf die Alterungsrückstellung ergibt sich aber immer dann, wenn bei einer Beitragsänderung dem ursprünglichen Eintrittsalter mit Anrechnung der rechnungsmäßigen Alterungsrückstellung auf den neuen Beitrag Rechnung getragen wird. Damit ergibt sich ein individueller, aber doch zweckgebundener Anspruch des Versicherten auf die Alterungsrückstellung (nach SCHMID-GROTJOHANN (1995), S. 80).

Wie bereits oben erwähnt, können alle Rechnungsgrundlagen im Zeitablauf vom Änderungsrisiko betroffen sein. Am nachhaltigsten schlägt sich dabei erfahrungsgemäß das Änderungsrisiko auf die Entwicklung der Kopfschäden durch. Als sinnvolles Signal einer durch das Änderungsrisiko verursachten Störung des Äquivalenzprinzips betrachtet die PKV-Mathematik ein bestimmtes Verhältnis von sogenannten erforderlichen Versicherungsleistungen zu den rechnungsmäßigen Versicherungsleistungen innerhalb eines Jahres für ein bestimmtes Alterskollektiv, meist das Alterskollektiv des Grundkopfschadens, also zum Alter von 28 bzw. 43 Jahren. Die erforderlichen Versicherungsleistungen werden ermittelt auf der Basis realisierter Versicherungsleistungen der Vergangenheit (meist der letzten drei Jahre), die mittels Inter- und Extrapolation auf einen zukünftigen Zeitraum hochgerechnet werden. Mit der Auswahl der notwendigen mathematisch-statistischen Verfahren wird somit eine Aussage über den angenommenen Schadenentwicklungstrend im Zeitablauf getroffen. Das Signal zur Anpassung wirkt damit antizipativ.

Liegt der Quotient aus erforderlichen und rechnungsmäßigen Versicherungsleistungen (man spricht auch vom sogenannten auslösenden Faktor) über 1,1 oder unter 0,9, so wird ein Anpassungsbedarf der Prämie an die Versicherungsleistung vermutet. Diese Untersuchungen im Tarif erfolgen jährlich und in der Regel getrennt nach den Beobachtungseinheiten Frauen, Männer und Kinder. Das Signal ist Auslöser für weitergehende Untersuchungen, an deren Ende die Entscheidung stehen muss, ob es sich bei den Leistungssteigerungen um eine Erscheinung des Zufallsrisikos handelt oder ob die Veränderung dauerhaft und insofern Ausdruck eines Änderungsrisikos ist. Die Prämienanpassung ist – sieht man von der Möglichkeit einer Leistungseinschränkung mittels Erhöhung des Selbstbehalts ab – die vorgesehene Reaktion der Versicherungstechnik auf die leistungssteigernde Wirkung des Änderungsrisikos während der Laufzeit des Vertrages und ist eine Maßnahme der Sekundärregulation des Systems PKV. Mit der Angleichung der Prämien an die Leistungen wird die versicherungstechnische Äquivalenz im Tarif gewahrt. Voraussetzung für eine Prämienanpassung ist deren rechtlich einwandfreie Verankerung mittels einer sogenannten Prämienanpassungsklausel in den Allgemeinen Versicherungsbedingungen entsprechend §12b Abs. 1 und 2 VAG. Der Versicherer muss einem mit Alter x eingetretenen Versicherungsnehmer bei einer Prämienanpassung im Alter $x+m$ seine bis dahin angesammelte Alterungsrückstellung nach §12 Abs. 1 Nr. 4 VAG anrechnen. Bei einer bis zum rechnungsmäßigen Endzeitpunkt verrenteten Alterungsrückstellung ${}_mV_x(\text{alt})$ ergibt sich somit ein Anrechnungsbetrag von $\frac{{}_mV_x(\text{alt})}{\ddot{a}_{x+m}}$. Dies entspricht genau

dem jährlichen Nettoprämiennachlass durch die geforderte Anrechnung der Alterungsrückstellung. Die Jahresnettoprämie $P_{x+m}(\text{neu})$ eines zum Zeitpunkt $x+m$ neu Eintretenden ergibt sich aus den zu diesem Zeitpunkt beobachteten Kopfschäden der künftigen Altersgruppen.

Damit ist die Jahresnettoprämie $P_{x|x+m}$ eines zum Zeitpunkt x Eintretenden nach einer Prämienanpassung in $x+m$ insgesamt gegeben durch

I Personenversicherungsmathematik

$$P_{x|x+m} = P_{x+m}(\text{neu}) - \frac{{}_mV_x(\text{alt})}{\ddot{a}_{x+m}}.$$

Gedanklich kann man den Beitrag $P_{x|x+m}$ nach Anpassung auch zerlegen in den bisherigen Nettobeitrag $P_x(\text{alt})$, der auch so weiterzuführen ist und einen Zusatzbeitrag $g_{x|x+m}$ auf die durch die gestiegenen Kopfschäden hervorgerufene Mehrleistung ab $x+m$:

$$g_{x|x+m} := P_{x+m}(\text{neu}) - \frac{{}_mV_x(\text{alt})}{\ddot{a}_{x+m}} - P_x(\text{alt})$$

(nach SCHMID-GROTJOHANN (1995), S. 84). Hieraus folgt:

$$g_{x|x+m} := P_{x+m}(\text{neu}) - \frac{A_{x+m}(\text{alt}) - \ddot{a}_{x+m}P_x(\text{alt})}{\ddot{a}_{x+m}} - P_x(\text{alt}) = P_{x+m}(\text{neu}) - P_{x+m}(\text{alt}),$$

d.h. der Erhöhungsbetrag ist in diesem Fall für Versicherte desselben Alterskollektivs stets gleich, unabhängig davon, wie lange sie bereits vorher versichert sind. Lediglich der Sockelbetrag, von dem aus die Erhöhung erfolgt, ist individuell verschieden und hängt u.a. vom Eintrittsalter x sowie den gegebenenfalls zwischenzeitlich erfolgten Beitragsanpassungen ab.

Wir wollen die Auswirkungen einer solchen Beitragserhöhung an dem (typischen) Beispiel schildern, bei dem sich die Kopfschäden zum Zeitpunkt $x+m$ erstmalig von denjenigen zum Zeitpunkt x um einen Faktor $\phi > 1$ unterscheiden. Für die neuen Rechnungsgrößen folgt dann

$$A_{x+m}(\text{neu}) = \phi A_{x+m}(\text{alt}),$$

woraus

$$g_{x|x+m} = P_{x+m}(\text{neu}) - P_{x+m}(\text{alt}) = \frac{A_{x+m}(\text{neu})}{\ddot{a}_{x+m}} - P_{x+m}(\text{alt}) = \phi \frac{A_{x+m}(\text{alt})}{\ddot{a}_{x+m}} - P_{x+m}(\text{alt}) = (\phi - 1)P_{x+m}(\text{alt})$$

resultiert. Die relative Beitragserhöhung beträgt somit

$$\frac{g_{x|x+m}}{P_x(\text{alt})} = (\phi - 1) \frac{P_{x+m}(\text{alt})}{P_x(\text{alt})} = (\phi - 1) + (\phi - 1) \frac{{}_mV_x(\text{alt})}{P_x(\text{alt})\ddot{a}_{x+m}},$$

wohingegen die relative Kostensteigerung lediglich $(\phi - 1)$ beträgt.

Dies bedeutet, dass bei positiver Alterungsrückstellung die Beitragserhöhung prozentual – bezogen auf die bisher bezahlte Prämie – stets größer als die prozentuale Kostensteigerung ist und dabei um so höher ausfällt, desto höher die angesammelte Alterungsrückstellung ist.

Eine Prämienhöhung wird damit bei Personen mit längerer PKV-Zugehörigkeit und entsprechend höheren Rückstellungen stärker ausfallen als bei Personen mit einer vergleichsweise kurzen Zugehörigkeit. Eine längere Zugehörigkeit ist aber in der Tendenz meist bei den älteren Versicherten gegeben. Der Effekt wird noch verstärkt, wenn man auf die Prämisse des konstanten Profils verzichtet und stattdessen von der realistischeren Annahme eines im Zeitablauf steigenden Profils ausgeht.

Versicherungstechnisch kann die durch die Beitragserhöhung erhaltene Prämie als Prämie einer Neuversicherung zum Zeitpunkt $x + m$ angesehen werden, wobei aufgrund der Anrechnung der bis dahin erworbenen Alterungsrückstellung de facto eine Beitragssenkung erfolgt. Die neue Prämie beinhaltet somit selbst wieder Anteile, die der Bildung einer neuen Alterungsrückstellung zugeführt werden. Formal erhält man so für die entsprechende Alterungsrückstellung ${}_nV_{x|x+m}$ zur Zeit $x + n > x + m$ die Beziehung

$$\begin{aligned} {}_nV_{x|x+m} &= A_{x+n}(\text{neu}) - P_{x|x+m}\ddot{a}_{x+n} = A_{x+n}(\text{neu}) - \left[\underbrace{P_{x+m}(\text{neu}) - P_{x+m}(\text{alt})}_{=g_{x|x+m}} + P_x(\text{alt}) \right] \ddot{a}_{x+n} \\ &= \phi \cdot A_{x+n}(\text{alt}) - [(\phi - 1)P_{x+m}(\text{alt}) + P_x(\text{alt})] \ddot{a}_{x+n} \\ &= A_{x+n}(\text{alt}) - P_x(\text{alt})\ddot{a}_{x+n} + (\phi - 1)[A_{x+n}(\text{alt}) - P_{x+m}(\text{alt})\ddot{a}_{x+n}] \\ &= {}_nV_x(\text{alt}) + (\phi - 1) {}_{n-m}V_{x+m}(\text{alt}). \end{aligned}$$

Der Unterschiedsbetrag zwischen der alten und neuen Alterungsrückstellung, bezogen auf den Zeitpunkt $x + n$,

$${}_nV_{x|x+m} - {}_nV_x(\text{alt}) = (\phi - 1) {}_{n-m}V_{x+m}(\text{alt}),$$

entspricht dabei exakt der Alterungsrückstellung für das $n - m$ -te Versicherungsjahr bei Neuabschluss einer Versicherung lediglich über die ab dem Jahr $x + m$ anfallenden Mehrkosten. Dies zeigt noch einmal aus anderer Sicht, dass die versicherungstechnische Behandlung der Beitragsanpassung wie eine Neuversicherung ab dem Zeitpunkt $x + m$ nichts anderes bedeutet, als die alte Versicherung unverändert fortzuführen und lediglich die ab dem Zeitpunkt $x + m$ anfallenden Mehrkosten durch eine entsprechende Zusatzversicherung abzudecken.

Diese Zusammenhänge werden bezeichnet als "schicksalsbedingte Folgen der von der PKV angewandten Kalkulationsmethoden" und sind in der Literatur lange bekannt (nach SCHMID-GROTJOHANN (1995), S. 88).

Zur Milderung der Beitragssteigerung für langjährig Versicherte hat es in der Literatur eine Reihe von Vorschlägen gegeben, von denen der Ansatz der sogenannten dynamischen Äquivalenz wohl am vielversprechendsten ist (vgl. SCHMID-GROTJOHANN (1995), S. 100). Hier wird die voraussichtliche Kostensteigerung in Form von zeitabhängigen Faktoren für die augenblicklich bekannten Kopfschäden berücksichtigt, und damit die erwartete spätere Kostensteigerung durch entsprechende "Vorauszahlungen" – wieder in Form konstanter Prämien – aufgefangen oder zumindest gemildert. Eine Umsetzung solcher Vorschläge zur Entlastung älterer Versicherter ist inzwischen durch den Gesetzgeber erfolgt. Details hierzu sowie zu dem gesamten angesprochenen Themenkreis findet man in dem aktuellen und überaus detaillierten Text von MILBRODT (2005).

II. Sachversicherungsmathematik

Im Gegensatz zur Personenversicherungsmathematik, die im wesentlichen auf Sterbe- und anderen Ausscheidewahrscheinlichkeiten beruht, die aufgrund großer Datenbestände regelmäßig in repräsentativem Umfang erhoben und gepflegt werden können, sind die wahrscheinlichkeitstheoretischen Grundlagen der Sachversicherung erheblich weniger gut gesichert. Deshalb kommt hier der mathematischen Statistik und Modellierung als Hilfsmittel zur Erfassung und Beschreibung der potenziellen Risiken eine große Bedeutung zu. Zwei wesentlichen Komponenten begegnet man hier aber in jedem Fall immer wieder: zum einen der *Schadenfrequenz*, d.h. der Häufigkeit, mit der Versicherungsschäden oder Schäden auslösende Ereignisse innerhalb einer Police in einem befristeten Zeitraum – in der Regel ein Jahr – auftreten (etwa die Anzahl der Sturmschäden in der Gebäudeversicherung, oder die Anzahl der Verkehrsunfälle in der Kraftfahrzeug-Haftpflichtversicherung), zum anderen die *Einzelschadenhöhe*, d.h. der geldwerte Schaden, der durch das jeweilige Schadenereignis ausgelöst wird (z.B. die Reparaturkosten für verlorene Dachpfannen bei einem Sturm oder beschädigte Fahrzeuge bei einem Verkehrsunfall). Diese beiden Komponenten sind wesentliche Bestandteile des so genannten *kollektiven Modells der Risikotheorie*, das nun als erstes im Detail vorgestellt wird.

II. 1. Das kollektive Modell der Risikotheorie

Die dem Modell zu Grunde liegenden mathematischen Annahmen sind:

- Die Anzahl N der Schäden innerhalb einer bestimmten Periode ist eine nicht-negative, ganzzahlige Zufallsvariable.
- Die in dieser Periode anfallenden Einzelschadenhöhen X_1, X_2, \dots sind stochastisch unabhängige, identisch verteilte, *positive* Zufallsvariablen, die auch von N stochastisch unabhängig sind.

Unter dem *Gesamtschaden* (für die betrachtete Periode) versteht man dann die Zufallsvariable

$$S := \sum_{k=1}^N X_k$$

mit der Konvention, dass die "leere Summe" als Null verstanden wird.

Es handelt sich bei S also um eine zufällige Summe von Zufallsvariablen. Strenggenommen muss hier noch überprüft werden, ob S tatsächlich auch eine Zufallsvariable ist, d.h. ob Messbarkeit vorliegt. Dies sieht man aber leicht an der folgenden Darstellung:

$$\{S \in B\} = \bigcup_{n=0}^{\infty} \{N = n\} \cap \left\{ \sum_{k=1}^n X_k \in B \right\} \text{ für jede Borel-Menge } B \in \mathcal{B}.$$

Hierbei bezeichne \mathcal{B} die Borel'sche σ -Algebra über \mathbb{R} .

Für die folgenden Ausführungen wollen wir als General-Annahme voraussetzen, dass alle stetigen Verteilungen Dichten (bez. des Lebesgue-Maßes) besitzen.

II. Sachversicherungsmathematik

Von zentralem Interesse für die Sachversicherungsmathematik ist die Verteilung des Gesamtschadens. Hierbei spielt der Begriff der *Faltung* von Verteilungen eine große Rolle.

Definition 2 (Faltung). Es seien X und Y reellwertige, stochastisch unabhängige Zufallsvariablen. Dann heißt die Verteilung der Summe $X + Y$ die Faltung der Verteilungen von X und Y , in Zeichen:

$$P^{X+Y} = P^X * P^Y.$$

Die Faltung von Verteilungen kann auf verschiedene Weise dargestellt werden:

Lemma 8. Es seien X und Y stochastisch unabhängige, reellwertige Zufallsvariablen. Ihre Verteilungsfunktionen seien mit F_X und F_Y bezeichnet, bei stetigen Verteilungen ferner ihre Dichten entsprechend mit f_X und f_Y . Die Verteilungsfunktion der Faltung $P^X * P^Y$ sei mit $F_X * F_Y$ bezeichnet, entsprechend die Dichte mit $f_X * f_Y$. Dann gilt:

$$F_X * F_Y(z) = F_Y * F_X(z) = \int_{-\infty}^{\infty} F_X(z-v) f_Y(v) dv$$

$$f_X * f_Y(z) = f_Y * f_X(z) = \int_{-\infty}^{\infty} f_X(z-v) f_Y(v) dv, \quad z \in \mathbb{R}.$$

Beweis: Es ist

$$\begin{aligned} F_X * F_Y(z) &= P(X + Y \leq z) = \int_{-\infty}^{\infty} P(X + Y \leq z | Y = y) P^Y(dy) = \int_{-\infty}^{\infty} P(X + y \leq z) P^Y(dy) \\ &= \int_{-\infty}^{\infty} F_X(z - y) P^Y(dy) = \int_{-\infty}^{\infty} F_X(z - y) f_Y(y) dy, \end{aligned}$$

woraus durch Differentiation folgt

$$f_X * f_Y(z) = \frac{d}{dz} F_X * F_Y(z) = \frac{d}{dz} \int_{-\infty}^{\infty} F_X(z - y) f_Y(y) dy = \int_{-\infty}^{\infty} f_X(z - v) f_Y(v) dv.$$

Die Verteilung des Gesamtschadens lässt sich damit auf folgende Weise ausdrücken:

Lemma 9. Für die Verteilungsfunktion F_S des Gesamtschadens gilt:

$$P(S \leq z) = F_S(z) = \begin{cases} p_0 + \sum_{n=1}^{\infty} p_n F^{n*}(z), & \text{falls } z \geq 0 \\ 0, & \text{falls } z < 0. \end{cases}$$

II. Sachversicherungsmathematik

Dabei bezeichne $p_n := P(N = n)$ für $n = 0, 1, \dots$ die Elementarwahrscheinlichkeiten von N und F^{n*} die n -fache Faltung von F , der Verteilungsfunktion der Einzelschadenhöhen.

Ferner gilt für die bedingte Verteilung unter $S > 0$:

$$P(S \leq z | S > 0) = \frac{F_S(z) - F_S(0)}{1 - F_S(0)} = \begin{cases} \frac{1}{1 - p_0} \sum_{n=1}^{\infty} p_n F^{n*}(z), & \text{falls } z \geq 0 \\ 0, & \text{falls } z < 0. \end{cases}$$

Sind die Einzelschadenhöhen stetig verteilt, so auch der Gesamtschaden S unter $S > 0$ mit

$$f_S(z | S > 0) = \frac{1}{1 - p_0} \sum_{n=1}^{\infty} p_n f^{n*}(z) \quad \text{für } z \geq 0.$$

Hierbei bezeichnet analog f^{n*} die n -fache Faltung von f , der Dichte der Einzelschadenhöhenverteilung.

Beweis: Es ist nach obigem

$$\begin{aligned} P(S \leq z) = F_S(z) &= P\left(\bigcup_{n=0}^{\infty} \{N = n\} \cap \left\{\sum_{k=1}^n X_k \leq z\right\}\right) = p_0 + \sum_{n=1}^{\infty} p_n P\left(\sum_{k=1}^n X_k \leq z\right) \\ &= p_0 + \sum_{n=1}^{\infty} p_n F^{n*}(z), \quad z \in \mathbb{R}. \end{aligned}$$

Die angegebene Fallunterscheidung resultiert daraus, dass nach Voraussetzung der Positivität der Einzelschadenhöhen $F(z) = 0$ ist für $z < 0$.

Die zweite Beziehung ergibt sich sofort hieraus wegen $P(S > 0) = 1 - F_S(0) = 1 - p_0$.

Man beachte, dass bei Stetigkeit der Einzelschadenhöhen in der Regel die Verteilung von S nicht mehr stetig ist, sondern – falls $p_0 > 0$ gilt – ein Atom im Nullpunkt besitzt. Allerdings bleibt in diesem Fall die bedingte Verteilung von S unter $S > 0$ stetig mit

$$f_S(z | S > 0) = \frac{d}{dz} \left[\frac{1}{1 - p_0} \sum_{n=1}^{\infty} p_n F^{n*}(z) \right] = \frac{1}{1 - p_0} \sum_{n=1}^{\infty} p_n f^{n*}(z), \quad z \geq 0.$$

In speziellen Fällen lässt sich die Gesamtschadenverteilung explizit angeben:

Lemma 10. Die Verteilung von N sei geometrisch mit

$$p_n = pq^n, \quad n = 0, 1, \dots \quad \text{mit einem Parameter } p = 1 - q \in (0, 1),$$

die Verteilung der Einzelschadenhöhen sei exponentiell $\mathcal{E}(\lambda)$ mit einem Parameter $\lambda > 0$. Dann gilt:

II. Sachversicherungsmathematik

$$P(S \leq z) = F_S(z) = \begin{cases} p + (1-p)(1 - e^{-\lambda pz}), & \text{falls } z \geq 0 \\ 0, & \text{falls } z < 0. \end{cases}$$

Ferner ist der Gesamtschaden S unter der Bedingung $S > 0$ wieder exponentialverteilt, und zwar nach $\mathcal{E}(\lambda p)$.

Beweis: Die Verteilung einer Summe unabhängiger, exponentialverteilter Zufallsvariablen ist vom Erlang-Typ (spezieller Gamma-Typ), d.h. es gilt hier für $n \in \mathbb{N}$:

$$f^{n*}(z) = \frac{\lambda^n z^{n-1}}{(n-1)!} e^{-\lambda z}, \quad z \geq 0.$$

Es ergibt sich zunächst

$$f_S(z | S > 0) = \frac{1}{q} \sum_{n=1}^{\infty} pq^n f^{n*}(z) = \sum_{n=1}^{\infty} pq^{n-1} \frac{\lambda^n z^{n-1}}{(n-1)!} e^{-\lambda z} = \lambda p e^{-\lambda z} \underbrace{\sum_{n=1}^{\infty} \frac{(\lambda q z)^{n-1}}{(n-1)!}}_{=\exp(\lambda q z)} = \lambda p e^{-\lambda pz}, \quad z \geq 0,$$

d.h. es ist S unter der Bedingung $S > 0$ nach $\mathcal{E}(\lambda p)$ -verteilt. Hieraus folgt:

$$\begin{aligned} F_S(z) &= P(S \leq z) = P(S \leq z | S = 0) \cdot P(S = 0) + P(S \leq z | S > 0) \cdot P(S > 0) \\ &= 1 \cdot P(N = 0) + \left(\int_0^z f_S(u | S > 0) du \right) \cdot P(N > 0) = \begin{cases} p + (1-p)(1 - e^{-\lambda pz}), & \text{falls } z \geq 0 \\ 0, & \text{falls } z < 0. \end{cases} \end{aligned}$$

Die Verteilung von S kann also aufgefasst werden als eine *Mischung* zweier Verteilungen, und zwar einer Einpunktverteilung ε_0 im Nullpunkt mit Gewicht p und einer Exponentialverteilung $\mathcal{E}(\lambda p)$ mit Gewicht $1-p$.

Für die Darstellung der Gesamtschadenverteilung sind manchmal gewisse Transformationen von Wahrscheinlichkeitsverteilungen nützlich.

Definition 3 (erzeugende Funktionen). Es sei X eine reellwertige Zufallsvariable derart, dass für eine Teilmenge $I \subseteq \mathbb{R}$ der Ausdruck

$$\psi_X(t) := E(e^{tX}), \quad t \in I$$

für alle $t \in I$ endlich ist. Dann heißt die auf I definierte Abbildung ψ_X die *momenterzeugende Funktion* zu X bzw. zu der Verteilung P^X .

Die durch

$$\varphi_X(s) := \psi_X(\ln s) = E(s^X), \quad s \in e^I := \{e^t \mid t \in I\}$$

II. Sachversicherungsmathematik

definierte Funktion heißt die *wahrscheinlichkeitserzeugende* Funktion zu X bzw. zu der Verteilung P^X .

Die momenterzeugende Funktion ψ_X charakterisiert die Verteilung P^X eindeutig, wenn die Menge I ein Intervall $[-\delta, \delta]$ mit $\delta > 0$ enthält. Der zugehörige Beweis wird üblicherweise mit Hilfsmitteln der Fourier-Analyse geführt; wir verweisen deshalb hier auf die einschlägige Literatur zur Wahrscheinlichkeitstheorie (z.B. BILLINGSLEY (1986), Theorem 30.1).

Satz 1. Es sei X eine reellwertige Zufallsvariable derart, dass für eine Teilmenge $I \subseteq \mathbb{R}$ die momenterzeugende Funktion ψ_X existiert. Dann gilt:

- a) Es ist stets $\psi_X(0) = \varphi_X(1) = 1$. Existiert ferner $\psi_X(t)$ für ein $t = t^* > 0$ bzw. $t = t_* < 0$, so auch für alle $t \in [0, t^*]$ bzw. $t \in [t_*, 0]$. Bezeichnet speziell

$$t^+ := \sup \{t \in \mathbb{R} \mid \psi_X(t) < \infty\}, \quad t^- := \inf \{t \in \mathbb{R} \mid \psi_X(t) < \infty\},$$

so existiert $\psi_X(t)$ für alle $t \in (t^-, t^+)$, und $\varphi_X(s)$ existiert für alle $s \in (e^{t^-}, e^{t^+})$ (mit der Konvention $e^{-\infty} = 0$, $e^{\infty} = \infty$). Gilt insbesondere $X \geq 0$ bzw. $X \leq 0$ (mit Wahrscheinlichkeit 1), so ist $t^- = -\infty$ bzw. $t^+ = +\infty$.

- b) Es sei $0 < \delta < \min \{t^+, -t^-\}$. Dann existieren sämtliche Momente $E(|X|^k)$, $k \in \mathbb{N}$, ψ_X ist im Nullpunkt beliebig oft differenzierbar, und es gilt

$$\begin{aligned} \psi_X^{(k)}(0) &= E(X^k), \quad k \in \mathbb{N} \quad \text{und} \\ \psi_X(t) &= \sum_{k=0}^{\infty} \frac{E(X^k)}{k!} t^k, \quad |t| \leq \delta. \end{aligned}$$

Insbesondere ist

$$E(X) = \psi_X'(0), \quad \text{Var}(X) = \psi_X''(0) - \{\psi_X'(0)\}^2.$$

Ferner ist $\varphi_X(s)$ für $s = 1$ differenzierbar, und es gilt

$$\begin{aligned} \varphi_X^{(k)}(1) &= E\left(\prod_{i=0}^{k-1} (X - i)\right), \quad k \in \mathbb{N} \quad \text{und} \\ \varphi_X(s) &= \sum_{k=0}^{\infty} \frac{E\left(\prod_{i=0}^{k-1} (X - i)\right)}{k!} (s-1)^k, \quad |s-1| \leq 1 - e^{-\delta}. \end{aligned}$$

Insbesondere ist

II. Sachversicherungsmathematik

$$E(X) = \varphi'_X(1), \text{ Var}(X) = \varphi''_X(1) + \varphi'_X(1)\{1 - \varphi'_X(1)\}.$$

- c) Gilt $P(X \in \mathbb{Z}^+) = 1$ mit $\mathbb{Z}^+ := \{0, 1, 2, \dots\}$, so lässt sich φ_X fortsetzen, d.h. es existiert $\varphi_X(s) = E(s^X)$ auch für alle $|s| \leq 1$, und es gilt

$$\frac{\varphi_X^{(k)}(0)}{k!} = P(X = k), \quad k \in \mathbb{N} \text{ und}$$

$$\varphi_X(s) = \sum_{k=0}^{\infty} P(X = k) s^k, \quad |s| \leq 1.$$

- d) Sind X und Y stochastisch unabhängige, reellwertige Zufallsvariablen mit momenterzeugenden Funktionen ψ_X und ψ_Y , die beide in derselben Menge $I \subseteq \mathbb{R}$ existieren, so besitzt dort auch die Zufallsvariable $Z = X + Y$ eine momenterzeugende Funktion, und es gilt

$$\psi_{X+Y}(t) = \psi_X(t) \cdot \psi_Y(t), \quad t \in I$$

bzw. auch

$$\varphi_{X+Y}(s) = \varphi_X(s) \cdot \varphi_Y(s), \quad s \in e^I.$$

Beweis:

- a) Die erste Aussage ist wegen $e^0 = 1$ trivial. Mit der *Hölder-Ungleichung*⁷ erhält man ferner für $0 \leq t \leq t^*$

$$\psi_X(t) = E(e^{tX}) \leq \left\{ E(e^{t^*X}) \right\}^{t/t^*} = \left\{ \psi_X(t^*) \right\}^{t/t^*}$$

(mit der Wahl $p = \frac{t^*}{t}$ und $Y \equiv 1$), wenn man dort X durch e^X ersetzt. Mit analoger Argumentation folgt für $t_* \leq t \leq 0$

$$\psi_X(t) = E(e^{tX}) = E(e^{(-t)(-X)}) \leq \left\{ E(e^{(-t_*)(-X)}) \right\}^{t/t_*} = \left\{ \psi_X(t_*) \right\}^{t/t_*}.$$

Somit existiert $\psi_X(t)$ für alle $t \in (t^-, t^+)$, also auch $\varphi_X(s)$ für alle $s \in (e^{t^-}, e^{t^+})$. Für $X \geq 0$ (mit Wahrscheinlichkeit 1) ist stets $e^{tX} \leq 1$ für $t \leq 0$ (mit Wahrscheinlichkeit 1), d.h. es folgt $t^- = -\infty$. Analog folgt $t^+ = +\infty$, falls $X \leq 0$ (mit Wahrscheinlichkeit 1).

- b) Für $|t| \leq \delta$ seien die Abbildungen $G_n(t, \bullet)$ definiert durch

⁷ Für Zufallsvariablen X, Y , für die $|X|^p$ und $|Y|^q$ integrierbar sind mit $\frac{1}{p} + \frac{1}{q} = 1$, $p, q > 0$, ist das Produkt $X \cdot Y$

absolut integrierbar, und es gilt (Hölder-Ungleichung) $E(|X \cdot Y|) \leq \left\{ E(|X|^p) \right\}^{1/p} \left\{ E(|Y|^q) \right\}^{1/q}$.

II. Sachversicherungsmathematik

$$G_n(t; x) := \sum_{k=0}^n \frac{x^k}{k!} t^k, \quad x \in \mathbb{R}, \quad n \in \mathbb{N}.$$

Für $n \in \mathbb{N}$ und $|t| \leq \delta$ ist dann

$$|G_n(t; X)| \leq e^{|tX|} \leq e^{\delta X} + e^{-\delta X} =: Z,$$

d.h. Z ist integrierbar mit $E(Z) = \psi_X(\delta) + \psi_X(-\delta)$. Ferner ist

$$|X|^n \leq \frac{n!}{\delta^n} G_n(\delta; |X|) \leq \frac{n!}{\delta^n} Z \quad \text{für alle } n \in \mathbb{N},$$

d.h. es existieren sämtliche Momente von X . Wegen

$$e^{tX} = \lim_{n \rightarrow \infty} G_n(t; X) \quad \text{für } |t| \leq \delta$$

sind die Voraussetzungen des Satzes von der *majorisierten Konvergenz*⁸ erfüllt, also gilt

$$\begin{aligned} \psi_X(t) &= E(e^{tX}) = \lim_{n \rightarrow \infty} E(G_n(t; X)) = \lim_{n \rightarrow \infty} E\left(\sum_{k=0}^n \frac{X^k}{k!} t^k\right) \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{E(X^k)}{k!} t^k = \sum_{k=0}^{\infty} \frac{E(X^k)}{k!} t^k, \quad |t| \leq \delta, \end{aligned}$$

wie behauptet. Die Differenzierbarkeit von ψ_X im Nullpunkt ergibt sich sofort durch gliedweise Differentiation dieser Reihe, mit $\psi_X^{(k)}(0) = E(X^k)$, $k \in \mathbb{N}$.

Die letzte Behauptung ergibt sich analog unter Beachtung der Reihenentwicklung (verallgemeinerte binomische Formel)

$$s^x = (1 + (s-1))^x = \sum_{k=0}^{\infty} \frac{\prod_{i=0}^{k-1} (x-i)}{k!} (s-1)^k, \quad |s-1| < 1.^9$$

c) Unter der angegebenen Bedingung lässt sich φ_X darstellen als

$$\varphi_X(s) = E(s^X) = \sum_{k=0}^{\infty} s^k P(X = k) = \sum_{k=0}^{\infty} \frac{\varphi_X^{(k)}(0)}{k!} s^k,$$

⁸ Ist Y eine integrierbare Zufallsvariable und gilt $|X_n| \leq Y$ für alle $n \in \mathbb{N}$, so folgt aus $X_n \rightarrow X$ (mit Wahrscheinlichkeit 1): $E(|X_n - X|) \rightarrow 0$ und damit auch $E(X_n) = E(X)$ für $n \rightarrow \infty$.

⁹ Für ganzzahlige x ist die Reihe endlich, da $\prod_{i=0}^{k-1} (x-i) = 0$ ist für $k > x$.

II. Sachversicherungsmathematik

wobei die Reihe wegen $\sum_{k=0}^{\infty} P(X = k) = 1$ in jedem Fall für $|s| \leq 1$ (absolut) konvergiert. Hieraus folgt die Behauptung.

- d) Mit X und Y sind auch die Zufallsvariablen e^{tX} und e^{tY} für alle $t \in \mathbb{R}$ stochastisch unabhängig, woraus sofort

$$\psi_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX} \cdot e^{tY}) = E(e^{tX})E(e^{tY}) = \psi_X(t)\psi_Y(t), \quad t \in I$$

folgt. Die andere Aussage ergibt sich völlig analog.

Die folgenden Tabellen geben einen Überblick über die erzeugenden Funktionen einiger für die Versicherungsmathematik wichtigen Verteilungen.

Diskrete Verteilungen:

P^X	Name	Zähldichte $f(k) = P(X = k)$	$\varphi_X(s)$	$E(X)$	$Var(X)$
U_n	Gleichverteilung	$\frac{1}{n}, k = 1, \dots, n$	$\frac{s \cdot s^n - 1}{n \cdot s - 1}$	$\frac{n+1}{2}$	$\frac{n^2 - 1}{12}$
$B(n, p)$	Binomialverteilung	$\binom{n}{k} p^k (1-p)^{n-k}, k = 0, \dots, n$	$(1-p + ps)^n$	np	$np(1-p)$
$NB(\beta, p)$	negative Binomialverteilung	$\binom{\beta+k-1}{k} p^k (1-p)^{\beta-k}, k = 0, 1, 2, \dots$	$\left(\frac{p}{1-(1-p)s}\right)^\beta$	$\beta \frac{1-p}{p}$	$\beta \frac{1-p}{p^2}$
$\mathcal{P}(\lambda)$	Poisson-Verteilung	$e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, \dots$	$e^{\lambda(s-1)}$	λ	λ

Man beachte, dass die negative Binomialverteilung $NB(\beta, p)$ auch für nicht-ganzzahlige Parameter $\beta > 0$ definiert ist; es gilt dann

$$\binom{\beta+k-1}{k} = \frac{(\beta+k-1)(\beta+k-2)\cdots\beta}{k!}, \quad k = 0, 1, 2, \dots$$

Als Spezialfall für $\beta = 1$ ergibt sich hieraus die geometrische Verteilung $\mathcal{G}(p)$.

Stetige Verteilungen:

P^X	Name	Dichte $f(x)$	$\psi_X(t)$	$E(X)$	$Var(X)$
$\mathcal{U}[a, b]$	Gleichverteilung	$\frac{1}{b-a}, a \leq x \leq b$	$\frac{e^{bt} - e^{at}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$\mathcal{E}(\lambda)$	Exponentialverteilung	$\lambda e^{-\lambda x}, x \geq 0$	$\frac{\lambda}{\lambda - t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\Gamma(\alpha, \lambda)$	Gamma-Verteilung	$\lambda^\alpha \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}, x > 0$	$\left(\frac{\lambda}{\lambda - t}\right)^\alpha$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
$\mathcal{N}(\mu, \sigma^2)$	Normalverteilung	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$\exp\left(\frac{\sigma^2 t^2}{2} + \mu t\right)$	μ	σ^2

II. Sachversicherungsmathematik

Die Exponentialverteilung ist ein Spezialfall der Gamma-Verteilung für $\alpha = 1$.

Einige der genannten Verteilungen besitzen spezielle Faltungseigenschaften; so gilt etwa

$$\begin{aligned}
 B(n, p) * B(m, p) &= B(n + m, p) && [n, m \in \mathbb{N}, 0 < p < 1] \\
 NB(\beta, p) * NB(\gamma, p) &= NB(\beta + \gamma, p) && [\beta, \gamma > 0, 0 < p < 1] \\
 \mathcal{P}(\lambda) * \mathcal{P}(\mu) &= \mathcal{P}(\lambda + \mu) && [\lambda, \mu > 0] \\
 \\
 \Gamma(\alpha, \lambda) * \Gamma(\delta, \lambda) &= \Gamma(\alpha + \delta, \lambda) && [\alpha, \delta, \lambda > 0] \\
 \mathcal{N}(\mu, \sigma^2) * \mathcal{N}(\nu, \tau^2) &= \mathcal{N}(\mu + \nu, \sigma^2 + \tau^2) && [\mu, \nu \in \mathbb{R}, \sigma^2, \tau^2 > 0].
 \end{aligned}$$

Satz 2. Im kollektiven Modell der Risikotheorie gilt: existiert die wahrscheinlichkeitserzeugende Funktion $\varphi_N(s)$ für $0 \leq s < \eta$ mit $\eta > 1$ und die momenterzeugende Funktion $\psi_X(t)$ der die Einzelschadenhöhen repräsentierenden Zufallsvariablen X für $0 \leq t < \delta$ mit $\delta > 0$, so gilt

$$\psi_S(t) = \varphi_N(\psi_X(t)), \quad t \in I,$$

wobei I ein geeignetes, die Null enthaltendes Intervall bezeichnet mit der Eigenschaft, dass $\psi_X(I) \subseteq [0, \eta]$ ist. Ist X diskret mit Werten in \mathbb{N} , so gilt auch

$$\varphi_S(t) = \varphi_N(\varphi_X(t)), \quad t \in e^I \cup [0, 1].$$

Insbesondere existieren alle Momente des Gesamtschadens S , und speziell gilt

$$E(S) = E(N) \cdot E(X), \quad \text{Var}(S) = E(N) \cdot \text{Var}(X) + \text{Var}(N) \cdot \{E(X)\}^2.$$

Beweis: Wir wählen $\tau > 0$ so klein, dass $0 \leq \psi_X(\tau) < \eta$ ausfällt; das ist wegen $\lim_{t \rightarrow 0} \psi_X(t) = 1$ nach dem Satz von der majorisierten Konvergenz möglich. Mit der Wahl $I \subseteq (-\infty, \tau)$ ist ein geeignetes Intervall I gefunden, da nach Voraussetzung über die Einzelschadenhöhen $X > 0$ gewählt werden kann. Für $t \in I$ folgt nun aufgrund der gemachten Unabhängigkeitsvoraussetzungen:

$$\begin{aligned}
 \psi_S(t) &= E \left(\exp \left(t \sum_{k=1}^N X_k \right) \right) = E \left(\prod_{k=1}^N e^{tX_k} \right) = P(N=0) + \sum_{n=1}^{\infty} P(N=n) \cdot E \left(\prod_{k=1}^n e^{tX_k} \right) \\
 &= P(N=0) + \sum_{n=1}^{\infty} P(N=n) \cdot \prod_{k=1}^n E(e^{tX_k}) = P(N=0) + \sum_{n=1}^{\infty} P(N=n) \cdot \{\psi_X(t)\}^n \\
 &= \sum_{n=0}^{\infty} P(N=n) \cdot \{\psi_X(t)\}^n = \varphi_N(\psi_X(t)).
 \end{aligned}$$

Bei diskretem X ergibt sich entsprechend nach Satz 1

$$\varphi_S(t) = \psi_S(\ln t) = \varphi_N(\psi_X(\ln t)) = \varphi_N(\varphi_X(t)), \quad t \in e^I.$$

II. Sachversicherungsmathematik

Für $0 \leq t \leq 1$ ist ferner auch $0 \leq \varphi_X(t) = E(t^X) \leq 1$, so dass $\varphi_S(t)$ auch für $t \in [0,1]$ existiert.

Die Existenz der Momente ist nach Satz 1 unter den gemachten Annahmen gesichert; speziell ist

$$\begin{aligned} E(S) &= \psi'_S(0) = \varphi'_N(\psi_X(0)) \cdot \psi'_X(0) = \varphi'_N(1) \cdot \psi'_X(0) = E(N) \cdot E(X), \\ \text{Var}(S) &= \psi''_S(0) - \{\psi'_S(0)\}^2 = \varphi''_N(\psi_X(0)) \cdot \{\psi'_X(0)\}^2 + \varphi'_N(\psi_X(0)) \cdot \psi''_X(0) - \{E(S)\}^2 \\ &= \varphi''_N(1) \cdot \{E(X)\}^2 + \varphi'_N(1) \cdot \psi''_X(0) - \{E(S)\}^2 \\ &= E(N(N-1)) \cdot \{E(X)\}^2 + E(N) \cdot E(X^2) - \{E(N)E(X)\}^2 \\ &= [E(N^2) - \{E(N)\}^2] \cdot \{E(X)\}^2 + E(N) \cdot [E(X^2) - \{E(X)\}^2] \\ &= \text{Var}(N) \cdot \{E(X)\}^2 + E(N) \cdot \text{Var}(X), \end{aligned}$$

wie behauptet.

Ohne die (einschränkenden) Voraussetzungen an die Existenz der erzeugenden Funktionen lässt sich mit anderen Mitteln noch zeigen, dass diese letzte Beziehung auch dann gültig bleibt, wenn lediglich die Existenz der Varianzen (also der ersten beiden Momente) von N und X gesichert ist.

Für die Gesamtschadenverteilung aus Lemma 10 ergibt sich so etwa

$$\begin{aligned} E(S) &= E(N) \cdot E(X) = \frac{1-p}{\lambda p}, \\ \text{Var}(S) &= E(N) \cdot \text{Var}(X) + \text{Var}(N) \cdot \{E(X)\}^2 = \frac{1-p^2}{\lambda^2 p^2}. \end{aligned}$$

Mit Hilfe von Satz 2 lässt sich Lemma 10 auch insgesamt alternativ beweisen. Für die momenterzeugende Funktion des Gesamtschadens erhalten wir dort nämlich

$$\begin{aligned} \psi_S(t) &= \varphi_N(\psi_X(t)) = \frac{p}{1-q \cdot \psi_X(t)} = \frac{p}{1-q \cdot \frac{\lambda}{\lambda-t}} = p \frac{\lambda-t}{\lambda-q\lambda-t} = p \frac{\lambda-q\lambda-t+q\lambda}{\lambda-q\lambda-t} \\ &= p + q \frac{p\lambda}{p\lambda-t} = p + q \cdot \psi_Y(t), \quad t < \lambda, \end{aligned}$$

d.h. mit $0 \in I = (-\infty, \lambda)$ ergibt sich eindeutig die momenterzeugende Funktion einer Mischverteilung aus ε_0 (mit Gewicht p) und einer Exponentialverteilung $\mathcal{E}(\lambda p)$ (mit Gewicht $q = 1 - p$).

Satz 1 erlaubt eine aus numerischer Sicht interessante Anwendung auf die Berechnung der Gesamtschadenverteilung, wenn man die (als stetig verteilt angenommenen) Einzelschadenhöhen geeignet diskretisiert, z.B. durch Aufrundung auf Vielfache von 1000 € o.ä. Bezeichnet man diese "Schrittweite" mit $\Delta > 0$, so wird aus der zu Grunde liegenden Zufallsvariablen X für die Einzelschadenhöhen die diskretisierte Zufallsvariable

II. Sachversicherungsmathematik

$$X_{\Delta} := \min \{k \in \mathbb{N} \mid k\Delta \geq X\}.$$

Der hieraus entstehende "Gesamtschaden" S_{Δ} besitzt nun die wahrscheinlichkeitserzeugende Funktion

$$\varphi_{S_{\Delta}}(s) = \varphi_N(\varphi_{X_{\Delta}}(s)), \quad |s| \leq 1.$$

Durch Entwicklung dieser Funktion in eine Taylorreihe um den Nullpunkt erhält man die Wahrscheinlichkeiten $P(S_{\Delta} = k)$ für alle $k = 0, 1, 2, \dots$ über die Koeffizienten von s^k zurück. Damit lässt sich die Gesamtschadenverteilung – zumindest näherungsweise – auch in solchen Fällen numerisch berechnen, in denen ein geschlossener Ausdruck nicht zur Verfügung steht.

Beispiel 13. Wir nehmen an, dass die Schadenszahl N Poisson-verteilt ist mit Parameter $\lambda = 3$ und die diskretisierten Einzelschadenhöhen einer $NB(\beta, p)$ -Verteilung mit den Parametern $\beta = 10$ und $p = 0,3$ genügen mit $\Delta = 10000$. Die zugehörige wahrscheinlichkeitserzeugende Funktion lautet dann

$$\begin{aligned} \varphi_{S_{\Delta}}(s) &= \varphi_N(\varphi_{X_{\Delta}}(s)) = \exp(\lambda(\varphi_{X_{\Delta}}(s) - 1)) = \exp\left(\lambda\left(\left(\frac{p}{1 - (1-p)s}\right)^{\beta} - 1\right)\right) \\ &= \exp\left(3\left(\left(\frac{0,3}{1 - 0,7s}\right)^{10} - 1\right)\right), \quad |s| \leq 1 \end{aligned}$$

mit $E(S_{\Delta}) = 70$, $Var(S_{\Delta}) = 1866\frac{2}{3}$, was eine zugehörige Standardabweichung von 43,20 ergibt.

Der erwartete Gesamtschaden $E(S)$ liegt also in der Größenordnung von 700000 €

Mit Hilfe eines Computer-Algebra-Systems (z.B. MAPLE) erhält man dann folgende Reihenentwicklung bis zum 100. Term:

II. Sachversicherungsmathematik

```

> phi:=s->exp(3*((0.3/(1-0.7*s))^10-1));
                                (0.0000177147
                                (1-0.7s)^10 - 3)
                                phi = s -> e
> series(phi(s), s=0, 101);
0.04978795035 + 0.6173850228 10^-5 s + 0.00002376970617 s^2 + 0.00006655705296 s^3 + 0.0001514245168 s^4 + 0.0002968153065 s^5 +
0.0005194918976 s^6 + 0.0008313498184 s^7 + 0.001237003202 s^8 + 0.001732582249 s^9 + 0.002305858932 s^10 + 0.002937543289 s^11 +
0.003603415190 s^12 + 0.004276882791 s^13 + 0.004931572676 s^14 + 0.005543629210 s^15 + 0.006093502837 s^16 + 0.006567114180 s^17 +
0.006956376067 s^18 + 0.007259129673 s^19 + 0.007478599698 s^20 + 0.007622499471 s^21 + 0.007701922534 s^22 + 0.007730148488 s^23 +
0.007721472041 s^24 + 0.007690140424 s^25 + 0.007649458391 s^26 + 0.007611095919 s^27 + 0.007584611513 s^28 + 0.007577186292 s^29 +
0.007593550148 s^30 + 0.007636071960 s^31 + 0.007704979867 s^32 + 0.007798675945 s^33 + 0.007914109448 s^34 + 0.008047176189 s^35 +
0.008193115416 s^36 + 0.008346881009 s^37 + 0.008503469341 s^38 + 0.008658191745 s^39 + 0.008806885286 s^40 + 0.008946059709 s^41 +
0.009072983026 s^42 + 0.009185710976 s^43 + 0.009283068211 s^44 + 0.009364590254 s^45 + 0.009430436048 s^46 + 0.009481281007 s^47 +
0.009518199698 s^48 + 0.009542546379 s^49 + 0.009555840454 s^50 + 0.009559662147 s^51 + 0.009555562577 s^52 + 0.009544990505 s^53 +
0.009529237099 s^54 + 0.009509398468 s^55 + 0.009486355191 s^56 + 0.009460766838 s^57 + 0.009433079345 s^58 + 0.009403542654 s^59 +
0.009372235901 s^60 + 0.009339097300 s^61 + 0.009303956890 s^62 + 0.009266569181 s^63 + 0.009226644612 s^64 + 0.009183877878 s^65 +
0.009137972626 s^66 + 0.009088661281 s^67 + 0.009035720429 s^68 + 0.008978981150 s^69 + 0.008918335000 s^70 + 0.008853735812 s^71 +
0.008785198167 s^72 + 0.008712793010 s^73 + 0.008636641125 s^74 + 0.008556905329 s^75 + 0.008473781732 s^76 + 0.008387490845 s^77 +
0.008298269115 s^78 + 0.008206360718 s^79 + 0.008112010740 s^80 + 0.008015459089 s^81 + 0.007916935850 s^82 + 0.007816657735 s^83 +
0.007714825708 s^84 + 0.007611623816 s^85 + 0.007507218822 s^86 + 0.007401760665 s^87 + 0.007295383595 s^88 + 0.007188207709 s^89 +
0.007080340780 s^90 + 0.006971880219 s^91 + 0.006862915187 s^92 + 0.006753528337 s^93 + 0.006643797726 s^94 + 0.006533798280 s^95 +
0.006423603032 s^96 + 0.006313284221 s^97 + 0.006202913840 s^98 + 0.006092564225 s^99 + 0.005982308276 s^100 + O(s^101)

```

Aufgrund der hohen Standardabweichung werden also relativ viele Terme benötigt, um die Gesamtschadenverteilung einigermaßen realistisch abzubilden. In diesem Beispiel ist es sicher empfehlenswert, Terme bis zur Ordnung 200 oder mehr zu verwenden. Mit demselben MAPLE-Worksheet ergibt sich etwa

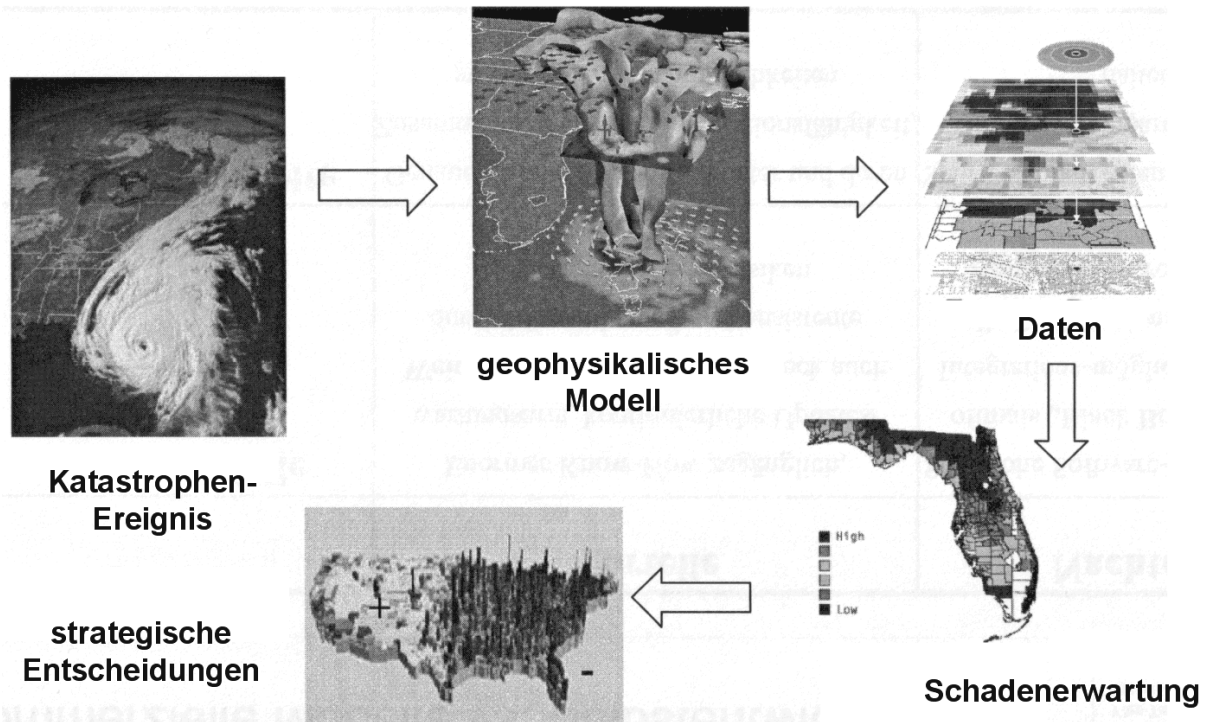
$$\varphi_{S_{\Delta}}(s) = \dots + 0,00028787 s^{200} + \dots + 0,00000281 s^{300} + \dots$$

Weitere numerische Verfahren zur Berechnung der Gesamtschadenverteilung werden im nachfolgenden Abschnitt vorgestellt.

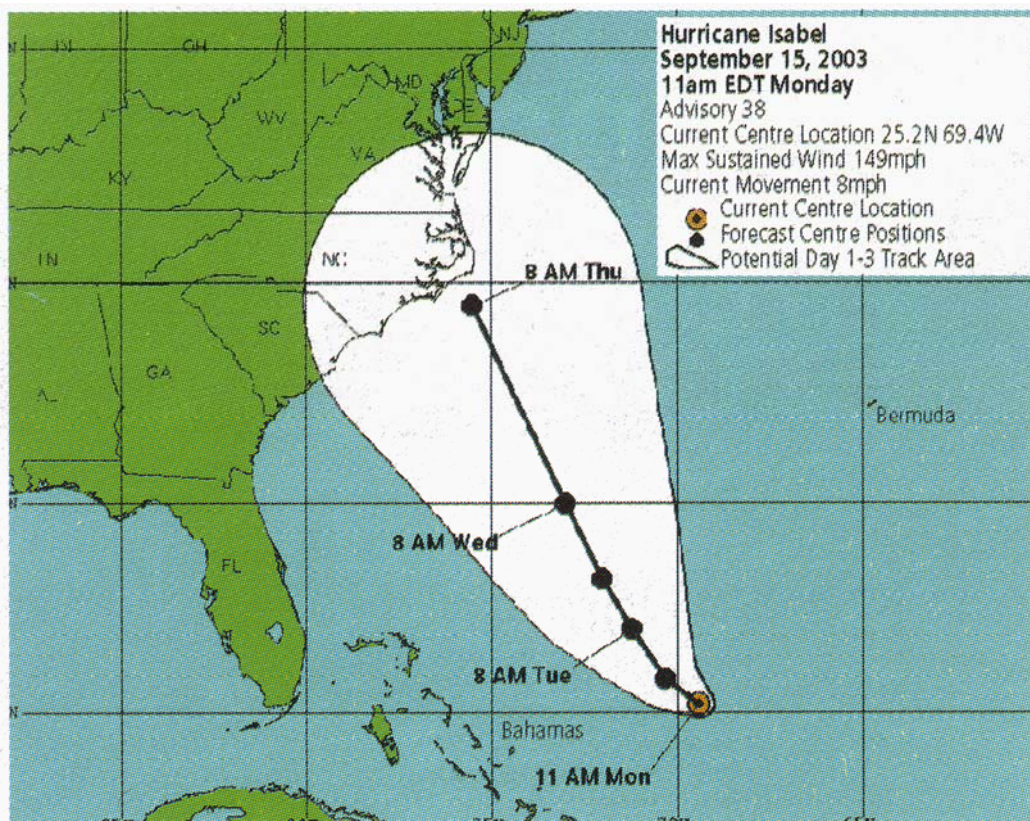
Ähnlich wie in der Personenversicherungsmathematik lassen sich auch in der Sachversicherungsmathematik (Netto-)Prämien als Erwartungswert $E(S)$ des (Jahres-)Gesamtschadens berechnen. Die Sätze 1 und 2 zeigen, wie dies konkret in einem kollektiven Modell der Risikotheorie geschehen kann.

Wir wollen zum Ende dieses Abschnitts noch auf eine wichtige Anwendung von Satz 2 in der Naturkatastrophenmodellierung eingehen (vgl. DONG (2001)). Ein beträchtlicher Teil der weltweit verursachten Sachschäden geht auf Naturgefahren wie Stürme, Überschwemmungen oder Erdbeben zurück. Dies ist ein Grund dafür, dass sowohl große Rückversicherer als auch unabhängige Consulting-Firmen mit geophysikalischen Modellen versuchen, die möglichen Risiken "in vitro", also auf dem Computer zu simulieren, um damit die Gefährdungspotenziale konkreter Portfolios der Erstversicherer bestimmen zu können. Datengrundlage bilden zunächst die historischen Ereignisse mit ihren bekannten Parametern wie Windgeschwindigkeiten und -richtungen, Wasserständen oder Magnituden bei Erdbeben (sog. *Event Set*). Durch zufällige Perturbationen dieser Parameter wird die Datenbasis geeignet vergrößert, so dass sich bis zu 20000 und mehr "Ereignisse" ergeben (sog. *Stochastic Event Set*). Die folgende Graphik zeigt das prinzipielle Vorgehen bei solchen geophysikalischen Modellierungen.

II. Sachversicherungsmathematik



Die folgende Graphik zeigt eine entsprechende Vorausberechnung der Sturmbahn des Hurricanes Isabel vor der Ostküste der USA¹⁰.



¹⁰ Quelle: **reinsurance**, November 2003, S. 53

II. Sachversicherungsmathematik

Als Ergebnis von solchen Portfolio-Analysen stellen die kommerziellen Produkte eine sog. *Event Loss Table* (Ereignisschaden-Tabelle) bereit, die etwa folgendermaßen aussieht (exemplarisch sind hier nur 7 Zeilen der Tabelle wiedergegeben):

Analysis Name	Scenario	Modelled Loss	Standard Deviation	Exposed SI	Rate
Example Wind Analysis	3656	1,940,550,920	36,794,128	68,947,100,000	0.0000062953
Example Wind Analysis	3968	1,563,781,833	49,352,347	95,221,396,000	0.0000129744
Example Wind Analysis	7264	1,482,396,982	41,468,066	69,668,353,333	0.0000113048
Example Wind Analysis	7219	1,461,229,040	43,029,488	72,023,880,000	0.0000113048
Example Wind Analysis	3665	1,431,950,171	47,062,942	73,402,510,667	0.0000047371
Example Wind Analysis	7222	1,332,616,058	40,221,122	78,780,377,333	0.0000113048
Example Wind Analysis	6283	1,169,279,403	35,134,601	74,784,286,000	0.0000468744

Die unter "Scenario" angegebenen Nummern beziehen sich auf diejenigen Ereignisse im *Stochastic Event Set*, die Auswirkungen auf das zu Grunde liegende Portfolio haben (können). (Die vorige Graphik zeigt, dass z.B. Gebäude in Florida von dem Sturm-Ereignis nicht betroffen sind, während Gebäude an der Atlantikküste der Staaten Virginia und North Carolina mit hoher Wahrscheinlichkeit betroffen sind.) "Exposed S(um) I(nsured)" bezeichnet hierbei die dem Ereignis mit der betreffenden Szenarien-Nummer ausgesetzte Versicherungssumme (das ist im Sinne der vorigen Bemerkung nicht immer das gesamte Portfolio). Mit "Modelled Loss" wird der erwartete Verlust bezeichnet, mit "Standard Deviation" die zugehörige Standardabweichung. In der Regel werden – je nach Anbieter unterschiedliche – Verteilungstypen für den Verlust angegeben, z.B. Lognormal- oder Betaverteilungen, auf die sich diese Parameter beziehen. In den meisten Fällen wird für die Anzahl der (Sturm-)Ereignisse je Jahr eine Poisson-Verteilung angenommen, manchmal auch eine negative Binomialverteilung. Mit "Rate" ist der Erwartungswert der Anzahl der Ereignisse im Poisson-Modell gemeint, was in diesem Fall mit dem Parameter der Verteilung identisch ist.

Wenn n die Anzahl der relevanten Szenarien (= Anzahl der Zeilen in der Tabelle) bezeichnet und wir mit N_1, N_2, \dots, N_n die (als voneinander unabhängig angenommenen) Ereignisfrequenzen sowie mit X_{ij} , $1 \leq i \leq n$, $j \in \mathbb{N}$ die (ebenfalls als unabhängig angenommenen) positiven Einzelschadenshöhen bezeichnen, wobei alle X_{ij} dieselbe Verteilung Q_i besitzen mögen, so erhalten wir folgende Summenschäden:

$$S_i := \sum_{j=1}^{N_i} X_{ij}, \quad i = 1, \dots, n \quad (\text{Ereignis-Gesamtschaden, Scenario Loss})$$

$$S := \sum_{i=1}^n S_i = \sum_{i=1}^n \sum_{j=1}^{N_i} X_{ij} \quad (\text{Jahres-Gesamtschaden, Aggregate Loss}).$$

Aus Sicht der Rückversicherung, die wir später noch ausführlicher besprechen werden, sind beide Schadenssummen von Interesse. Für den Ereignis-Gesamtschaden erhalten wir also gerade für jedes i ein kollektives Modell, so dass diese Gesamtschadenverteilung mit Lemma 9 oder mit Hilfe von Satz 2 berechnet werden kann über die erzeugenden Funktionen. Die Jahres-Gesamtschadenverteilung ist demnach die Faltung der Ereignis-Gesamtschadenverteilungen. Für das in der Praxis meist verwendete Poisson-Modell lässt sich dies aber geschickt vereinfachen, so dass sich auch für den Jahres-Gesamtschaden ein (äquivalentes) kollektives Modell ergibt. Dieser Ansatz wird in der Praxis auch tatsächlich verwendet (vgl. DONG (2001), Part I).

II. Sachversicherungsmathematik

Zur Vorbereitung benötigen wir noch zwei Hilfsresultate.

Lemma 11. Es seien Z_1, \dots, Z_n (nicht notwendig stochastisch unabhängige) Zufallsvariablen mit Verteilungen Q_1, \dots, Q_n . Ferner sei J eine von den Z_i stochastisch unabhängige Zufallsvariable mit Werten in der Menge $\{1, \dots, n\}$. Dann ist auch Z_J eine Zufallsvariable, deren Verteilung durch die folgende *Mischung* der Verteilungen Q_1, \dots, Q_n gegeben ist:

$$P^{Z_J} = \sum_{i=1}^n P(J = i) Q_i.$$

Beweis: Zunächst gilt

$$\{Z_J \in B\} = \bigcup_{i=1}^n \{J = i\} \cap \{Z_i \in B\} \text{ für jede Borel-Menge } B \in \mathcal{B}.$$

Dies zeigt, dass Z_J eine Zufallsvariable (d.h. messbar) ist. Für die Verteilung folgt entsprechend:

$$P(Z_J \in B) = P\left(\bigcup_{i=1}^n \{J = i\} \cap \{Z_i \in B\}\right) = \sum_{j=1}^n P(\{J = i\} \cap \{Z_i \in B\}) = \sum_{i=1}^n P(J = i) P(Z_i \in B)$$

für jede Borel-Menge $B \in \mathcal{B}$, woraus die Behauptung folgt.

Bemerkung: Die Aussage von Lemma 11 bleibt auch für eine *Folge* von Zufallsvariablen richtig; J nimmt dann entsprechend Werte in \mathbb{N} an.

Eine Mischung von Verteilungen Q_1, \dots, Q_n kann also offensichtlich dadurch erzeugt werden, dass zunächst ein Index i zufällig durch die Zufallsvariable J ausgewählt und anschließend die entsprechende Zufallsvariable Z_i realisiert wird.

Für die zugehörigen Verteilungsfunktionen F_1, \dots, F_n bzw. im Fall der Existenz der Dichten f_1, \dots, f_n gilt analog:

$$F_{Z_J}(x) = \sum_{i=1}^n P(J = i) F_i(x)$$
$$f_{Z_J}(x) = \sum_{i=1}^n P(J = i) f_i(x), \quad x \in \mathbb{R}.$$

Lemma 12. In der Situation von Lemma 11 gilt: Falls die Z_i momenterzeugende Funktionen auf einem gemeinsamen Intervall I besitzen, existiert dort auch die momenterzeugende Funktion von Z_J , und es gilt

$$\psi_{Z_J}(t) = \sum_{i=1}^n P(J = i) \psi_{Z_i}(t), \quad t \in I.$$

II. Sachversicherungsmathematik

Beweis: Es ist aufgrund der Unabhängigkeitsannahme

$$\psi_{Z_j}(t) = E(e^{tZ_j}) = \sum_{i=1}^n P(J=i) \underbrace{E(e^{tZ_j} | J=i)}_{=E(\exp\{tZ_i\})} = \sum_{i=1}^n P(J=i) \psi_{Z_i}(t), \quad t \in I.$$

Wir kommen jetzt zu dem angekündigten Satz.

Satz 3. Es seien N_1, N_2, \dots, N_n stochastisch unabhängige, Poisson-verteilte Zufallsvariablen mit Parametern $\lambda_1, \lambda_2, \dots, \lambda_n > 0$. Ferner seien X_{ij} , $1 \leq i \leq n$, $j \in \mathbb{N}$ untereinander und auch von den N_i stochastisch unabhängige, positive Zufallsvariablen (Schäden), so dass alle $X_{i\cdot}$ dieselbe Verteilung Q_i besitzen. Dann kann die Verteilung des Gesamtschadens

$$S := \sum_{i=1}^n \sum_{j=1}^{N_i} X_{ij}$$

alternativ auch über ein kollektives Modell dargestellt werden, d.h. S und die Summe

$$\tilde{S} := \sum_{k=1}^{\tilde{N}} \tilde{X}_k$$

besitzen dieselbe Verteilung, wobei \tilde{N} eine Poisson-verteilte Zufallsvariable ist mit Parameter

$$\tilde{\lambda} = \sum_{i=1}^n \lambda_i$$

und die \tilde{X}_i (auch von \tilde{N}) unabhängige positive Zufallsvariablen sind mit der Verteilung \tilde{Q} , die eine Mischung der Q_i darstellt mit

$$\tilde{Q} = \sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}} Q_i.$$

II. Sachversicherungsmathematik

Beweis: Wir verwenden hier erzeugende Funktionen; die Aussage ist aber auch allgemeiner gültig. Es ist nach Satz 2 und Lemma 12

$$\begin{aligned}\psi_{S_i}(t) &= \varphi_{N_i}(\psi_{X_i}(t)) = \exp\{\lambda_i(\psi_{X_i}(t) - 1)\} \text{ und} \\ \psi_S(t) &= \prod_{i=1}^n \psi_{S_i}(t) = \exp\left\{\sum_{i=1}^n \lambda_i(\psi_{X_i}(t) - 1)\right\} = \exp\left\{\tilde{\lambda}\left[\sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}}(\psi_{X_i}(t) - 1)\right]\right\} \\ &= \exp\left\{\tilde{\lambda}\left[\sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}}\psi_{X_i}(t) - 1\right]\right\} = \exp\{\tilde{\lambda}(\psi_{\tilde{X}}(t) - 1)\} = \varphi_{\tilde{N}}(\psi_{\tilde{X}}(t))\end{aligned}$$

für $t \in I$ geeignet. Hierzu ist nur zu beachten, dass die Zahlen $\frac{\lambda_i}{\tilde{\lambda}}$ als Wahrscheinlichkeiten $P(J = i)$ für eine diskrete Zufallsvariable J wie in Lemmata 11 und 12 gefordert aufgefasst werden können.

Neben dem jeweiligen Ereignis-Gesamtschaden ist manchmal auch der Ereignis-*Maximalschaden* $M_i := \max\{X_{ij} \mid 1 \leq j \leq N_i\}$ von Interesse (bei DONG (2001) irreführender Weise mit *Occurrence Loss* bezeichnet). Auch seine Verteilung kann in einem Poisson-Modell recht einfach berechnet werden.

Lemma 13. Im kollektiven Modell der Risikotheorie sei mit

$$M := \max\{X_i \mid 1 \leq i \leq N\}$$

der *Maximalschaden* bezeichnet. Dann gilt:

$$P(M \leq z) = F_M(z) = \varphi_N(F(z)) = \begin{cases} \sum_{n=0}^{\infty} p_n F^n(z), & \text{falls } z \geq 0 \\ 0, & \text{falls } z < 0. \end{cases}$$

Beweis: Es ist ähnlich dem Beweis von Lemma 9 (vgl. Lemma 7)

$$\begin{aligned}P(M \leq z) &= F_M(z) = P\left(\bigcup_{n=0}^{\infty} \{N = n\} \cap \{\max\{X_k \mid 1 \leq k \leq n\} \leq z\}\right) = p_0 + \sum_{n=1}^{\infty} p_n P\left(\bigcap_{k=1}^n \{X_k \leq z\}\right) \\ &= p_0 + \sum_{n=1}^{\infty} p_n F^n(z) = \sum_{n=0}^{\infty} p_n F^n(z) = \varphi_N(F(z)), \quad z \in \mathbb{R}.\end{aligned}$$

Der Unterschied zwischen der Verteilung des Gesamt- und des Maximalschadens ist also lediglich der, dass im ersten Fall *Faltungspotenzen*, im zweiten Fall *gewöhnliche* Potenzen der Verteilungsfunktionen auftreten.

II. Sachversicherungsmathematik

Bemerkung: Für das Poisson-Modell, d.h. $P^N = \mathcal{P}(\lambda)$ mit $\lambda > 0$ bedeutet dies gerade:

$$P(M \leq z) = F_M(z) = \varphi_N(F(z)) = e^{\lambda\{F(z)-1\}} = e^{-\lambda\{1-F(z)\}} \quad \text{für } z \geq 0,$$

das an die entsprechende Formel für die momenterzeugende Funktion des Gesamtschadens erinnert.

Für die *Event Loss Table* bedeutet dies, dass der maximale Ereignisschaden M_i die folgende Verteilungsfunktion besitzt:

$$P(M_i \leq z) = F_{M_i}(z) = e^{-\lambda_i\{1-F_i(z)\}}, \quad z \geq 0.$$

Hiermit lässt sich analog zu Lemma 7 sofort auch die Verteilungsfunktion des Jahres-Maximalschadens $M = \max\{M_i \mid 1 \leq i \leq n\}$ (wegen der angenommenen Unabhängigkeit aller Einzelschäden) angeben:

$$\begin{aligned} P(M \leq z) &= P\left(\bigcap_{i=1}^n \{M_i \leq z\}\right) = \prod_{i=1}^n P(M_i \leq z) = \prod_{i=1}^n e^{-\lambda_i\{1-F_i(z)\}} = \exp\left\{-\sum_{i=1}^n [\lambda_i\{1-F_i(z)\}]\right\} \\ &= \exp\left\{-\tilde{\lambda}[1-\tilde{F}(z)]\right\}, \quad z \geq 0, \end{aligned}$$

wobei

$$\tilde{F}(z) = \sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}} F_i(z), \quad z \in \mathbb{R}$$

die Verteilungsfunktion zur Mischverteilung \tilde{Q} bezeichnet.

Wenn man davon ausgeht, dass ein Schaden nicht wesentlich höher ausfallen kann als die Versicherungssumme des versicherten Objekts, dann besitzen die Verteilungen Q_i bzw. die Verteilungsfunktionen F_i explizit bekannte endliche rechte Endpunkte (vgl. Abschnitt I. 2)

$$\varpi_i = \sup\{t \in \mathbb{R} \mid F_i(t) < 1\}, \quad i = 1, \dots, n.$$

Für die nachfolgenden Überlegungen wollen wir jetzt o.B.d.A. annehmen, dass diese strikt der Größe nach geordnet sind, d.h. dass gilt

$$\varpi_1 < \varpi_2 < \dots < \varpi_n.$$

Es ist dann $F_i(\varpi_k) = 1$ für $1 \leq i \leq k$, so dass hieraus

$$P(\tilde{X} \leq \varpi_k) = \tilde{F}(\varpi_k) = \sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}} F_i(\varpi_k) = \sum_{i=1}^k \frac{\lambda_i}{\tilde{\lambda}} + \sum_{i=k+1}^n \frac{\lambda_i}{\tilde{\lambda}} F_i(\varpi_k) \geq \sum_{i=1}^k \frac{\lambda_i}{\tilde{\lambda}}, \quad k = 1, \dots, n$$

folgt. Für die Verteilung des Jahres-Maximalschadens M bedeutet dies:

II. Sachversicherungsmathematik

$$P(M > \varpi_k) = 1 - \exp\left\{-\tilde{\lambda}\left[1 - \tilde{F}(\varpi_k)\right]\right\} \leq 1 - \exp\left\{-\sum_{i=k+1}^n \lambda_i\right\}, \quad k = 1, \dots, n.$$

Nimmt man zur Vereinfachung und in einer ersten Näherung an die „Wirklichkeit“ an, dass die Einzelschäden einer Einpunktverteilung in ϖ_i genügen, so gilt sogar noch $F_i(\varpi_k) = 0$ für $i > k$, so dass dann

$$P(\tilde{X} \leq \varpi_k) = \tilde{F}(\varpi_k) = \sum_{i=1}^k \frac{\lambda_i}{\tilde{\lambda}}, \quad k = 1, \dots, n$$

folgt mit

$$P(M > \varpi_k) = 1 - \exp\left\{-\tilde{\lambda}\left[1 - \tilde{F}(\varpi_k)\right]\right\} = 1 - \exp\left\{-\sum_{i=k+1}^n \lambda_i\right\}, \quad k = 1, \dots, n.$$

Die sich so ergebende Funktion wird in der geophysikalischen Modellierungswelt auch als OEP curve (Occurrence Loss Exceeding Probability) bezeichnet, obwohl dies nach den Bemerkungen vor Lemma 13 missverständlich ist (vgl. DONG (2001), S. 15). Die dahinter stehende Abschätzung der möglichen Schäden durch deterministische Einzelschäden wird in der dortigen Literatur auch als *ignorance of secondary uncertainties* bezeichnet.

Der Vorteil dieser Form der OEP curve besteht in ihrer Einfachheit, allerdings müssen dazu vorher in der *Event Loss Table* die entsprechenden ϖ_i der Größe nach geordnet werden.

Für den Jahres-Gesamtschaden S liefern Lemma 9 und Satz 3 für das geophysikalische Poisson-Modell entsprechend

$$P(S \leq z) = F_S(z) = e^{-\tilde{\lambda}} + e^{-\tilde{\lambda}} \sum_{k=1}^n \frac{\tilde{\lambda}^k}{k!} \tilde{F}^{k*}(z), \quad z \geq 0.$$

Die sich hieraus ergebende Abbildung $P(S > \varpi_k) = 1 - F_S(\varpi_k)$ für $k = 1, \dots, n$ heißt analog AEP curve (Aggregate Loss Exceeding Probability) (vgl. DONG (2001), S. 16). Bei geeigneter Diskretisierung im Sinne von Beispiel 13 kann eine Berechnung der zugehörigen Wahrscheinlichkeiten wieder über die wahrscheinlichkeitserzeugende Funktion φ_S (etwa mit der dort beschriebenen Reihenentwicklung) erfolgen. Aufgrund der in der Praxis recht umfangreichen *Event Loss Tables* sind aber die im nachfolgenden Abschnitt besprochenen numerischen Verfahren in der Regel effizienter. Alternativ ist eine statistische Bestimmung der AEP curve mittels Monte-Carlo-Simulation möglich, indem eine große Zahl von zufälligen Beobachtungen aus dem Modell gezogen und die zugehörige *empirische Verteilung* bestimmt wird. Allerdings sind hier je nach Parameter-Situation bis zu einigen 100000 Stichproben notwendig.

Die obigen Formeln zeigen, dass der Einbezug von *secondary uncertainties*, also der Berücksichtigung von tatsächlich zufälligen Einzelschäden, ausgedrückt durch die X_i bzw. durch \tilde{X} , nichts Wesentliches an der Darstellung ändert. Wir fassen diese allgemeineren Darstellungen hier noch einmal abschließend zusammen (vgl. DONG (2001), S. 46ff):

II. Sachversicherungsmathematik

$$P(M > z) = 1 - \exp\left\{-\tilde{\lambda}[1 - \tilde{F}(z)]\right\}, \quad z \geq 0: \quad \text{OEP curve (Occurrence Loss Exceeding Probability)}$$

$$P(S > z) = 1 - e^{-\tilde{\lambda}} \left\{1 + \sum_{k=1}^n \frac{\tilde{\lambda}^k}{k!} \tilde{F}^{k*}(z)\right\}, \quad z \geq 0: \quad \text{AEP curve (Aggregate Loss Exceeding Probability).}$$

Hierbei ist wie oben

$$\tilde{\lambda} = \sum_{i=1}^n \lambda_i \quad \text{und} \quad \tilde{F}(z) = \sum_{i=1}^n \frac{\lambda_i}{\tilde{\lambda}} F_i(z), \quad z \in \mathbb{R}.$$

Tabellen mit den notwendigen erweiterten Angaben (wie Erwartungswerte und Standard-Abweichungen für die Einzelschäden) heißen entsprechend *Extended Event Loss Tables*.

II. 2. Die Panjer-Rekursion

In den beiden folgenden Abschnitten wollen wir weiterhin das kollektive Modell unterstellen und dabei annehmen, dass die Einzelschadenhöhenverteilung wieder mit einer "Schrittweite" $\Delta > 0$ diskretisiert ist. Für umfangreichere Probleme – etwa wie im Fall der geophysikalischen Modellierung im vorigen Abschnitt – ist dann aber die Reihenentwicklung der wahrscheinlichkeitserzeugenden Funktion $\varphi_{S_\Delta}(s) = \varphi_N(\varphi_{X_\Delta}(s))$, $|s| \leq 1$ um den Nullpunkt nicht effizient. Gehört die Verteilung von N jedoch zur Familie der Poisson-, Binomial- oder negativen Binomialverteilungen, so lässt sich die zu S_Δ gehörige Verteilung auf elegante und effiziente Weise rekursiv berechnen. Wir verwenden für diesen Fall die folgende Darstellung der wahrscheinlichkeitserzeugenden Funktion:

$$\varphi_N(s) = \begin{cases} \left(\frac{1-A}{1-As}\right)^{\frac{B}{A}} & \text{für } A, B \neq 0 \\ e^{B(s-1)} & \text{für } A = 0, B \neq 0 \end{cases}$$

mit geeigneten reellen Zahlen $A, B \in \mathbb{R}$. Durch Vergleich mit der Tabelle auf S. 59 ergibt sich:

$$\begin{aligned} \varphi_N(s) &= e^{\lambda(s-1)} && \text{für } A = 0, B = \lambda > 0 \\ \varphi_N(s) &= ((1-p) + ps)^n && \text{für } A < 0, p := \frac{A}{A-1}, B := -nA \\ \varphi_N(s) &= \left(\frac{p}{1-(1-p)s}\right)^\beta && \text{für } 0 < A < 1, p := 1-A, B := \beta A = \beta(1-p) \end{aligned}$$

für $|s| \leq 1$. Damit bilden die wahrscheinlichkeitserzeugenden Funktionen der Poisson-, Binomial- und negativen Binomialverteilungen Spezialfälle dieser Klasse, die wir *Panjer-Klasse* nennen wollen.

Durch Differenzieren dieser Funktionen und Zusammenfassen der Terme erhält man sofort

II. Sachversicherungsmathematik

$$\frac{d}{ds}\varphi_N(s) = \frac{B}{1-A}s\varphi_N(s), \quad |s| < 1.$$

Mit Hilfe dieses Ergebnisses lässt sich jetzt folgendes fundamentale Resultat formulieren.

Satz 4 (Panjer-Rekursion). Gehört die wahrscheinlichkeitserzeugende Funktion φ_N der Schaden-
zahl zur Panjer-Klasse und bezeichnet man mit

$$f_k := P(X_\Delta = k), \quad g_k := P(S_\Delta = k), \quad k = 0, 1, 2, \dots,$$

die Elementarwahrscheinlichkeiten der (diskretisierten) positiven Einzelschadenhöhen bzw. des
(diskretisierten) Gesamtschadens, so gilt:

$$\begin{aligned} g_0 &= \varphi_N(0) \\ g_{k+1} &= \frac{1}{k+1} \sum_{j=0}^k g_{k-j} f_{j+1} [A(k-j) + B(j+1)], \quad k = 0, 1, 2, \dots \end{aligned}$$

Beweis: Es ist für $|s| < 1$

$$\frac{d}{ds}\varphi_{S_\Delta}(s) = \frac{d}{ds}\left\{\varphi_N(\varphi_{X_\Delta}(s))\right\} = \varphi'_N(\varphi_{X_\Delta}(s)) \cdot \varphi'_{X_\Delta}(s) = \frac{B}{1-A\varphi_{X_\Delta}(s)}\varphi_N(\varphi_{X_\Delta}(s)) \cdot \varphi'_{X_\Delta}(s),$$

also

$$\varphi'_{S_\Delta}(s)(1 - A\varphi_{X_\Delta}(s)) = B \cdot \varphi_{S_\Delta}(s) \cdot \varphi'_{X_\Delta}(s) \quad \text{bzw.} \quad \varphi'_{S_\Delta}(s) = A \cdot \varphi'_{S_\Delta}(s) \cdot \varphi_{X_\Delta}(s) + B \cdot \varphi_{S_\Delta}(s) \cdot \varphi'_{X_\Delta}(s).$$

Aus der Reihendarstellung der wahrscheinlichkeitserzeugenden Funktionen folgt also:

$$\varphi'_{S_\Delta}(s) = \left(\sum_{k=0}^{\infty} (k+1) g_{k+1} s^k \right)$$

sowie

$$B \cdot \varphi_{S_\Delta}(s) \cdot \varphi'_{X_\Delta}(s) = B \cdot \left(\sum_{k=0}^{\infty} g_k s^k \right) \cdot \left(\sum_{k=0}^{\infty} (k+1) f_{k+1} s^k \right) = B \cdot \sum_{k=0}^{\infty} \left(\sum_{j=0}^k (j+1) f_{j+1} g_{k-j} \right) s^k$$

und wegen $f_0 = 0$ nach Voraussetzung:

$$\begin{aligned} A \cdot \varphi'_{S_\Delta}(s) \cdot \varphi_{X_\Delta}(s) &= A \cdot \left(\sum_{k=0}^{\infty} (k+1) g_{k+1} s^k \right) \cdot \left(\sum_{k=0}^{\infty} f_k s^k \right) = A \cdot \sum_{k=0}^{\infty} \left(\sum_{j=1}^k (k+1-j) f_j g_{k+1-j} \right) s^k \\ &= A \cdot \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k-1} (k-j) f_{j+1} g_{k-j} \right) s^k. \end{aligned}$$

Hieraus folgt

II. Sachversicherungsmathematik

$$\begin{aligned}
 \left(\sum_{k=0}^{\infty} (k+1) g_{k+1} s^k \right) &= \varphi'_{S_{\Delta}}(s) = A \cdot \varphi'_{S_{\Delta}}(s) \cdot \varphi_{X_{\Delta}}(s) + B \cdot \varphi_{S_{\Delta}}(s) \cdot \varphi'_{X_{\Delta}}(s) \\
 &= A \cdot \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k-1} (k-j) f_{j+1} g_{k-j} \right) s^k + B \cdot \sum_{k=0}^{\infty} \left(\sum_{j=0}^k (j+1) f_{j+1} g_{k-j} \right) s^k \\
 &= \sum_{k=0}^{\infty} \left(A \sum_{j=0}^{k-1} (k-j) f_{j+1} g_{k-j} + B \sum_{j=0}^k (j+1) f_{j+1} g_{k-j} \right) s^k, \quad |s| < 1.
 \end{aligned}$$

Da beide Reihen im Bereich $|s| < 1$ übereinstimmen, müssen ihre Koeffizienten identisch sein:

$$(k+1) g_{k+1} = A \sum_{j=0}^{k-1} (k-j) f_{j+1} g_{k-j} + B \sum_{j=0}^k (j+1) f_{j+1} g_{k-j} = \sum_{j=0}^k g_{k-j} f_{j+1} [A(k-j) + B(j+1)]$$

für $k = 0, 1, 2, \dots$. Hieraus folgt die Rekursionsformel und wegen

$$g_0 = P(S_{\Delta} = 0) = P(N = 0) = \varphi_N(0)$$

insgesamt die Behauptung.

Bemerkung.

- a) Für einige Anwendungen etwa im Zusammenhang mit Rückversicherung ist es vorteilhaft, die die "Schäden" beschreibenden Zufallsvariablen bzw. die Zufallsvariable X_{Δ} als lediglich nicht-negativ anzunehmen, d.h. den Wert Null zu zulassen. In diesem Fall ergibt sich nach analoger Rechnung folgende modifizierte Panjer-Rekursion:

$$\begin{aligned}
 g_0 &= \varphi_N(f_0) \\
 g_{k+1} &= \frac{1}{(k+1)(1-Af_0)} \sum_{j=0}^k g_{k-j} f_{j+1} [A(k-j) + B(j+1)], \quad k = 0, 1, 2, \dots
 \end{aligned}$$

- b) Besitzt die Verteilung der Zufallsvariablen X_{Δ} einen rechten Endpunkt $\varpi_{\Delta} \in \mathbb{N}$, so ist die Anzahl der Summanden in der Rekursion nach oben beschränkt, und es gilt wegen $f_j = 0$ für $j > \varpi_{\Delta}$:

$$\begin{aligned}
 g_0 &= \varphi_N(f_0) \\
 g_{k+1} &= \frac{1}{(k+1)(1-Af_0)} \sum_{j=0}^{\min\{k, \varpi_{\Delta}-1\}} g_{k-j} f_{j+1} [A(k-j) + B(j+1)], \quad k = 0, 1, 2, \dots,
 \end{aligned}$$

entsprechend in der obigen Rekursion in Satz 4.

II. Sachversicherungsmathematik

Die Panjer-Rekursion kann leicht in beliebigen Sprachen programmiert werden. Wir geben hier beispielhaft ein MAPLE-Worksheet wieder, das als universelle Vorlage für solche Rechnungen dienen kann.

Beispiel 14. Die Schadenzahl sei negativ-binomialverteilt nach $NB(\beta, p)$ mit $\beta = 2$ und $p = 0,25$. Die Verteilung der (diskretisierten) Einzelschadenhöhen sei gegeben durch

k	0	1	2	3	4	5	6
f_k	0	0,1	0,15	0,2	0,25	0,2	0,1

Mit MAPLE kann man die zugehörige Panjer-Rekursion z.B. wie folgt programmieren (mit den Parametern $A = 1 - p = 0,75$, $B = \beta A = 1,5$, Berechnung der ersten 16 Terme):

```
[> restart;
> p:=0.25;beta:=2;A:=1-p;B:=beta*A;
                                     p := 0.25
                                     β := 2
                                     A := 0.75
                                     B := 1.50
> phi:=t->((1-A)/(1-A*t))^(B/A);
                                     (B)
                                     (A)
                                     (1-A)
                                     (1-A*t)
                                     φ = t → -----
> f:=[0.1,0.15,0.2,0.25,0.2,0.1];
                                     f := [0.1, 0.15, 0.2, 0.25, 0.2, 0.1]
> m:=nops(f)-1;
                                     m := 5
> g[0]:=phi(0);
                                     g0 := 0.06250000000
> for n from 0 to 14 do g[n+1]:=sum(g[n-j]*f[j+1]*(A*(n-j)+B*(j+1)),j=0..min(n,m))/(n+1) od;
>
>
                                     g1 := 0.009375000000
                                     g2 := 0.01511718750
                                     g3 := 0.02201953125
                                     g4 := 0.03051379395
                                     g5 := 0.03175650512
                                     g6 := 0.02898740392
                                     g7 := 0.02529763434
                                     g8 := 0.02869401690
                                     g9 := 0.03014012038
                                     g10 := 0.02970935336
                                     g11 := 0.02838951724
                                     g12 := 0.02766760794
                                     g13 := 0.02752733499
                                     g14 := 0.02731233223
                                     g15 := 0.02652398488
```

Man beachte, dass das Feldelement für die Wahrscheinlichkeiten f_k für $k = 0$ nicht definiert werden muss, da Listenindizes in MAPLE standardmäßig bei 1 beginnen. (Die Funktion $\text{nops}(f)$ zählt die Länge der Liste, also hier 6, das ist zugleich ϖ_Δ).

II. 3. Die diskrete Fourier-Transformation

Die Panjer-Rekursion ist aufgrund ihrer einschränkenden Voraussetzungen an die Form der Verteilung für die Schadenzahl N nicht in allen relevanten Fällen verwendbar. Das in diesem Abschnitt vorgestellte Verfahren vermeidet diese Schwierigkeiten. Dazu wollen wir zunächst einige theoretische Grundlagen erarbeiten.

Die zentrale Idee der diskreten Fourier-Transformation besteht wie im vorigen Abschnitt in einer effizienten Auswertung der Funktion $\varphi_{S_\Delta}(s) = \varphi_N(\varphi_{X_\Delta}(s))$. Allerdings betrachten wir jetzt nicht *reelle* Argumente s , sondern *komplexe* Argumente in der Form $s = e^{-it}$ mit $t \in \mathbb{R}$. Dabei bezeichnet i wie in der Analysis üblich die imaginäre Einheit $i = \sqrt{-1}$. Formal erhält man die Verteilung von S_Δ dann über die "Inversion" der Fourier-Transformierten über das Integral

$$\frac{1}{2\pi} \int_0^{2\pi} \varphi_{S_\Delta}(e^{-it}) e^{ikt} dt = P(S_\Delta = k), \quad k = 0, 1, 2, \dots$$

Die Fourier-Transformation lässt sich allgemein für Folgen f reeller Zahlen definieren. Dabei wollen wir die Notation $f = \{f(k)\}_{k=0}^\infty$ verwenden. Die Menge aller solcher Folgen soll mit \mathfrak{F} bezeichnet werden.

Definition 4. Es bezeichne

$$\|f\| := \sum_{k=0}^{\infty} |f(k)| \quad \text{für } f \in \mathfrak{F}$$

und

$$\ell^1 = \{f \in \mathfrak{F} \mid \|f\| < \infty\}.$$

Für $f \in \ell^1$ heißt $\|f\|$ auch die (ℓ^1) -Norm von f . Ferner ist die *Faltung* $f * g$ zweier Folgen $f, g \in \ell^1$ definiert durch

$$(f * g)(k) := \sum_{j=0}^k f(j) g(k-j) = \sum_{j=0}^k f(k-j) g(j), \quad k = 0, 1, 2, \dots$$

Die Menge ℓ^1 bildet im Sinne der Funktionalanalysis einen *Banachraum*, d.h. einen vollständigen¹¹, normierten Vektorraum (hier: von Folgen reeller Zahlen als Elemente). Dabei ist die Vektorraumstruktur gegeben durch die Verknüpfungen

¹¹ Vollständigkeit bedeutet, dass jede Cauchy-Folge $\{f_n\}$ von Elementen des Raumes in der Norm konvergiert. Vgl. hierzu auch das Skript ANALYTISCHE PRINZIPIEN DER STOCHASTIK.

II. Sachversicherungsmathematik

$$\begin{aligned}(f + g)(k) &:= f(k) + g(k) \\ (\alpha f)(k) &:= \alpha f(k),\end{aligned}\quad f, g \in \ell^1, \alpha \in \mathbb{R}.$$

Lemma 14. Der Raum ℓ^1 ist abgeschlossen gegenüber der Faltungsbildung, d.h. es gilt

$$f * g \in \ell^1 \text{ für alle } f, g \in \ell^1$$

mit

$$\|f * g\| \leq \|f\| \cdot \|g\|, \quad f, g \in \ell^1.$$

Beweis: Es ist für $k = 0, 1, 2, \dots$

$$|(f * g)(k)| = \left| \sum_{j=0}^k f(j) g(k-j) \right| \leq \sum_{j=0}^k |f(j)| \cdot |g(k-j)|$$

und somit

$$\begin{aligned}\sum_{k=0}^{\infty} |(f * g)(k)| &\leq \sum_{k=0}^{\infty} \sum_{j=0}^k |f(j)| \cdot |g(k-j)| = \sum_{0 \leq j \leq k} |f(j)| \cdot |g(k-j)| = \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} |f(j)| \cdot |g(k-j)| \\ &= \sum_{j=0}^{\infty} |f(j)| \sum_{k=j}^{\infty} |g(k-j)| = \sum_{j=0}^{\infty} |f(j)| \cdot \|g\| = \|g\| \sum_{j=0}^{\infty} |f(j)| = \|f\| \cdot \|g\|\end{aligned}$$

für alle $f, g \in \ell^1$.

Definition 5 (diskrete Fourier-Transformation). Die durch

$$\hat{f}(t) := \sum_{k=0}^{\infty} f(k) e^{-ikt}, \quad t \in \mathbb{R}; \quad f \in \ell^1$$

definierte Funktion auf \mathbb{R} heißt (diskrete) *Fourier-Transformierte* der Folge $f \in \ell^1$.

Lemma 15. Für die Fourier-Transformierte \hat{f} einer Folge $f \in \ell^1$ gilt folgende Abschätzung:

$$|\hat{f}(t)| \leq \|f\|, \quad t \in \mathbb{R}.$$

Beweis: Es gilt

$$|\hat{f}(t)| \leq \sum_{k=0}^{\infty} |f(k)| \cdot \underbrace{|e^{-ikt}|}_{=1} = \|f\|, \quad t \in \mathbb{R}.$$

II. Sachversicherungsmathematik

Bemerkung: Ist Z eine Zufallsvariable mit Werten in $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ und bezeichnet man mit f_Z die Folge der zugehörigen Elementarwahrscheinlichkeiten, d.h. es ist $f_Z(k) = P(Z = k)$, $k \in \mathbb{Z}^+$, so gilt gerade

$$\hat{f}_Z(t) = \sum_{k=0}^{\infty} e^{-ikt} P(Z = k) = E(e^{-itZ}) = \psi_Z(-it) = \varphi_Z(e^{-it}), \quad t \in \mathbb{R}.$$

Entsprechend gilt: bezeichnet man mit

$$f := \{P(X_{\Delta} = k)\}_{k=0}^{\infty}, \quad g := \{P(S_{\Delta} = k)\}_{k=0}^{\infty}$$

die Folgen der Elementarwahrscheinlichkeiten für die (diskretisierten) Einzelschadenhöhen bzw. den (diskretisierten) Gesamtschaden, so ist

$$\hat{g}(t) = \varphi_N(\hat{f}(t)), \quad t \in \mathbb{R}.$$

Dies folgt völlig analog wie im Beweis zu Satz 2.

Satz 5 (Fourier-Inversionstheorem I). Die Fourier-Transformierte \hat{f} einer Folge $f \in \ell^1$ ist integrierbar über dem Intervall $[0, 2\pi]$, und es gilt folgende Umkehrformel:

$$\frac{1}{2\pi} \int_0^{2\pi} \hat{f}(t) e^{ikt} dt = f(k) \quad \text{für alle } k = 0, 1, 2, \dots$$

Beweis: Die Integrierbarkeit folgt aus Lemma 15. Zur Vereinfachung des Beweisgangs nehmen wir zunächst an, dass die Folge f endlich ist, d.h. es existiert ein $M \in \mathbb{N}$ mit der Eigenschaft

$$f(k) = 0 \quad \text{für } k > M.$$

Dann gilt

$$\begin{aligned} \int_0^{2\pi} \hat{f}(t) e^{ikt} dt &= \int_0^{2\pi} \left\{ \sum_{j=0}^{\infty} f(j) e^{-ijt} \right\} e^{ikt} dt = \int_0^{2\pi} \left\{ \sum_{j=0}^M f(j) e^{-ijt} \right\} e^{ikt} dt = \int_0^{2\pi} \sum_{j=0}^M f(j) e^{i(k-j)t} dt \\ &= \sum_{j=0}^M f(j) \int_0^{2\pi} e^{i(k-j)t} dt = 2\pi f(k), \quad k \in \mathbb{Z}^+ \end{aligned}$$

wegen

$$\int_0^{2\pi} e^{i(k-j)t} dt = \int_0^{2\pi} \cos((k-j)t) dt + i \int_0^{2\pi} \sin((k-j)t) dt = \begin{cases} 2\pi, & \text{falls } j = k \\ 0, & \text{falls } j \neq k \end{cases}$$

II. Sachversicherungsmathematik

(man beachte, dass sich deshalb für $k > M$ in dem vorletzten Ausdruck der oberen Gleichung immer der Wert Null ergibt). Die Vertauschung von Integral und Summe ist dabei aufgrund der Integrierbarkeit von f und der Endlichkeit der Summe gerechtfertigt.

Sei nun f eine beliebige Folge aus ℓ^1 . Wir definieren für jedes $M \in \mathbb{N}$ die Folge f_M durch

$$f_M(k) := \begin{cases} f(k), & k \leq M \\ 0, & k > M. \end{cases}$$

Dann sind auch f_M und $g_M := f - f_M$ Elemente von ℓ^1 wegen

$$\|f_M\| = \sum_{k=0}^M |f(k)| \leq \|f\| \quad \text{und} \quad \|g_M\| = \|f - f_M\| = \sum_{k=M+1}^{\infty} |f(k)| =: c_M \leq \|f\|.$$

Wegen $f \in \ell^1$ strebt die Folge $\{c_M\}_{M \in \mathbb{N}}$ für $M \rightarrow \infty$ gegen Null, und es ist

$$\left| \int_0^{2\pi} \hat{g}_M(t) e^{ikt} dt \right| \leq \int_0^{2\pi} |\hat{g}_M(t) e^{ikt}| dt \leq \int_0^{2\pi} |\hat{g}_M(t)| dt \leq 2\pi \|g_M\| = 2\pi c_M.$$

Es folgt

$$\int_0^{2\pi} \hat{f}(t) e^{ikt} dt = \int_0^{2\pi} \hat{f}_M(t) e^{ikt} dt + \int_0^{2\pi} \hat{g}_M(t) e^{ikt} dt = 2\pi f_M(k) + \int_0^{2\pi} \hat{g}_M(t) e^{ikt} dt$$

mit

$$\left| \int_0^{2\pi} \hat{g}_M(t) e^{ikt} dt \right| \leq 2\pi c_M \rightarrow 0 \quad \text{für } M \rightarrow \infty.$$

Hieraus folgt die Behauptung wegen $f_M \rightarrow f$ für $M \rightarrow \infty$.

Satz 6 (Fourier-Inversionstheorem II). Es sei $f \in \ell^1$ eine endliche Folge, d.h. es existiert ein $M \in \mathbb{N}$ mit der Eigenschaft

$$f(k) = 0 \quad \text{für } k \geq M.$$

Dann gilt exakt:

$$\frac{1}{2\pi} \int_0^{2\pi} \hat{f}(t) e^{ikt} dt = \frac{1}{M} \sum_{j=0}^{M-1} \hat{f}\left(\frac{2\pi j}{M}\right) e^{2\pi ijk/M} \quad \text{für alle } k = 0, 1, 2, \dots$$

Beweis: Es ist für $k = 0, 1, 2, \dots$

II. Sachversicherungsmathematik

$$\begin{aligned}
 \sum_{j=0}^{M-1} \hat{f}\left(\frac{2\pi j}{M}\right) e^{2\pi ijk/M} &= \sum_{j=0}^{M-1} \left(\sum_{m=0}^{M-1} f(m) e^{-2\pi imj/M} \right) e^{2\pi ijk/M} = \sum_{m=0}^{M-1} f(m) \sum_{j=0}^{M-1} \left(\exp\left\{\frac{2\pi i(k-m)}{M}\right\} \right)^j \\
 &= f(k) \cdot M + \sum_{\substack{m=0 \\ m \neq k}}^{M-1} f(m) \frac{\left(\exp\left\{\frac{2\pi i(k-m)}{M}\right\} \right)^M - 1}{\left(\exp\left\{\frac{2\pi i(k-m)}{M}\right\} \right) - 1} \\
 &= f(k) \cdot M + \sum_{\substack{m=0 \\ m \neq k}}^{M-1} f(m) \frac{\exp\{2\pi i(k-m)\} - 1}{\left(\exp\left\{\frac{2\pi i(k-m)}{M}\right\} \right) - 1} = f(k) \cdot M,
 \end{aligned}$$

woraus mit Satz 5 die Behauptung folgt.

Bemerkung: Die einschränkende Annahme in Satz 6 ist wesentlich, da für *unendliche* Folgen f i. Allg. nur gilt:

$$\frac{1}{2\pi} \int_0^{2\pi} \hat{f}(t) e^{ikt} dt = \int_0^1 \hat{f}(2\pi s) e^{2\pi iks} ds \approx \frac{1}{M} \sum_{j=0}^{M-1} \hat{f}\left(\frac{2\pi j}{M}\right) e^{2\pi ijk/M} \quad \text{für alle } k = 0, 1, 2, \dots^{12}$$

Diese Approximationsformel entspricht gerade der Rechteckregel der numerischen Integration. Eine Fehlerabschätzung ist möglich über die Zerlegung $f = f_{M-1} + g_{M-1}$ mit den im Beweis von Satz 5 definierten Hilfsfunktionen. Es ergibt sich dann:

$$\begin{aligned}
 \left| \frac{1}{2\pi} \int_0^{2\pi} \hat{f}(t) e^{ikt} dt - \frac{1}{M} \sum_{j=0}^{M-1} \hat{f}\left(\frac{2\pi j}{M}\right) e^{2\pi ijk/M} \right| &= \left| \frac{1}{2\pi} \int_0^{2\pi} (\hat{f}_{M-1}(t) + \hat{g}_{M-1}(t)) e^{ikt} dt - \frac{1}{M} \sum_{j=0}^{M-1} \hat{f}_{M-1}\left(\frac{2\pi j}{M}\right) e^{2\pi ijk/M} \right| \\
 &\leq \left| \frac{1}{2\pi} \int_0^{2\pi} \hat{g}_{M-1}(t) e^{ikt} dt \right| \leq \frac{1}{2\pi} \int_0^{2\pi} |\hat{g}_{M-1}(t) e^{ikt}| dt \leq \frac{1}{2\pi} \int_0^{2\pi} |\hat{g}_{M-1}(t)| dt \\
 &\leq \frac{1}{2\pi} \int_0^{2\pi} \|g_{M-1}\| dt = \sum_{k=M}^{\infty} |f(k)|.
 \end{aligned}$$

Eine wesentliche Konsequenz der Sätze 5 und 6 besteht darin, dass unter den Voraussetzungen von Satz 6 die Hin- und Rücktransformation mit Hilfe von Vektoren und Matrizen beschrieben werden kann. Es gilt nämlich:

Lemma 16. Es bezeichne $v_{rs} := \exp\left\{-\frac{2\pi i}{M} rs\right\}$, $w_{rs} = \frac{1}{v_{rs}} = \exp\left\{\frac{2\pi i}{M} rs\right\}$ für $0 \leq r, s \leq M-1$ und $M \in \mathbb{N}$ sowie $V := [v_{rs}] \in \mathbb{R}^{M \times M}$. Ferner bezeichne

$$\underline{f}_{M-1} := [f(0), f(1), \dots, f(M-1)]^T \in \mathbb{R}^M, \quad \hat{\underline{f}}_{M-1} := \left[\hat{f}(0), \hat{f}\left(\frac{2\pi}{M}\right), \hat{f}\left(\frac{4\pi}{M}\right), \dots, \hat{f}\left(\frac{2\pi(M-1)}{M}\right) \right]^T \in \mathbb{R}^M.$$

¹² Der hierbei entstehende Fehler wird in der Informationstechnik auch als *Aliasing* bezeichnet; er tritt u.a. auf bei der Konvertierung von analogen zu digitalen Bild- und Tonsignalen und umgekehrt.

II. Sachversicherungsmathematik

Dann gilt: $V^{-1} = W := \frac{1}{M} [w_{rs}]$ und unter den Voraussetzungen von Satz 6:

$$\hat{f}_{M-1} = V \cdot \underline{f}_{M-1} \quad \text{sowie} \quad \underline{f}_{M-1} = W \cdot \hat{f}_{M-1}.$$

Beweis: Es ist für $0 \leq r, t \leq M-1$

$$\sum_{s=0}^{M-1} v_{rs} w_{st} = \sum_{s=0}^{M-1} \exp\left\{\frac{2\pi i}{M} s(t-r)\right\} = \begin{cases} \frac{\exp\{2\pi i(t-r)\} - 1}{\exp\left\{\frac{2\pi i}{M}(t-r)\right\} - 1} = 0, & r \neq t \\ M, & r = t, \end{cases}$$

d.h. es gilt $V \cdot W = E$ (Einheitsmatrix). Ebenso folgt $W \cdot V = E$, d.h. es ist $V^{-1} = W = \frac{1}{M} [w_{rs}]$. Die vektorielle Darstellung von \hat{f}_{M-1} und \underline{f}_{M-1} folgt nun unmittelbar aus der Definition der Fourier-Transformierten sowie den Sätzen 5 und 6.

Bemerkung: Es ist möglich, die erforderlichen Rechenschritte für die Matrixmultiplikation aufgrund der besonderen Struktur der Matrizen V und W technisch erheblich zu beschleunigen. Das dem zu Grunde liegende Verfahren heißt die *schnelle Fourier-Transformation (Fast Fourier Transform, FFT)* und ist z.B. in MAPLE über eigene Befehle implementiert.

Lemma 16 erlaubt nun auf elegante Weise die angekündigte Variante der numerischen Berechnung der Gesamtschadenverteilung im diskretisierten kollektiven Modell der Risikotheorie. Dazu wollen wir wieder mit

$$f := \{P(X_{\Delta} = k)\}_{k=0}^{\infty}, \quad g := \{P(S_{\Delta} = k)\}_{k=0}^{\infty}$$

die Folgen der Elementarwahrscheinlichkeiten für die (diskretisierten) Einzelschadenhöhen bzw. den (diskretisierten) Gesamtschaden bezeichnen. Das Verfahren beinhaltet die folgenden drei Schritte:

1. Fourier-Transformation: wähle $M \in \mathbb{N}$ geeignet und

$$\text{berechne } \hat{f}_{M-1} = V \cdot \underline{f}_{M-1} \in \mathbb{R}^M.$$

2. Transformiere das Ergebnis koordinatenweise mit der wahrscheinlichkeitserzeugenden Funktion φ_N der Schadenzahl:

$$\text{berechne } \varphi_N(\hat{f}_{M-1}) \in \mathbb{R}^M.$$

3. Rücktransformation:

$$\text{berechne } \underline{g}_{M-1} = W \cdot \varphi_N(\hat{f}_{M-1}) \in \mathbb{R}^M.$$

II. Sachversicherungsmathematik

Der Vektor g_{M-1} enthält dann näherungsweise die ersten M Komponenten von g . Falls die Verteilung des Gesamtschadens einen endlichen rechten Endpunkt $\varpi \leq M-1$ besitzt, ist dieses Ergebnis sogar exakt, anderenfalls entsteht ein *Aliasing-Fehler*, der mit der Markoff-Ungleichung durch den Ausdruck

$$\sum_{k=M}^{\infty} |g(k)| = \sum_{k=M}^{\infty} P(S_{\Delta} = k) = P(S_{\Delta} > M-1) = P(t^{S_{\Delta}} \geq t^M) \leq \frac{\varphi_{S_{\Delta}}(t)}{t^M} = \frac{\varphi_N(\varphi_{X_{\Delta}}(t))}{t^M}$$

für jedes $t > 0$ abgeschätzt werden kann.

Für das bei der Panjer-Rekursion besprochene Beispiel 14 ergibt sich mit MAPLE für die Fourier-Transformation nach obigem Muster:

```

[> restart;
[> with(linalg):
Warning, the protected names norm and trace have been redefined and unprotected
[> p:=0.25;beta:=2;A:=1-p;B:=beta*A;

                                     p = 0.25
                                     β = 2
                                     A = 0.75
                                     B = 1.50

[> phi:=t->((1-A)/(1-A*t))^(B/A);

                                     ϕ = t → (1 - A)^(B/A) / (1 - A t)

[> f:= [0,0.1,0.15,0.2,0.25,0.2,0.1,seq(0,j=1..43)];
                                     f = [0,0.1,0.15,0.2,0.25,0.2,0.1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]
[> M:=nops(f);
                                     M = 50
[> v:=(r,s,M)->exp(-2*Pi*r*s*I/M);
                                     v = (r, s, M) → e^((-2 I π r s) / M)
[> V:=array(1..M,1..M,[seq([seq(evalf(v(r-1,s-1,M)),s=1..M)],r=1..M)]):
[> W:=array(1..M,1..M,[seq([seq(evalf(1/v(r-1,s-1,M))/M,s=1..M)],r=1..M)]):
[> fv:=transpose(f):z:=transpose(array(1..1,1..M,[seq(phi(h ln,l),n=1..M)])):h:=evalm(V&*fv):gv:=evalm(W&*z):
[> for i from 0 to 15 do g[i]:=Re(gv[i+1,1]) od;

                                     g0 = 0.06751239145
                                     g1 = 0.01409452873
                                     g2 = 0.01955956582
                                     g3 = 0.02619976319
                                     g4 = 0.03444619326
                                     g5 = 0.03545471458
                                     g6 = 0.03246441556
                                     g7 = 0.02856580993
                                     g8 = 0.03176510929
                                     g9 = 0.03302529526
                                     g10 = 0.03241921122
                                     g11 = 0.03093411509
                                     g12 = 0.03005648139
                                     g13 = 0.02976951960
                                     g14 = 0.02941638498
                                     g15 = 0.02849800483

```

Der Aliasing-Fehler beträgt hier maximal 0,07869 (berechnet z.B. durch Summation der ersten 50 Terme der Panjer-Rekursion oder Reihenentwicklung von $\varphi_{S_{\Delta}}$).

II. 4. Rückversicherung

Unter Rückversicherung versteht man im Allgemeinen eine "Versicherung der Versicherer". Sie basiert auf einem frei vereinbarten Vertrag zwischen dem (Erst-)Versicherer, der hier *Zedent* genannt wird, und dem Rückversicherer, genannt *Zessionär*. Rückversicherung dient dazu, große Risiken des (Erst-)Versicherers und damit eventuell verbundene hohe finanzielle, u.U. sogar ruinöse Belastungen – etwa aus Naturkatastrophen – zu mildern. Rückversicherung kann damit auch als eine Form von Eigenkapitalersatz betrachtet werden.

Bei der *Form* der Rückversicherung unterscheidet man grundsätzlich zwischen *obligatorischer* und *fakultativer* Rückversicherung. Erstere bedeutet, dass die Abgabe des Risikos vom VU bzw. die Annahme des Risikos durch den Rückversicherer (RVU) verpflichtend bzw. freiwillig (auch bzgl. der Risikoübernahme!) ist. Als Kombination beider Formen gibt es noch die *fakultativ-obligatorische* Rückversicherung (engl.: *open cover*), bei der der Zedent die freie Wahl hat, Risiken in Rückdeckung zu geben, der Zessionär aber in diesem Fall die Haftung übernehmen muss.

Bei der *Art* der Rückversicherung unterscheidet man zwischen *proportionaler* und *nicht-proportionaler* Rückversicherung. Bei der ersteren übernimmt der Zessionär einen bestimmten Anteil am versicherungstechnischen Risiko des Zedenten und erhält dafür auch einen entsprechenden (proportionalen) Anteil (plus ggf. weiterer Zuschläge) an der Originalprämie des VU. Die Bemessungsgrundlage für die Bestimmung des Beteiligungsverhältnisses des RVU bildet entweder die Versicherungssumme oder der sog. "wahrscheinliche Höchstschaden" (engl.: *probable maximum loss*, kurz PML). Daher wird diese Art der Rückversicherung auch als *Summenrückversicherung* (bzw. *Summenexzedenten-Rückversicherung*) bezeichnet. Aus mathematischer Sicht ergeben sich hier gegenüber der klassischen Versicherung (Äquivalenzprinzip) keine wesentlich neuen Aspekte, so dass wir hier auf diese Art der Rückversicherung nicht weiter eingehen wollen.

Bei der nicht-proportionalen Rückversicherung versichert das RVU seinen Zedenten gegen Schäden, die dieser durch den Eintritt fest definierter Schadenereignisse erleidet (z.B. Schäden aus Stürmen, die meteorologisch einer klar abgrenzbaren Wetterlage zugeordnet werden können wie *Hurricane Floyd* oder die Stürme *Lothar* und *Martin*, die im Dezember 1999 große Teile Europas verwüsteten). Die Höhe der Leistung des RVU wird ausschließlich durch die Höhe des Schadens bestimmt, weswegen diese Art der Rückversicherung auch "Schadenrückversicherung" genannt wird. Wenn sich die Rückversicherung auf Schäden aus Einzelereignissen bezieht, spricht man auch von einem *Excess-of-Loss-* (bzw. *XL-*)Vertrag, wenn sie sich auf den Jahresgesamtschaden bezieht, von einem *Stop-Loss-* (bzw. *SL-*)Vertrag.

Ein Einzelschadenexzedent wird beschrieben durch die "Haftung *excess* (kurz: *xs*) of Priorität". Die Priorität (oder der Selbstbehalt) gibt an, bis zu welcher Höhe das VU selbst Anteile am Schaden trägt. Das RVU übernimmt dann alle Kosten aus Schäden zwischen der Priorität und dem sog. Plafond; die Differenz aus Plafond und Priorität heißt *Haftungsstrecke*, das zugehörige Intervall wird meist als *Layer* bezeichnet. Bei einem XL-Vertrag mit "6 Mio. €xs 2 Mio. €" trägt also das VU am eventuellen Schaden einen Kostenanteil bis zu 2 Mio. €, das RVU trägt das, was darüber hinausgeht, bis zu einer Höhe von 8 Mio. € mit einer Haftungsstrecke von 6 Mio. €. Die folgende Tabelle zeigt exemplarisch die technischen Aspekte dieses Rückversicherungsvertrags.

Einzelschadenhöhe	das VU trägt hiervon	das RVU trägt hiervon	nicht versichert sind
1,5 Mio. €	1,5 Mio. €	---	---
2,8 Mio. €	2,0 Mio. €	0,8 Mio. €	---
9,4 Mio. €	2,0 Mio. €	6,0 Mio. €	1,4 Mio. €

II. Sachversicherungsmathematik

In der Praxis werden in der Regel mehrere, aufeinander aufbauende Layer, meist bei verschiedenen RVU's rückversichert.

Auch bei derartigen Vertragsformen orientiert sich die Tarifierung (d.h. Prämienfindung für das RVU) zunächst am Äquivalenzprinzip. Beschreibt die Zufallsvariable X ein definiertes Einzelschadenrisiko und $a > 0$ eine Priorität, so wäre die Netto-Prämie, die das VU an das RVU bei Rückversicherung mit XL-Vertrag und unbeschränkter Haftstrecke zu entrichten hätte, gegeben durch

$$E((X - a)^+) = E(\max\{X - a, 0\}) = E((X - a) \cdot \mathbb{1}_{(a, \infty)}(X)).$$

Lemma 17. Es sei Z eine beliebige reelle Zufallsvariable mit Verteilungsfunktion F und ggf. Dichte f . Dann gilt:

$$E(Z^+) = E(\max\{Z, 0\}) = E(Z \cdot \mathbb{1}_{(0, \infty)}(Z)) = \int_0^{\infty} (1 - F(z)) dz$$

$$\text{bzw. bei Existenz einer Dichte } E(Z^+) = \int_0^{\infty} z \cdot f(z) dz.$$

Beweis: Es ist mit dem Satz von Fubini¹³ und dem Satz von der majorisierten Konvergenz

$$\begin{aligned} E(Z^+) &= E(\max\{Z, 0\}) = E\left(\lim_{K \rightarrow \infty} Z \cdot \mathbb{1}_{(0, K]}(Z)\right) = \lim_{K \rightarrow \infty} E(Z \cdot \mathbb{1}_{(0, K]}(Z)) = \lim_{K \rightarrow \infty} \int_{(0, K]} z P^Z(dz) \\ &= \lim_{K \rightarrow \infty} \iint_{0 < u < z \leq K} du P^Z(dz) = \lim_{K \rightarrow \infty} \int_0^K \int_u^K P^Z(dz) du = \lim_{K \rightarrow \infty} \int_0^K P(u < Z \leq K) du = \int_0^{\infty} (1 - F(u)) du \end{aligned}$$

$$\text{bzw. bei Existenz einer Dichte } E(Z^+) = E(Z \cdot \mathbb{1}_{(0, \infty)}(Z)) = \int_0^{\infty} z \cdot f(z) dz.$$

Bemerkung: Aus Lemma 17 folgt für die Situation der Rückversicherung mit $Z = X - a$:

$$\begin{aligned} E((X - a)^+) &= \int_0^{\infty} (1 - F_{X-a}(z)) dz = \int_0^{\infty} P(X - a > z) dz = \int_0^{\infty} P(X > z + a) dz \\ &= \int_0^{\infty} (1 - F_X(z + a)) dz = \int_a^{\infty} (1 - F_X(x)) dx \end{aligned}$$

$$\text{bzw. analog bei Existenz einer Dichte } E((X - a)^+) = \int_a^{\infty} (x - a) \cdot f(x) dx = \int_0^{\infty} x \cdot f(x + a) dx.$$

¹³ Ist eine reellwertige Produkt-messbare Abbildung g absolut integrierbar bez. eines Produktmaßes $\mu \otimes \nu$, so gilt

$$\int \int g(u, v) \mu \otimes \nu(du, dv) = \iint g(u, v) \nu(dv) \mu(du) = \iint g(u, v) \mu(du) \nu(dv).$$

II. Sachversicherungsmathematik

Bei einem XL-Vertrag mit "c Mio. €xs a Mio. €", also einem Vertrag mit Priorität $a > 0$ und Haftstrecke $c > 0$ (d.h. Plafond $b = a + c$), beträgt das vom RVU zu tragende Risiko also gerade

$$X_{a|b} := (X - a)^+ - (X - b)^+ = \begin{cases} 0, & X \leq a \\ X - a, & a < X \leq b \\ c = b - a, & X > b. \end{cases}$$

Für die Netto-Prämie, die das VU an das RVU für diesen Layer zu zahlen hätte, ergibt sich demnach

$$E(X_{a|b}) = E((X - a)^+) - E((X - b)^+) = \int_a^{\infty} (1 - F_X(x)) dx - \int_b^{\infty} (1 - F_X(x)) dx = \int_a^b (1 - F_X(x)) dx$$

$$\text{bzw. analog bei Existenz einer Dichte } E(X_{a|b}) = \int_0^{\infty} x \cdot (f(x + a) - f(x + b)) dx.$$

Beispiel 15. In der Situation von Lemma 10 (kollektives Modell der Risikotheorie mit geometrisch verteilter Schadenzahl und exponentialverteilten Einzelschadenshöhen) mit den Parametern $p = 0,9$ und $\lambda = 0,000003$ soll ein Stop-Loss-Vertrag quotiert werden mit Priorität $a = 100000 \text{ €}$ und Haftstrecke $c = 350000 \text{ €}$. Es folgt mit $b = 450000 \text{ €}$:

$$\begin{aligned} E(S_{a|b}) &= \int_a^b (1 - F_S(x)) dx = \int_a^b (1 - p)e^{-\lambda px} dx = \frac{1-p}{\lambda p} (e^{-\lambda pa} - e^{-\lambda pb}) \\ &= \frac{1-p}{\lambda p} e^{-\lambda pa} (1 - e^{-\lambda pc}) = 17283,95 \text{ €} \end{aligned}$$

Die Netto-Prämie beträgt also 17283,95 € das sind 4,94% der Haftstrecke. Dieser Prozentsatz wird auch als *Rate-on-Line* (ROL) bezeichnet. Die Wahrscheinlichkeit dafür, dass der Gesamtschaden oberhalb von b liegt, beträgt dabei

$$1 - F_S(b) = (1 - p)e^{-\lambda pb} = 0,029671,$$

d.h. in etwa 3% aller Jahre ist der Gesamtschaden nicht über die Rückversicherung gedeckt. Der erwartete Verlust beträgt in dieser Situation aber nur

$$E(S_{b|\infty}) = \frac{1-p}{\lambda p} e^{-\lambda pb} = 10989,26 \text{ €}$$

Sind die Kreditzinsen am Markt höher als die Rate-on-Line und will man das Unterdeckungsrisiko von etwa 3% beibehalten, ist es also günstiger, eine entsprechende Rückversicherung abzuschließen als etwa das für die Haftstrecke benötigte Eigenkapital auf dem Finanzmarkt zu leihen. Es ist u.U. sogar günstiger, vorhandenes Eigenkapital in der Höhe der Haftstrecke auf dem Kapitalmarkt anzulegen, nämlich wenn auch der Guthabenzins höher als die ROL ist. In diesem Sinne kann Rückversicherung – wie schon oben erwähnt – in manchen Fällen als Eigenkapitalersatz dienen.

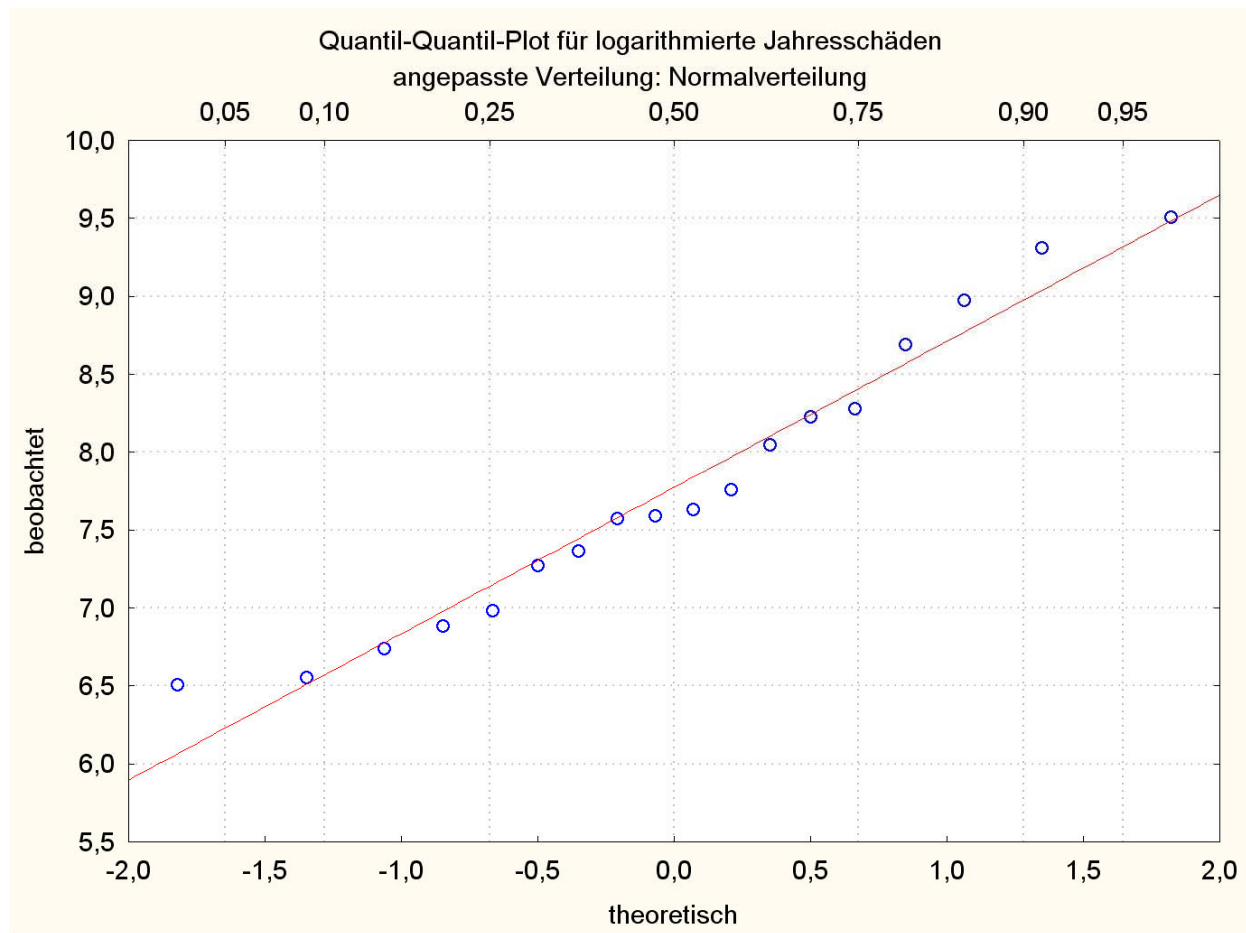
II. Sachversicherungsmathematik

Beispiel 16. Bei konkreten Portfolios besteht fast immer das Problem, dass kein "natürliches" mathematisches Modell zur Beschreibung der Risiken zur Verfügung steht. In diesem Fall muss man zunächst mit statistischen Methoden eine geeignete (Gesamt-)Schadenverteilung bestimmen. Die folgende Tabelle enthält trendbereinigte¹⁴ Sturm-Jahresschäden aus einem Gebäudeversicherungsportfolio in 1000 €(vgl. PFEIFER (2001)).

Jahr	1980	1981	1982	1983	1984	1985	1986	1987	1988
Gesamtschaden	978	2065	1949	5964	3946	669	7920	1438	1077

Jahr	1989	1990	1991	1992	1993	1994	1995	1996	1997
Gesamtschaden	3123	13496	847	3748	1982	2344	11063	703	1582

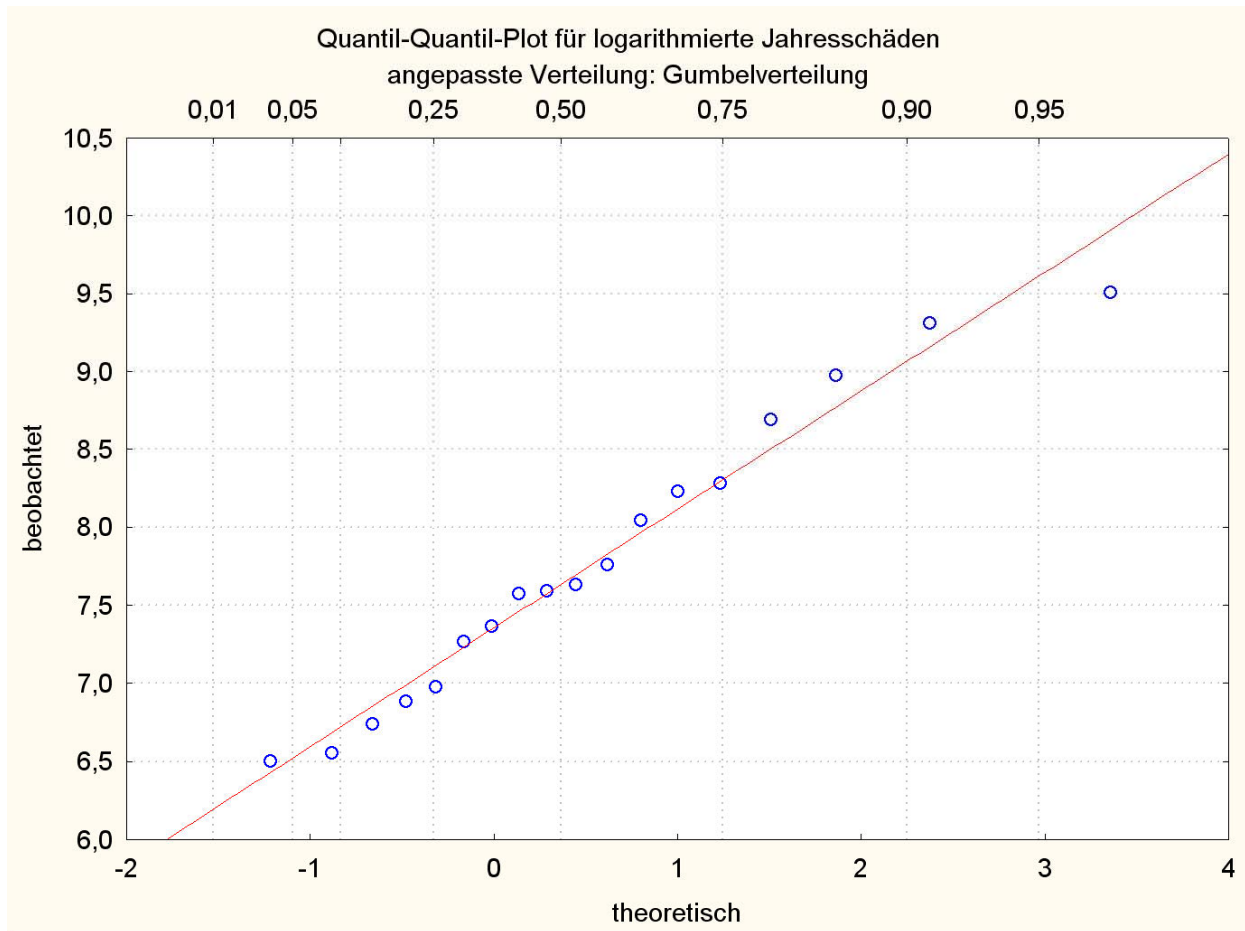
Die folgenden beiden Graphiken zeigen Q-Q-Plots¹⁵ für die logarithmierten Werte, einmal für die Normalverteilung, zum anderen für die Gumbel-Verteilung.



¹⁴ d.h. mit Inflationsrate, Baukostenindex usw. indizierte Schäden

¹⁵ vgl. REISS UND THOMAS (2001), Kapitel 2.4 und die nachfolgenden Ausführungen

II. Sachversicherungsmathematik



Für die Verteilungsfunktion des Jahres-Gesamtschadens S erhält man im ersten Fall mit der Verteilungsfunktion Φ für die Standard-Normalverteilung:

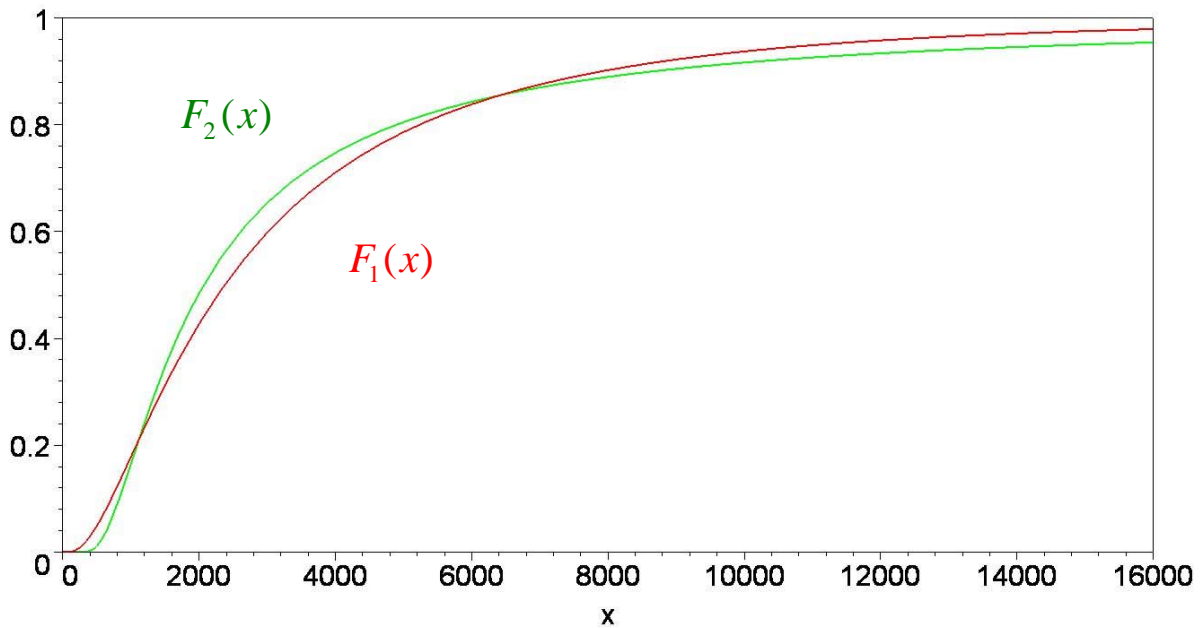
$$P(S \leq x) = F_1(x) = \Phi\left(\frac{\ln x - \mu_1}{\sigma_1}\right), \quad x > 0 \text{ mit } \mu_1 = 7,7731 \text{ und } \sigma_1 = 0,9382$$

und im zweiten Fall mit $G(x) = e^{-e^{-x}}$, $x \in \mathbb{R}$ (Gumbel-Verteilung):

$$P(S \leq x) = F_2(x) = G\left(\frac{\ln x - \mu_2}{\sigma_2}\right) = \exp\left(-\left\{\frac{e^{\mu_2}}{x}\right\}^{1/\sigma_2}\right), \quad x > 0 \text{ mit } \mu_2 = 7,3560 \text{ und } \sigma_2 = 0,7603.$$

Die erste Gesamtschadenverteilung ist also eine Lognormal-Verteilung, die zweite eine so genannte *Fréchet-Verteilung*. Die folgende Graphik zeigt die beiden Verteilungsfunktionen im Vergleich (F_1 : rot, F_2 : grün).

II. Sachversicherungsmathematik



Wir wollen jetzt unter beiden Modellen einen SL-Vertrag quotieren mit Priorität $a = 4$ Mio. € und Haftstrecke $c = 7$ Mio. €, also Plafond $b = 11$ Mio. €. Im ersten Fall erhält man:

$$E(S_{ab}) = \int_a^b (1 - F_1(x)) dx = 902\,284,80 \text{ €}$$

im zweiten

$$E(S_{ab}) = \int_a^b (1 - F_2(x)) dx = 929\,869,76 \text{ €}$$

also zwei relativ nahe beieinander liegende Ergebnisse mit 12,9% ROL im ersten und 13,3% ROL im zweiten Fall. Größere Unterschiede ergeben sich aber in höheren Layern; z.B. erhalten wir für einen zweiten Layer mit Priorität $a = 11$ Mio. € und Haftstrecke $c = 5$ Mio. €, also $b = 16$ Mio. €

$$E(S_{ab}) = \int_a^b (1 - F_1(x)) dx = 166\,821,44 \text{ €} \quad \text{im ersten und}$$

$$E(S_{ab}) = \int_a^b (1 - F_2(x)) dx = 290\,277,50 \text{ €} \quad \text{im zweiten Fall,}$$

das sind 3,3% ROL im ersten bzw. 5,8% ROL im zweiten Fall. Im ersten Modell beträgt die Wahrscheinlichkeit, einen Jahresschaden von 16 Mio. € oder mehr zu erleiden, ca. 2,1%, im zweiten dagegen schon ca. 4,6%, also mehr als doppelt so viel, woraus sich auch die höhere Netto-Prämie erklärt.

Problematisch ist dabei immer die Extrapolation von Daten in bisher nicht beobachtete Schadenhöhen. Allerdings sind im Laufe von 18 Jahren bereits zweimal Schäden oberhalb von 10 Mio. € aufgetreten, das ist empirisch *eine* Überschreitung alle 9 Jahre (Wiederkehrperiode), wobei im ersten Modell die entsprechende theoretische Größe für *eine* Überschreitung 15,9 Jahre, im zweiten Modell die etwas realistischere Zahl 11,9 Jahre ergibt. Für Schäden der Größenordnung 13 Mio. € beträgt die empirische Wiederkehrperiode 18 Jahre, im ersten Modell (unrealistische) 28,6 Jahre, im zweiten Modell (realistische) 16,7 Jahre. Ein 100-Jahres-Schaden beträgt nach dem ersten Modell ca. 21,1 Mio. € nach dem zweiten schon 51,7 Mio. €

II. Sachversicherungsmathematik

Wir wollen hier noch kurz auf die in Beispiel 16 verwendete Methode der Q-Q-Plots eingehen (vgl. hierzu auch das Skript zur STOCHASTIK, Kapitel III.4). Wir betrachten dazu zunächst eine Zufallsvariable X mit einer Verteilungsfunktion F_X , die im Bereich $\{x \in \mathbb{R} \mid 0 < F_X(x) < 1\}$ streng monoton wachsend und stetig sein möge.

Mit Hilfe der positiv-linearen Transformation

$$Z = \frac{X - \mu}{\sigma} \text{ mit } \mu, \sigma \in \mathbb{R}, \sigma > 0$$

wird eine verschobene und reskalierte Zufallsvariable Z definiert, die folgende Verteilungsfunktion F_Z besitzt:

$$F_Z(x) = P(Z \leq x) = P\left(\frac{X - \mu}{\sigma} \leq x\right) = P(X \leq \mu + \sigma x) = F_X(\mu + \sigma x), \quad x \in \mathbb{R}.$$

Ist beispielsweise X nach $\mathcal{N}(\mu, \sigma^2)$ (normal-)verteilt, so ist Z nach $\mathcal{N}(0, 1)$ verteilt. Ist die Verteilungsfunktion F_Z der Standard-Typ, so heißt die Klasse der Verteilungen F_X die durch F_Z induzierte *Lage- und Skalenfamilie* mit den Parametern μ und σ .

Die *Quantilfunktion* Q_X ist nun definiert durch die Inverse der Verteilungsfunktion:

$$Q_X(u) = F_X^{-1}(u) \text{ bzw. } F_X(Q_X(u)) = u, \quad 0 < u < 1.$$

Für die Standard-Zufallsvariable Z ergibt sich daraus folgender Zusammenhang:

$$Q_X(u) = \mu + \sigma Q_Z(u), \quad 0 < u < 1,$$

wie man durch Vergleich der Argumente feststellen kann:

$$u = F_X(Q_X(u)) = F_Z(Q_Z(u)) = F_X(\mu + \sigma Q_Z(u)), \quad 0 < u < 1.$$

Trägt man also in einem Koordinatensystem die Paare $(Q_Z(u), Q_X(u))$ gegeneinander auf, so erhält man eine Gerade mit (positiver) Steigung σ und Achsenabschnitt μ . Diese Beobachtung kann man sich zu Nutze machen, um zu prüfen, ob einem Datensatz (X_1, \dots, X_n) eine durch eine Verteilungsfunktion induzierte Lage- und Skalenfamilie, wie z.B. eine Familie von Normalverteilungen mit Erwartungswert μ und Standardabweichung σ , zu Grunde liegt, wobei wir hier annehmen, dass alle Beobachtungen paarweise verschieden sind (das ist theoretisch bei stetigen Verteilungsfunktionen gegeben). Dazu trägt man in ein Koordinatensystem für geeignete Werte u_1, \dots, u_n die aus den Beobachtungen abgeleiteten Paare $(Q_Z(u_k), \hat{Q}_X(u_k))$, $k = 1, \dots, n$ gegeneinander auf, wobei \hat{Q}_X die *empirische Quantilfunktion*¹⁶ bezeichnet. Ordnet man die Daten (X_1, \dots, X_n) der Größe nach mit den Werten (sog. *Ordnungsstatistiken*) $X_{(1)} < \dots < X_{(n)}$ und wählt man z.B.

¹⁶ das ist die Pseudo-Inverse der empirischen Verteilungsfunktion

II. Sachversicherungsmathematik

$u_k = \frac{k}{n+1}$ für $k=1, \dots, n$, so ist gerade $\hat{Q}_X(u_k) = X_{(k)}$, d.h. man trägt die geordneten Beobachtungswerte ("beobachtete Quantile") gegen die Werte $Q_Z(u_k)$ ("theoretische Quantile") auf. Mit Hilfe einer Ausgleichsgeraden durch diese n Punktepaare (z.B. durch lineare Regression) lässt sich dann zunächst durch visuelle Überprüfung abschätzen, ob die getroffene Verteilungsannahme haltbar ist. Dazu sollten die Punktepaare nicht "zu weit" von der Ausgleichsgeraden entfernt sein. Findet man mit dieser Methode die Verteilungsannahme gerechtfertigt, lassen sich anschließend durch den Achsenabschnitt und die Steigung der Geraden die unbekannt Parameter μ und σ schätzen.

Grundsätzlich haftet das RVU für jeden rückversicherten Schaden während der Laufzeit des Rückversicherungsvertrages (in der Regel 1 Jahr) im Rahmen seiner Haftung ab Überschreitung der Priorität. Dieser universelle Schutz wird allerdings meist durch zusätzliche Haftungsbegrenzungen eingeschränkt, z.B. durch die Vereinbarung einer maximalen Gesamtleistung während der Vertragslaufzeit ("Jahreshöchsthafung", engl. *annual aggregate limit*, AAL). Wenn vereinbart ist, dass das RVU für mehrere gleichartige Schäden im Jahr (z.B. aus der Sturmversicherung) haftet, spricht man auch von "Wiederauffüllungen". Diese kann der Anzahl nach ebenfalls vertraglich nach oben begrenzt sein. Üblich ist häufig auch die Vorschaltung eines Selbstbehalts (engl. *annual aggregate deductible*, AAD). Schäden, die eigentlich nach Abzug der Priorität an das rückversicherte VU zu bezahlen wären, sind dann solange zu saldieren, bis dieser Selbstbehalt verbraucht ist. Danach werden alle (weiteren) Schäden wie üblich abgerechnet (vgl. auch SCHWEPKE (2001), Kapitel 3).

Insbesondere bei Wiederauffüllungen für XL-Verträge ist es wichtig, zu wissen, wie sich das kollektive Modell der Risikotheorie hier bez. der Quotierung des Rückversicherungsvertrags auswirkt. Der Gesamtschaden $S_{a|b}$, der in einem solchen Modell entsteht, wenn eine Priorität $a > 0$ mit Plafond $b > a$ und unbegrenzten Wiederauffüllungen vereinbart ist, beträgt gemäß den Ausführungen aus Abschnitt II.1 gerade

$$S_{a|b} = \sum_{k=1}^N \left\{ (X_k - a)^+ - (X_k - b)^+ \right\} = \sum_{k=1}^N \left\{ (X_k - a) \cdot \mathbb{1}_{(a,b]}(X_k) + (b - a) \cdot \mathbb{1}_{(b,\infty)}(X_k) \right\}.$$

Lemma 18. Im üblichen kollektiven Modell der Risikotheorie gilt: ist g eine messbare Funktion und B eine Borel-Menge, so dass gilt:

$$g(x) > 0, \text{ falls } x \in B,$$

so kann das transformierte Modell mit

$$\hat{N} := \sum_{k=1}^N \mathbb{1}_B(X_k) \text{ und Zufallsvariablen } \hat{X}_k > 0 \text{ mit } P^{\hat{X}_k} = P^{g(X_k)}(\cdot | X_k \in B), k \in \mathbb{N}$$

wieder als ein kollektives Modell der Risikotheorie angesehen werden, für das der Gesamtschaden

$$\hat{S} := \sum_{k=1}^{\hat{N}} \hat{X}_k$$

dieselbe Verteilung besitzt wie der modifizierte "Gesamtschaden"

II. Sachversicherungsmathematik

$$\bar{S} = \sum_{k=1}^N g(X_k) \cdot \mathbb{1}_B(X_k).$$

Beweis: Wie verwenden wieder erzeugende Funktionen. Wie im Beweis zu Satz 2 ergibt sich:

$$\psi_{\bar{S}}(t) = \varphi_N(\psi_Y(t)), \quad t \in I \text{ geeignet,}$$

wobei die Zufallsvariable Y verteilt ist wie $g(X_k) \cdot \mathbb{1}_B(X_k)$ (diese Folge bleibt stochastisch unabhängig). Sei zusätzlich X verteilt wie die X_k , $k \in \mathbb{N}$. Dann kann Y dargestellt werden als

$$Y = \begin{cases} g(X), & X \in B \\ 0, & X \notin B \end{cases}$$

mit der (Misch-)Verteilung

$$P^Y = (1-\alpha) \cdot \varepsilon_0 + \alpha \cdot P^{g(X)}(\cdot | X \in B) \text{ mit } \alpha = P(X \in B).$$

Dies sieht man so: Für eine beliebige Borel-Menge A ist

$$\begin{aligned} P(Y \in A) &= P(Y \in A | X \in B) \cdot P(X \in B) + P(Y \in A | X \notin B) \cdot P(X \notin B) \\ &= \alpha \cdot P \left(\underbrace{g(X) \cdot \mathbb{1}_B(X)}_{=1 \text{ unter } X \in B} \in A | X \in B \right) + (1-\alpha) \cdot P \left(\underbrace{g(X) \cdot \mathbb{1}_B(X)}_{=0 \text{ unter } X \notin B} \in A | X \notin B \right) \\ &= \alpha \cdot P(g(X) \in A | X \in B) + (1-\alpha) \cdot \mathbf{1}_A(0) = \alpha \cdot P(g(X) \in A | X \in B) + (1-\alpha) \cdot \varepsilon_0(A), \end{aligned}$$

woraus die Behauptung folgt. Damit ist also nach obigem

$$P^Y = (1-\alpha) \cdot \varepsilon_0 + \alpha \cdot P^{\bar{X}} \text{ mit } \alpha = P(X \in B),$$

wobei

$$P^{\bar{X}} = P^{g(X)}(\cdot | X \in B).$$

Gemäß Lemma 12 kann die momenterzeugende Funktion von Y also auch dargestellt werden als

$$\psi_Y(t) = \sum_{i=0}^1 P(J=i) \cdot \{\psi_{\bar{X}}(t)\}^i = \varphi_J(\psi_{\bar{X}}(t))$$

mit einer von Z unabhängigen Zufallsvariablen J mit den Werten 0 und 1 und

$$P(J=1) = P(X \in B) = \alpha, \quad P(J=0) = P(X \notin B) = 1-\alpha.$$

Damit ist

II. Sachversicherungsmathematik

$$\psi_{\bar{S}}(t) = \varphi_N(\psi_Y(t)) = \varphi_N(\varphi_J(\psi_{\hat{X}}(t))) = (\varphi_N \circ \varphi_J)(\psi_{\hat{X}}(t)) = \varphi_{\hat{N}}(\psi_{\hat{X}}(t)), \quad t \in I,$$

woraus die Behauptung folgt.

Für das obige Rückversicherungsproblem bedeutet dies: setzen wir

$$B := (a, \infty) \quad \text{und} \quad g(x) := \begin{cases} 0, & x \leq a \\ x - a, & a < x \leq b \\ b, & x > b, \end{cases}$$

so ist $g(x) > 0$ für $x \in B$, und mit den Original-Schäden X_k folgt

$$g(X_k) \cdot \mathbb{1}_B(X_k) = (X_k - a)^+ - (X_k - b)^+ = (X_k - a) \cdot \mathbb{1}_{(a,b]}(X_k) + (b - a) \cdot \mathbb{1}_{(b,\infty)}(X_k),$$

d.h. $\bar{S} = \sum_{k=1}^N g(X_k) \cdot \mathbb{1}_B(X_k)$ entspricht genau dem auf das RVU entfallenden Schadenanteil. Andererseits kann $\hat{N} = \sum_{k=1}^N \mathbb{1}_B(X_k)$ als die Anzahl der für das RVU relevanten (in das Intervall (a, ∞))

fallenden) Schäden und \hat{X}_k , $k \in \mathbb{N}$ als Folge stochastischer Repräsentanten der vom RVU an das versicherte VU im Vertragsfall zu bezahlenden Regulierungsbeträge aufgefasst werden. Auf diese Weise kann das RVU bei XL-Verträgen seine Schadenstruktur auch in einem "echten" kollektiven Modell der Risikotheorie abbilden. Dieses Modell wollen wir von nun an als das "kollektive Modell der Rückversicherung" bezeichnen. Dabei ist es interessant, zu untersuchen, welche Verteilungen sich für die neue Schadenfrequenz $\hat{N} = \sum_{k=1}^N \mathbb{1}_B(X_k)$ ergeben können.

Lemma 19. Gehört im kollektiven Modell der Rückversicherung die originale Schadenzahl N der Panjer-Klasse an, so auch die modifizierte Schadenzahl \hat{N} . Genauer gilt (mit den Bezeichnungen von oben, d.h. $\alpha = P(X \in B)$):

Name der Verteilung	P^N	$P^{\hat{N}}$
Binomialverteilung	$B(n, p)$	$B(n, \alpha p)$
negative Binomialverteilung	$NB(\beta, p)$	$NB\left(\beta, \frac{p}{p + \alpha - \alpha p}\right)$
Poisson-Verteilung	$\mathcal{P}(\lambda)$	$\mathcal{P}(\alpha \lambda)$

Beweis: Wir zeigen über die wahrscheinlichkeitserzeugenden Funktionen:

II. Sachversicherungsmathematik

$$\varphi_N(\varphi_J(s)) = \varphi_N(1 - \alpha + \alpha s) = \dots$$

$$\dots = \begin{cases} (1 - p + p(1 - \alpha + \alpha s))^n = (1 - \alpha p + \alpha ps)^n & \text{(Binomialverteilung)} \\ \left(\frac{p}{1 - (1 - p)(1 - \alpha + \alpha s)} \right)^\beta = \left(\frac{\frac{p}{p + \alpha - \alpha p}}{1 - \frac{\alpha - \alpha p}{p + \alpha - \alpha p} s} \right)^\beta & \text{(negative Binomialverteilung)} \\ \exp\{\lambda((1 - \alpha + \alpha s) - 1)\} = \exp\{\alpha\lambda(s - 1)\} & \text{(Poisson-Verteilung).} \end{cases}$$

Lemma 20 (Ergänzung zu Lemma 10). Gegeben sei ein kollektives Modell der Risikotheorie mit geometrisch verteilter Schadenzahl N gemäß

$$P(N = n) = pq^n, \quad n = 0, 1, \dots \text{ mit einem Parameter } p = 1 - q \in (0, 1),$$

die Verteilung der Einzelschadenshöhen $\{X_k\}_{k \in \mathbb{N}}$ sei exponentiell $\mathcal{E}(\lambda)$ mit einem Parameter $\lambda > 0$. Dann gilt:

Das zugehörige kollektive Modell der Rückversicherung ist für $a > 0$ und $b = \infty$ (sog. *Illimitée-Deckung*) vom selben Typ mit geometrisch verteilter Schadenzahl \hat{N} gemäß

$$P(\hat{N} = n) = \frac{p}{p + \alpha - \alpha p} \left(\frac{\alpha(1 - p)}{p + \alpha - \alpha p} \right)^n, \quad n = 0, 1, \dots$$

mit

$$\alpha = P(X > a) = e^{-\lambda a}$$

und exponentiell $\mathcal{E}(\lambda)$ -verteilten Einzelschadenshöhen $\{\hat{X}_k\}_{k \in \mathbb{N}}$.

Ist der Plafond b endlich, so ist die Verteilung der Einzelschadenshöhen $\{\hat{X}_k\}_{k \in \mathbb{N}}$ die Mischung einer nach oben abgeschnittenen Exponentialverteilung und einer Einpunktverteilung im Punkt $b - a$, mit Verteilungsfunktion

$$F_{a|b}(x) = \begin{cases} 1 - e^{-\lambda x}, & 0 \leq x < b - a \\ 1, & x \geq b - a. \end{cases}$$

Beweis: Der erste Teil ist klar nach Lemma 19. Der zweite Teil folgt zum einen aus Lemma 18 und der daran anschließenden Bemerkung wegen

$$P(X - a \leq x | X > a) = \frac{P(a < X \leq x + a)}{P(X > a)} = \frac{e^{-\lambda a} - e^{-\lambda(x+a)}}{e^{-\lambda a}} = 1 - e^{-\lambda x}, \quad x > 0,$$

zum anderen hieraus analog den Ausführungen in Beispiel 5.

II. Sachversicherungsmathematik

Bemerkung: Das Erstaunliche an der Aussage in Lemma 20 ist, dass in dem betrachteten Modell die Einzelschadenhöhen im kollektiven Modell der Rückversicherung bei Illimitée-Deckung dieselbe Exponentialverteilung besitzen wie im Original-Modell und die Schadenfrequenz-Verteilung vom selben Typ ist mit einem lediglich anderen Parameter.

Es ist prinzipiell möglich, jede Form der Erstversicherung mit Rückversicherung zu versehen, so z.B. auch in der Lebens- oder Krankenversicherung. Wir wollen dies hier exemplarisch an einem Beispiel verdeutlichen.

Beispiel 17. Wir betrachten ein Portfolio aus Todesfallversicherungen für n Personen gleichen Geschlechts und Altern x_1, \dots, x_n mit Versicherungssummen VS_1, \dots, VS_n . Der jährliche "Gesamtschaden" S_i für den Versicherten i kann dann exakt beschrieben werden durch ein kollektives Modell mit Schadenfrequenz N_i mit Binomialverteilung $P^{N_i} = B(1, q_{x_i})$ und (konstanter) Schadenhöhe VS_i . Da für kleine Werte von q_{x_i} die Binomialverteilung $B(1, q_{x_i})$ recht gut durch eine Poisson-Verteilung $\mathcal{P}(q_{x_i})$ mit demselben Parameter approximiert werden kann, wollen wir zur Vereinfachung hier diese Verteilungsannahme für die N_i machen.¹⁷ (Das Modell gleicht insofern dem entsprechenden Ansatz in der geophysikalischen Modellierung.) Mit Satz 3 folgt dann, dass der Portfolio-Gesamtschaden S dargestellt werden kann als

$$S = \sum_{k=1}^N X_k,$$

wobei $N = \sum_{i=1}^n N_i$ jetzt $\mathcal{P}\left(\sum_{i=1}^n q_{x_i}\right)$ -verteilt ist und die Zufallsvariablen X_k einer Mischverteilung

genügen, die der Versicherungssumme VS_i die Wahrscheinlichkeit $\frac{q_{x_i}}{\lambda}$ mit $\lambda = \sum_{i=1}^n q_{x_i}$ zuweist.

(Bei gleichen Versicherungssummen werden diese Wahrscheinlichkeiten entsprechend der Vielfachheit summiert.) Wir wollen jetzt konkret einen SL-Vertrag quotieren für den folgenden Bestand (männlich; Sterbetafel DAV T94m):

Alter x_i	q_{x_i}	nominale Versicherungssumme	VS_i	Anzahl
20	0,00148	€ 50000	1	100
30	0,00148	€100000	2	500
40	0,00257	€150000	3	300
50	0,00675	€200000	4	100

Zur Vereinfachung betrachten wir hier die Versicherungssummen VS_i als Vielfache von €50000. Es folgt

¹⁷ Es ist dann $P(N_i \geq 2) = 1 - (1 + q_{x_i})e^{-q_{x_i}} \leq \frac{1}{2}q_{x_i}^2$ und damit $P\left(\bigcup_{i=1}^n \{N_i \geq 2\}\right) \leq \frac{1}{2} \sum_{i=1}^n q_{x_i}^2$. In dem nachfolgend betrachteten Rechenbeispiel beträgt diese Fehlerwahrscheinlichkeit maximal 0,008.

II. Sachversicherungsmathematik

$$\lambda = 2,334$$

und für die "Einzelschadenhöhen"

k	1	2	3	4
$P(X = k)$	0,06341	0,31705	0,33033	0,28921

Analog zu Beispiel 13 erhalten wir über die Reihenentwicklung der wahrscheinlichkeitserzeugenden Funktion des Gesamtschadens

```

> restart:with(PolynomialTools):
> f:=t->0.06341*t+0.31705*t^2+0.33033*t^3+0.28921*t^4;
                                f:=t -> 0.06341 t + 0.31705 t^2 + 0.33033 t^3 + 0.28921 t^4
> phi:=t->exp(2.334*(t-1));
                                phi:=t -> e^(2.334 t - 2.334)
> g:=convert(series(phi(f(t)),t=0,51),polynom):L:=CoefficientList(g,t):
> for k from 0 to 14 do g[k]:=L[k+1] od;
                                g_0 := 0.09690734143
                                g_1 := 0.01434218381
                                g_2 := 0.07277223304
                                g_3 := 0.08538011025
                                g_4 := 0.1037918570
                                g_5 := 0.06975352102
                                g_6 := 0.09298428546
                                g_7 := 0.08394219675
                                g_8 := 0.07395270939
                                g_9 := 0.05984304645
                                g_10 := 0.05635253455
                                g_11 := 0.04496427158
                                g_12 := 0.03567900989
                                g_13 := 0.02798068838
                                g_14 := 0.02236444222
> 1-sum(g[i],i=0..14);
                                0.0589895689
> sum(g[i]*(i-7),i=8..14)+(1-sum(g[i],i=0..14))*7;
                                1.458310749

```

Die Quotierung eines SL-Vertrags mit Priorität $a = 7$ (entsprechend 350000 €) und Plafond 14 (entsprechend 700000 €) kann dann gemäß Bemerkung zu Lemma 17 erfolgen über

$$E(S_{ab}) = \sum_{k=8}^{14} (k-7) \cdot P(S=k) + 7 \cdot P(S > 14) = \sum_{k=8}^{14} (k-7) \cdot P(S=k) + 7 \cdot (1 - P(S \leq 14)) = 1,4583$$

(das entspricht einer Netto-Prämie von 72915 € oder 20,83% ROL). Nicht versichert ist also ein über 700000 € hinausgehender "Schaden"; dies geschieht allerdings nur mit Wahrscheinlichkeit $P(X > 14) = 1 - P(X \leq 14) = 0,059$, also in ca. 5,9% aller Jahre. (Man beachte, dass theoretisch auch ein Totalverlust in Höhe von 120 Mio. € eintreten kann, nämlich wenn alle 1000 Versicherten gleichzeitig versterben.)

II. Sachversicherungsmathematik

Zum Abschluss dieses Abschnitts wollen wir noch kurz auf einen Aspekt des Alternativen Risiko-transfers (ART) im Zusammenhang mit Rückversicherung eingehen. Es ist nämlich seit einigen Jahren möglich, Risiken aus Naturkatastrophen auch auf den Finanzmarkt über bestimmte Options-geschäfte abzuwälzen. Das "Underlying" ist hier beispielsweise ein Marktschadenindex für ein be-stimmtes geographisches Gebiet (z.B. der Südosten der USA). Ein VU, das in dem betreffenden Gebiet versicherte Gebäude besitzt, kann sich dann über den Erwerb einer passenden Call-Option bzw. eines Spreads auf den Index eine Art Rückversicherung kaufen. Ein Indexpunkt entspricht dabei grundsätzlich einem festen Schadenbetrag L , z.B. $L = 10$ Mio. US \$. Eine Call-Option mit Ausübungspreis X auf einen solchen Index greift also dann, wenn der Index zu einem bestimmten, vorher festgelegten Zeitpunkt (z.B. einem Quartalsende) den Wert X erreicht oder überschreitet. Die damit verbundene Auszahlung durch den Verkäufer der Call-Option wird in sog. Ticks ver-rechnet, d.h. pro Indexpunkt Unterschied zwischen Indexwert W und Ausübungspreis X wird im Ausübungsfall eine feste Summe Q , z.B. $Q = 200$ US \$, vergütet. Erreicht der Indexwert W den Ausübungspreis X dagegen nicht, ist die Option wertlos; es findet dann keine Auszahlung statt.

In dieser Konstruktion entspricht der Ausübungspreis X der Priorität a der Rückversicherung bei einer Illimitée-Deckung, allerdings ist hierbei nur der "Markt" als Ganzes versichert, nicht unbed-ingt ein einzelnes VU. Durch Kauf einer Call-Option mit Ausübungspreis X_1 auf den Index und gleichzeitigem Verkauf einer Call-Option mit höherem Ausübungspreis X_2 auf den Index (sog. Bull-Call-Spread) erhält man ein Konstrukt, das der klassischen Rückversicherung mit Priorität a (bezogen auf X_1) und Plafond b (bezogen auf X_2) entspricht. Das folgende Beispiel soll verdeut-lichen, wie dies technisch funktioniert.

Beispiel 18. In einem Gebäudeversicherungsmarkt sind 5 VU vertreten mit folgenden Marktantei-len α_i :

VU i	1	2	3	4	5
α_i	0,10	0,05	0,35	0,20	0,30

VU 1 möchte sich auf dem Finanzmarkt "rückversichern" über einen Bull-Call-Spread, der einer klassischen Rückversicherung mit Priorität $a = 50$ Mio. US \$ und Plafond $b = 100$ Mio. US \$ ent-spricht. Da der Marktanteil 10% beträgt, entspricht dies Indexwerten zwischen 50 Punkten (= 500 Mio. US \$) und 100 Punkten (= 1 Mrd. US \$). Die Ausübungspreise sind hier also gegeben durch $X_1 = 50$ und $X_2 = 100$. Da jeder Indexpunkt 10 Mio. \$ / 200 \$ = 50000 Ticks entspricht und der Marktanteil 10% beträgt, müssen also 5000 Spreads erworben werden. Die Call-Option zum Aus-übungspreis $X_1 = 50$ koste nun 600 US \$, die zum Ausübungspreis $X_2 = 100$ koste 350 US \$. Dann beträgt der finanzielle Aufwand für das VU genau $250 \times 5000 = 1,25$ Mio. US \$, oder in der Sprache der klassischen Rückversicherung 2,5% ROL (für die Haftstrecke von 50 Mio. US \$).

II. Sachversicherungsmathematik

Wir betrachten jetzt drei Szenarien für mögliche Schadenentwicklungen.

Szenario 1	VU i	1	2	3	4	5
	Schäden (Mio. US \$)	55	27	195	110	123

Szenario 2	VU i	1	2	3	4	5
	Schäden (Mio. US \$)	45	28	204	122	111

Szenario 3	VU i	1	2	3	4	5
	Schäden (Mio. US \$)	55	27	195	110	103

In den ersten beiden Szenarien beträgt der Marktschaden 510 Mio. US \$ mit Indexwert $W = 51$, d.h. die Call-Option mit Ausübungspreis $X_1 = 50$ greift in beiden Fällen. Die Auszahlung an VU 1 beläuft sich in beiden Szenarien auf $5000 \times 1 \times 200 = 1$ Mio. US \$. Für das erste Szenario wäre eine klassische Rückversicherung mit der fälligen Zahlung von 5 Mio. US \$ günstiger gewesen, im zweiten Szenario hätte die Rückversicherung wegen Unterschreiten der Priorität gar nicht gezahlt. Kritisch ist Szenario 3, da hier der Marktschaden nur 490 Mio. US \$ mit Indexwert $W = 49$ beträgt, also beide Optionen wertlos sind und damit keine Auszahlung an VU 1 erfolgt, wogegen die klassische Rückversicherung wieder 5 Mio. US \$ gezahlt hätte. Dies zeigt, dass ein solcher ART in der Regel kein exaktes Äquivalent für die Rückversicherung sein kann; manchmal ist er trotz gleichen Schadenaufkommens günstiger (Szenario 2), manchmal ungünstiger (Szenarien 1 und 3).

Die obige Rechnung zeigt, wie die Anzahl n der benötigten Spreads und die Ausübungspreise X_1 und X_2 durch die Größen L, Q, a, b und α (Marktanteil) allgemein bestimmt werden können, nämlich durch

$$X_1 = \frac{a}{\alpha L}, \quad X_2 = \frac{b}{\alpha L}, \quad n = \frac{\alpha L}{Q}.$$

Sind die Optionspreise entsprechend gegeben durch C_1 und C_2 , so kostet der "Rückversicherungsschutz" also genau $n \cdot (C_2 - C_1)$ mit einer ROL von

$$\frac{n(C_2 - C_1)}{b - a} = \frac{1}{Q} \frac{C_2 - C_1}{X_2 - X_1}.$$

Für weitere Details siehe BANNISTER (1998), Kapitel 7 oder LIEBWEIN (2009), Abschnitt VI, speziell Beispiel 20.3.3.

II. 5. Spätschadenreservierung

Ein zentrales, in der Versicherungspraxis immer wieder zu beobachtendes Problem stellt die verzögerte Abwicklung von Schäden dar, die z.B. in der Kfz-Haftpflichtversicherung oder in der Krankenversicherung durch aufwändige Gutachten, Gerichtsprozesse usw. verursacht werden können. Manchmal können sogar mehrere Jahre oder Jahrzehnte zwischen dem Auftreten des Schadens und seiner abschließenden Regulierung liegen, z.B. bei Produkthaftpflichtschäden (Stichworte: Asbest, Contergan). Aus diesem Grund müssen die Versicherungsunternehmen frühzeitig geeignete Rückstellungen bilden, um die Konsequenzen möglicher "Spätschäden" ausreichend finanziell tragen zu können. Mathematisch stellt sich dabei das Problem, aus nur teilweise beobachteten "Anfangsschäden" auf die Größenordnung der "Spätschäden" zu schließen. Dies wird formal durch die Betrachtung sogenannter *Abwicklungsdreiecke* geregelt, wobei noch zwischen den *Schadenzuwächsen* und den *Schadenständen* unterschieden wird. Recht ausführliche Darstellungen zu diesem Thema, die über den hier gewählten Rahmen deutlich hinausgehen, findet man z.B. in TAYLOR (2000), KAAS ET AL. (2001) oder RADTKE UND SCHMIDT (2004).

Beispiel 19. Einem VU liegen aus den letzten fünf Jahren folgende Informationen zu Schadenzahlungen aus der Kfz-Haftpflichtversicherung vor (in Mio. €):

Anfalljahr	Abwicklungsjahr (Schadenzuwächse)				
	1999	2000	2001	2002	2003
1999	255	354	199	153	34
2000		312	427	155	88
2001			165	201	123
2002				178	204
2003					148

Man sieht, dass die Zahlungen im zweiten Jahr nach Eintritt des Schadens am höchsten ausfallen, um dann in den Folgejahren langsam abzunehmen. Diesen Effekt sieht man allerdings bei der hier gewählten Form der tabellarischen Darstellung nicht sehr deutlich. Aus diesem Grund werden *Abwicklungsdreiecke* fast immer als *relative Abwicklungsdreiecke* behandelt, d.h. in der Form

Anfalljahr	Abwicklungsjahr (Schadenzuwächse)				
	0	1	2	3	4
0	255	354	199	153	34
1	312	427	155	88	
2	165	201	123		
3	178	204			
4	148				

(engl.: *run-off-triangle*). Der genannte Effekt wird jetzt wesentlich deutlicher sichtbar. Für die Schadenstände erhält man analog

II. Sachversicherungsmathematik

Anfalljahr	Abwicklungsjahr (Schadenstände)				
	0	1	2	3	4
0	255	609	808	961	995
1	312	739	894	982	
2	165	366	489		
3	178	382			
4	148				

Ziel der mathematischen Analyse ist es nun, diese Dreiecke in der unteren Hälfte "sinngemäß" zu vervollständigen, um somit eine Übersicht über die zukünftig zu erwartenden Schadenlasten zu erhalten. Es ist naheliegend, hierfür die Verhältniszahlen zwischen den einzelnen Spalten in geeigneter Weise zu benutzen. Da im Abwicklungsjahr 4 nur eine Beobachtung zur Verfügung steht, wird man hier das Verhältnis $995:961 = 1,0354$ ansetzen; beim Übergang von Abwicklungsjahr 2 nach Abwicklungsjahr 3 stehen aber schon vier Werte (graue Felder) zur Verfügung, so dass hier die Betrachtung des Verhältnisses der *Jahresgesamtschadenstände* sinnvoll scheint, also der Quotient $(961 + 982) : (808 + 894) = 1943 : 1702 = 1,1416$. Dieses (als *Chain-Ladder* bezeichnete) Verfahren wird im Folgenden noch ausführlicher besprochen.

Für den Rest dieses Abschnitts wollen wir folgende allgemeine Bezeichnungen für Abwicklungsdreiecke vereinbaren:

Anfalljahr	Abwicklungsjahr (Schadenstände)								
	0	1	...	k	...	$n-i$...	$n-1$	n
0	$S_{0,0}$	$S_{0,1}$...	$S_{0,k}$...	$S_{0,n-i}$...	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$...	$S_{1,k}$...	$S_{1,n-i}$...	$S_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$S_{i,0}$	$S_{i,1}$...	$S_{i,k}$		$S_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$...	$S_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
n	$S_{n,0}$								

Die Größen $S_{i,j}$ seien dabei positiv und (schwach) monoton wachsend in j (kumulierte Jahres-Schäden).

Im schon erwähnten *Chain-Ladder Verfahren* werden nun sogenannte *Abwicklungsfaktoren* betrachtet, die formal definiert sind durch

II. Sachversicherungsmathematik

$$F_k := \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}, \quad k = 1, \dots, n.$$

Das Abwicklungsdreieck wird dann in der unteren Hälfte ergänzt vermöge der Vorschrift

$$\hat{S}_{i,k} := S_{i,n-i} \cdot \prod_{m=n-i+1}^k F_m, \quad i = 0, \dots, n, \quad k = n-i, \dots, n.$$

Beispiel 20 (Fortsetzung von Beispiel 19). Als Vervollständigung nach dem Chain-Ladder Verfahren ergibt sich (graue Felder):

Anfalljahr	Abwicklungsjahr (Schadenstände)					benötigte Reserve
	0	1	2	3	4	
0	255	609	808	961	995	–
1	312	739	894	982	1017	35
2	165	366	489	558	578	89
3	178	382	488	557	577	195
4	148	341	436	498	516	368
F_k	–	2,3033	1,2783	1,1416	1,0354	Summe: 687

Das (heuristische) Chain-Ladder Verfahren erscheint allerdings nur dann sinnvoll, wenn sich die Schadenstandsentwicklungen innerhalb eines Anfalljahres über die Abwicklungsjahre hinweg (also zeilenweise) "ähnlich" sind. Man kann das Chain-Ladder Verfahren aber aus statistischer Sicht rechtfertigen, wenn man folgende Annahmen trifft (sogenanntes *multiplikatives Modell*):

Die Schadenzuwächse

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

mit $i = 1, \dots, n$ und $k = 0, \dots, n$ sind strikt positive Zufallsvariablen mit der Eigenschaft

$$E(Z_{i,k}) = \alpha_i \cdot \vartheta_k, \quad i, k = 0, 1, \dots, n \quad \text{und} \quad \sum_{k=0}^n \vartheta_k = 1$$

für geeignete Parameter $\alpha_0, \dots, \alpha_n > 0$ und $\vartheta_0, \dots, \vartheta_n \in (0, 1)$.

II. Sachversicherungsmathematik

Man kann zeigen, dass dieser Ansatz auch äquivalent ist zu einem multiplikativen Modell für Schadenstände. Hier gilt entsprechend

$$E(S_{i,k}) = \alpha_i \cdot \gamma_k, \quad i, k = 0, 1, \dots, n$$

für geeignete Parameter $\alpha_0, \dots, \alpha_n > 0$ und monoton wachsende $\gamma_0, \dots, \gamma_n > 0$ mit $\gamma_n = 1$. Im multiplikativen Modell für Schadenzuwächse gilt ja

$$E(S_{i,k}) = \sum_{j=0}^k E(Z_{i,j}) = \sum_{j=0}^k \alpha_i \cdot \vartheta_j = \alpha_i \cdot \sum_{j=0}^k \vartheta_j = \alpha_i \cdot \gamma_k \quad \text{mit} \quad \gamma_k = \sum_{j=0}^k \vartheta_j \quad \text{für} \quad i, k = 0, 1, \dots, n$$

(was insbesondere $\gamma_n = \sum_{j=0}^n \vartheta_j = 1$ impliziert). Im multiplikativen Modell für Schadenstände ergibt sich umgekehrt

$$E(Z_{i,k}) = E(S_{i,k}) - E(S_{i,k-1}) = \alpha_i \cdot (\gamma_k - \gamma_{k-1}) = \alpha_i \cdot \vartheta_k \quad \text{für} \quad i = 0, 1, \dots, n, \quad k = 1, \dots, n$$

und

$$E(Z_{i,0}) = E(S_{i,0}) = \alpha_i \cdot \gamma_0 = \alpha_i \cdot \vartheta_0 \quad \text{für} \quad i = 0, 1, \dots, n$$

mit $\sum_{j=0}^n \vartheta_j = \gamma_0 + \sum_{j=1}^n (\gamma_j - \gamma_{j-1}) = \gamma_n = 1$. Für Einzelheiten sei hier z.B. auf SCHMIDT (2002), Abschnitt 11.5 oder RADTKE UND SCHMIDT (2004), Abschnitt *Marginalsummenverfahren* (S. 145 ff) verwiesen.

Das Chain-Ladder Verfahren als solches beruht zunächst auf keiner expliziten stochastische Struktur. Man kann aber die Schadenstände z.B. wie im multiplikativen Modell als Zufallsvariablen $S_{i,k}$ auffassen, deren Verteilungen $P^{S_{i,k}}$ durch die Parameter $\alpha_0, \dots, \alpha_n > 0$ und $\gamma_0, \dots, \gamma_n > 0$ (i. Allg. ohne die Restriktion der Monotonie) charakterisiert sind.

Ein Beispiel eines solchen stochastischen Ansatzes erhält man z.B. durch folgende Annahmen:

- Die Schadenstände $S_{i,k}$ für $i, k = 0, 1, \dots, n$ sind sämtlich positive Zufallsvariablen.
- Die logarithmierten Schadenstände $\ln(S_{i,k})$ sind normalverteilt für $i, k = 0, 1, \dots, n$ und besitzen unabhängige Zuwächse.

Dies bedeutet, dass die Schadenstände $S_{i,k}$ sämtlich lognormalverteilt sind und ihre sukzessiven

Quotienten $F_{i,k} = \frac{S_{i,k}}{S_{i,k-1}}$ stochastisch unabhängig und ebenfalls lognormalverteilt sind für

$i = 0, 1, \dots, n, \quad k = 1, \dots, n$. Damit erhält man folgende stochastische Struktur des Modells:

$$S_{i,k} = S_{i,0} \cdot \prod_{j=1}^k F_{i,j} \quad \text{für} \quad i = 0, 1, \dots, n, \quad k = 1, \dots, n$$

mit einer Familie stochastisch unabhängiger, lognormalverteilter Zufallsvariablen $\{F_{i,j}\}_{i,j=0,\dots,n}$, für die gilt:

II. Sachversicherungsmathematik

$$F_{i,0} = S_{i,0} \text{ mit } E(F_{i,0}) = E(S_{i,0}) = \alpha_i \cdot \gamma_0 \text{ für } i = 0, 1, \dots, n \text{ und}$$

$$E(F_{i,k}) = \varphi_k = \frac{\gamma_k}{\gamma_{k-1}} > 0 \text{ für } i = 0, 1, \dots, n, k = 1, \dots, n.$$

Es folgt dann nämlich wunschgemäß

$$E(S_{i,k}) = E(S_{i,0}) \cdot \prod_{j=1}^k E(F_{i,j}) = \alpha_i \cdot \gamma_0 \cdot \prod_{j=1}^k \varphi_j = \alpha_i \cdot \gamma_0 \cdot \prod_{j=1}^k \frac{\gamma_j}{\gamma_{j-1}} = \alpha_i \cdot \gamma_k, \quad i, k = 0, 1, \dots, n.$$

Wenn man zusätzlich annimmt, dass die Quotienten $F_{i,k}$ – unabhängig von i – identisch verteilt sind, dann lassen sich die beiden Verteilungs-Parameter μ_k und σ_k^2 der Lognormal-Verteilung für $k \geq 1$ spaltenweise aus den logarithmierten Daten für $i = 0, \dots, n - k$ über den jeweiligen Mittelwert bzw. die empirische Varianz schätzen. Lediglich für die letzten Quotienten $F_{i,n}$ ist diese Methode bezüglich der Varianz nicht anwendbar, da hier für die Schätzung nur eine einzige Beobachtung vorliegt. Man muss sich daher mit einer Behelfslösung begnügen (z.B. indem man die Varianz der vorletzten Spalte übernimmt oder anderweitig geeignet festsetzt).

Für das obige Beispiel erhält man auf diese Weise folgende Schätzer:

k	1	2	3	4
$\hat{\mu}_k$	0,8233	0,2543	0,1336	0,0348
$\hat{\sigma}_k^2$	0,0027	0,0031	0,0032	(0,0032)

Hieraus lassen sich die Schätzungen $\tilde{\mu}_i$ und $\tilde{\sigma}_i^2$ für die entsprechenden Verteilungs-Parameter der Faktoren-Produkte $G_i = \prod_{k=n-i+1}^n F_{i,k}$, $i = 1, \dots, n$, die die (letzten beobachteten) Schadenstände auf der Diagonalen auf den Schadenendstand hochrechnen, durch Addition bestimmen, mit

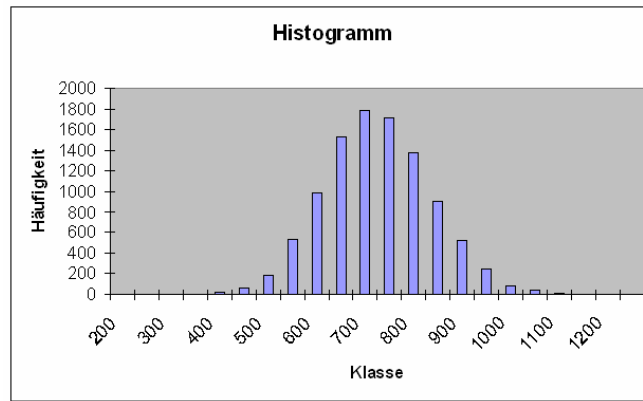
$$\tilde{\mu}_i = \sum_{k=n-i+1}^n \hat{\mu}_k, \quad \tilde{\sigma}_i^2 = \sum_{k=n-i+1}^n \hat{\sigma}_k^2, \quad i = 1, \dots, n.$$

Für das obige Beispiel ergibt das:

i	1	2	3	4
$\tilde{\mu}_i$	0,0348	0,1684	0,4227	1,2460
$\tilde{\sigma}_i^2$	0,0032	0,0064	0,0094	0,0121

Mit Hilfe eines Monte-Carlo-Verfahrens kann man so etwa die Schadenendstände simulieren und daraus die empirische Verteilung der Reserve bestimmen. Mit EXCEL ergibt sich beispielsweise mit 10.000 Simulationen folgendes Histogramm für die Reserve:

II. Sachversicherungsmathematik



mit einem empirischen Erwartungswert von 700,11 und einer empirischen Standardabweichung von 112,58 (zum Vergleich: mit der Chain-Ladder Methode ergab sich eine geschätzte erwartete Reserve von 687). Der aus den Schätzern *berechnete* Erwartungswert beträgt 701,22, die Standardabweichung 109,16.

II. 6. Prämienprinzipien

In den bisherigen Betrachtungen haben wir fast immer nur *Netto-Prämien* für Versicherungsleistungen in der Sachversicherung betrachtet. Diese leiten sich aus dem schon mehrfach erwähnten *Äquivalenzprinzip* ab, das auf dem Gesetz der großen Zahlen beruht und daher aus der Sicht der Versicherten "die" faire Prämie schlechthin darstellt. Allerdings muss die Versichertengemeinschaft auch Interesse daran haben, das das VU wirtschaftlich arbeiten und langfristig bestehen kann, d.h. dass ein möglicher finanzieller Ruin – ausgelöst z.B. durch "zu viele" Versicherungsfälle zur gleichen Zeit – nur mit einer sehr geringen Wahrscheinlichkeit eintritt. Wir wollen diese Problematik einmal an der folgenden einfachen Überlegung verdeutlichen:

Die – als stochastisch unabhängig und identisch verteilt vorausgesetzten – nicht-negativen Zufallsvariablen $\{X_n\}_{n \in \mathbb{N}}$ mögen die jährlichen Schäden (bzw. Versicherungsleistungen) eines Kollektivs von Versicherten bezeichnen, mit den Erwartungswerten und Varianzen $E(X_k) = \mu < \infty$, $Var(X_k) = \sigma^2 < \infty$ für alle $k \in \mathbb{N}$. Dann bezeichnet

$$S_n := \sum_{k=1}^n X_k \text{ für } n \in \mathbb{N}$$

die insgesamt bis zum Jahr n geleisteten Zahlungen; bei Zugrundelegung von Nettoprämien stehen dem gegenüber insgesamt Einnahmen in Höhe von $n\mu$. Die Differenz

$$L_n := S_n - n\mu \text{ für } n \in \mathbb{N}$$

gibt also den Verlust des VU im Jahr n an. Nach dem Zentralen Grenzwertsatz gilt nun aber:

$$\lim_{n \rightarrow \infty} P\left(\frac{L_n}{\sigma\sqrt{n}} \leq x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} P\left(\frac{L_n}{\sigma\sqrt{n}} > x\right) = 1 - \Phi(x)$$

für alle $x \in \mathbb{R}$; für genügend große n erhält man so etwa approximativ

II. Sachversicherungsmathematik

$$P(L_n > \sigma\sqrt{n}) \approx 1 - \Phi(1) = \Phi(-1) = 0,1587;$$

d.h. die Wahrscheinlichkeit dafür, dass im 25. Jahr ein Verlust von mehr als der fünffachen Standardabweichung des Risikos eintritt, beträgt etwa 16%!

Noch genauer kann diese Aussage mit Hilfe des berühmten *Gesetzes vom iterierten Logarithmus* präzisiert werden (vgl. etwa BILLINGSLEY (1986), Theorem 9.5), das für die hier vorliegende Situation folgendes aussagt:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{L_n}{\sqrt{2n \ln \ln n}} &= +\sigma \\ \liminf_{n \rightarrow \infty} \frac{L_n}{\sqrt{2n \ln \ln n}} &= -\sigma \end{aligned} \quad \text{mit Wahrscheinlichkeit 1;}$$

dies impliziert insbesondere, dass mit Wahrscheinlichkeit 1 im Laufe der Jahre unendlich oft ein Verlust *beliebiger Höhe* eintritt! Aus diesem Grund werden auf die Nettoprämien in der Regel sogenannte *Sicherheitszuschläge* erhoben. Verlangt man etwa statt der Prämie μ für das Risiko X die höhere Prämie $\mu + \delta$ mit $\delta > 0$, so erhält man wieder mit dem Zentralen Grenzwertsatz approximativ für den (modifizierten) Verlust

$$L_n^\delta := S_n - n(\mu + \delta) = L_n - n\delta \quad \text{für } n \in \mathbb{N}$$

die Abschätzung

$$P(L_n^\delta > z) = P(L_n > z + n\delta) = P\left(\frac{L_n}{\sigma\sqrt{n}} > \frac{z + n\delta}{\sigma\sqrt{n}}\right) \approx 1 - \Phi\left(\frac{z + n\delta}{\sigma\sqrt{n}}\right) = 1 - \Phi\left(\frac{z}{\sigma\sqrt{n}} + \sqrt{n} \frac{\delta}{\sigma}\right)$$

mit

$$\lim_{n \rightarrow \infty} P(L_n^\delta > z) = 0 \quad \text{für alle } z > 0.$$

Der Ausdruck auf der rechten Seite der Approximation wird dabei maximal $1 - \Phi\left(\frac{2\sqrt{\delta z}}{\sigma}\right)$. Will

man also gleichmäßig über alle Jahre eine "Ruinwahrscheinlichkeit" von approximativ $\alpha \in (0,1)$ nicht überschreiten, so kann man etwa

$$\delta = \left(\sigma \frac{\Phi^{-1}(1-\alpha)}{2\sqrt{z}}\right)^2$$

wählen. Dabei geht man vereinbarungsgemäß von "Ruin" aus, wenn ein Verlust größer als der Wert $z > 0$ ausfällt.

Beispiel 21 (Fortsetzung von Beispiel 16). Aus den Daten schätzen wir $\mu = 3605$ und $\sigma = 3710$. Wenn wir davon ausgehen, dass eine finanzielle Reserve von $z = 10000$ vorhanden ist und $\alpha = 5\%$ gewählt wird, d.h. es ist $\Phi^{-1}(1-\alpha) = 1,645$, so ergibt sich $\delta = 931$. Der Sicherheitszuschlag beträgt dann also ca. 26% auf die Nettoprämie.

Es gibt offensichtlich keine eindeutig bestimmte Vorgehensweise, wie die Nettoprämie geeignet zu modifizieren ist, um den genannten Wirtschaftlichkeitsüberlegungen Rechnung zu tragen. Man

II. Sachversicherungsmathematik

kann aber versuchen, durch eine Reihe von Anforderungen an eine solche Modifizierung die Wahlmöglichkeiten geeignet einzuschränken. In diesem Zusammenhang kommen nun die sogenannten *Prämienprinzipien* ins Spiel.

Definition 6. Wir bezeichnen mit \mathcal{Z} die Menge aller nicht-negativen Zufallsvariablen. Ein Prämi-
enprinzip ist dann eine Abbildung $H : \mathcal{Z} \rightarrow \mathbb{R}^+$ mit der Eigenschaft¹⁸

$$P^X = P^Y \Rightarrow H(X) = H(Y) \quad \text{für alle } X, Y \in \mathcal{Z}.$$

Diese Bedingung besagt, dass ein Prämiensprinzip nur von der *Verteilung* der Zufallsvariablen (in der Sprache der Versicherungsmathematik: des Risikos) abhängt, nicht aber von der konkreten Definition als (messbare) Abbildung.

Offensichtlich ist die Nettoprämie ein Prämiensprinzip, da der Erwartungswert einer Zufallsvariablen die obige Eigenschaft besitzt.

Neben dieser (simply) definierenden Eigenschaft unterscheidet man in der Praxis u.a. die folgenden Eigenschaften, die Prämiensprinzipien haben können bzw. haben sollen.

Definition 7. Es sei H ein Prämiensprinzip im Sinne von Definition 6. Man sagt, H sei

- a) *erwartungswertübersteigend (eü)*, wenn gilt: $H(X) \geq E(X)$ für alle $X \in \mathcal{Z}$;
- b) *positiv homogen (ph)*, wenn gilt: $H(cX) = cH(X)$ für alle $c \geq 0$ und $X \in \mathcal{Z}$;
- c) *additiv (ad)*, wenn gilt: $H(X + Y) = H(X) + H(Y)$ für alle $X, Y \in \mathcal{Z}$, die *stochastisch unabhängig* sind;
- d) *maximalschadenbegrenzt (ms)*, wenn gilt: $H(X) \leq \varpi_x := \sup\{x \in \mathbb{R} \mid F_x(x) < 1\}$ für alle $X \in \mathcal{Z}$, wobei F_x wie üblich die Verteilungsfunktion von X bezeichne;
- e) *stochastisch monoton (sm)*, wenn gilt: $H(X) \leq H(Y)$ für alle $X, Y \in \mathcal{Z}$, die *stochastisch geordnet* sind, d.h. für die gilt: $F_x(x) \geq F_y(x)$ für alle $x \in \mathbb{R}$.¹⁹

Die Eigenschaften a) und d) sollte jedes Prämiensprinzip aus praktischen Gründen mindestens besitzen; zu a) haben wir bereits oben einiges ausgeführt, d) bedeutet, dass die Versicherung überhaupt ein Risiko trägt (anderenfalls wäre es für den Versicherungsnehmer besser, erst gar keine Versicherung abzuschließen). Die Eigenschaft b) der positiven Homogenität bedeutet, dass die monetäre Skala keine Rolle spielt, also die Prämie in € oder US-\$ äquivalent ist. Eigenschaft c) ist selbsterklärend, und Eigenschaft e) bedeutet, dass eine Versicherung von Risiken Y , die häufiger bestimmte Grenzen x überschreiten als Risiken X , entsprechend teurer ist.

Es ist offensichtlich, dass die Nettoprämie formal allen fünf Eigenschaften genügt; allerdings ist sie aus den bekannten Gründen wirtschaftlich nicht vertretbar. Es ist daher wichtig, auch andere Prämiensprinzipien auf ihre Anwendbarkeit zu untersuchen. Im Folgenden sollen deshalb einige praxisrelevante Prämiensprinzipien genauer untersucht werden.

¹⁸ Diese Definition ist in der Literatur nicht eindeutig; vgl. etwa SCHMIDT (2002), Abschnitt 10.1.

¹⁹ anschaulich bedeutet dies, dass die Zufallsvariable Y "häufiger" große Werte annimmt als X .

II. Sachversicherungsmathematik

Definition 8. Es werden die Bezeichnungen der Definitionen 6 und 7 verwendet.

a) Es sei $\delta \geq 0$. Dann heißt H mit

$$H(X) = (1 + \delta) E(X), \quad X \in \mathcal{Z}$$

Erwartungswertprinzip (EwP) mit Sicherheitszuschlag δ .

b) Es sei $\delta \geq 0$. Falls die Varianz $Var(X)$ existiert, heißt H mit

$$H(X) = E(X) + \delta Var(X), \quad X \in \mathcal{Z}$$

Varianzprinzip (VaP) mit Zuschlagsfaktor δ .

c) Es sei $\delta \geq 0$. Falls die Varianz $Var(X)$ existiert, heißt H mit

$$H(X) = E(X) + \delta \sqrt{Var(X)}, \quad X \in \mathcal{Z}$$

Standardabweichungsprinzip (StP) mit Zuschlagsfaktor δ .

d) Es sei $g: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ eine streng monoton wachsende, konvexe Funktion. Dann heißt (im Fall der Existenz des Erwartungswerts) H mit

$$H(X) = g^{-1}(E(g(X))), \quad X \in \mathcal{Z}$$

Mittelwertprinzip (MwP) bezüglich g .

e) Es sei $g: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ eine (schwach) monoton wachsende Funktion. Dann heißt (im Fall der Existenz des Erwartungswerts) H mit

$$H(X) = \frac{E[X \cdot g(X)]}{E[g(X)]}, \quad X \in \mathcal{Z}$$

Esscherprinzip (EsP) bezüglich g .

f) Es sei $\varepsilon \in (0,1)$. Dann heißt H mit

$$H(X) = F_x^{-1}(1 - \varepsilon) = \inf \{x \in \mathbb{R}^+ \mid F_x(x) \geq 1 - \varepsilon\}, \quad X \in \mathcal{Z}$$

Perzentilprinzip (PzP) zur Risiko-Wahrscheinlichkeit ε .

Lemma 21. Die folgende Tabelle gibt eine Übersicht über die Eigenschaften nach Definition 7 für die Prämienprinzipien nach Definition 8.

Eigenschaft	<i>eü</i>	<i>ph</i>	<i>ad</i>	<i>ms</i>	<i>sm</i>
Prinzip					
<i>EwP</i>	ja	ja	ja	nein	ja
<i>VaP</i>	ja	nein	ja	nein	nein
<i>StP</i>	ja	ja	nein	nein	nein
<i>MwP</i>	ja	nein	nein	ja	ja
<i>EsP</i>	ja	nein	nein	ja	nein
<i>PzP</i>	nein	ja	nein	ja	ja

II. Sachversicherungsmathematik

Beweis: Wir zeigen die Aussagen zeilenweise.

EwP: Es ist $H(X) = (1 + \delta)E(X) = E(X) + \delta E(X) \geq E(X)$, also gilt (*eü*).

Für $\delta \geq 0$ ist $H(cX) = (1 + \delta)E(cX) = c(1 + \delta)E(X) = cH(X)$, also gilt (*ph*).

Es ist $H(X + Y) = (1 + \delta)E(X + Y) = (1 + \delta)E(X) + (1 + \delta)E(Y) = H(X) + H(Y)$
(auch ohne die Voraussetzung der stochastischen Unabhängigkeit), also gilt (*ad*).

Falls ϖ_x endlich und $E(X) > 0$ ist, verletzt H für jede Wahl $\delta > \frac{\varpi_x - E(X)}{E(X)} \left[\geq 0 \right]$ die Eigenschaft (*ms*).

Es ist $H(X) = (1 + \delta)E(X) = (1 + \delta) \int_0^{\infty} (1 - F_X(x)) dx \leq (1 + \delta) \int_0^{\infty} (1 - F_Y(x)) dx$
 $= (1 + \delta)E(Y) = H(Y)$, also ist (*sm*) erfüllt.

VaP: Es ist $H(X) = E(X) + \delta \text{Var}(X) \geq E(X)$, also gilt (*eü*).

Für $c > 1$ und $\text{Var}(X) > 0$ ist

$$H(cX) = cE(X) + \delta c^2 \text{Var}(X) > cE(X) + \delta c \text{Var}(X) = cH(X),$$

also gilt (*ph*) nicht.

Es ist $H(X + Y) = E(X + Y) + \delta \text{Var}(X + Y) = E(X) + E(Y) + \delta \text{Var}(X) + \delta \text{Var}(Y)$
 $= H(X) + H(Y)$

(hier ist die Voraussetzung der stochastischen Unabhängigkeit wesentlich), also gilt (*ad*).

Falls ϖ_x endlich und $\text{Var}(X) > 0$ ist, verletzt H für jede Wahl $\delta > \frac{\varpi_x - E(X)}{\text{Var}(X)} \left[\geq 0 \right]$ die Eigenschaft (*ms*).

Gegenbeispiel: $P^X = R[0,1]$, $Y = \frac{1}{3}X + 1 \geq X$, $\delta = 12$, dann ist

$$H(X) = E(X) + \delta \text{Var}(X) = \frac{1}{2} + 12 \cdot \frac{1}{12} = \frac{3}{2} > \frac{23}{18} = \frac{7}{6} + \frac{1}{9} = E(Y) + \delta \text{Var}(Y),$$

also gilt (*sm*) nicht.

StP: Es ist $H(X) = E(X) + \delta \sqrt{\text{Var}(X)} \geq E(X)$, also gilt (*eü*).

Für $c \geq 0$ ist $H(cX) = cE(X) + \delta c \sqrt{\text{Var}(X)} = cH(X)$, also gilt (*ph*).

(*ad*) gilt nicht, weil i.Allg. $\sqrt{a+b} \neq \sqrt{a} + \sqrt{b}$ für $a, b \geq 0$.

Falls ϖ_x endlich und $\text{Var}(X) > 0$ ist, verletzt H für jede Wahl $\delta > \frac{\varpi_x - E(X)}{\sqrt{\text{Var}(X)}} \left[\geq 0 \right]$

die Eigenschaft (*ms*).

Gleiches Gegenbeispiel wie für *VaP* impliziert: (*sm*) gilt nicht.

II. Sachversicherungsmathematik

MwP: Mit der Jensen'schen Ungleichung²⁰ folgt:

$$H(X) = g^{-1}(E(g(X))) \geq g^{-1}(g(E(X))) = E(X), \text{ also gilt (eü).}$$

Gegenbeispiel: $g(x) = e^x$; hier ist für $c > 0$ i.Allg.

$$H(cX) = \ln(E(e^{cX})) = \ln \psi_X(c) \neq c \ln \psi_X(1) = c \ln(E(e^X)); \text{ so ist etwa bei } \mathcal{E}(2)\text{-Ex-}$$

ponentialverteilung: $\psi_X(c) = \frac{1}{2-c}$ für $0 < c < 2$, d.h. (*ph*) ist verletzt.

Gegenbeispiel: $g(x) = x^2$; i.Allg. gilt bei $\text{Var}(X) > 0$, $\text{Var}(Y) > 0$ mit der Jensen'schen Ungleichung:

$$\begin{aligned} H^2(X+Y) &= E(X+Y)^2 = E(X^2) + E(Y^2) + 2E(X)E(Y) \\ &= H^2(X) + H^2(Y) + 2E(X)E(Y) \\ &= (H(X) + H(Y))^2 + 2(E(X)E(Y) - H(X)H(Y)) < (H(X) + H(Y))^2 \end{aligned}$$

wegen $H(X) = \sqrt{E(X^2)} > \sqrt{(E(X))^2} = E(X)$, also gilt (*ad*) nicht.

Es ist $H(X) = g^{-1}(E(g(X))) \leq g^{-1}(E(g(\varpi_X))) = g^{-1}(g(\varpi_X)) = \varpi_X$, falls $\varpi_X < \infty$, sonst trivial, also gilt (*ms*).

(*sm*) gilt, weil

$$\begin{aligned} E(g(X)) &= \int_0^\infty (1 - F_{g(X)}(x)) dx = \int_0^\infty P(X > g^{-1}(x)) dx = \int_0^\infty (1 - F_X(g^{-1}(x))) dx \\ &\leq \int_0^\infty (1 - F_Y(g^{-1}(x))) dx = E(g(Y)) \text{ und damit} \end{aligned}$$

$$H(X) = g^{-1}(E(g(X))) \leq g^{-1}(E(g(Y))) = H(Y) \text{ ist.}$$

EsP:

Mit der Abkürzung $\mu = E(X)$ ist $H(X) = \mu + \frac{E[(X - \mu)(g(X) - g(\mu))]}{E[g(X)]} \geq \mu$ (folgt durch

Ausmultiplizieren und der Tatsache, dass $X \geq \mu \Leftrightarrow g(X) \geq g(\mu)$). Also gilt (*eü*).

Gegenbeispiel: $g(x) = 1 + x$; dann ist für $c > 0$, $c \neq 1$ und $P^X = \mathcal{E}(1)$

$$H(cX) = \frac{cE(X + cX^2)}{E(1 + cX)} = c \frac{1 + 2c}{1 + c} \neq \frac{3c}{2} = \frac{cE(X + X^2)}{E(1 + X)} = cH(X), \text{ also gilt (ph) nicht.}$$

Gegenbeispiel: $g(x) = x$, $P^X = P^Y = \mathcal{P}(\lambda)$ mit $\lambda > 0$:

$$H(X+Y) = \frac{E((X+Y)^2)}{E(X+Y)} = \frac{2\lambda + 4\lambda^2}{2\lambda} = 1 + 2\lambda < 2 + 2\lambda = H(X) + H(Y), \text{ also gilt (ad)}$$

nicht.

$$\text{Es ist } H(X) = \frac{E[X \cdot g(X)]}{E[g(X)]} \leq \frac{E[\varpi_X \cdot g(X)]}{E[g(X)]} = \varpi_X, \text{ also gilt (ms).}$$

²⁰ Für konvexe Abbildungen h gilt: $E(h(X)) \geq h(E(X))$, sofern alle Erwartungswerte existieren. Bei Konkavität gilt das umgekehrte Relationszeichen.

II. Sachversicherungsmathematik

Gegenbeispiel: $g(x) = x$; hier ist $H(X) = \frac{E(X^2)}{E(X)} = E(X) + \frac{\text{Var}(X)}{E(X)}$. Wähle X so, dass $E(X) = \frac{1}{2}$, $\text{Var}(X) = 1$, z.B. $P^X = \mathcal{LN}(\mu, \sigma^2)$ mit $\mu = -\frac{1}{2} \ln 20$, $\sigma^2 = \ln 5$. Mit der Wahl $Y := X + \frac{1}{2}$ ist dann $X < Y$, aber $H(X) = \frac{5}{2} > 2 = H(Y)$. Also gilt *(sm)* nicht.

PzP: Für nicht-konstantes X und genügend große $\varepsilon > 0$ ist $H(X) = F_X^{-1}(1-\varepsilon)$ nahe Null und somit kleiner als $E(X)$, also gilt *(eü)* nicht in jedem Fall.

Die Gültigkeit von *(ph)* folgt aus $P(X \leq x) = P(cX \leq cx)$, $c > 0$; $x \in \mathbb{R}^+$.

Gegenbeispiel: $P^X = P^Y = \mathcal{P}(\lambda)$ (Poisson-Verteilungen) mit Parameter $\lambda > 0$; für $e^{-2\lambda} < 1-\varepsilon < e^{-\lambda}$ ist dann $H(X) = H(Y) = 0$, $H(X+Y) \geq 1$, also gilt *(ad)* nicht.

Wegen $H(X) = F_X^{-1}(1-\varepsilon) \leq F_X^{-1}(1) = \varpi_x$ folgt *(ms)*.

Die Eigenschaft *(sm)* ergibt sich aus der Definition der (Pseudo-)Inversen F_X^{-1} .

Die globale Verletzung einiger Kriterien wie *(eü)* oder *(ms)* durch gewisse Prämienprinzipien ist nicht notwendig kritisch, da meist durch entsprechende Wahl der Parameter (in Abhängigkeit der Verteilung des Risikos) die Einhaltung erzwungen werden kann. So lässt sich beim Perzentilprinzip das $\varepsilon \in (0,1)$ bei Kenntnis von P^X immer so wählen, dass $H(X)$ erwartungswertübersteigend ist. Die Zuschlagsfaktoren beim Erwartungswert-, Varianz- und Standardabweichungsprinzip lassen sich entsprechend klein wählen, so dass $H(X)$ dann maximalschadenbegrenzt ist.

Bemerkung. Wählt man im Esscher-Prinzip speziell $g(x) = x^\beta$, $x \in \mathbb{R}^+$ mit $\beta > 0$, so erhält man das sogenannte *Karlsruhe-Prinzip (KaP)*. Dieses ist im Gegensatz zum allgemeinen Fall positiv-homogen. Für $\beta = 1$ ergibt sich speziell noch

$$H(X) = \frac{E(X^2)}{E(X)} = E(X) + \frac{\text{Var}(X)}{E(X)}.$$

Für die Wahl $g(x) = e^{\beta x}$, $x \in \mathbb{R}^+$ mit $\beta > 0$ erhält man das sogenannte *spezielle Esscher-Prinzip*. Allerdings setzt dies die Existenz der momenterzeugenden Funktion voraus, weil dann

$$H(X) = \frac{E(Xe^{\beta X})}{E(e^{\beta X})} = \frac{\psi'_X(\beta)}{\psi_X(\beta)} = \frac{d}{d\beta} \ln \psi_X(\beta)$$

gilt. Für das Mittelwertprinzip ergibt sich bei derselben Wahl $g(x) = e^{\beta x}$, $x \in \mathbb{R}^+$ mit $\beta > 0$ noch das *Exponentialprinzip (ExpP)* vermöge

$$H(X) = \frac{1}{\beta} \ln \left(E(e^{\beta X}) \right) = \frac{1}{\beta} \ln \psi_X(\beta).$$

II. Sachversicherungsmathematik

Das eingangs betrachtete Beispiel mit dem "Sicherheitszuschlag" $\delta = \left(\sigma \frac{\Phi^{-1}(1-\alpha)}{2\sqrt{z}} \right)^2$ beinhaltet

Aspekte verschiedener Prämienprinzipien. Wählt man z und α unabhängig von σ , so ergibt sich ein Varianzprinzip; wählt man z (die finanzielle Reserve) proportional zu σ , so erhält man ein Standardabweichungsprinzip; von der Intention her entspricht die gesamte Prämienberechnung einem Perzentilprinzip (Kontrolle der Ruinwahrscheinlichkeit).

Beispiel 22. Es sei X exponentialverteilt mit Parameter $\lambda = 1$ und $\beta < 1$. Dann gilt:

Prinzip	<i>EwP</i>	<i>VaP</i>	<i>StP</i>	<i>ExP</i>	<i>EsP²¹</i>	<i>KaP</i>	<i>PzP</i>
$H(X)$	$1 + \delta$	$1 + \delta$	$1 + \delta$	$-\frac{1}{\beta} \ln(1 - \beta)$	$\frac{1}{1 - \beta}$	$1 + \beta$	$-\ln(1 - \epsilon)$

Beim Karlsruhe-Prinzip ist nur zu beachten, dass $E(X^\vartheta) = \Gamma(\vartheta + 1)$ für $\vartheta > 0$ gilt und somit

$$\frac{E(X^{\beta+1})}{E(X^\beta)} = \frac{\Gamma(\beta+2)}{\Gamma(\beta+1)} = \beta + 1 \text{ ist.}$$

Wir wollen uns zum Abschluß dieses Abschnitts noch mit einer anderen Klasse von Prämienprinzipien beschäftigen, die nutzentheoretisch begründet sind und ihren Ursprung in betriebswirtschaftlichen Fragestellungen haben.

Definition 9. Eine zweimal stetig differenzierbare Abbildung $u: \mathbb{R} \rightarrow \mathbb{R}$ mit den Eigenschaften

$$\begin{aligned} u(0) &= 0 \\ u'(x) &> 0 \text{ und } u'(0) = 1 \\ u''(x) &< 0, \quad x \in \mathbb{R} \end{aligned}$$

heißt Standard-Nutzenfunktion²². Insbesondere ist u also monoton wachsend und konkav.

Mit Nutzenfunktionen werden z.B. Kapitalien bewertet. Die (aus der dritten Bedingung folgende) Konkavität bedeutet dabei, dass ein Kapitalzuwachs von 1 € bei einem hohem Grundkapital weniger Nutzenzuwachs mit sich bringt als bei einem kleinen Grundkapital. Bei negativen Kapitalien (Schulden) verhält es sich entsprechend umgekehrt. Definitionsgemäß besitzt "kein" Kapital auch keinen Nutzen (erste Bedingung), und der Nutzenzuwachs bei Null Kapital Ausgangssituation ist näherungsweise proportional zu einem geringen Kapitalzuwachs (Bedingung 2).

Definition 10. Es sei X ein Risiko und u eine Standard-Nutzenfunktion. Dann heißt das durch

$$E[u(H(X) - X)] = 0$$

(implizit) definierte Prämienprinzip *Nullnutzenprinzip* zur Nutzenfunktion u .

²¹ hier: spezielles Esscher-Prinzip.

²² In der Literatur wird ansonsten nur Monotonie und Konkavität vorausgesetzt; vgl. SCHMIDT (2002), Abschnitt 10.4.

II. Sachversicherungsmathematik

Anschaulich bedeutet das Nullnutzenprinzip, dass sich der durchschnittliche Nutzen aus Verlusten (Versicherungsleistungen, die die Nettoprämie übersteigen, stärker gewichtet) und Gewinnen (Versicherungsleistungen, die unter der Nettoprämie bleiben, schwächer gewichtet) die Waage hält.

Lemma 22. Ein Nullnutzenprinzip zu einer Nutzenfunktion u ist erwartungswertübersteigend.

Beweis: Mit der (schon oben erwähnten) Jensen'schen Ungleichung ergibt sich

$$0 = E[u(H(X) - X)] \leq u(E[H(X) - X]) = u(H(X) - E(X)),$$

und damit wegen der Monotonie von u auch $0 \leq H(X) - E(X)$, also $H(X) \geq E(X)$.

Leider sind Nullnutzenprinzipien i.Allg. weder positiv-homogen noch additiv noch stochastisch monoton, wodurch ihre Relevanz für die Versicherungsmathematik eher gering ist.

Beispiel 23. Für den sogenannten *Exponentialnutzen*

$$u(x) = \frac{1}{\beta}(1 - e^{-\beta x}), \quad x \in \mathbb{R}$$

erhält man $u''(0) = -\beta$ und als Nullnutzenprinzip das *Exponentialprinzip* wegen

$$0 = E[u(H(X) - X)] = \frac{1}{\beta} E[1 - e^{-\beta(H(X) - X)}], \text{ also } 1 - e^{-\beta H(X)} E(e^{\beta X}) = 0 \text{ und damit}$$

$$H(X) = \frac{1}{\beta} \ln(E(e^{\beta X})).$$

Nullnutzenprinzipien weisen eine gewisse Ähnlichkeit zu Varianzprinzipien auf, wie das folgende Ergebnis zeigt.

Lemma 23. Es sei \mathcal{U} eine Familie von Nutzenfunktionen u_β mit Parameter $\beta = -u''_\beta(0)$, $\beta > 0$. Dann gilt approximativ für das Nullnutzenprinzip H :

$$H(X) \approx E(X) + \frac{\beta}{2} \text{Var}(X) \text{ für kleine Werte von } \beta,$$

d.h. für diese Situation ist das Nullnutzenprinzip approximativ ein Varianzprinzip mit Zuschlagsfaktor $\beta/2$.

Beweis: Zur Abkürzung setzen wir $\mu = E(X)$, $\sigma^2 = \text{Var}(X)$. Mit der Taylor-Entwicklung von u_β

um den Nullpunkt²³ ergibt sich für $H^*(X) := \mu + \frac{\beta}{2} \sigma^2$

²³ d.h. $u_\beta(x) \approx u_\beta(0) + x \cdot u'_\beta(0) + \frac{x^2}{2} \cdot u''_\beta(0) = x - \frac{\beta}{2} x^2$

II. Sachversicherungsmathematik

$$\begin{aligned} E[u_\beta(H^*(X) - X)] &\approx E\left[H^*(X) - X - \frac{\beta}{2}(H^*(X) - X)^2\right] = H^*(X) - \mu - \frac{\beta}{2}E(H^*(X) - X)^2 \\ &= \frac{\beta}{2}\sigma^2 - \frac{\beta}{2}\left((H^*(X) - \mu)^2 + \sigma^2\right) = -\frac{\beta^3}{8}\sigma^4 \approx 0 \end{aligned}$$

für kleine Werte von β , woraus $H(X) \approx H^*(X)$, also die Behauptung folgt.

Alternativ kann man auch mit der Substitution $z := H^*(X) - \mu$ die quadratische Gleichung

$$E\left[H^*(X) - X - \frac{\beta}{2}(H^*(X) - X)^2\right] = H^*(X) - \mu - \frac{\beta}{2}E(H^*(X) - X)^2 = z - \frac{\beta}{2}(z^2 + \sigma^2) = 0$$

lösen; man erhält hieraus zunächst die beiden Lösungen $z_{1,2} = \frac{1}{\beta} \pm \frac{1}{\beta} \sqrt{1 - \beta^2 \sigma^2}$, von denen hier aber nur die zweite sinnvoll ist mit

$$z_2 = \frac{1}{\beta} - \frac{1}{\beta} \sqrt{1 - \beta^2 \sigma^2} \approx \frac{1}{\beta} - \frac{1}{\beta} \left(1 - \frac{\beta^2}{2} \sigma^2\right) = \frac{\beta}{2} \sigma^2,$$

was wieder zur obigen „Lösung“ $H^*(X) = \mu + \frac{\beta}{2} \sigma^2$ führt.

Beispiel 24 (Fortsetzung von Beispiel 22). Für den Exponentialnutzen erhalten wir das Exponentialprinzip als Nullnutzenprinzip, d.h. hier, für $\beta < 1$ (siehe oben)

$$H(X) = -\frac{1}{\beta} \ln(1 - \beta) = \frac{1}{\beta} \sum_{k=1}^{\infty} \frac{\beta^k}{k} = \sum_{k=0}^{\infty} \frac{\beta^k}{k+1} = 1 + \frac{\beta}{2} + \frac{\beta^2}{3} + \dots$$

Das approximative Varianzprinzip liefert zum Vergleich

$$H^*(X) = 1 + \frac{\beta}{2},$$

also den Reihenanfang mit zwei Termen und einem Fehler von maximal

$$|H(X) - H^*(X)| \leq \frac{\beta^2}{3} \sum_{k=0}^{\infty} \beta^k = \frac{\beta^2}{3(1-\beta)}.$$

Eine weiterführende Diskussion von Prämienprinzipien findet man z.B. in SCHMIDT (2002) oder HEILMANN (1987).

III Credibility-Theorie

Bei der bisherigen Modellbildung – vor allem im Sachversicherungsbereich – sind wir davon ausgegangen, dass die einzelnen Risiken durch eine einzige, geschlossene Verteilung dargestellt werden können. Häufig ist es aber so, dass das Gefährdungspotenzial z.B. durch menschliches Handeln aktiv beeinflusst wird, etwa in der Kfz-Haftpflichtversicherung, wo diesbezüglich Fahrstil, Fahrpraxis, Lebensalter, Wohnort (Stadt / Land) usw. eine große Rolle spielen. Aber auch in der Gebäudeversicherung können quantifizierbare Umwelteinflüsse Auswirkungen auf das Gefährdungspotenzial haben, z.B. die geographische Lage (Flussnähe für Überschwemmungsrisiken, Hanglage im Gebirge für Erdbeben usw.). Man ist daher im Versicherungsgeschäft durchaus an geeigneten Differenzierungsmöglichkeiten interessiert, die sich i.a. auch in entsprechenden Prämienendifferenzierungen niederschlagen (etwa Bonus-Malus-Systeme in der Kfz-Haftpflichtversicherung). Die mathematische Grundannahme in der Credibility-Theorie (auch *Erfahrungstari-fierung* genannt) ist dabei die folgende:

1. Das der Versicherung zu Grunde liegende Risiko X besitzt eine gemischte Verteilung der Form

$$P^X(A) = \int P^X(A | \Theta = \vartheta) P^\Theta(d\vartheta)$$

für jede Borel-Menge A . Die Zufallsvariable Θ heißt *Strukturparameter* für das Kollektiv. Die Verteilung $P^X(\cdot | \Theta = \vartheta)$ entspricht dabei der Verteilung der individuellen Risiken in einem (homogenen) Teilkollektiv, das durch den Parameter ϑ charakterisiert ist.

2. Die (über die „Zeit“ oder im Teilkollektiv beobachteten) Risiken X_1, \dots, X_n sind *bedingt unabhängig* und *bedingt identisch verteilt* unter $\Theta = \vartheta$.

Beispiel 25. Wir gehen von einer vereinfachten Situation für ein Kfz-Haftpflichtversicherungsportfolio aus, indem wir nur zwischen risikoaversiven und risikobetonen Fahrern unterscheiden. Hier ist die Zufallsvariable Θ natürlicherweise zweiwertig; o.B.d.A. können wir also annehmen, dass $P^\Theta = B(1, p)$ gilt mit einem Parameter $p \in (0, 1)$. Für die bedingte Verteilung des Risikos X unter $\Theta = \vartheta$ wollen wir (fiktiv) annehmen, dass

$$P^X(\cdot | \Theta = \vartheta) = \mathcal{G}(\rho / (1 + \vartheta)), \quad \vartheta \in \{0, 1\}$$

gilt mit einem Parameter $\rho \in (0, 1)$. Die Risikoverteilung ist also für risikoaverse Fahrer gegeben durch eine $\mathcal{G}(\rho)$ -geometrische Verteilung, für risikobetonen Fahrer dagegen durch eine $\mathcal{G}(\rho/2)$ -geometrische Verteilung. Die ganzzahligen Werte von X können dabei wie im Kapitel über Schadenversicherung als Vielfache einer monetären Grundeinheit Δ angesehen werden (z.B. 1000 €). Wegen

$$E(X | \Theta = \vartheta) = \frac{1 + \vartheta - \rho}{\rho}, \quad \vartheta \in \{0, 1\}$$

produzieren demnach in diesem Modell risikobetonen Fahrer im Mittel höhere Schäden als risikoaverse Fahrer. Dementsprechend sollten risikoaverse Fahrer auch eine geringere Versicherungsprämie zahlen. Das Problem ist allerdings, dass nicht von vornherein klar ist, ob ein Fahrer zur ersten oder zur zweiten Gruppe gehört. Eine entsprechende Klassifizierung kann nur mit statistischen Mitteln durch langfristige Schadenbeobachtungen erfolgen – daher auch der Name *Erfahrungstari-fierung*. Nehmen wir an, dass ein Fahrer "Schäden" X_1, \dots, X_n verursacht hat (hier kann wegen der geometrischen Verteilungsannahme auch ein "Schaden" von Null auftreten). Man kann dann die

III. Credibility-Theorie

(bedingte) Wahrscheinlichkeit dafür, dass der Fahrer risikoaversiv ist oder nicht, auf der Basis dieser Schäden berechnen als

$$\begin{aligned}
 P(\Theta = \vartheta | X_1 = k_1, \dots, X_n = k_n) &= \frac{P(\Theta = \vartheta, X_1 = k_1, \dots, X_n = k_n)}{P(X_1 = k_1, \dots, X_n = k_n)} \\
 &= \frac{P(X_1 = k_1, \dots, X_n = k_n | \Theta = \vartheta) P(\Theta = \vartheta)}{\sum_{\eta \in \{0,1\}} P(X_1 = k_1, \dots, X_n = k_n | \Theta = \eta) P(\Theta = \eta)} \\
 &= \frac{P(\Theta = \vartheta) \prod_{i=1}^n P(X_i = k_i | \Theta = \vartheta)}{\sum_{\eta \in \{0,1\}} P(\Theta = \eta) \prod_{i=1}^n P(X_i = k_i | \Theta = \eta)}
 \end{aligned}$$

mit $\prod_{i=1}^n P(X_i = k_i | \Theta = \vartheta) = \left(\frac{\rho}{1+\vartheta}\right)^n \left(1 - \frac{\rho}{1+\vartheta}\right)^{\sum_{i=1}^n k_i}$. Es ergibt sich konkret:

$$\begin{aligned}
 P(\Theta = 0 | X_1 = k_1, \dots, X_n = k_n) &= \frac{(1-p)\rho^n(1-\rho)^{\sum_{i=1}^n k_i}}{(1-p)\rho^n(1-\rho)^{\sum_{i=1}^n k_i} + p\left(\frac{\rho}{2}\right)^n \left(1 - \frac{\rho}{2}\right)^{\sum_{i=1}^n k_i}} \\
 P(\Theta = 1 | X_1 = k_1, \dots, X_n = k_n) &= \frac{p\left(\frac{\rho}{2}\right)^n \left(1 - \frac{\rho}{2}\right)^{\sum_{i=1}^n k_i}}{(1-p)\rho^n(1-\rho)^{\sum_{i=1}^n k_i} + p\left(\frac{\rho}{2}\right)^n \left(1 - \frac{\rho}{2}\right)^{\sum_{i=1}^n k_i}}.
 \end{aligned}$$

Wie man sieht, können sich bereits in relativ einfachen Modellen recht komplizierte Formeln für die entsprechenden bedingten Wahrscheinlichkeiten ergeben.

Zur Veranschaulichung wollen wir einmal annehmen, dass $\rho = 0,8$ gilt und $p = 0,3$ – d.h. 70% der Fahrer verhalten sich risikoaversiv. Ein Fahrer möge nun in den letzten 10 Jahren folgende "Schäden" X_1, \dots, X_n verursacht haben:

Jahr i	1	2	3	4	5	6	7	8	9	10
Schaden k_i	0	0	0	0	0	2	1	0	0	1

Damit ergibt sich:

$$\begin{aligned}
 P(\Theta = 0 | X_1 = k_1, \dots, X_n = k_n) &= \frac{0,7 \cdot 0,8^{10} \cdot 0,2^4}{0,7 \cdot 0,8^{10} \cdot 0,2^4 + 0,3 \cdot 0,4^{10} \cdot 0,6^4} = 0,9672 \\
 P(\Theta = 1 | X_1 = k_1, \dots, X_n = k_n) &= \frac{0,3 \cdot 0,4^{10} \cdot 0,6^4}{0,7 \cdot 0,8^{10} \cdot 0,2^4 + 0,3 \cdot 0,4^{10} \cdot 0,6^4} = 0,0328,
 \end{aligned}$$

III. Credibility-Theorie

d.h. es ist sehr wahrscheinlich, dass der Fahrer – im Rahmen unseres Modells – risikoaversiv ist.

Würden für einen anderen Fahrer folgende Daten beobachtet:

Jahr i	1	2	3	4	5	6	7	8	9	10
Schaden k_i	0	3	2	1	0	0	0	2	0	2

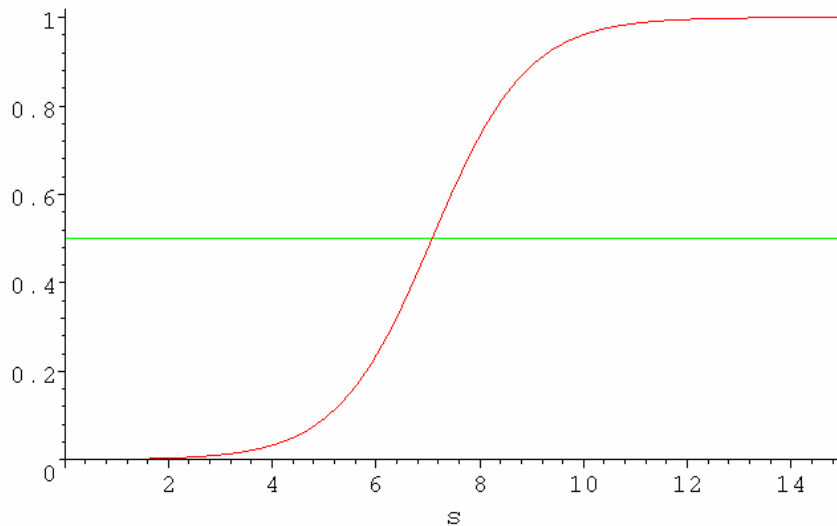
so wäre

$$P(\Theta = 0 | X_1 = k_1, \dots, X_n = k_n) = \frac{0,7 \cdot 0,8^{10} \cdot 0,2^{10}}{0,7 \cdot 0,8^{10} \cdot 0,2^{10} + 0,3 \cdot 0,4^{10} \cdot 0,6^{10}} = 0,0389$$

$$P(\Theta = 1 | X_1 = k_1, \dots, X_n = k_n) = \frac{0,3 \cdot 0,4^{10} \cdot 0,6^{10}}{0,7 \cdot 0,8^{10} \cdot 0,2^{10} + 0,3 \cdot 0,4^{10} \cdot 0,6^{10}} = 0,9611,$$

d.h. der Fahrer würde auch entsprechend als risikofreudig eingestuft.

Man kann die Wahrscheinlichkeit $P(\Theta = 1 | X_1 = k_1, \dots, X_n = k_n)$, einen Fahrer aufgrund seiner Schäden korrekt als risikofreudig zu klassifizieren, auch als Funktion von $s = \sum_{i=1}^n k_i$ graphisch darstellen, denn nur von diesem Wert, d.h. dem Gesamtschaden in n Jahren, hängt diese Wahrscheinlichkeit funktional ab. So erhält man etwa für das obige Beispiel:



wobei die "kritische" Grenze bei einem Gesamtschaden von 7 liegt.

Die zentrale Idee der Erfahrenstarifizierung und des daraus abgeleiteten Bonus-Malus-Systems besteht nun darin, dass zu Beginn der Versicherung, wo noch keine Schadenerfahrung vorliegt, der versicherte Fahrer eine Netto-Prämie zahlt, die dem Erwartungswert $E(X)$ des Gesamtkollektivs entspricht, also in unserem letzten Fall

III. Credibility-Theorie

$$\begin{aligned}
 E(X) &= \int E(X | \Theta = \vartheta) P^\Theta(d\vartheta) = \sum_{\vartheta \in \{0,1\}} E(X | \Theta = \vartheta) P(\Theta = \vartheta) \\
 &= \sum_{\vartheta \in \{0,1\}} \frac{1 + \vartheta - \rho}{\rho} P(\Theta = \vartheta) = \frac{1 - \rho}{\rho} (1 - p) + \frac{2 - \rho}{\rho} p = \frac{1 - \rho + p}{\rho} = 0,625.
 \end{aligned}$$

Mit fortschreitender Schadenerfahrung wird diese Prämie dann graduell angepasst, so dass sie sich langfristig entweder der Prämie

$$E(X | \Theta = 0) = \frac{1 - \rho}{\rho} = 0,25 \quad (\text{entsprechend } 250 \text{ €})$$

für die risikoaversive Fahrergruppe oder der Prämie

$$E(X | \Theta = 1) = \frac{2 - \rho}{\rho} = 1,50 \quad (\text{entsprechend } 1500 \text{ €})$$

für die risikofreudige Fahrergruppe annähert.

Der Strukturparameter ϑ charakterisiert also in solchen Modellen das Risikoverhalten eines einzelnen Versicherten; daher rührt auch die obige zweite Grundannahme der bedingten Unabhängigkeit, die sich in homogenen Portfolios mit unseren früheren Annahmen deckt.

Um die Darstellungen im folgenden etwas zu systematisieren, nehmen wir an, dass die Verteilung $P^X(\cdot | \Theta = \vartheta)$ eine (bedingte) Dichte $f_X(\cdot | \Theta = \vartheta)$ besitzt, und zwar entweder bezüglich des Lebesgue-Maßes (stetiger Fall) oder bezüglich des abzählenden Maßes (diskreter Fall). In beiden Fällen wollen wir das jeweilige Maß mit dem Symbol μ bezeichnen. Die Abbildung f_Θ bezeichne entsprechend die Dichte des Strukturparameters Θ . Ferner wollen wir bei Funktionen mehrerer Variablen x_1, \dots, x_n abkürzend den Vektor $\mathbf{x} = (x_1, \dots, x_n)$ bzw. für Zufallsvariablen X_1, \dots, X_n entsprechend $\mathbf{X} = (X_1, \dots, X_n)$ verwenden. Auf die jeweilige genaue Angabe der Definitionsbereiche wird zur Beibehaltung der Übersichtlichkeit verzichtet, diese sind aber in den konkreten Anwendungsfällen aus der Form der jeweiligen Verteilung ableitbar.

Lemma 24. Im Grundmodell der Credibility-Theorie gilt:

- a) Die gemeinsame Dichte g des Risikovektors \mathbf{X} und des Strukturparameters Θ ist gegeben durch

$$g(\mathbf{x}, \vartheta) = f_\Theta(\vartheta) \prod_{i=1}^n f_X(x_i | \Theta = \vartheta).$$

- b) Die bedingte Dichte von Θ unter den beobachteten Schäden $\mathbf{X} = \mathbf{x}$ ist gegeben durch

$$f_\Theta(\vartheta | \mathbf{X} = \mathbf{x}) = \frac{g(\mathbf{x}, \vartheta)}{\int g(\mathbf{x}, \eta) \mu(d\eta)}.$$

III. Credibility-Theorie

Beweis:

a) folgt aus der Beziehung $\frac{g(\mathbf{x}, \vartheta)}{f_{\Theta}(\vartheta)} = f_{\mathbf{X}}(\mathbf{x} | \Theta = \vartheta) = \prod_{i=1}^n f_X(x_i | \Theta = \vartheta)$ (nach Grundannahme 2).

b) folgt aus der Beziehung $f_{\Theta}(\vartheta | \mathbf{X} = \mathbf{x}) = \frac{g(\mathbf{x}, \vartheta)}{f_{\mathbf{X}}(\mathbf{x})}$ und $f_{\mathbf{X}}(\mathbf{x}) = \int g(\mathbf{x}, \eta) \mu(d\eta)$.

In den nachfolgenden Tabellen sind noch einmal wichtige Verteilungen, die zum Teil auch schon vorher betrachtet wurden, zusammengefasst. Der Strukturparameter ist hier jedoch einheitlich mit dem Symbol ϑ gekennzeichnet.

Diskrete Verteilungen:

P^X	Name	Zähldichte $f(k) = P(X = k)$	Definitionsbereich	Parameter
$B(m, \vartheta)$	Binomialverteilung	$\binom{m}{k} \vartheta^k (1 - \vartheta)^{m-k}$	$\{0, 1, \dots, m\}$	$\vartheta \in (0, 1), m \in \mathbb{N}$
$NB(\beta, \vartheta)$	negative Binomialverteilung	$\binom{\beta + k - 1}{k} \vartheta^k (1 - \vartheta)^{\beta}$	\mathbb{Z}^+	$\vartheta \in (0, 1), \beta > 0$
$\mathcal{P}(\vartheta)$	Poisson-Verteilung	$e^{-\vartheta} \frac{\vartheta^k}{k!}$	\mathbb{Z}^+	$\vartheta > 0$

Stetige Verteilungen:

P^X	Name	Dichte $f(x)$	Definitionsbereich	Parameter
$\mathcal{B}(a, b)$	Beta-Verteilung	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$	$[0, 1]$	$a, b > 0$
$\mathcal{E}(\lambda)$	Exponentialverteilung	$\lambda e^{-\lambda x}$	\mathbb{R}^+	$\lambda > 0$
$\Gamma(\alpha, \lambda)$	Gamma-Verteilung	$\lambda^{\alpha} \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}$	\mathbb{R}^+	$\alpha, \lambda > 0$

Die schon früher betrachtete stetige Gleichverteilung über dem Intervall $[0, 1]$ ergibt sich als Spezialfall der Beta-Verteilung für die Parameterwahl $a = b = 1$.

Das folgende Resultat zeigt die wichtigsten Beziehungen zwischen diesen Verteilungen in Bezug auf Lemma 24.

Lemma 25. Im Grundmodell der Credibility-Theorie gilt, mit der Abkürzung $\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n x_i$, in tabellarischer Zusammenfassung:

III. Credibility-Theorie

$P^X(\cdot \Theta = \vartheta)$	P^Θ	$P^\Theta(\cdot \mathbf{X} = \mathbf{x})$
$B(m, \vartheta)$	$\mathcal{B}(a, b)$	$\mathcal{B}(a + n\bar{x}, b + mn - n\bar{x})$
$NB(\beta, \vartheta)$	$\mathcal{B}(a, b)$	$\mathcal{B}(a + n\beta, b + n\bar{x})$
$\mathcal{P}(\vartheta)$	$\Gamma(\alpha, \lambda)$	$\Gamma(\alpha + n\bar{x}, \lambda + n)$
$\mathcal{E}(\vartheta)$	$\Gamma(\alpha, \lambda)$	$\Gamma(\alpha + n, \lambda + n\bar{x})$

Beweis: durch Nachrechnen mit Lemma 24:

1. Fall: Es ist

$$g(\mathbf{x}, \vartheta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \vartheta^{a-1} (1-\vartheta)^{b-1} \left\{ \prod_{i=1}^n \binom{m}{x_i} \right\} \vartheta^{n\bar{x}} (1-\vartheta)^{mn-n\bar{x}} = \left\{ \prod_{i=1}^n \binom{m}{x_i} \right\} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \vartheta^{n\bar{x}+a-1} (1-\vartheta)^{mn-n\bar{x}+b-1}$$

mit

$$\begin{aligned} \int g(\mathbf{x}, \eta) d\eta &= \left\{ \prod_{i=1}^n \binom{m}{x_i} \right\} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+n\bar{x})\Gamma(b+mn-n\bar{x})}{\Gamma(mn+a+b)} \times \dots \\ &\dots \times \int_0^1 \underbrace{\frac{\Gamma(mn+a+b)}{\Gamma(a+n\bar{x})\Gamma(b+mn-n\bar{x})} \eta^{n\bar{x}+a-1} (1-\eta)^{mn-n\bar{x}+b-1} d\eta}_{\text{Dichte der } \mathcal{B}(a+n\bar{x}, b+mn-n\bar{x})\text{-Verteilung}}, \end{aligned}$$

also

$$f_\Theta(\vartheta | \mathbf{X} = \mathbf{x}) = \frac{g(\mathbf{x}, \vartheta)}{\int g(\mathbf{x}, \eta) d\eta} = \frac{\Gamma(mn+a+b)}{\Gamma(a+n\bar{x})\Gamma(b+mn-n\bar{x})} \vartheta^{a+n\bar{x}-1} (1-\vartheta)^{b+mn-n\bar{x}-1},$$

woraus die erste Behauptung folgt.

2. Fall: Es ist

$$\begin{aligned} g(\mathbf{x}, \vartheta) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \vartheta^{a-1} (1-\vartheta)^{b-1} \left\{ \prod_{i=1}^n \binom{\beta+x_i-1}{x_i} \right\} \vartheta^{n\beta} (1-\vartheta)^{n\bar{x}} \\ &= \left\{ \prod_{i=1}^n \binom{\beta+x_i-1}{x_i} \right\} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \vartheta^{n\beta+a-1} (1-\vartheta)^{n\bar{x}+b-1} \end{aligned}$$

mit

III. Credibility-Theorie

$$\int g(\mathbf{x}, \eta) d\eta = \left\{ \prod_{i=1}^n \binom{\beta + x_i - 1}{x_i} \right\} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+n\beta)\Gamma(b+n\bar{\mathbf{x}})}{\Gamma(a+b+n\beta+n\bar{\mathbf{x}})} \times \dots$$

$$\dots \times \underbrace{\int_0^1 \frac{\Gamma(a+b+n\beta+n\bar{\mathbf{x}})}{\Gamma(a+n\beta)\Gamma(b+n\bar{\mathbf{x}})} \eta^{n\beta+a-1} (1-\eta)^{n\bar{\mathbf{x}}+b-1} d\eta}_{\text{Dichte der } \mathcal{B}(a+n\beta, b+n\bar{\mathbf{x}})\text{-Verteilung}}$$

also

$$f_{\Theta}(\vartheta | \mathbf{X} = \mathbf{x}) = \frac{g(\mathbf{x}, \vartheta)}{\int g(\mathbf{x}, \eta) d\eta} = \frac{\Gamma(a+b+n\beta+n\bar{\mathbf{x}})}{\Gamma(a+n\beta)\Gamma(b+n\bar{\mathbf{x}})} \vartheta^{a+n\beta-1} (1-\vartheta)^{b+n\bar{\mathbf{x}}-1},$$

woraus die zweite Behauptung folgt.

3. Fall: Es ist

$$g(\mathbf{x}, \vartheta) = e^{-n\vartheta} \left\{ \prod_{i=1}^n \frac{\vartheta^{x_i}}{x_i!} \right\} \frac{\lambda^\alpha}{\Gamma(\alpha)} \vartheta^{\alpha-1} e^{-\lambda\vartheta} = \frac{1}{\prod_{i=1}^n x_i!} \frac{\lambda^\alpha}{\Gamma(\alpha)} \vartheta^{n\bar{\mathbf{x}}+\alpha-1} e^{-(\lambda+n)\vartheta}$$

mit

$$\int g(\mathbf{x}, \eta) d\eta = \frac{1}{\prod_{i=1}^n x_i!} \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(n\bar{\mathbf{x}} + \alpha)}{(\lambda + n)^{n\bar{\mathbf{x}} + \alpha}} \int_0^\infty \underbrace{\frac{(\lambda + n)^{n\bar{\mathbf{x}} + \alpha}}{\Gamma(n\bar{\mathbf{x}} + \alpha)} \eta^{n\bar{\mathbf{x}} + \alpha - 1} e^{-(\lambda + n)\eta}}_{\text{Dichte der } \Gamma(\alpha + n\bar{\mathbf{x}}, \lambda + n)\text{-Verteilung}} d\eta,$$

also

$$f_{\Theta}(\vartheta | \mathbf{X} = \mathbf{x}) = \frac{g(\mathbf{x}, \vartheta)}{\int g(\mathbf{x}, \eta) d\eta} = \frac{(\lambda + n)^{n\bar{\mathbf{x}} + \alpha}}{\Gamma(n\bar{\mathbf{x}} + \alpha)} \vartheta^{n\bar{\mathbf{x}} + \alpha - 1} e^{-(\lambda + n)\vartheta},$$

woraus die dritte Behauptung folgt.

4. Fall: Es ist

$$g(\mathbf{x}, \vartheta) = \vartheta^n e^{-n\bar{\mathbf{x}}\vartheta} \frac{\lambda^\alpha}{\Gamma(\alpha)} \vartheta^{\alpha-1} e^{-\lambda\vartheta} = \frac{\lambda^\alpha}{\Gamma(\alpha)} \vartheta^{n+\alpha-1} e^{-(\lambda+n\bar{\mathbf{x}})\vartheta}$$

mit

$$\int g(\mathbf{x}, \eta) d\eta = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(n + \alpha)}{(\lambda + n\bar{\mathbf{x}})^{n + \alpha}} \int_0^\infty \underbrace{\frac{(\lambda + n\bar{\mathbf{x}})^{n + \alpha}}{\Gamma(n + \alpha)} \eta^{n + \alpha - 1} e^{-(\lambda + n\bar{\mathbf{x}})\eta}}_{\text{Dichte der } \Gamma(\alpha + n, \lambda + n\bar{\mathbf{x}})\text{-Verteilung}} d\eta,$$

also

III. Credibility-Theorie

$$f_{\Theta}(\vartheta | \mathbf{X} = \mathbf{x}) = \frac{g(\mathbf{x}, \vartheta)}{\int g(\mathbf{x}, \eta) d\eta} = \frac{(\lambda + n\bar{\mathbf{x}})^{n+\alpha}}{\Gamma(n+\alpha)} \vartheta^{n+\alpha-1} e^{-(\lambda+n\bar{\mathbf{x}})\vartheta},$$

woraus auch die letzte Behauptung folgt.

III. 1. Die Credibility-Prämie

Wenn alle Verteilungen im Grundmodell der Credibility-Theorie bekannt sind, kann die Netto-Prämie für das Risiko X nach den verschiedenen, durch den Strukturparameter ϑ charakterisierten Risiko-Klassen differenziert werden vermöge der Beziehung

$$H(\vartheta) = E(X | \Theta = \vartheta).$$

Die übliche Netto-Prämie ergibt sich hieraus durch Integration zu

$$H = E(X) = E[H(\Theta)] = \int E(X | \Theta = \vartheta) P^{\Theta}(d\vartheta).$$

Dieses Vorgehen entspricht genau demjenigen in Beispiel 25 oben. Wenn die Verteilungen allerdings nicht genauer bekannt sind – was der Situation in der Praxis eher entspricht – dann liegt es nahe, ersatzweise eine möglichst gute Approximation an die differenzierte Prämie unter Verwendung der Schadendaten \mathbf{X} zu verwenden, d.h. die Lösung des Optimierungsproblems

$$\min_{H^*} E\left[\left(H^* - H(\Theta)\right)^2 | \mathbf{X}\right]$$

zu betrachten, wobei nur Lösungen H^* zugelassen sind, die messbar sind bezgl. der von \mathbf{X} erzeugten σ -Algebra $\mathcal{F} = \sigma(\mathbf{X})$. Nach dem Faktorisierungssatz⁷ sind dies also geeignete messbare Funktionen von \mathbf{X} .

Satz 7. Die Lösung des Optimierungsproblems $\min_{H^*} E\left[\left(H^* - H(\Theta)\right)^2 | \mathbf{X}\right]$ ist gegeben durch

$$H^* = E[H(\Theta) | \mathbf{X}], \quad \text{mit } E(H^*) = E(H(\Theta)) = H = E(X).$$

Diese Lösung heißt auch *Credibility-Prämie* für das Netto-Risiko-Prinzip. Der minimale Wert des Optimierungsproblems ist ferner gegeben durch

$$\min_{H^*} E\left[\left(H^* - H(\Theta)\right)^2 | \mathbf{X}\right] = \text{Var}(H(\Theta) | \mathbf{X}).$$

Beweis: Durch Umformen und Entklammern erhält man mit der Faktorregel für bedingte Erwartungen⁸

⁷ Ist $\mathcal{F} = \sigma(\mathbf{X})$ eine von einem Zufallsvektor \mathbf{X} erzeugte σ -Algebra und Z eine \mathcal{F} -messbare Zufallsvariable, so existiert eine messbare Abbildung h mit der Eigenschaft $Z = h(\mathbf{X})$.

III. Credibility-Theorie

$$\begin{aligned}
 E\left[\left(H^* - H(\Theta)\right)^2 \mid \mathbf{X}\right] &= E\left[\left(H^* - E[H(\Theta) \mid \mathbf{X}] + E[H(\Theta) \mid \mathbf{X}] - H(\Theta)\right)^2 \mid \mathbf{X}\right] \\
 &= \text{Var}(H(\Theta) \mid \mathbf{X}) + E\left[\left(H^* - E[H(\Theta) \mid \mathbf{X}]\right)^2 \mid \mathbf{X}\right] \\
 &= \text{Var}(H(\Theta) \mid \mathbf{X}) + \left(H^* - E[H(\Theta) \mid \mathbf{X}]\right)^2,
 \end{aligned}$$

weil für den "gemischten" Term gilt:

$$\begin{aligned}
 E\left[\underbrace{\left(H^* - E[H(\Theta) \mid \mathbf{X}]\right)}_{\sigma(\mathbf{X})\text{-messbar}} \cdot \left(E[H(\Theta) \mid \mathbf{X}] - H(\Theta)\right) \mid \mathbf{X}\right] &= \left(H^* - E[H(\Theta) \mid \mathbf{X}]\right) \cdot E\left[\left(E[H(\Theta) \mid \mathbf{X}] - H(\Theta)\right) \mid \mathbf{X}\right] \\
 &= \left(H^* - E[H(\Theta) \mid \mathbf{X}]\right) \cdot \left(E[H(\Theta) \mid \mathbf{X}] - E[H(\Theta) \mid \mathbf{X}]\right) = 0.
 \end{aligned}$$

Die Lösung besteht also in der Minimierung des Terms $\left(H^* - E[H(\Theta) \mid \mathbf{X}]\right)^2$, die gerade durch $H^* = E[H(\Theta) \mid \mathbf{X}]$ gegeben ist, mit $E\left[\left(H^* - H(\Theta)\right)^2 \mid \mathbf{X}\right] = \text{Var}(H(\Theta) \mid \mathbf{X})$. Ferner ist

$$E(H^*) = E(E[H(\Theta) \mid \mathbf{X}]) = E(H(\Theta)) = H = E(X),$$

wie behauptet.

Beispiel 26 (Fortsetzung von Beispiel 25). Nach Modellansatz ergibt sich

$$H(\Theta) = \frac{1 + \Theta - \rho}{\rho}$$

und somit für die Credibility-Prämie der Wert

$$\begin{aligned}
 H^* = E[H(\Theta) \mid \mathbf{X}] &= \frac{1 + E[\Theta \mid \mathbf{X}] - \rho}{\rho} = \frac{1 + P[\Theta = 1 \mid \mathbf{X}] - \rho}{\rho} \\
 &= \frac{1 - \rho}{\rho} + \frac{1}{\rho} \cdot \frac{p \left(\frac{\rho}{2}\right)^n \left(1 - \frac{\rho}{2}\right)^{\sum_{i=1}^n X_i}}{(1 - p)\rho^n (1 - \rho)^{\sum_{i=1}^n X_i} + p \left(\frac{\rho}{2}\right)^n \left(1 - \frac{\rho}{2}\right)^{\sum_{i=1}^n X_i}}.
 \end{aligned}$$

Für den ersten Fahrer ergibt sich damit eine jährliche Credibility-Prämie in Höhe von $H^* = 0,291$ (entsprechend 291 €), für den zweiten eine in Höhe von $H^* = 1,451$ (entsprechend 1451 €).

⁸ Ist $\mathcal{F} = \sigma(\mathbf{X})$ eine von einem Zufallsvektor \mathbf{X} erzeugte σ -Algebra und ist Y eine \mathcal{F} -messbare und Z eine beliebige Zufallsvariable, so gilt $E(Y \cdot Z \mid \mathbf{X}) = Y \cdot E(Z \mid \mathbf{X})$ fast sicher.

III. Credibility-Theorie

Lemma 26 (Fortsetzung von Lemma 25). Im Grundmodell der Credibility-Theorie gilt, mit der Abkürzung $\bar{\mathbf{X}} = \frac{1}{n} \sum_{i=1}^n X_i$ und $H = E(X)$, in tabellarischer Zusammenfassung:

$P^{\mathbf{X}}(\cdot \Theta = \vartheta)$	P^{Θ}	$H(\Theta)$	H	$H^* = E[H(\Theta) \mathbf{X}]$
$B(m, \vartheta)$	$\mathcal{B}(a, b)$	$m\Theta$	$\frac{am}{a+b}$	$\frac{mn}{a+b+mn} \bar{\mathbf{X}} + \frac{a+b}{a+b+mn} H$
$NB(\beta, \vartheta)$	$\mathcal{B}(a, b)$	$\beta \frac{1-\Theta}{\Theta}$	$\frac{\beta b}{a-1}$	$\frac{n\beta}{a+n\beta-1} \bar{\mathbf{X}} + \frac{a-1}{a+n\beta-1} H$
$\mathcal{P}(\vartheta)$	$\Gamma(\alpha, \lambda)$	Θ	$\frac{\alpha}{\lambda}$	$\frac{n}{\lambda+n} \bar{\mathbf{X}} + \frac{\lambda}{\lambda+n} H$
$\mathcal{E}(\vartheta)$	$\Gamma(\alpha, \lambda)$	$\frac{1}{\Theta}$	$\frac{\lambda}{\alpha-1}$	$\frac{n}{\alpha+n-1} \bar{\mathbf{X}} + \frac{\alpha-1}{\alpha+n-1} H$

Beweis: durch Nachrechnen mit Lemma 25:

1. Fall: Es ist

$$H(\vartheta) = E(X | \Theta = \vartheta) = m\vartheta, \text{ also } H(\Theta) = m\Theta, \text{ also } H = mE(\Theta) = \frac{am}{a+b}$$

mit

$$\begin{aligned} H^* &= E[H(\Theta) | \mathbf{X}] = mE(\Theta | \mathbf{X}) = m \frac{a+n\bar{\mathbf{X}}}{a+b+mn} = \frac{mn}{a+b+mn} \bar{\mathbf{X}} + \frac{am}{a+b+mn} \\ &= \frac{mn}{a+b+mn} \bar{\mathbf{X}} + \frac{a+b}{a+b+mn} H, \end{aligned}$$

woraus die erste Behauptung folgt.

2. Fall: Es ist

$$H(\vartheta) = E(X | \Theta = \vartheta) = \beta \frac{1-\vartheta}{\vartheta}, \text{ also } H(\Theta) = \beta \frac{1-\Theta}{\Theta}, \text{ also } H = \beta E\left(\frac{1}{\Theta}\right) - \beta = \beta \frac{b}{a-1} \text{ wegen}$$

$$E\left(\frac{1}{\Theta}\right) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \vartheta^{a-2} (1-\vartheta)^{b-1} d\vartheta = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a-1)\Gamma(b)}{\Gamma(a+b-1)} = \frac{a+b-1}{a-1},$$

mit

III. Credibility-Theorie

$$\begin{aligned}
 H^* &= E[H(\Theta) | \mathbf{X}] = \beta \left\{ E\left(\frac{1}{\Theta} \mid \mathbf{X}\right) - 1 \right\} = \beta \left\{ \frac{a + n\beta + b + n\bar{\mathbf{X}} - 1}{a + n\beta - 1} - 1 \right\} = \beta \frac{b + n\bar{\mathbf{X}}}{a + n\beta - 1} \\
 &= \frac{n\beta}{a + n\beta - 1} \bar{\mathbf{X}} + \frac{b\beta}{a + n\beta - 1} = \frac{n\beta}{a + n\beta - 1} \bar{\mathbf{X}} + \frac{a-1}{a + n\beta - 1} H,
 \end{aligned}$$

woraus die zweite Behauptung folgt.

3. Fall: Es ist

$$H(\vartheta) = E(X | \Theta = \vartheta) = \vartheta, \text{ also } H(\Theta) = \Theta, \text{ also } H = E(\Theta) = \frac{\alpha}{\lambda}$$

mit

$$H^* = E[H(\Theta) | \mathbf{X}] = mE(\Theta | \mathbf{X}) = \frac{\alpha + n\bar{\mathbf{X}}}{\lambda + n} = \frac{n}{\lambda + n} \bar{\mathbf{X}} + \frac{\lambda}{\lambda + n} H,$$

woraus die dritte Behauptung folgt.

4. Fall: Es ist

$$H(\vartheta) = E(X | \Theta = \vartheta) = \frac{1}{\vartheta}, \text{ also } H(\Theta) = \frac{1}{\Theta}, \text{ also } H = E(\Theta) = \frac{\alpha}{\lambda} \text{ wegen}$$

$$E\left(\frac{1}{\Theta}\right) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \vartheta^{\alpha-2} e^{-\lambda\vartheta} d\vartheta = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha-1)}{\lambda^{\alpha-1}} = \frac{\lambda}{\alpha-1},$$

mit

$$H^* = E[H(\Theta) | \mathbf{X}] = E\left(\frac{1}{\Theta} \mid \mathbf{X}\right) = \frac{\lambda + n\bar{\mathbf{X}}}{\alpha + n - 1} = \frac{n}{\alpha + n - 1} \bar{\mathbf{X}} + \frac{\lambda}{\alpha + n - 1} = \frac{n}{\alpha + n - 1} \bar{\mathbf{X}} + \frac{\alpha - 1}{\alpha + n - 1} H,$$

woraus auch die letzte Behauptung folgt.

III. 2. Die linearisierte Credibility-Prämie

Es fällt auf, dass die Ergebnisse von Lemma 26 sämtlich von der Form

$$H^* = E[H(\Theta) | \mathbf{X}] = z_n \bar{\mathbf{X}} + (1 - z_n) H$$

sind, wobei der positive Faktor z_n monoton wachsend ist mit Limes 1. Die Credibility-Prämie ist also hier ein gewichtetes Mittel aus dem beobachteten Schadendurchschnitt und der theoretischen Nettoprämie für das Gesamtkollektiv, mit

$$H^* = E[H(\Theta) | \mathbf{X}] \approx \bar{\mathbf{X}}$$

für wachsenden Beobachtungsumfang n . Dass dies aber nicht immer so ist, zeigt Beispiel 26. Wir können jedoch einen solchen funktionalen Ansatz hernehmen und als weiteres Optimierungsproblem

III. Credibility-Theorie

$$\min_{z_n}! E \left[\left(z_n \bar{\mathbf{X}} + (1 - z_n) H - H^* \right)^2 \right] \quad \text{bzw.} \quad \min_{H^{**}}! E \left[\left(H^{**} - H^* \right)^2 \right]$$

mit

$$H^{**} = z_n \bar{\mathbf{X}} + (1 - z_n) H$$

lösen. Das sich hieraus ergebende Resultat heißt entsprechend *linearisierte Credibility-Prämie*. Durch einfache Umformung ergibt sich

$$\begin{aligned} E \left[\left(z_n \bar{\mathbf{X}} + (1 - z_n) H - H^* \right)^2 \right] &= E \left[\left(z_n (\bar{\mathbf{X}} - H) + (H - H^*) \right)^2 \right] \\ &= z_n^2 \text{Var}(\bar{\mathbf{X}} - H) + 2 z_n E \left[(\bar{\mathbf{X}} - H)(H - H^*) \right] + E \left[(H - H^*)^2 \right] \end{aligned}$$

mit Minimalstelle

$$z_n^* = \frac{E \left[(\bar{\mathbf{X}} - H)(H^* - H) \right]}{\text{Var}(\bar{\mathbf{X}} - H)}.$$

Zur Vorbereitung des nächsten Satzes benötigen wir noch folgendes

Lemma 27 (bedingte Varianz-Formel). Es seien Z und Θ beliebige Zufallsvariablen derart, dass $\text{Var}(Z) < \infty$. Dann gilt:

$$\text{Var}(Z) = E \left[\text{Var}(Z | \Theta) \right] + \text{Var} \left[E(Z | \Theta) \right].$$

Beweis: Es ist

$$\begin{aligned} \text{Var}(Z) &= E(Z^2) - [E(Z)]^2 = E \left[E(Z^2 | \Theta) \right] - E \left[\{E(Z | \Theta)\}^2 \right] + E \left[\{E(Z | \Theta)\}^2 \right] - [E(Z)]^2 \\ &= E \left[E(Z^2 | \Theta) - \{E(Z | \Theta)\}^2 \right] + E \left[\{E(Z | \Theta)\}^2 \right] - \{E[E(Z | \Theta)]\}^2 \\ &= E \left[\text{Var}(Z | \Theta) \right] + \text{Var} \left[E(Z | \Theta) \right]. \quad \blacksquare \end{aligned}$$

Satz 8. Für die linearisierte Credibility-Prämie $H^{**} = z_n \bar{\mathbf{X}} + (1 - z_n) H$ gilt:

$$z_n = \frac{\text{Var}(H(\Theta))}{\text{Var}(\bar{\mathbf{X}})} = \frac{\text{Var}(H(\Theta))}{\frac{1}{n} E[\text{Var}(X | \Theta)] + \text{Var}(H(\Theta))}, \quad n \in \mathbb{N} \text{ mit } z_n \uparrow 1 \text{ für } n \rightarrow \infty.$$

Beweis: Es ist nach Lemma 27

$$\begin{aligned} \text{Var}(\bar{\mathbf{X}} - H) &= \text{Var}(\bar{\mathbf{X}}) = E \left[\text{Var}(\bar{\mathbf{X}} | \Theta) \right] + \text{Var} \left[E(\bar{\mathbf{X}} | \Theta) \right] \\ &= E \left[\frac{1}{n} \text{Var}(X | \Theta) \right] + \text{Var} \left[E(X | \Theta) \right] = \frac{1}{n} E \left[\text{Var}(X | \Theta) \right] + \text{Var}(H(\Theta)) \end{aligned}$$

und

III. Credibility-Theorie

$$\begin{aligned}
 E[(\bar{\mathbf{X}} - H)(H^* - H)] &= E[E\{(\bar{\mathbf{X}} - H)(H(\Theta) - H) | \mathbf{X}\}] = E[(\bar{\mathbf{X}} - H)(H(\Theta) - H)] \\
 &= E[E\{(\bar{\mathbf{X}} - H)(H(\Theta) - H) | \Theta\}] = E[(H(\Theta) - H)E\{(\bar{\mathbf{X}} - H) | \Theta\}] \\
 &= E[(H(\Theta) - H)E\{(X - H) | \Theta\}] = E[(H(\Theta) - H)^2] = \text{Var}(H(\Theta))
 \end{aligned}$$

auf Grund der $\sigma(\mathbf{X})$ -Messbarkeit von $\bar{\mathbf{X}}$ und $H^* = E[H(\Theta) | \mathbf{X}]$. Hieraus folgt die Behauptung. ■

Beispiel 27 (Fortsetzung von Beispiel 26). Nach Voraussetzung gilt:

$$\begin{aligned}
 \text{Var}(H(\Theta)) &= \frac{1}{\rho^2} \text{Var}(\Theta) = \frac{p(1-p)}{\rho^2}, \quad \text{Var}(X | \Theta) = \frac{1+\Theta}{\rho^2}(1+\Theta-\rho), \\
 H &= \frac{1+p-\rho}{\rho}, \quad E[\text{Var}(X | \Theta)] = \frac{(1-\rho)(1+p)+2p}{\rho^2},
 \end{aligned}$$

also

$$\begin{aligned}
 z_n &= \frac{\text{Var}(H(\Theta))}{\frac{1}{n} E[\text{Var}(X | \Theta)] + \text{Var}(H(\Theta))} = \frac{\frac{p(1-p)}{\rho^2}}{\frac{1}{n} \frac{(1-\rho)(1+p)+2p}{\rho^2} + \frac{p(1-p)}{\rho^2}} \\
 &= \frac{np(1-p)}{(1-\rho)(1+p)+2p+np(1-p)}, \quad n \in \mathbb{N}.
 \end{aligned}$$

Für die konkrete Situation ergibt sich damit

$$H^{**} = z_n \bar{\mathbf{X}} + (1 - z_n) H = 0,70945 \cdot \bar{\mathbf{X}} + 0,29054 \cdot 0,625 = 0,70945 \cdot \bar{\mathbf{X}} + 0,1815878378,$$

entsprechend einer Prämie von € 465,37 für den ersten und € 891,04 für den zweiten Fahrer.

Man kann zeigen, dass mit wachsendem Beobachtungsumfang n mit Wahrscheinlichkeit 1 sowohl die gewöhnliche als auch die linearisierte Credibility-Prämie gegen $H(\Theta)$, also die dem Individuum entsprechende "richtige" Prämie konvergiert. Insbesondere gilt hier

$$\lim_{n \rightarrow \infty} \bar{\mathbf{X}} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = E(X | \Theta) = H(\Theta) \text{ mit Wahrscheinlichkeit 1,}$$

d.h. im Gegensatz zum "gewöhnlichen" Gesetz der großen Zahlen, wo unter vollständiger Unabhängigkeit der X_i das arithmetische Mittel der Beobachtungen fast sicher gegen eine Konstante (nämlich den zugehörigen Erwartungswert) konvergiert, strebt hier das arithmetische Mittel fast sicher gegen die Zufallsvariable $H(\Theta) = E[X | \Theta]$. Dies kann man heuristisch so erklären, dass unter Θ die Folge $\{X_n\}_{n \in \mathbb{N}}$ bedingt stochastisch unabhängig ist, also auch nur

$$P\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = E[X | \Theta] \mid \Theta\right) = 1$$

III. Credibility-Theorie

gilt. Setzen wir $A := \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = E[X | \Theta] \right\} = \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = H(\Theta) \right\}$, so gilt also $P(A | \Theta) = 1$ und somit auch, in unbedingter Form,

$$P(A) = E[P(A | \Theta)] = 1,$$

was der obigen Grenzwertaussage für das arithmetische Mittel entspricht. Aufgrund der besonderen Form der linearisierten Credibility-Prämie ergibt sich hieraus auch unmittelbar die fast sichere Konvergenz der linearisierten Credibility-Prämie gegen $H(\Theta)$. Die andere Konvergenzaussage lässt sich über die Forderung

$$\min_{z_n}! E \left[\left(z_n \bar{\mathbf{X}} + (1 - z_n) H - H^* \right)^2 \right] \quad \text{bzw.} \quad \min_{H^{**}}! E \left[\left(H^{**} - H^* \right)^2 \right]$$

heuristisch begründen, d.h. mit wachsendem Beobachtungsumfang wird $E \left[\left(H^{**} - H^* \right)^2 \right]$ immer kleiner, so dass H^* und H^{**} dasselbe asymptotische Verhalten zeigen.

Bemerkung: Die linearisierte Credibility-Prämie hängt über den Faktor z_n bzw. den Index n nur von der bedingten Varianz $\text{Var}(X | \Theta)$ ab. Man kann daher zur Bestimmung der linearisierten Credibility-Prämie die Modellannahmen zu Beginn dieses Kapitels folgendermaßen abschwächen:

Die individuellen Risiken X_1, \dots, X_n sind *bedingt unabhängig* unter $\Theta = \vartheta$ mit

- $H(\vartheta) = E(X_i | \Theta = \vartheta)$
- $\sigma^2(\vartheta) = \text{Var}(X_i | \Theta = \vartheta)$

für alle $i \in \{1, \dots, n\}$.

Diese Variante der Credibility-Theorie wird auch das *einfache Credibility-Modell* genannt (vgl. etwa BÜHLMANN UND GISLER (2005), Abschnitt 3.1).

III. 3. Die empirische Credibility-Prämie

Eine weitere Möglichkeit, die Schaden-Vorerfahrung angemessen für eine Credibility-Prämie zu berücksichtigen, besteht in der Betrachtung des Ausdrucks

$$H^{***} := E(X_{n+1} | \mathbf{X}),$$

d.h. des bedingten Erwartungswerts des den ersten n Beobachtungen unmittelbar folgenden Schadens. Dieser Ausdruck heißt auch *empirische Credibility-Prämie* und kann z.B. formelmäßig berechnet werden unter Verwendung der Beziehung

$$E(X_{n+1} | \mathbf{X} = \mathbf{x}) = \int_0^{\infty} (1 - F_{X_{n+1}}(z | \mathbf{X} = \mathbf{x})) dz,$$

III. Credibility-Theorie

wobei $F_{X_{n+1}}(z | \mathbf{X} = \mathbf{x})$ die bedingte Verteilungsfunktion von X_{n+1} unter $\mathbf{X} = \mathbf{x}$ bezeichnet, die im Fall der Existenz von Lebesgue-Dichten dargestellt werden kann über

$$F_{X_{n+1}}(z | \mathbf{X} = \mathbf{x}) = \int_0^z f_{X_{n+1}}(u | \mathbf{X} = \mathbf{x}) du, \quad z \geq 0$$

mit

$$f_{X_{n+1}}(z | \mathbf{X} = \mathbf{x}) = \frac{f_{(X_{n+1}, \mathbf{X})}(z, \mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})} = \frac{\int f_{(X_{n+1}, \mathbf{X}, \Theta)}(z, \mathbf{x}, \vartheta) \mu_{\Theta}(d\vartheta)}{\int f_{(\mathbf{X}, \Theta)}(\mathbf{x}, \vartheta) \mu_{\Theta}(d\vartheta)} \quad \text{für } f_{\mathbf{X}}(\mathbf{x}) > 0$$

(bedingte Dichte von X_{n+1} unter $\mathbf{X} = \mathbf{x}$).

Für einen Vergleich der bisher betrachteten Credibility-Prämien betrachten wir zunächst

Beispiel 28 (Ergänzung zu Lemma 26, Teil 4). Hier gilt (vgl. den Beweis zu Lemma 25, Teil 4):

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \int f_{(\mathbf{X}, \Theta)}(\mathbf{x}, \vartheta) \mu_{\Theta}(d\vartheta) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(n+\alpha)}{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}} \int_0^{\infty} \frac{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}}{\Gamma(n+\alpha)} \vartheta^{n+\alpha-1} e^{-(\lambda+n\bar{\mathbf{x}})\vartheta} d\vartheta \\ &= \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(n+\alpha)}{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}}, \end{aligned}$$

also

$$f_{X_{n+1}}(z | \mathbf{X} = \mathbf{x}) = \frac{f_{(X_{n+1}, \mathbf{X})}(z, \mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})} = \frac{\frac{\lambda^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(n+1+\alpha)}{(\lambda+z+n\bar{\mathbf{x}})^{n+1+\alpha}}}{\frac{\lambda^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(n+\alpha)}{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}}} = (n+\alpha) \frac{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}}{(\lambda+z+n\bar{\mathbf{x}})^{n+1+\alpha}}$$

mit

$$\begin{aligned} F_{X_{n+1}}(z | \mathbf{X} = \mathbf{x}) &= \int_0^z f_{X_{n+1}}(u | \mathbf{X} = \mathbf{x}) du = (n+\alpha) \int_0^z \frac{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}}{(\lambda+u+n\bar{\mathbf{x}})^{n+1+\alpha}} du \\ &= 1 - \frac{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}}{(\lambda+z+n\bar{\mathbf{x}})^{n+\alpha}}, \quad z \geq 0. \end{aligned}$$

Es folgt

$$E(X_{n+1} | \mathbf{X} = \mathbf{x}) = \int_0^{\infty} (1 - F_{X_{n+1}}(z | \mathbf{X} = \mathbf{x})) dz = \int_0^{\infty} \frac{(\lambda+n\bar{\mathbf{x}})^{n+\alpha}}{(\lambda+z+n\bar{\mathbf{x}})^{n+\alpha}} dz = \frac{\lambda+n\bar{\mathbf{x}}}{n+\alpha-1},$$

also

$$H^{***} = E(X_{n+1} | \mathbf{X}) = \frac{\lambda+n\bar{\mathbf{X}}}{n+\alpha-1} = \frac{n}{\alpha+n-1} \bar{\mathbf{X}} + \frac{\alpha-1}{\alpha+n-1} H = H^*,$$

d.h. die empirische stimmt hier mit der gewöhnlichen Credibility-Prämie überein.

Dass dieses Ergebnis nicht überraschend ist, zeigt der nachfolgende

III. Credibility-Theorie

Satz 9. Im Standard-Modell der Credibility-Theorie gilt stets:

$$H^{***} = H^*,$$

d.h. empirische und gewöhnliche Credibility-Prämie sind identisch.

Beweis: Es ist

$$\begin{aligned} E(X_{n+1} | \mathbf{X}) &= E[E(X_{n+1} | \mathbf{X}) | (\mathbf{X}, \Theta)] = E[E(X_{n+1} | (\mathbf{X}, \Theta)) | \mathbf{X}] \\ &= E[E(X_{n+1} | \Theta) | \mathbf{X}] = E[H(\Theta) | \mathbf{X}] = H^*, \end{aligned}$$

da aufgrund der vorausgesetzten bedingten Unabhängigkeit gilt:

$$E(X_{n+1} | (\mathbf{X}, \Theta)) = E(X_{n+1} | \Theta) = H(\Theta). \quad \blacksquare$$

III. 4. Modellerweiterungen

Bisher haben wir im Wesentlichen nur Vergangenheitsschäden eines einzelnen VN betrachtet, um „seine“ risikogerechte Tarifierung zu ermöglichen. Dabei war aber in jedem Fall noch die Kenntnis der (theoretischen) Kollektivprämie H erforderlich. In diesem Abschnitt wollen wir abschließend zwei Modellansätze vorstellen, die eine Schätzung auch dieser Prämie über das Kollektiv ermöglichen und damit den Modellrahmen geeignet erweitern.

Im so genannten *einfachen Bühlmann-Modell* wird angenommen, dass ein Kollektiv aus m VN vorliegt, von denen jeder einen Strukturparameter ϑ_i besitzt und für den die individuellen Risiken $\{X_{ij} | 1 \leq j \leq n\}$ unter ϑ_i wieder bedingt unabhängig sind mit

- $H(\vartheta) = E(X_{ij} | \Theta_i = \vartheta)$
- $\sigma^2(\vartheta) = \text{Var}(X_{ij} | \Theta_i = \vartheta)$

für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$.

Zusätzlich wird angenommen, dass die ϑ_i Realisationen von stochastisch unabhängigen, identisch wie Θ verteilten Zufallsvariablen Θ_i sind, und auch die „Paare“ $(\Theta_i, \mathbf{X}_{in})$ für $i = 1, \dots, m$ stochastisch unabhängig sind mit $\mathbf{X}_m = \{X_{ij} | 1 \leq j \leq n\}$. Insbesondere folgt aus diesen Annahmen, dass

$$E[H(\Theta_i)] = E[H(\Theta)] =: H$$

unabhängig von $i \in \{1, \dots, m\}$ ist und damit auch

$$H = E[E(X_{ij} | \Theta_i)] = E(X_{ij}) \text{ für alle } i \in \{1, \dots, m\} \text{ und } j \in \{1, \dots, n\}$$

gilt. Es ist also nahe liegend, die Kollektivprämie H aus bekannten Schäden der Vergangenheit durch Mittelbildung zu schätzen durch

III. Credibility-Theorie

$$\hat{H} = \bar{\mathbf{X}} = \frac{1}{nm} \sum_{i=1}^m \sum_{j=1}^n X_{ij} = \frac{1}{m} \sum_{i=1}^m \bar{\mathbf{X}}_i,$$

wenn n zurückliegende Schadenjahre des Kollektivs beobachtet wurden, wobei sinngemäß wieder $\bar{\mathbf{X}}_i = \frac{1}{n} \sum_{j=1}^n X_{ij}$ für $i \in \{1, \dots, m\}$ zu setzen ist. Die individuelle linearisierte Credibility-Prämie würde sich dann durch

$$\hat{H}_i^{**} = z_{in} \bar{\mathbf{X}}_i + (1 - z_{in}) \hat{H}$$

schätzen lassen, sowie

$$z_{in} = \frac{\text{Var}(H(\Theta_i))}{\frac{1}{n} E[\text{Var}(X_{ij} | \Theta_i)] + \text{Var}(H(\Theta_i))}, \text{ unabhängig von } j \in \{1, \dots, n\}.$$

Man beachte, dass auf Grund der Annahmen des einfachen Bühlmann-Modells diese Faktoren auch nicht von $i \in \{1, \dots, m\}$ abhängen.

Man kann ferner zeigen, dass sich dieses Ergebnis auch aus einem allgemeineren Ansatz für die linearisierte Credibility-Prämie ergibt, bei dem die Schäden der Vergangenheit *aller* VN berücksichtigt werden. Das zugehörige Lagrange-Minimierungsproblem für eine linearisierte Credibility-

Prämie der Form $H_i^{**} = \sum_{k=1}^m \sum_{j=1}^n z_{kj}(i) X_{kj}$ lautet:

$$\min! E \left[\left(H_i^{**} - H_i^* \right)^2 \right] = E \left[\left(\sum_{k=1}^m \sum_{j=1}^n z_{kj}(i) X_{kj} - H_i^* \right)^2 \right] \text{ unter der Nebenbedingung } \sum_{k=1}^m \sum_{j=1}^n z_{kj}(i) = 1$$

für $i \in \{1, \dots, m\}$.

Die Nebenbedingung ist dabei erforderlich, weil $E(H_i^*) = E(H_i^{**}) = H = E(X_{ij})$ gelten muss für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$.

Die auf diese Weise gewonnene Credibility-Prämie heißt *homogene* linearisierte Credibility-Prämie, weil sie – anders als üblicherweise bei linearen Regressionsproblemen – auf ein Absolutglied verzichtet. Schließt man ein solches Absolutglied im Minimierungsproblem mit ein, ergibt sich abweichend für jeden VN die ursprüngliche (*inhomogene*) linearisierte Credibility-Prämie

$$H_i^{**} = z_{in} \bar{\mathbf{X}}_i + (1 - z_{in}) H \text{ für alle } i \in \{1, \dots, m\}.$$

Für Einzelheiten, die wir hier nicht ausführen, sei auf BÜHLMANN UND GISLER (2005), Abschnitt 3.1.4 verwiesen.

Als zweite Modellerweiterung, die vor allem bei der Kraftfahrzeug-Haftpflicht-Versicherung Anwendung findet, wollen wir abschließend noch das so genannte *Bühlmann-Straub-Modell* vorstellen, welches durch folgende Modellspezifikationen charakterisiert ist.

III. Credibility-Theorie

Es wird angenommen, dass ein Kollektiv aus m VN vorliegt, von denen jeder einen Strukturparameter ϑ_i besitzt und für den die individuellen Risiken $\{X_{ij} | 1 \leq j \leq n\}$ unter ϑ_i wieder bedingt unabhängig sind mit

- $H(\vartheta) = E(X_{ij} | \Theta_i = \vartheta)$
- $Var(X_{ij} | \Theta_i = \vartheta) = \frac{\sigma^2(\vartheta)}{w_{ij}}$

für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$. Hierbei sind die w_{ij} positive Gewichte (meist als *Volumenmaße* bezeichnet). Zusätzlich wird wieder angenommen, dass die ϑ_i Realisationen von stochastisch unabhängigen, identisch wie Θ verteilten Zufallsvariablen Θ_i sind, und auch die „Paare“ $(\Theta_i, \mathbf{X}_{in})$ für $i = 1, \dots, m$ stochastisch unabhängig sind mit $\mathbf{X}_{in} = \{X_{ij} | 1 \leq j \leq n\}$.

Für eine inhaltliche Erläuterung dieser Modellspezifikationen aus „praktischer Sicht“ sei hier auf BÜHLMANN UND GISLER (2005), Abschnitt 4.2 verwiesen.

Der Unterschied des Bühlmann-Straub-Modells zum einfachen Bühlmann-Modell besteht also im Wesentlichen darin, dass die bedingten Varianzen der individuellen Risiken nicht konstant zu sein brauchen.

Als inhomogene linearisierte Credibility-Prämie erhält man hier

$$H_i^{**} = z_{in} \bar{\mathbf{X}}_i + (1 - z_{in}) H$$

mit

$$z_{in} := \frac{Var(H(\Theta_i))}{\frac{1}{w_{i\bullet}} E[\sigma^2(\Theta_i)] + Var(H(\Theta_i))} \quad \text{und} \quad w_{i\bullet} := \sum_{j=1}^n w_{ij} \quad \text{für alle } i \in \{1, \dots, m\}.$$

Im Fall gleicher Gewichte $w_{ij} = 1$ (woraus $w_{i\bullet} = n$ folgt) ergibt sich offensichtlich die alte Formel des einfachen Bühlmann-Modells.

Für die homogene linearisierte Credibility-Prämie ergibt sich aber – abweichend von der Intuition –

$$\hat{H}_i^{**} = z_{in} \bar{\mathbf{X}}_i + (1 - z_{in}) \bar{\bar{\mathbf{X}}}, \quad \text{mit} \quad \bar{\bar{\mathbf{X}}} := \sum_{i=1}^m \frac{z_{in}}{z_{\bullet n}} \sum_{j=1}^n \frac{w_{ij}}{w_{i\bullet}} X_{ij} \quad \text{und} \quad z_{\bullet n} = \sum_{j=1}^n z_{jn} \quad \text{für alle } i \in \{1, \dots, m\}.$$

Diese Prämie hat eine im Allgemeinen kleinere Varianz als die intuitive Alternative

$$\hat{H}_i^{**} = z_{in} \bar{\mathbf{X}}_i + (1 - z_{in}) \bar{\mathbf{X}}.$$

Für Details siehe BÜHLMANN UND GISLER (2005), Abschnitt 4.3.

III. Credibility-Theorie

Beispiel 29. Wir betrachten der Einfachheit halber ein (fiktives) Portfolio aus $m = 6$ m VN mit einem Beobachtungsumfang von $n = 10$ Jahren. Es wird ein einfaches Bühlmann-Modell unterstellt. Die Schadendaten seien wie folgt tabellarisch in € gegeben (alle Rechnungen gerundet):

VN\Jahr	1	2	3	4	5	6	7	8	9	10	\bar{X}_i	V_i	
1	170	141	104	190	100	160	235	163	275	146	168	2.938	
2	124	118	117	173	75	197	133	55	107	295	139	4.691	
3	344	161	126	423	156	228	283	247	186	78	223	10.998	
4	89	97	251	276	242	148	205	332	226	306	217	6.909	
5	93	82	136	107	81	48	76	170	209	119	112	2.337	
6	192	113	151	218	143	196	132	450	233	383	221	12.319	
											$\hat{H} = \bar{X} \quad \quad \bar{V}$	180	6.699
											emp. Varianz s^2	2.786	----

Die beiden rechten Spalten enthalten die Mittelwerte \bar{X}_i der (bedingt unabhängigen) Schäden der einzelnen VN als Schätzer für die $H(\Theta_i) = E(X_{ij}|\Theta_i)$ sowie die empirischen Varianzen V_i der (bedingt unabhängigen) Schäden als Schätzer für die $Var(X_{ij}|\Theta_i)$. Wenn keine weiteren Verteilungsannahmen getroffen werden, muss man die für die Credibility-Faktoren benötigten Größen $Var(H(\Theta_i))$ und $E[Var(X_{ij}|\Theta_i)]$ je VN i unter Verwendung dieser Informationen schätzen. Für die Credibility-Faktoren ergibt sich damit

$$\hat{z}_{in} = \frac{s^2}{\frac{1}{n}\bar{V} + s^2} = 0,772 \quad \text{für } i \in \{1, \dots, m\},$$

wobei \bar{V} den Mittelwert der V_i und s^2 die empirische Varianz der \bar{X}_i bezeichnet. Man erhält damit auf der Basis von 10 beobachteten Schadenjahren die folgenden Credibility-Prämien für die einzelnen VN:

VN	1	2	3	4	5	6
\hat{H}_i^{**}	171	149	213	209	128	212

Zum Vergleich: die Daten wurden simuliert aus Lognormal-Verteilungen, mit

$$P^{X_{ij}/100}(\cdot|\Theta_i = \vartheta) = \mathcal{LN}(\vartheta, 1/2) \quad \text{und} \quad P^{\Theta_i} = \mathcal{U}[0,1].$$

Damit ergibt sich:

$$H(\Theta_i) = 100 \cdot \exp(\Theta_i + 1/4), \quad Var(X_{ij}|\Theta_i) = 10.000 \cdot \exp(2\Theta_i + 1/2) \cdot (e^{1/2} - 1)$$

sowie $Var(H(\Theta_i)) = 3990$, $E[Var(X_{ij}|\Theta_i)] = 34.167$, $z_{in} = 0,5387$ mit den Credibility-Prämien

VN	1	2	3	4	5	6
H_i^{**}	174	158	203	200	144	202

IV Ruintheorie

Schon sehr früh in der Entwicklung der modernen Versicherungsmathematik spielte das Ruinproblem, wie wir es am Anfang des Abschnitts II.6 skizziert haben, eine zentrale Rolle. Es wurde vor allem in Skandinavien durch Harald Cramér und Filip Lundberg Anfang des 20. Jahrhunderts behandelt, wobei die Einbettung des Problems in den sogenannten *Risiko-Prozess* eine große Rolle spielt. Man gelangt damit in das größere Gebiet der (insbesondere zeitstetigen) *Stochastischen Prozesse*, die allerdings im Rahmen dieser Veranstaltung nicht vertieft besprochen werden.

IV. 1. Zeitdiskrete Ruinprobleme

Bevor wir uns dem zeitstetigen Ruin-Problem widmen, wollen wir zunächst einige elementarere Überlegungen im diskreten Rahmen anstellen.

Als Ausgangssituation betrachten wir wie auch schon zuvor in der Schadenversicherungsmathematik eine Familie von Risiken $\{X_n\}_{n \in \mathbb{N}}$, die stochastisch unabhängig und wie ein "typisches" Risiko X (etwa der Jahres-Gesamtschaden) verteilt sein mögen, für das die momenterzeugende Funktion $\psi_X(t)$ für $0 < t \leq \delta$ mit einem geeigneten $\delta > 0$ existiere. Wir nehmen ferner an, dass jedes Risiko mit einer (konstanten) Prämie H behaftet ist, die z.B. nach einem der obigen Prämien-Prinzipien berechnet sei. Dann entspricht

$$\sum_{i=1}^n X_i \text{ dem } \textit{Schadenaufkommen} \text{ und } nH \text{ dem } \textit{Prämienaufkommen}$$

der ersten n Perioden. Als zusätzlichen Parameter führen wir die sogenannte *Anfangsreserve* $s \geq 0$ in das Modell ein. Wir sagen, dass in der Periode n ein *technischer Ruin* eintritt, wenn zu diesem Zeitpunkt das Schadenaufkommen erstmalig das Prämienaufkommen plus die Anfangsreserve übersteigt, d.h. wenn gilt:

$$\sum_{i=1}^n X_i > nH + s \quad \text{und} \quad \sum_{i=1}^k X_i \leq kH + s \quad \text{für alle } k < n.$$

Die Größe

$$\phi_m(s) := P\left(\bigcup_{n=1}^m \left\{ \sum_{i=1}^n X_i > nH + s \right\}\right), \quad 1 \leq m \leq \infty$$

gibt dann die Wahrscheinlichkeit dafür an, dass wenigstens einmal innerhalb von m Perioden ein technischer Ruin eintritt. Ein Ziel dieses Abschnitts ist es, geeignete Abschätzungen für diese "Ruin-Wahrscheinlichkeit" herzuleiten.

IV. Ruintheorie

Lemma 28. Unter den Bedingungen $\psi_X(t) < \infty$ und $H > \frac{1}{t} \ln(\psi_X(t))$ für $0 < t \leq T$ gilt:

$$\begin{aligned} \phi_m(s) &\leq e^{-st} e^{-tH} \psi_X(t) \frac{1 - \{e^{-tH} \psi_X(t)\}^m}{1 - e^{-tH} \psi_X(t)}, \quad m \in \mathbb{N} \quad \text{und} \\ \phi_\infty(s) &\leq e^{-st} \frac{e^{-tH} \psi_X(t)}{1 - e^{-tH} \psi_X(t)} \quad \text{für alle } 0 < t \leq T. \end{aligned}$$

Beweis: Mit der Markoff-Ungleichung⁷ erhalten wir:

$$\begin{aligned} \phi_m(s) &\leq \sum_{n=1}^m P\left(\sum_{i=1}^n X_i > nH + s\right) = \sum_{n=1}^m P\left(e^{\sum_{i=1}^n X_i} > e^{st} e^{nH}\right) \leq \sum_{n=1}^m \frac{E\left(e^{\sum_{i=1}^n X_i}\right)}{e^{st} e^{nH}} = \\ &e^{-st} \sum_{n=1}^m \{e^{-tH} \psi_X(t)\}^n = e^{-st} e^{-tH} \psi_X(t) \frac{1 - \{e^{-tH} \psi_X(t)\}^m}{1 - e^{-tH} \psi_X(t)}. \end{aligned}$$

Der Grenzübergang $m \rightarrow \infty$ ergibt den zweiten Teil der Aussage. ■

Die zusätzliche Bedingung an die Prämie H impliziert dabei, dass H erwartungswertübersteigend ist wegen

$$\frac{1}{t} \ln(\psi_X(t)) = \frac{1}{t} \ln(E(e^{tX})) \geq \frac{1}{t} \ln e^{E(tX)} = E(X) \quad \text{für alle } 0 < t \leq T$$

nach der Jensen'schen Ungleichung. Für $H = \frac{1}{t} \ln(\psi_X(t))$ läge dabei gerade eine Prämie nach dem Exponential-Prinzip vor. Für $t \rightarrow 0$ ergibt sich noch

$$\lim_{t \rightarrow 0} \frac{1}{t} \ln(\psi_X(t)) = \frac{d}{dt} \ln(\psi_X(t)) \Big|_{t=0} = \frac{\psi'_X(0)}{\psi_X(0)} = \psi'_X(0) = E(X).$$

Es ist interessant zu untersuchen, ob es einen Wert für t gibt, bei dem die rechte Seite in Lemma 28 möglichst klein wird. Diese Frage beantwortet

Lemma 29. Unter den Voraussetzungen von Lemma 28 existiert stets ein $t_0 > 0$, für das der Ausdruck $g(t) := e^{-tH} \psi_X(t)$ für $0 < t \leq T$ minimal wird. Gilt ferner $\psi'_X(T) > H \psi_X(T)$, so ist t_0 Lösung der Gleichung

$$\tilde{H} := \frac{E(X e^{t_0 X})}{E(e^{t_0 X})} = \frac{\psi'_X(t_0)}{\psi_X(t_0)} = H,$$

⁷ Für integrierbare Zufallsvariablen Z gilt: $P(Z > z) \leq \frac{E(|Z|)}{z}$ für $z > 0$.

IV. Ruintheorie

d.h. t_0 entspricht dann demjenigen Parameter der Esscher-Prämie \tilde{H} , für den diese und die Original-Prämie zusammenfallen. Es gilt dann

$$\phi_m(s) \leq e^{-st_0} g(t_0) \frac{1 - g^m(t_0)}{1 - g(t_0)}, \quad m \in \mathbb{N} \quad \text{und} \quad \phi_\infty(s) \leq e^{-st_0} \frac{g(t_0)}{1 - g(t_0)}.$$

Beweis: Es gilt:

$$g'(t) = E\left((X - H) e^{t(X-H)}\right) \quad \text{mit} \quad g'(0) = E(X) - H < 0 \quad \text{und}$$

$$g''(t) = E\left((X - H)^2 e^{t(X-H)}\right) \geq 0, \quad 0 \leq t \leq T,$$

d.h. g ist konvex mit $g(0) = 1$ und $g'(0) < 0$, woraus die Existenz eines g minimierenden $t_0 \in (0, T]$ folgt. Unter der obigen Zusatzbedingung, die $g'(T) > 0$ impliziert, ist dann t_0 Lösung der Gleichung

$$0 = g'(t_0) = e^{-t_0 H} E\left((X - H) e^{t_0 X}\right) = e^{-t_0 H} \left(\psi_X'(t_0) - H \psi_X(t_0)\right),$$

woraus die Behauptung wegen $\phi_m(s) \leq e^{-st_0} \sum_{n=1}^m g^n(t_0) = e^{-st_0} g(t_0) \frac{1 - g^m(t_0)}{1 - g(t_0)}$ für alle $m \in \mathbb{N}$ und

$$\phi_\infty(s) \leq e^{-st_0} \sum_{n=1}^{\infty} g^n(t_0) = e^{-st_0} \frac{g(t_0)}{1 - g(t_0)} \quad \text{folgt.} \quad \blacksquare$$

Beispiel 30. Das Risiko X sei $\Gamma(\alpha, \lambda)$ -verteilt mit $\alpha, \lambda > 0$. Dann ist $\psi_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha$ mit $T < \lambda$.

Nach Voraussetzung muss $H > \frac{1}{t} \ln(\psi_X(t)) = \frac{\alpha}{t} \ln\left(\frac{\lambda}{\lambda - t}\right) =: f_{\alpha, \lambda}(t)$ sein. Die Funktion $f_{\alpha, \lambda}(t)$ ist aber für alle $\alpha, \lambda > 0$ in t streng monoton wachsend, so dass sich die Bedingung zu

$$H > \frac{1}{T} \ln(\psi_X(T)) = \frac{\alpha}{T} \ln\left(\frac{\lambda}{\lambda - T}\right) =: f_{\alpha, \lambda}(T)$$

reduziert. Aus der Ungleichung $\ln(1 + z) \leq z$ für $z \geq 0$ ergibt sich die Ungleichung

$$\frac{\alpha}{T} \ln\left(\frac{\lambda}{\lambda - T}\right) \leq \frac{\alpha}{\lambda - T} = \frac{\psi_X'(T)}{\psi_X(T)},$$

so dass für die (stets mögliche) Wahl der Prämie H zwischen diesen Grenzen, also

$$\frac{\alpha}{T} \ln\left(\frac{\lambda}{\lambda - T}\right) < H \leq \frac{\alpha}{\lambda - T}$$

folgt: $t_0 = \lambda - \frac{\alpha}{H}$. Damit ergibt sich etwa

IV. Ruintheorie

$$\phi_\infty(s) \leq e^{-st_0} \frac{g(t_0)}{1-g(t_0)} = e^{-s\left(\lambda - \frac{\alpha}{H}\right)} \frac{g(t_0)}{1-g(t_0)} \quad \text{mit } g(t_0) = e^{-\lambda H} \left(\frac{e\lambda H}{\alpha}\right)^\alpha.$$

Für die konkrete Situation mit $\alpha = 2$ und $\lambda = \frac{1}{25}$ (monetäre Einheit: 1 Mio €) ergibt sich bei $T = \frac{1}{50}$ und der Prämie $H = 75$ der Wert $t_0 = \frac{1}{75} = 0,01\bar{3}$ sowie die Abschätzung

$$\phi_\infty(s) \leq e^{-st_0} \frac{g(t_0)}{1-g(t_0)} = e^{-s\left(\lambda - \frac{\alpha}{H}\right)} \frac{g(t_0)}{1-g(t_0)} = 4,8048 \cdot e^{-0,0133s}.$$

Will man diese Ruinwahrscheinlichkeit unter 1% halten, benötig man eine Anfangsreserve von (maximal) 463,12 Mio €

Das folgende Resultat ergänzt Lemma 29 um eine (einfache) Abschätzung der Ruinwahrscheinlichkeit nach unten.

Lemma 30. Es sei $G_X := 1 - F_X$, H die Prämie für das Risiko X und s die Anfangsreserve. Dann gilt (auch ohne die Voraussetzung der Existenz der momenterzeugenden Funktion):

$$\begin{aligned} \phi_m(s) &\geq 1 - \exp\left\{-\frac{1}{H} \int_{H+s}^{(m+1)H+s} G_X(z) dz\right\}, \quad m \in \mathbb{N} \quad \text{und} \\ \phi_\infty(s) &\geq 1 - \exp\left\{-\frac{1}{H} \int_{H+s}^{\infty} G_X(z) dz\right\}. \end{aligned}$$

Für den Fall, dass $\mu := E(X) < \infty$ ist, gilt alternativ auch

$$\phi_\infty(s) \geq 1 - \exp\left\{-\frac{1}{H} \left(\mu - \int_0^{H+s} G_X(z) dz\right)\right\}.$$

Beweis: Es ist

$$\begin{aligned} \phi_m(s) &\geq P\left(\bigcup_{n=1}^m \{X_n > nH + s\}\right) = 1 - P\left(\bigcap_{n=1}^m \{X_n \leq nH + s\}\right) = 1 - \prod_{n=1}^m F_X(nH + s) \\ &= 1 - \prod_{n=1}^m \underbrace{(1 - G_X(nH + s))}_{\leq \exp(-G_X(nH+s))} \geq 1 - \exp\left\{-\sum_{n=1}^m G_X(nH + s)\right\} \geq 1 - \exp\left\{-\frac{1}{H} \int_{H+s}^{(m+1)H+s} G_X(z) dz\right\} \end{aligned}$$

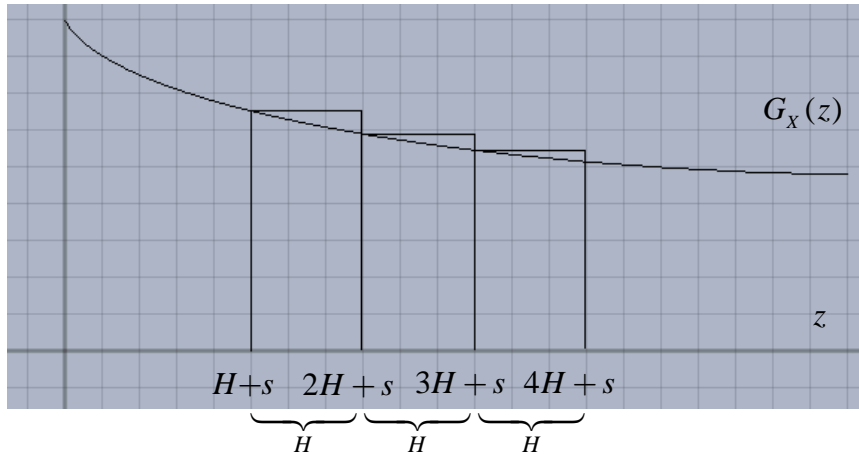
wegen

$$H \cdot \sum_{n=1}^m G_X(nH + s) \geq \int_{H+s}^{(m+1)H+s} G_X(z) dz;$$

IV. Ruintheorie

vgl. nachfolgende Skizze. Der Rest der ersten Aussage ergibt sich durch den Grenzübergang $m \rightarrow \infty$. Die alternative Darstellung folgt aus

$$\mu = E(X) = \int_0^{\infty} G_X(z) dz. \quad \blacksquare$$



Im obigen Beispiel 30 ergibt sich rechnerisch mit $H = 75$, $\mu = 50$ und $s = 463,12$ hierfür noch

$$\phi_{\infty}(s) \geq 1 - e^{-\mu/H} \exp \left\{ \frac{1}{H} \int_0^{H+s} G_X(z) dz \right\} = 3,62 \cdot 10^{-9}.$$

Die folgende Tabelle zeigt, wie sich diese untere Schranke für verschiedene Werte von H und s ändert:

H	50	55	60	65
s				
0	0,2371	0,1907	0,1532	0,1231
50	0,0535	0,0414	0,0322	0,0252
100	0,0098	0,0075	0,0058	0,0045
150	0,0017	0,0013	0,0010	0,0008

Es ist interessant, dass Lemma 30 für *nicht-versicherbare* Risiken X , also solche mit $E(X) = \infty$, noch die Aussage

$$\phi_{\infty}(s) = 1 \text{ für alle } H, s > 0$$

liefert, d.h. es tritt im Laufe der Zeit mit Sicherheit Ruin ein, unabhängig davon, wie hoch die Prämie H oder die Anfangsreserve s gewählt werden! Dies folgt aus der Beobachtung, dass im vorliegenden Fall stets

IV. Ruintheorie

$$\infty = E(X) = \int_0^{\infty} G_X(z) dz = \int_{H+s}^{\infty} G_X(z) dz$$

gilt. Mit Hilfe des Borel-Cantelli-Lemmas⁸ kann diese Aussage noch dahingehend verschärft werden, dass für ein nicht-versicherbares Risiko X mit Wahrscheinlichkeit 1 sogar unendlich oft ein *Einzelschaden* (!) das gesamte bisherige Prämienaufkommen samt Anfangsreserve überschreitet, weil offenbar mit dem Beweisargument von oben

$$\sum_{n=1}^{\infty} P(X_n > nH + s) = \sum_{n=1}^{\infty} G_X(nH + s) \geq \frac{1}{H} \int_{H+s}^{\infty} G_X(z) dz = \infty$$

gilt. Die Ungleichung aus Lemma 30 ist daher vor allem für "gefährliche" Risiken, etwa vom Pareto-Typ, geeignet.

Beispiel 31 (Pareto-Verteilung). In der Rückversicherung wird gern die so genannte *Pareto-Verteilung* zur Modellierung von Großschäden benutzt; ihre Verteilungsfunktion ist in der Standard-Form gegeben durch

$$F_X(x) = 1 - \frac{1}{(1+x)^\alpha}, \quad x \geq 0; \alpha > 0.$$

Für $\alpha \leq 1$ ergibt sich ein nicht-versicherbares Risiko, für $\alpha > 1$ ist der Erwartungswert gegeben durch

$$\mu = E(X) = \int_0^{\infty} \frac{1}{(1+x)^\alpha} dx = \frac{1}{\alpha-1}.$$

Eine ähnliche Rechnung zeigt

$$\int_{H+s}^{\infty} G_X(z) dz = \int_{H+s}^{\infty} \frac{1}{(1+x)^\alpha} dx = \frac{1}{(\alpha-1)(1+H+s)^{\alpha-1}}$$

und somit

$$\phi_\infty(s) \geq 1 - \exp\left[-\frac{1}{(\alpha-1)H(1+H+s)^{\alpha-1}}\right] = 1 - \exp\left[-\frac{\mu}{H(1+H+s)^{\alpha-1}}\right].$$

Die nachfolgenden Überlegungen werden es ermöglichen, die bisherigen Abschätzungen für Ruin-Wahrscheinlichkeiten erheblich zu verbessern.

⁸ Gilt für eine unabhängige Ereignisfolge $\{A_n\}_{n \in \mathbb{N}}$ die Beziehung $\sum_{n=1}^{\infty} P(A_n) = \infty$, so ist $P\left(\limsup_{n \rightarrow \infty} A_n\right) = 1$, d.h. es treten unendlich viele der A_n mit Wahrscheinlichkeit 1 ein.

IV. Ruintheorie

Lemma 31. $\{X_n\}_{n \in \mathbb{N}}$ sei eine Folge unabhängiger, identisch wie X verteilter reeller Zufallsvariablen mit Verteilung $Q = P^X$. Ferner sei

$$U_m(s) := P\left(\bigcap_{n=1}^m \left\{ \sum_{i=1}^n X_i \leq nH + s \right\}\right)$$

Für $m \in \mathbb{N}$, $H, s \in \mathbb{R}$. Dann gilt:

$$U_{m+1}(s) = \int_{-\infty}^{H+s} U_m(H + s - y) Q(dy).$$

Beweis: Es gilt

$$\begin{aligned} U_{m+1}(s) &= P\left(\bigcap_{n=1}^{m+1} \left\{ \sum_{i=1}^n X_i \leq nH + s \right\}\right) = P\left(\bigcap_{n=1}^{m+1} \left\{ \sum_{i=2}^n X_i \leq (n-1)H + s + H - X_1 \right\}\right) \\ &= P\left(\bigcap_{n=2}^{m+1} \left\{ \sum_{i=2}^n X_i \leq (n-1)H + s + H - X_1 \right\} \cap \{X_1 \leq H + s\}\right) \\ &= \int_{-\infty}^{H+s} P\left(\bigcap_{n=2}^{m+1} \left\{ \sum_{i=2}^n X_i \leq (n-1)H + s + H - y \right\}\right) Q(dy) \\ &= \int_{-\infty}^{H+s} \underbrace{P\left(\bigcap_{r=1}^m \left\{ \sum_{j=1}^r X_{j+1} \leq rH + s + H - y \right\}\right)}_{=U_m(H+s-y)} Q(dy) = \int_{-\infty}^{H+s} U_m(H + s - y) Q(dy). \quad \blacksquare \end{aligned}$$

Lemma 32 (Anpassungskoeffizient). Es sei $X \geq 0$ eine Zufallsvariable, für die die momenterzeugende Funktion $\psi_X(t)$ existiere an einer Stelle $t = T > 0$. Dann ist insbesondere X integrierbar, etwa mit Erwartungswert $E(X) = \mu$. Ferner gelte die Zusatzbedingung aus Lemma 29 für die Prämie H , d.h. $\psi'_X(T) > H \psi_X(T)$.⁹ Dann existiert genau ein $R > 0$ mit der Eigenschaft

$$\psi_X(R) = e^{RH}.$$

R wird auch als *Cramér-Lundberg-Anpassungskoeffizient* bezeichnet.

Beweis: Sei $g(t) := e^{-tH} \psi_X(t) = E\left(e^{t(X-H)}\right)$, $0 \leq t \leq T$ wie im Beweis zu Lemma 29. Wegen

$$\begin{aligned} g'(t) &= E\left((X - H) e^{t(X-H)}\right) \text{ mit } g'(0) = E(X) - H < 0 \text{ und} \\ g''(t) &= E\left((X - H)^2 e^{t(X-H)}\right) \geq 0, \quad 0 \leq t \leq T \end{aligned}$$

⁹ Dies impliziert insbesondere, dass $H \geq \mu$ ist; vgl. Lemma 21 zum Esscher-Prinzip.

IV. Ruintheorie

ist g konvex mit $g(0)=1$ und $g'(0)<0$; die Zusatzbedingung impliziert $g'(T)>0$, woraus $\lim_{t \rightarrow \infty} g(t) = \infty$ und somit wegen der Stetigkeit von g die Behauptung folgt. ■

Wir kommen jetzt zum Hauptresultat dieses Abschnitts.

Satz 10 (Cramér-Lundberg). $\{X_n\}_{n \in \mathbb{N}}$ sei eine Folge unabhängiger, identisch wie X verteilter nicht-negativer Risiken. Die momenterzeugende Funktion $\psi_X(t)$ existiere an einer Stelle $t = T > 0$. Ferner gelte die Zusatzbedingung aus Lemma 29 für die Prämie H , d.h. $\psi'_X(T) > H \psi_X(T)$; R bezeichne den nach Lemma 31 eindeutig bestimmten Anpassungskoeffizienten. Dann gilt:

$$\phi_m(s) \leq \phi_\infty(s) \leq e^{-Rs}, \quad s \geq 0; \quad m \in \mathbb{N}.$$

Beweis: Wir zeigen den Satz mit vollständiger Induktion. Für $m=1$ können wir den Beweisgang für Lemma 28 verwenden; für diesen Fall ist nämlich die Zusatzbedingung $H > \frac{1}{t} \ln(\psi_X(t))$ für $0 < t \leq T$ nicht notwendig, und es ergibt sich

$$\phi_1(s) \leq e^{-st} e^{-tH} \psi_X(t) \quad \text{für alle } t \geq 0 \text{ mit } \psi_X(t) < \infty.$$

Mit der Wahl $t = R$ erhält man also

$$\phi_1(s) \leq e^{-Rs} \underbrace{e^{-RH} \psi_X(R)}_{=1} = e^{-Rs}, \quad \text{wie behauptet.}$$

Unter der Annahme, dass die linke Ungleichung für ein $m \in \mathbb{N}$ richtig ist, ergibt sich nun mit $Q = P^X$ und der Abkürzung

$$\phi_m^*(s) := \begin{cases} \phi_m(s), & s \geq 0 \\ 1, & s < 0 \end{cases}$$

gemäß Lemma 31:

$$\begin{aligned} \phi_{m+1}(s) &= 1 - U_{m+1}(s) = 1 - \int_{-\infty}^{H+s} U_m(H+s-y) Q(dy) = 1 - \int_{-\infty}^{H+s} \{1 - \phi_m(H+s-y)\} Q(dy) \\ &= 1 - \int_{-\infty}^{\infty} \{1 - \phi_m^*(H+s-y)\} Q(dy) = \int_{-\infty}^{\infty} \phi_m^*(H+s-y) Q(dy) \leq \int_{-\infty}^{\infty} e^{-R(H+s-y)} Q(dy) \\ &= e^{-Rs} e^{-RH} \int_{-\infty}^{\infty} e^{Ry} Q(dy) = e^{-Rs} \underbrace{e^{-RH} \psi_X(R)}_{=1} = e^{-Rs}, \end{aligned}$$

d.h. die Ungleichung gilt auch für $m+1$. Durch Grenzübergang $m \rightarrow \infty$ erhält man schließlich noch

$$\phi_m(s) \leq \phi_\infty(s) = \lim_{n \rightarrow \infty} \phi_n(s) \leq e^{-Rs} \quad \text{für alle } m \in \mathbb{N},$$

wie behauptet. Der Satz ist damit bewiesen. ■

IV. Ruintheorie

Beispiel 32. Es sei $Q = P^x = \mathcal{E}(\lambda)$ mit $\lambda > 0$; hier ist also $\mu = \frac{1}{\lambda}$. Der Anpassungskoeffizient R ist dann Lösung der Gleichung

$$e^{RH} = \frac{\lambda}{\lambda - R} = \frac{1}{1 - \mu R}.$$

Die für alle reellen z gültige Ungleichung $1 + z \leq e^z$ führt hier zu der Abschätzung

$$1 + RH \leq e^{RH} = \frac{1}{1 - \mu R} \text{ bzw. } 1 + (H - \mu)R - \mu HR^2 \leq 1 \text{ bzw. } R \geq \frac{H - \mu}{H\mu} = \frac{1}{\mu} - \frac{1}{H}.$$

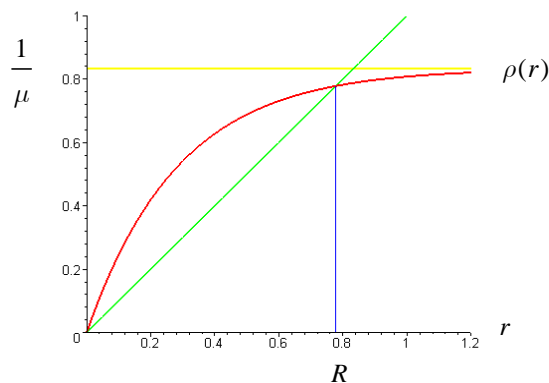
Satz 10 liefert somit die Abschätzung

$$\phi_m(s) \leq \phi_\infty(s) \leq \exp\left\{-\left(\frac{1}{\mu} - \frac{1}{H}\right)s\right\} = \exp\left\{-\left(\lambda - \frac{1}{H}\right)s\right\}, \quad s \geq 0; \quad m \in \mathbb{N}.$$

Eine verbesserte Näherungslösung ist iterativ erhältlich durch Betrachten der Fixpunktgleichung

$$r = \rho(r) := \frac{1}{\mu}(1 - e^{-rH}) \text{ mit } \rho'(0) = \frac{H}{\mu} > 1, \quad \rho''(r) = -\frac{H}{\mu^2}e^{-rH} < 0$$

(d.h. ρ ist konkav). Mit $R_1 := \frac{1}{\mu}$, $R_{n+1} := \rho(R_n)$, $n \in \mathbb{N}$ ergibt sich dann eine Folge, die gegen den Anpassungskoeffizienten konvergiert.



Als konkreten Fall betrachten wir eine Situation mit $\mu = 100$ (Mio €) und Prämie $H = 120$ (Mio €). Die ersten 5 Schritte der Iteration liefern

n	1	2	3	4	5
R_n	0,00698805	0,00567670	0,00493992	0,00447217	0,00415302

IV. Ruintheorie

(zum Vergleich die exakte Lösung, z.B. mit MAPLE: $R = 0,00313698$). Die obige untere Schranke für R ist hier gegeben durch $\frac{1}{\mu} - \frac{1}{H} = 0,001\bar{6}$.

Für eine Anfangsreserve von $s = 200$ (Mio €) erhält man dann als beste der obigen Abschätzungen

$$\phi_m(200) \leq \phi_\infty(200) \leq e^{-200R} = 0,533980, \quad m \in \mathbb{N}.$$

Ist das Risiko X allgemeiner $\Gamma(\alpha, \lambda)$ -verteilt mit $\alpha, \lambda > 0$, so ist der Anpassungskoeffizient zu bestimmen aus der Gleichung

$$e^{RH} = \left(\frac{\lambda}{\lambda - R} \right)^\alpha = \frac{1}{(1 - \mu R / \alpha)^\alpha} \quad \text{bzw.} \quad e^{RH/\alpha} = \frac{\lambda}{\lambda - R} = \frac{1}{1 - \mu R / \alpha}.$$

Überlegungen analog zu oben liefern damit auch eine Abschätzung für R in dieser Situation, wenn man oben formal H durch H/α und μ durch μ/α ersetzt. Insbesondere gilt dann

$$R \geq \alpha \left(\frac{1}{\mu} - \frac{1}{H} \right).$$

In der Situation von Beispiel 32 bedeutet dies: $R \geq \frac{1}{75}$, also

$$\phi_\infty(s) \leq e^{-s/75} = e^{-0,0133s}.$$

Diese Ungleichung ist bereits um den Faktor 4,8 besser als diejenige, die auf der aus Lemma 28 hergeleiteten Ungleichung beruht. Man beachte, dass hier die Minimalstelle t_0 aus Beispiel 30 genau mit dem Faktor für s im Exponenten zusammenfällt. Die benötigte Anfangsreserve reduziert sich unter Verwendung der besseren Abschätzung bei maximaler Ruinwahrscheinlichkeit von 1% damit zu 345,39 Mio €

Der Cramér-Lundberg-Anpassungskoeffizient beträgt für dieses Beispiel nach numerischer Berechnung mit MAPLE $R = 0,02331$, wodurch sich die erforderliche Anfangsreserve bei gleicher Ruinwahrscheinlichkeit nochmals auf 197,56 Mio € verringert.

Eine andere interessante Aussage erhält man für den Anpassungskoeffizienten des Gesamtschadens im kollektiven Modell der Risikotheorie (vgl. Abschnitt II.1), wenn die Schadenfrequenz Poisson-verteilt ist.

Lemma 33. Im kollektiven Modell der Risikotheorie bezeichne wie üblich $S := \sum_{k=1}^N X_k$ den Gesamtschaden. Zusätzlich sei angenommen, dass die Schadenfrequenz N Poisson-verteilt ist mit Parameter $\lambda > 0$. Bezeichnet wieder $\mu = E(X)$ den Erwartungswert der Einzelschadenhöhen und gilt $H > \lambda\mu = E(S)$ für die Prämie, so erfüllt der Anpassungskoeffizient R die Bestimmungsgleichung

$$\psi_X(R) = 1 + \frac{H}{\lambda} R.$$

Beweis: Nach Satz 2 gilt

$$e^{RH} = \psi_s(R) = \varphi_N(\psi_x(R)) = \exp\{\lambda(\psi_x(R) - 1)\}, \text{ also } RH = \lambda(\psi_x(R) - 1),$$

woraus die Behauptung folgt. ■

Beispiel 33 (Fortsetzung von Beispiel 32). Sind die (jährlichen) Einzelschadenhöhen in Lemma 33 sämtlich exponentialverteilt wie in Beispiel 32 und ist die (jährliche) Schadenfrequenz Poissonverteilt, so erhält man die Bestimmungsgleichung

$$\psi_x(R) = \frac{1}{1 - \mu R} = 1 + \frac{H}{\lambda} R \text{ mit der exakten Lösung } R = \frac{\lambda}{\mu H} \left(\frac{H}{\lambda} - \mu \right) = \frac{1}{\mu} - \frac{\lambda}{H},$$

die für $\lambda = 1$ mit der in Beispiel 32 angegebenen approximativen (unteren) Schranke für das Standard-Modell zusammenfällt. Nimmt man also an, dass jedes Versicherungsjahr durch ein kollektives Modell beschrieben werden kann – ähnlich den geophysikalischen Modellansätzen, vgl. Satz 3 – so lassen sich wieder entsprechende Abschätzungen für die Ruinwahrscheinlichkeiten gemäß Satz 10 in der Form

$$\phi_m(s) \leq \phi_\infty(s) \leq e^{-Rs} = \exp\left\{-\left(\frac{1}{\mu} - \frac{\lambda}{H}\right)s\right\}, \quad s \geq 0; \quad m \in \mathbb{N}$$

machen.

IV. 2. Zeitstetige Ruinprobleme

Das bekannteste in der Risikotheorie betrachtete zeitstetige Modell ist eine Erweiterung des kollektiven Modells, in dem die Eintrittszeitpunkte von Schäden nicht mehr diskret (z.B. jahresbezogen), sondern gemäß eines homogenen oder inhomogenen Poisson-Prozesses modelliert werden. Man spricht in diesem Zusammenhang auch von einem *zusammengesetzten Poisson-Prozess*.

Wir werden uns daher zunächst mit einigen mathematischen Grundlagen der so genannten *Stochastischen Prozesse*, insbesondere der Poisson-Prozesse befassen, die wir hier konstruktiv einführen.

Definition 11. Eine Familie $\{N_t | t \geq 0\}$ von Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) heißt *Zählprozess*, wenn alle N_t nur Werte in \mathbb{Z}^+ annehmen und für jedes $\omega \in \Omega$ die „Pfade“ $N_t(\omega)$ als Funktionen von $t \geq 0$ schwach monoton wachsend und rechtsseitig stetig sind. Ein Zählprozess heißt *einfach*, wenn er fast sicher nur Sprünge der Höhe 1 zulässt.

Man interpretiert die Zufallsvariable N_t in der Risikotheorie häufig als „Anzahl von (bestimmten) Ereignissen im Zeitintervall $[0, t]$ “, z.B. der Anzahl von Schadenmeldungen aus Versicherungsverträgen innerhalb des Zeitintervalls $[0, t]$, $t \geq 0$. Die Einfachheit eines solchen Zählprozesses bedeutet dabei anschaulich, dass die „Ereignisse“ fast sicher nicht gleichzeitig eintreten (was bei Kumulschäden allerdings technisch möglich wäre).

IV. Ruintheorie

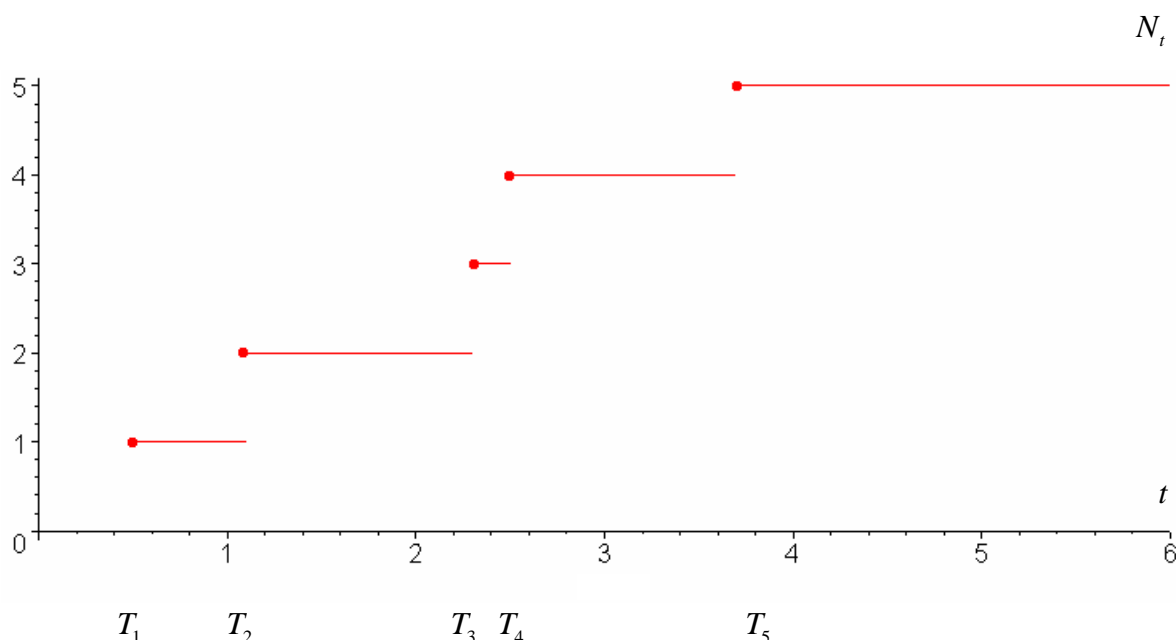
Für überabzählbare Familien von Zufallsvariablen – insbesondere in dem Fall, wo keine stochastische Unabhängigkeit vorliegt – erfordert die Charakterisierung ihrer gemeinsamen Verteilung in der Regel tieferliegende maßtheoretische Betrachtungen. Beispielsweise ist im Allgemeinen schon nicht gesichert, ob „einfache“ Ausdrücke wie $P\left(\bigcap_{t \in [a,b]} \{N_t \in A_t\}\right)$ für $0 < a < b$ und $A_t \in \mathcal{B}^1$ für $t \in [a,b]$ überhaupt Sinn machen, weil der auftretende Durchschnitt mit überabzählbar vielen Ereignissen gebildet wird und diese Operation nicht unbedingt wieder zu einem „Ereignis“, also einer messbaren Menge im Sinne der Maßtheorie führen muss. Solche Schwierigkeiten lassen sich aber umgehen, wenn die betrachteten Stochastischen Prozesse z.B. ein „diskretes Skelett“ besitzen. Für Zählprozesse wählt man daher häufig eine Konstruktion, bei der ein solches diskretes Skelett explizit vorgegeben ist, etwa in Form einer strikt monotonen Folge $\{T_n\}_{n \in \mathbb{N}}$ von Zufallsvariablen mit stetiger Verteilungsfunktion. Durch die Setzung

$$N_t := \max \{n \in \mathbb{N} | T_n \leq t\}, \quad t \geq 0$$

erhält man dann etwa einen einfachen Zählprozess im Sinne der Definition 11. Die T_n heißen üblicherweise *Sprungzeiten* oder *Ereigniszeiten* des Zählprozesses. Ist umgekehrt ein einfacher Zählprozess $\{N_t | t \geq 0\}$ gegeben, so lassen sich die Sprungzeiten darstellen als

$$T_n = \min \{t \geq 0 | N_t \geq n\} = N_{\bullet}^{-1}(n) \quad \text{für alle } n \in \mathbb{N},$$

wobei G^{-1} wie üblich die Pseudo-Inverse der (rechtsseitig stetigen, schwach monoton wachsenden) Funktion G bezeichnet.



Veranschaulichung eines einfachen Zählprozesses

IV. Ruintheorie

Definition 12. Ein einfacher Zählprozess $\{N_t | t \geq 0\}$ heißt Poisson-Prozess, wenn es eine nicht-negative stetige strikt monoton wachsende Funktion m auf \mathbb{R}^+ gibt mit $m(0) = 0$ und der Eigenschaft

$$P^{N_t} = \mathcal{P}(m(t)) \text{ für alle } t \geq 0.$$

Die Funktion m heißt auch Intensitätsfunktion des Poisson-Prozesses; insbesondere erhält man hier $E(N_t) = m(t)$ für alle $t \geq 0$.

Ist speziell $m(t) = \lambda t$ für ein $\lambda > 0$ und $t \geq 0$, so heißt der Poisson-Prozess auch *homogen* (mit Parameter λ).

Auf Grund eines bemerkenswerten Satzes von A. Rényi¹⁰ kann ein Poisson-Prozess im Sinne der Definition 12 kanonisch wie folgt dargestellt werden:

Satz 11. $\{W_n\}_{n \in \mathbb{N}}$ sei eine unabhängige Folge $\mathcal{E}(\lambda)$ -(exponential)verteilter Zufallsvariablen mit Parameter $\lambda > 0$ und $T_n := \sum_{k=1}^n W_k$ für alle $n \in \mathbb{N}$. Dann wird durch

$$N_t := \max \{n \in \mathbb{N} | T_n \leq t\}, t \geq 0$$

ein homogener Poisson-Prozess mit Parameter $\lambda > 0$ definiert. Ist umgekehrt $\{N_t | t \geq 0\}$ ein homogener Poisson-Prozess mit Parameter $\lambda > 0$, so besitzt die Folge der Sprungzeiten $\{T_n\}_{n \in \mathbb{N}} = \{N_{\bullet}^{-1}(n)\}_{n \in \mathbb{N}}$ unabhängige, $\mathcal{E}(\lambda)$ -verteilte Zuwächse $\{W_n\}_{n \in \mathbb{N}}$, definiert durch

$$W_1 := T_1, W_k := T_k - T_{k-1} \text{ für } k \geq 2.$$

Ist $\{N_t | t \geq 0\}$ ein homogener Poisson-Prozess mit Parameter 1 und m eine Intensitätsfunktion, so ist der Zeit-transformierte Zählprozess $\{N_{m(t)} | t \geq 0\}$ ein Poisson-Prozess mit dieser Intensitätsfunktion. Ist umgekehrt $\{N_t | t \geq 0\}$ ein beliebiger Poisson-Prozess mit Intensitätsfunktion m , so ist der Zeit-transformierte Zählprozess $\{N_{m^{-1}(t)} | t \geq 0\}$ ein homogener Poisson-Prozess mit Parameter 1.

Beweis: Wir zeigen hier nur die konstruktive Richtung der Aussagen, für die Umkehrung verweisen wir auf den zitierten Artikel von Rényi.

Aus den Eigenschaften der Folge $\{W_n\}_{n \in \mathbb{N}}$ folgt zunächst, dass jedes T_n eine $\mathcal{E}(n, \lambda)$ -(Erlang-) Verteilung besitzt. Für $n \in \mathbb{Z}^+, t \geq 0$ gilt dann unter Verwendung der Setzung $T_0 := 0$:

$$P(N_t = n) = P(T_n \leq t < T_{n+1}) = P(T_n \leq t) - P(T_{n+1} \leq t)$$

mit

¹⁰ A. Rényi (1967): *Remarks on the Poisson Process*, Stud. Sci. Math. Hung. 2, 119 - 123

IV. Ruintheorie

$$P(T_n \leq t) = \int_0^t \underbrace{\frac{\lambda^n}{(n-1)!} x^{n-1}}_{=u(x)} \underbrace{e^{-\lambda x}}_{=v(x)} dx = u(x)v(x)|_0^t + \int_0^t \frac{\lambda^{n+1}}{n!} x^n e^{-\lambda x} dx = \frac{(\lambda t)^n}{n!} e^{-\lambda t} + P(T_{n+1} \leq t)$$

für $n \in \mathbb{N}$ und

$$P(N_t = 0) = P(t < T_1) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t},$$

was gerade bedeutet, dass $P^{N_t} = \mathcal{P}(\lambda t)$ für alle $t \geq 0$ gilt. Dies zeigt den sich auf homogene Poisson-Prozesse beziehenden Teil des Satzes.

Auf Grund der Eigenschaften einer Intensitätsfunktion ist auch $\{N_{m(t)} | t \geq 0\}$ ein einfacher Zählprozess, mit $P^{N_{m(t)}} = \mathcal{P}(m(t))$ für alle $t \geq 0$, wenn $\{N_t | t \geq 0\}$ ein homogener Poisson-Prozess mit Parameter 1 ist. Die restliche Aussage folgt aus der Tatsache, dass mit m auch m^{-1} eine Intensitätsfunktion ist. ■

Die Idee der Umkehrung in dem genannten Satz von Rényi besteht darin, dass fast sicher jeder Pfad eines Zählprozesses eine maßerzeugende Funktion bildet, d.h. durch die Setzung

$$N_{(a,b]}(\omega) := N_b(\omega) - N_a(\omega) \text{ für } \omega \in \Omega \text{ und } 0 \leq a < b$$

wird pfadweise eine additive und σ -endliche Mengenfunktion auf den Standard-Intervallen $(a, b]$ mit $0 \leq a < b$ definiert, die über den Maßfortsetzungssatz zu einem (zufälligen) Maß ξ auf der Spur- σ -Algebra $\mathbb{R}^+ \cap \mathcal{B}^1$ fortgesetzt werden kann. Dieses (zufällige) Maß hat dann die Eigenschaft, dass für jede Folge paarweise disjunkter Mengen $\{B_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^+ \cap \mathcal{B}^1$ die Folge der Zufallsvariablen $\{\xi(B_n)\}_{n \in \mathbb{N}}$ stochastisch unabhängig ist, woraus sich auch die angegebene Struktur der Sprungzeiten ergibt. (Für einen etwas allgemeineren Zugang zu diesem Themenkreis vgl. etwa das Skript RÄUMLICHE STATISTIK.)

Man spricht in diesem Zusammenhang auch davon, dass ein Poisson-Prozess *unabhängige Zuwächse* besitzt, d.h. dass jede Folge $\{N_{t_{n+1}} - N_{t_n}\}_{n \in \mathbb{N}}$ für strikt monoton wachsende nicht-negative Folgen $\{t_n\}_{n \in \mathbb{N}}$ stochastisch unabhängig ist.

Abschließend soll noch bemerkt werden, dass in Verallgemeinerung von Satz 11 jeder beliebige Poisson-Prozess $\{N_t | t \geq 0\}$ mit Intensitätsfunktion m und Sprungzeitenfolge $\{T_n\}_{n \in \mathbb{N}}$ durch eine nicht-negative, strikt monotone und stetige Zeit-Transformation $z(t)$ für $t \geq 0$ mit $z(0) = 0$ in einen neuen Poisson-Prozess $\{N_t^* | t \geq 0\} = \{N_{z(t)} | t \geq 0\}$ übergeht, der die Intensitätsfunktion $m \circ z$ besitzt und dessen Sprungzeiten als $T_n^* = z^{-1}(T_n)$, $n \in \mathbb{N}$ dargestellt werden können.

Mit Hilfe von Zählprozessen kann das zeitdiskrete kollektive Modell der Risikotheorie, das wir ausführlicher in Kapitel II behandelt haben, recht einfach auf ein zeitstetiges Modell erweitert werden. Hierzu betrachten wir einen so genannten *zusammengesetzten Schadenprozess* $\{S_t | t \geq 0\}$, der folgendermaßen definiert ist:

IV. Ruintheorie

Definition 13. Es sei $\{N_t | t \geq 0\}$ ein einfacher Zählprozess und $\{X_n\}_{n \in \mathbb{N}}$ eine unabhängige Folge positiver, identisch verteilter Zufallsvariablen (Schäden), die auch von dem Zählprozess unabhängig ist¹¹. Dann wird der zusammengesetzte Schadenprozess durch

$$S_t := \sum_{k=0}^{N_t} X_k \quad \text{für } t \geq 0$$

definiert.

Der Unterschied zum klassischen kollektiven Modell der Risikotheorie besteht also lediglich darin, dass die (statische) Frequenzvariable N durch einen einfachen Zählprozess ersetzt wird. Analog zu Lemma 9 kann auch für diesen Fall für jedes $t > 0$ die jeweilige Verteilungsfunktion angegeben werden.

Lemma 34. Für die Verteilungsfunktionen F_{S_t} des zusammengesetzten Schadenprozesses gilt:

$$P(S_t \leq z) = F_{S_t}(z) = \begin{cases} p_0(t) + \sum_{n=1}^{\infty} p_n(t) F^{n*}(z), & \text{falls } z \geq 0 \\ 0, & \text{falls } z < 0 \end{cases} \quad \text{für alle } t \geq 0.$$

Dabei bezeichne wieder F^{n*} die n -fache Faltung von F , der Verteilungsfunktion der Einzelschadenshöhen und $p_n(t) = P(N_t = n)$ für alle $t \geq 0$ und $n \in \mathbb{Z}^+$.

Beweis: Folgt sofort aus Lemma 9. ■

Definition 14. Es sei $\{S_t | t \geq 0\}$ ein zusammengesetzter Schadenprozess gemäß Definition 13. Dann heißt der durch

$$\rho_t := s + H \cdot t - S_t, \quad t \geq 0$$

definierte Prozess *Risikoprozess mit der Anfangsreserve $s \geq 0$ und Prämienrate $H > 0$* .

Diesem Modell liegt die Annahme zu Grunde, dass die Prämien kontinuierlich über die Zeit eingenommen werden, mit einer Prämie der Höhe H bezogen auf eine Zeiteinheit (typischerweise 1 Jahr). Das Ereignis A eines technischen Ruins kann dann unter Verwendung der Sprungzeitenfolge $\{T_n\}_{n \in \mathbb{N}}$ des Zählprozesses beschrieben werden durch den Ausdruck

$$A = \bigcup_{t \geq 0} \{\rho_t < 0\} = \left(\bigcap_{t \geq 0} \{S_t \leq s + H \cdot t\} \right)^c$$

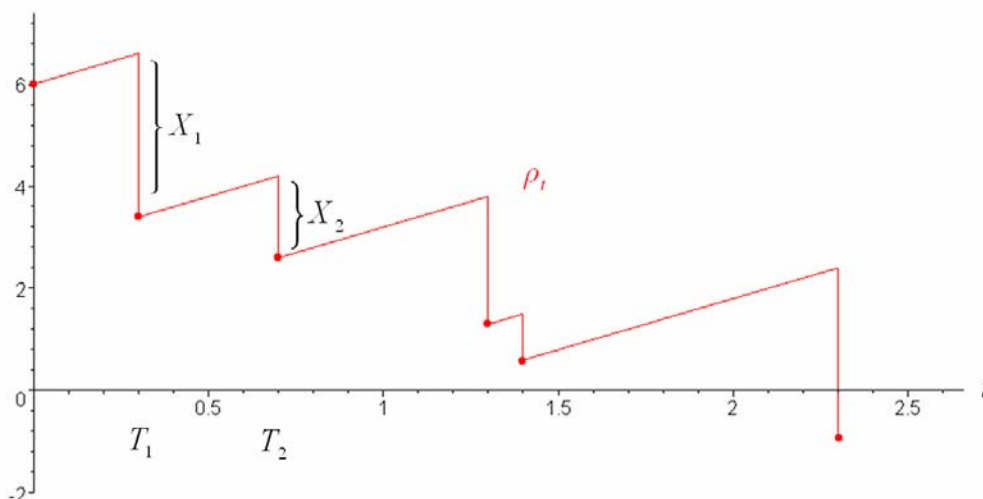
mit

$$\bigcap_{t \geq 0} \{S_t \leq s + H \cdot t\} = \bigcap_{n \in \mathbb{N}} \{S_{T_n} \leq s + H \cdot T_n\} \in \mathcal{A},$$

wobei noch $S_{T_n} = \sum_{k=1}^n X_k$ zu beachten ist, was für alle $n \in \mathbb{N}$ eine messbare Abbildung darstellt (vgl. in diesem Zusammenhang auch die Ausführungen auf S. 142 oben).

¹¹ Dies ist äquivalent dazu, dass alle Paare T_n, X_m für $n, m \in \mathbb{N}$ stochastisch unabhängig sind.

IV. Ruintheorie



Beispiel für den Verlauf eines Risikoprozesses mit Ruin; $s = 6$, $H = 2$

Eine Abschätzung oder gar Berechnung der Ruinwahrscheinlichkeit $P(A) = P\left(\bigcup_{t \geq 0} \{\rho_t < 0\}\right)$ gestaltet sich im zeitstetigen Kontext erheblich schwieriger als im diskreten Fall. Eine vergleichsweise einfache Abschätzung nach unten (vgl. Lemma 30) ergibt sich aber wie folgt.

Lemma 35. Für einen Risikoprozess im Sinne der Definition 14 gilt:

$$P(A) = P\left(\bigcup_{t \geq 0} \{\rho_t < 0\}\right) \geq \int (1 - F_X(s + H \cdot t)) P^{T_1}(dt).$$

Hierbei bezeichnet F_X die Verteilungsfunktion der Einzelschäden.

Beweis: Es gilt wegen der stochastischen Unabhängigkeit der Sprungzeiten und Einzelschäden:

$$\begin{aligned} P(A) &= P\left(\bigcup_{t \geq 0} \{\rho_t < 0\}\right) \geq P(\rho_{T_1} < 0) = P(X_1 > s + H \cdot T_1) \\ &= \int P(X_1 > s + H \cdot T_1 | T_1 = t) P^{T_1}(dt) = \int (1 - F_X(s + H \cdot t)) P^{T_1}(dt), \end{aligned}$$

d.h. $P(A)$ wird abgeschätzt durch die Wahrscheinlichkeit dafür, dass der Ruin bereits mit dem ersten Schaden eintritt. ■

Im Folgenden wollen wir in Verallgemeinerung von Lemma 35 zunächst eine Rekursionsformel analog zu Lemma 31 herleiten, aus der dann eine Satz 10 entsprechende Abschätzung der Ruinwahrscheinlichkeit nach oben folgt. Dafür werden wir formal auch den Fall einer negativen Anfangsreserve zulassen. Zusätzlich vereinbaren wir noch

$$T_0 := 0, S_0 := 0.$$

Für eine negative Anfangsreserve folgt dann wegen $\{\rho_0 < 0\}$: $P(A) = P\left(\bigcup_{t \geq 0} \{\rho_t < 0\}\right) = 1$.

IV. Ruintheorie

Lemma 36. Es sei $\{\rho_t | t \geq 0\}$ ein Risikoprozess mit der Anfangsreserve $s \in \mathbb{R}$ und Prämienrate $H > 0$. Der zu Grunde liegende Zählprozess sei ein homogener Poisson-Prozess mit Parameter $\lambda > 0$. Ferner bezeichne für $m \in \mathbb{Z}^+$

$$V_m(s) := 1 - P\left(\bigcap_{n=0}^m \{S_{T_n} \leq s + H \cdot T_n\}\right)$$

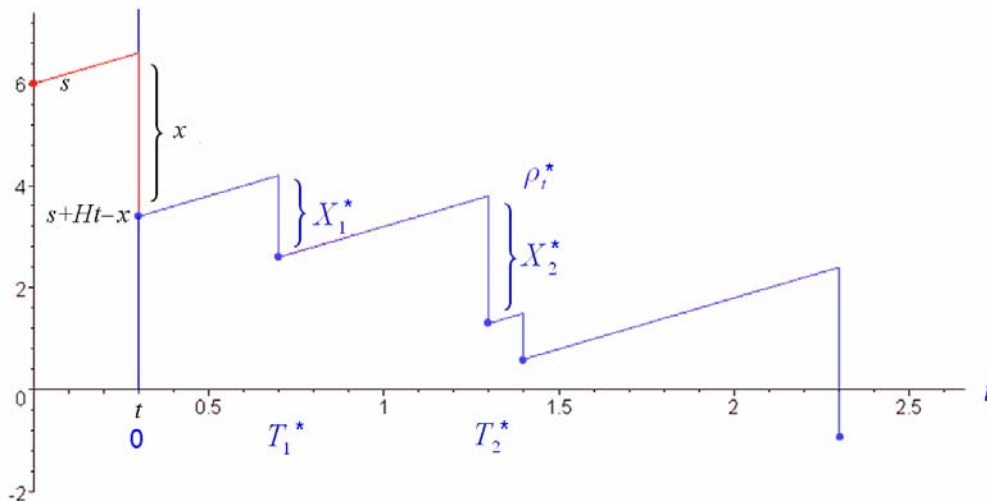
die Wahrscheinlichkeit dafür, dass ein Ruin innerhalb der Sprungzeiten T_0, \dots, T_m eintritt. Dann gilt:

$$V_0(s) = \mathbb{1}_{(-\infty, 0)}(s)$$

sowie

$$V_{m+1}(s) = \int_0^\infty \lambda e^{-\lambda t} \int_0^\infty V_m(s + H \cdot t - x) P^X(dx) dt \text{ für alle } m \in \mathbb{N}.$$

Beweis: Die erste Aussage ist klar, da ein Ruin mit Wahrscheinlichkeit 1 zur Zeit 0 genau dann vorliegt, wenn $s < 0$ ist. Die zweite Aussage ergibt sich aus einem Argument, das wir hier bildlich veranschaulichen (vgl. die Graphik auf S. 146):



Wir betrachten den in blau dargestellten Risikoprozess $\{\rho_t^* | t \geq 0\}$, der sich reskaliert bedingt ergibt gegeben $T_1 = t$ und $X_1 = x$. Dieser Prozess verhält sich stochastisch genau so wie der Risikoprozess $\{\rho_t | t \geq 0\}$, bei dem sich lediglich die Anfangsreserve zu $s + H \cdot t - x$ ändert. Daher kann der Ausdruck $V_m(s + H \cdot t - x)$ interpretiert werden als die (bedingte) Wahrscheinlichkeit dafür, dass in dem reskalierten Risikoprozess $\{\rho_t^* | t \geq 0\}$ ein Ruin innerhalb der Sprungzeiten T_0^*, \dots, T_m^* eintritt, woraus durch Integration nach der gemeinsamen Verteilung von X_1 und T_1 die angegebene Rekursionsformel folgt. ■

IV. Ruintheorie

Satz 12. Es sei $\{\rho_t | t \geq 0\}$ ein Risikoprozess mit der Anfangsreserve $s \geq 0$ und Prämienrate $H > 0$. Der zu Grunde liegende Zählprozess sei ein homogener Poisson-Prozess mit Parameter $\lambda > 0$. Dann gilt im Fall $H > \lambda \cdot E(X)$:

$$P(A) = P\left(\bigcup_{t \geq 0} \{\rho_t < 0\}\right) \leq e^{-Rs},$$

wobei R den Cramér-Lundberg-Anpassungskoeffizienten aus Lemma 33 bezeichnet.

Beweis: Wir zeigen durch vollständige Induktion, dass mit der Notation von Lemma 36 gilt:

$$V_m(s) \leq e^{-Rs} \text{ für alle } m \in \mathbb{Z}^+,$$

woraus die Aussage dann durch den Grenzübergang $m \rightarrow \infty$ folgt.

Für $m = 0$ gilt $V_0(s) = \mathbb{1}_{(-\infty, 0)}(s) \leq e^{-Rs}$ für alle $s \in \mathbb{R}$.

Es gelte nun die Ungleichung $V_m(s) \leq e^{-Rs}$ für ein $m \in \mathbb{Z}^+$ für $s \in \mathbb{R}$. Dann folgt nach Lemma 36 mit dem Satz von Fubini

$$\begin{aligned} V_{m+1}(s) &= \int_0^\infty \lambda e^{-\lambda t} \int_0^\infty V_m(s + H \cdot t - x) P^X(dx) dt \leq \int_0^\infty \lambda e^{-\lambda t} \int_0^\infty e^{-R(s+H \cdot t - x)} P^X(dx) dt \\ &= e^{-Rs} \int_0^\infty \int_0^\infty \lambda e^{-(\lambda+RH)t} dt e^{Rx} P^X(dx) = e^{-Rs} \frac{\lambda}{\lambda + RH} \psi_X(R) = e^{-Rs}, \end{aligned}$$

d.h. die Aussage gilt auch für $m + 1$. Der Satz ist damit bewiesen. ■

Für den Spezialfall, dass die Einzelschäden exponentialverteilt sind, kann man die Ruinwahrscheinlichkeit sogar exakt bestimmen. Hier gilt:

Satz 13. Unter den Voraussetzungen von Satz 12 gilt für $P^X = \mathcal{E}(\mu)$ mit $\mu > 0$, falls $H\mu > \lambda$:

$$P(A) = P\left(\bigcup_{t \geq 0} \{\rho_t < 0\}\right) = \frac{\lambda}{H\mu} \exp\left(-\left\{\mu - \frac{\lambda}{H}\right\}s\right).$$

Für den Beweis dieses Satzes sowie weiterführende Literatur verweisen wir auf HEILMANN (1987), Satz 7.3 oder GERBER (1979).

V Verallgemeinerte lineare Modelle

Die klassischen Linearen Modelle sind in der mathematischen Statistik ein langbewährtes Werkzeug im Bereich der Varianz- und (linearen) Regressionsanalyse (vgl. hierzu etwa PRUSCHA (2000), Kapitel III und das Skript zur STOCHASTIK, Kapitel III.3). Ihre nichtlinearen Verallgemeinerungen bieten demgegenüber eine größere Flexibilität und werden seit etwa 1990 verstärkt in der Versicherungsmathematik, insbesondere in der Kfz-Versicherung, aber auch anderen Sparten eingesetzt (vgl. etwa DE JONG UND HELLER (2008) oder OHLSSON UND JOHANSSON (2010)).

V. 1. Exponentialfamilien

Eine wesentliche statistische Grundlage verallgemeinerter linearer Modelle bilden Exponentialfamilien von Risikoverteilungen bezüglich eines dominierenden Maßes.

Definition 15 (eiparametrische Exponentialfamilie). Es sei μ ein σ -endliches Maß auf dem Messraum $(\mathcal{X}, \mathcal{B})$ und $\mathcal{Q} := \{Q_\gamma | \gamma \in \Gamma\}$ eine Familie von durch μ dominierten Verteilungen auf \mathcal{B} mit einer nicht-leeren Parametermenge Γ . \mathcal{Q} heißt *eiparametrische Exponentialfamilie* (in den Abbildungen $c(\cdot)$ und $t(\cdot)$), falls die zugehörigen μ -Dichten f_γ die Form

$$f_\gamma(x) = \exp(c(\gamma) \cdot t(x) + a(x) - b(\gamma)) \text{ für } x \in \mathcal{X} \text{ und } \gamma \in \Gamma$$

besitzen, mit geeigneten meßbaren reellwertigen Abbildungen $a(\cdot)$ und $t(\cdot)$ auf \mathcal{X} und geeigneten reellwertigen Abbildungen $b(\cdot)$ und $c(\cdot)$ auf Γ .

Bemerkung: Setzt man $h(x) := e^{a(x)}$ für $x \in \mathcal{X}$ und $d(\gamma) := e^{-b(\gamma)}$ für $\gamma \in \Gamma$, so geht die obige Darstellung über in

$$f_\gamma(x) = d(\gamma) \cdot \exp(c(\gamma) \cdot t(x)) \cdot h(x) \text{ für } x \in \mathcal{X} \text{ und } \gamma \in \Gamma.$$

Die Abbildungen $d(\cdot)$ und $h(\cdot)$ (und damit auch die Abbildungen $a(\cdot)$ und $b(\cdot)$) spielen also nur die Rolle von Normierungsfaktoren; definiert man nämlich das zu μ äquivalente Maß μ^* durch

$$\mu^*(B) := \int_B h(x) \mu(dx) \text{ für } B \in \mathcal{B},$$

so besitzt Q_γ offensichtlich die μ^* -Dichte

$$f_\gamma^*(x) = d(\gamma) \cdot \exp(c(\gamma) \cdot t(x)) \text{ für } x \in \mathcal{X} \text{ und } \gamma \in \Gamma.$$

Ferner gilt noch wegen der Dichte-Eigenschaft

$$e^{b(\gamma)} = \frac{1}{d(\gamma)} = \int \exp(c(\gamma) \cdot t(x)) \mu^*(dx) \text{ für } \gamma \in \Gamma.$$

V. Verallgemeinerte lineare Modelle

In konkreten risikotheorietischen Anwendungen ist in der Regel $\mathcal{X} \subseteq \mathbb{R}^d$ und $\mathcal{B} = \mathcal{X} \cap \mathcal{B}^d$ mit der Borel'schen σ -Algebra \mathcal{B}^d (mit der Dimension $d \in \mathbb{N}$); das Maß μ ist dann entweder das Lebesgue-Maß oder das (auf eine geeignete Teilmenge von \mathbb{Z}^d konzentrierte) abzählende Maß.

Gilt $\mathcal{X} = \mathbb{R}$ und $t(x) = x$ für $x \in \mathbb{R}$, so spricht man auch von einer *Exponentialfamilie in kanonischer Form*. Der reellwertige Parameter $\vartheta := c(\gamma)$ heißt auch *natürlicher Parameter* (der Exponentialfamilie), die Menge

$$\Theta := \left\{ \vartheta \in \mathbb{R} \mid \int \exp(\vartheta \cdot t(x)) \cdot h(x) \mu(dx) < \infty \right\}$$

der zugehörige *natürliche Parameterraum*. Prinzipiell kann man also eine Exponentialfamilie sogleich in der natürlichen Parametrisierung betrachten; in diesem Fall ist dann $\Gamma = \Theta$ und $c(\vartheta) = \vartheta$ für $\vartheta \in \Theta$.

Lemma 37. Die folgenden diskreten Verteilungen (der Panjer-Klasse) bilden kanonische einparametrische Exponentialfamilien bezüglich des abzählenden Maßes mit den Parametern / Parameterfunktionen

Verteilung	Γ	ϑ	Θ	γ	$t(x)$	$h(x)$	$b(\vartheta)$
$B(n, \gamma)$	$(0, 1)$	$\ln\left(\frac{\gamma}{1-\gamma}\right)$	\mathbb{R}	$\frac{e^\vartheta}{1+e^\vartheta}$	x	$\binom{n}{x}, x \in \{0, \dots, n\}$	$n \ln(1+e^\vartheta)$
$NB(\beta, \gamma)$	$(0, 1)$	$\ln(1-\gamma)$	$(-\infty, 0)$	$1-e^\vartheta$	x	$\binom{\beta+x-1}{x}, x \in \mathbb{Z}^+$	$-\beta \ln(1-e^\vartheta)$
$\mathcal{P}(\gamma)$	$(0, \infty)$	$\ln \gamma$	\mathbb{R}	e^ϑ	x	$\frac{1}{x!}, x \in \mathbb{Z}^+$	e^ϑ

mit festen $n \in \mathbb{N}$ und $\beta > 0$.

Beweis: Für $x \in \mathbb{Z}^+$ gilt:

Binomialverteilung:

$$\begin{aligned} f_\gamma(x) &= \binom{n}{x} \gamma^x (1-\gamma)^{n-x} = \exp(x \ln \gamma + (n-x) \ln(1-\gamma)) \cdot h(x) \\ &= \exp(\vartheta \cdot x + n \cdot \ln(1-\gamma)) \cdot h(x) = \exp(\vartheta \cdot x - b(\vartheta)) \cdot h(x) \end{aligned}$$

mit der Einschränkung $x \in \{0, \dots, n\}$.

Negative Binomialverteilung:

$$\begin{aligned} f_\gamma(x) &= \binom{\beta+x-1}{x} \gamma^\beta (1-\gamma)^x = \exp(x \ln(1-\gamma) + \beta \ln \gamma) \cdot h(x) \\ &= \exp(\vartheta \cdot x + \beta \ln \gamma) \cdot h(x) = \exp(\vartheta \cdot x - b(\vartheta)) \cdot h(x) \end{aligned}$$

V. Verallgemeinerte lineare Modelle

Poisson-Verteilung:

$$f_\gamma(x) = e^{-\gamma} \frac{\gamma^x}{x!} = \exp(x \ln \gamma - \gamma) \cdot h(x) = \exp(\vartheta \cdot x - \gamma) \cdot h(x) = \exp(\vartheta \cdot x - b(\vartheta)) \cdot h(x).$$

Lemma 38. Die folgenden stetigen Verteilungen bilden einparametrische Exponentialfamilien bezüglich des Lebesgue-Maßes mit den Parametern / Parameterfunktionen

Verteilung	Γ	ϑ	Θ	γ	$t(x)$	$a(x)$	$b(\vartheta)$
$\mathcal{E}(\gamma)$	$(0, \infty)$	$-\gamma$	$(-\infty, 0)$	$-\vartheta$	x	0	$-\ln(-\vartheta)$
$\mathcal{Pa}(\gamma)$	$(0, \infty)$	$-\gamma$	$(-\infty, 0)$	$-\vartheta$	$\ln(1+x)$	$-\ln(1+x)$	$-\ln(-\vartheta)$

Beweis: Für $x > 0$ gilt:

Exponentialverteilung:

$$f_\gamma(x) = \gamma e^{-\gamma x} = \exp(-\gamma \cdot x + \ln \gamma) = \exp(\vartheta \cdot x + \ln \gamma) = \exp(\vartheta \cdot x - b(\vartheta))$$

Paretoverteilung:

$$\begin{aligned} f_\gamma(x) &= \frac{\gamma}{(1+x)^{\gamma+1}} = \exp(-(\gamma+1) \ln(1+x) + \ln \gamma) \\ &= \exp(\vartheta \cdot \ln(1+x) - \ln(1+x) + \ln \gamma) = \exp(\vartheta \cdot t(x) + a(x) - b(\vartheta)). \end{aligned}$$

Bemerkung: Die Familie der Exponentialverteilungen bildet eine kanonische Exponentialfamilie, die der Paretoverteilungen dagegen nicht.

Eine naheliegende mehrdimensionale Verallgemeinerung enthält die folgende

Definition 16 (*k*-parametrische Exponentialfamilie). Es sei μ ein σ -endliches Maß auf dem Messraum $(\mathcal{X}, \mathcal{B})$ und $\mathfrak{Q} := \{Q_\gamma \mid \gamma \in \Gamma\}$ eine Familie von durch μ dominierten Verteilungen auf \mathcal{B} mit einer nicht-leeren Parametermenge Γ . Ferner sei $k \in \mathbb{N}$. \mathfrak{Q} heißt *k*-parametrische Exponentialfamilie (in den Abbildungen $\mathbf{c}(\bullet) = (c_1(\bullet), \dots, c_k(\bullet))^T$ und $\mathbf{t}(\bullet) = (t_1(\bullet), \dots, t_k(\bullet))^T$), falls die zugehörigen μ -Dichten f_γ die Form

$$f_\gamma(x) = \exp(\mathbf{c}(\gamma)^T \mathbf{t}(x) + a(x) - b(\gamma)) \text{ für } x \in \mathcal{X} \text{ und } \gamma \in \Gamma$$

besitzen, mit geeigneten meßbaren vektorwertigen Abbildungen $\mathbf{c}(\bullet)$ und $\mathbf{t}(\bullet)$ auf Γ bzw. auf \mathcal{X} und geeigneten reellwertigen Abbildungen $b(\bullet)$ und $a(\bullet)$ auf Γ bzw. auf \mathcal{X} , die zusätzlich folgende Anforderungen erfüllen:

- Die Abbildungen $1, c_1(\bullet), \dots, c_k(\bullet)$ sind linear unabhängig
- Die Abbildungen $1, t_1(\bullet), \dots, t_k(\bullet)$ sind μ -fast überall, d.h. eingeschränkt auf das Komplement einer beliebigen μ -Nullmenge, linear unabhängig.

V. Verallgemeinerte lineare Modelle

Bemerkung: Diese Forderungen stellen sicher, dass ein Zufallsvektor $\mathbf{X} = (X_1, \dots, X_k)$ mit der Verteilung \mathcal{Q}_γ , $\gamma \in \Gamma$ eine positiv-definite Varianz-Kovarianzmatrix besitzt.

Der k -dimensionale Parameter $\vartheta := \mathbf{c}(\gamma)$ heißt auch hier wieder *natürlicher Parameter* (der Exponentialfamilie), die Menge

$$\Theta := \left\{ \vartheta \in \mathbb{R}^k \mid \int \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) < \infty \right\}$$

wieder der zugehörige *natürliche Parameterraum*. Prinzipiell kann man also eine k -parametrische Exponentialfamilie auch gleich in der natürlichen Parametrisierung betrachten; in diesem Fall ist wieder $\Gamma = \Theta$ und $c(\vartheta) = \vartheta$ für $\vartheta \in \Theta$.

Lemma 39. Der natürliche Parameterraum Θ einer k -parametrischen Exponentialfamilie ist konvex und enthält ein nicht-ausgeartetes k -dimensionales Intervall.

Beweis: Seien $\vartheta_1, \vartheta_2 \in \Theta$ und $\alpha \in (0, 1)$. Setze $\vartheta = \alpha\vartheta_1 + (1-\alpha)\vartheta_2$. Die Exponentialfunktion ist konvex, also gilt

$$(e^x)^\alpha (e^y)^{1-\alpha} = \exp(\alpha x + (1-\alpha)y) \leq \alpha e^x + (1-\alpha)e^y \leq e^x + e^y \text{ für alle } x, y \in \mathbb{R}.$$

Damit ergibt sich

$$\exp(\vartheta^T \mathbf{t}(x)) = \exp(\vartheta_1^T \mathbf{t}(x))^\alpha \cdot \exp(\vartheta_2^T \mathbf{t}(x))^{1-\alpha} \leq \exp(\vartheta_1^T \mathbf{t}(x)) + \exp(\vartheta_2^T \mathbf{t}(x)) \text{ für alle } x \in \mathcal{X}$$

und somit

$$\int \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) \leq \int \exp(\vartheta_1^T \mathbf{t}(x)) \cdot h(x) \mu(dx) + \int \exp(\vartheta_2^T \mathbf{t}(x)) \cdot h(x) \mu(dx) < \infty,$$

d.h. auch $\vartheta \in \Theta$. Damit ist Θ konvex.

Zum Nachweis der zweiten Aussage nehmen wir gegenteilig an, dass Θ in einem $(k-1)$ -dimensionalen Teilraum von \mathbb{R}^k liegt. Es gibt dann $d_0 \in \mathbb{R}$, $\mathbf{d} \in \mathbb{R}^k$ mit $\mathbf{d} \neq \mathbf{0}$ und $d_0 + \mathbf{d}^T \vartheta = 0$ für alle $\vartheta \in \Theta$, im Widerspruch zur ersten Anforderung oben.

Damit ist alles bewiesen.

Lemma 40. Es sei $\mathfrak{Q} = \{\mathcal{Q}_\gamma \mid \gamma \in \Gamma\}$ eine k -parametrische Exponentialfamilie mit Dichten

$$f_\gamma(x) = \exp(\mathbf{c}(\gamma)^T \mathbf{t}(x) + a(x) - b(\gamma)) \text{ für } x \in \mathcal{X} \text{ und } \gamma \in \Gamma.$$

Dann ist die Familie der Produktverteilungen $\mathfrak{Q} := \left\{ \bigotimes_{i=1}^n \mathcal{Q}_\gamma \mid \gamma \in \Gamma \right\}$ für $n \in \mathbb{N}$ ebenfalls eine k -

parametrische Exponentialfamilie auf dem Produktraum $(\mathcal{X}, \mathcal{B}) = \left(\prod_{i=1}^n \mathcal{X}, \bigotimes_{i=1}^n \mathcal{B} \right)$ mit dem dominieren-

den Produkt-Maß $\bigotimes_{i=1}^n \mu$ und der Dichte

V. Verallgemeinerte lineare Modelle

$$f_\gamma(\mathbf{x}) = \exp(\mathbf{c}(\gamma)^T \mathbf{T}(\mathbf{x}) + A(\mathbf{x}) - B(\gamma))$$

für $\mathbf{x} = (x_1, \dots, x_n) \in \prod_{i=1}^n \mathcal{X}$ und $\gamma \in \Gamma$, wobei

$$\mathbf{T}(\mathbf{x}) = \sum_{i=1}^n \mathbf{t}(x_i) \in \mathbb{R}^k, \quad A(\mathbf{x}) = \sum_{i=1}^n a(x_i), \quad B(\gamma) = n \cdot b(\gamma).$$

Beweis: Für $\gamma \in \Gamma$ besitzt das Maß $\bigotimes_{i=1}^n \mathcal{Q}_\gamma$ die Produktdichte

$$\begin{aligned} f_\gamma(\mathbf{x}) &= \prod_{i=1}^n \exp(\mathbf{c}(\gamma)^T \mathbf{t}(x_i) + a(x_i) - b(\gamma)) = \exp\left(\sum_{i=1}^n \{\mathbf{c}(\gamma)^T \mathbf{t}(x_i) + a(x_i) - b(\gamma)\}\right) \\ &= \exp(\mathbf{c}(\gamma)^T \mathbf{T}(\mathbf{x}) + A(\mathbf{x}) - n \cdot b(\gamma)) = \exp(\mathbf{c}(\gamma)^T \mathbf{T}(\mathbf{x}) + A(\mathbf{x}) - B(\gamma)) \end{aligned}$$

für $\mathbf{x} = (x_1, \dots, x_n) \in \prod_{i=1}^n \mathcal{X}$. Man beachte dabei, dass die Komponenten von $\mathbf{T} \bigotimes_{i=1}^n \mu$ -fast überall linear unabhängig bleiben.

Bemerkung: Lemma 40 besagt in anderen Worten, dass stochastisch unabhängige Versuchswiederholungen die Eigenschaft einer Exponentialfamilie "vererben". Man kann das Resultat sogar verallgemeinern (vgl. etwa PRUSCHA (2000), Abschnitt 2.2), wie das folgende Beispiel zeigt.

Beispiel 34. Es sei $\mathcal{Q}_1 := \{\mathcal{P}(\gamma_1) | \gamma_1 > 0\}$ die (Exponential-)Familie der Poissonverteilungen und $\mathcal{Q}_2 := \{\mathcal{E}(\gamma_2) | \gamma_2 > 0\}$ die (Exponential-)Familie der Exponentialverteilungen. Mit dem Parameter $\gamma = (\gamma_1, \gamma_2) \in \Gamma := (0, \infty) \times (0, \infty)$ ergibt sich als Produktfamilie die neue Exponentialfamilie $\mathcal{Q} := \{\mathcal{P}(\gamma_1) \otimes \mathcal{E}(\gamma_2) | \gamma_1, \gamma_2 > 0\}$ mit der $\# \otimes m^1$ -Dichte

$$f_\gamma(\mathbf{x}) = \gamma_2 e^{-\gamma_1 - \gamma_2 x_2} \frac{\gamma_1^{x_1}}{x_1!} \cdot \mathbb{1}_{\mathbb{Z}^+}(x_1) \mathbb{1}_{(0, \infty)}(x_2) \quad \text{für } \mathbf{x} = (x_1, x_2) \in \mathbb{Z}^+ \times (0, \infty).$$

Hier ist $\mathbf{C}(\gamma) = (\ln \gamma_1, -\gamma_2)^T$, $\mathbf{T}(\mathbf{x}) = (x_1, x_2)^T$, $A(\mathbf{x}) = -\ln(x_1!)$, $B(\gamma) = \gamma_1 - \ln \gamma_2$.

Lemma 41. Die folgenden stetigen Verteilungen bilden zweiparametrische Exponentialfamilien bezüglich des Lebesgue-Maßes mit den Parametern / Parameterfunktionen

Verteilung	Γ	ϑ	Θ	$\mathbf{t}(x)$	$a(x)$	$b(\vartheta)$
$\mathcal{N}(\gamma_1, \gamma_2^2)$	$\mathbb{R} \times (0, \infty)$	$\begin{pmatrix} \frac{\gamma_1}{\gamma_2^2} \\ 1 \\ -\frac{1}{2\gamma_2^2} \end{pmatrix}$	$\mathbb{R} \times (-\infty, 0)$	$\begin{pmatrix} x \\ x^2 \end{pmatrix}$	0	$-\frac{\vartheta^2}{4\vartheta_2} - \frac{1}{2} \ln\left(-\frac{\vartheta_2}{\pi}\right)$
$\Gamma(\gamma_1, \gamma_2)$	$(0, \infty)^2$	$\begin{pmatrix} -\gamma_1 \\ \gamma_2 \end{pmatrix}$	$(-\infty, 0) \times (0, \infty)$	$\begin{pmatrix} x \\ \ln x \end{pmatrix}$	$-\ln x$	$-\vartheta_2 \ln(-\vartheta_1) + \ln \Gamma(\vartheta_2)$

V. Verallgemeinerte lineare Modelle

Beweis: Für $x > 0$ gilt:

Normalverteilung: Hier ist $\gamma = \left(-\frac{\vartheta_1}{2\vartheta_2}, \frac{1}{\sqrt{-2\vartheta_2}} \right)$ und

$$\begin{aligned} f_\gamma(x) &= \frac{1}{\gamma_2 \sqrt{2\pi}} \exp\left(-\frac{(x-\gamma_1)^2}{2\gamma_2^2}\right) = \exp\left(\frac{\gamma_1}{\gamma_2}x - \frac{1}{2\gamma_2^2}x^2 - \frac{\gamma_1^2}{2\gamma_2^2} - \frac{1}{2}\ln(2\pi\gamma_2^2)\right) \\ &= \exp\left(\vartheta_1 x + \vartheta_2 x^2 + \frac{\vartheta_1^2}{4\vartheta_2} + \frac{1}{2}\ln\left(-\frac{\vartheta_2}{\pi}\right)\right) = \exp(\vartheta^T \mathbf{t}(x) - b(\vartheta)) \end{aligned}$$

Gammaverteilung:

$$\begin{aligned} f_\gamma(x) &= \frac{(\gamma_1)^{\gamma_2}}{\Gamma(\gamma_2)} x^{\gamma_2-1} e^{-\gamma_1 x} = \exp(-\gamma_1 x + \gamma_2 \ln x - \ln x + \gamma_2 \ln \gamma_1 - \ln \Gamma(\gamma_2)) \\ &= \exp(\vartheta_1 x + \vartheta_2 \ln x - \ln x + \gamma_2 \ln \gamma_1 - \ln \Gamma(\gamma_2)) = \exp(\vartheta^T \mathbf{t}(x) + a(x) - b(\vartheta)). \end{aligned}$$

Ein großer Vorteil von Exponentialfamilien in natürlicher Parametrisierung besteht darin, dass gewisse Momente von Zufallsvariablen bzw. -vektoren mit einer solchen Verteilung sehr leicht durch Differentiation der Abbildung $b(\bullet)$ gewonnen werden können. Zur Vorbereitung benötigen wir dazu folgendes Resultat.

Lemma 42. Gegeben sei eine k -parametrische Exponentialfamilie in natürlicher Parametrisierung, mit μ -Dichten der Form

$$f_\vartheta(x) = \exp(\vartheta^T \mathbf{t}(x) + a(x) - b(\vartheta)) \text{ für } x \in \mathcal{X} \text{ und } \vartheta \in \Theta.$$

Ferner sei wieder $h = e^{a(\bullet)}$ und \mathbf{X} ein Zufallsvektor mit der Dichte f_ϑ . Ist dann G eine meßbare reellwertige Abbildung auf $(\mathcal{X}, \mathcal{B})$ mit $E_\vartheta(|G(\mathbf{X})|) < \infty$ für alle $\vartheta \in \Theta$, so gilt

$$E_\vartheta(G(\mathbf{X})) = e^{-b(\vartheta)} \int G(x) \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx),$$

und die Abbildung

$$K(\vartheta) = E_\vartheta(G(\mathbf{X})) \cdot e^{b(\vartheta)} = \int G(x) \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx), \quad \vartheta \in \overset{\circ}{\Theta}$$

ist auf dem Inneren $\overset{\circ}{\Theta}$ des Parameterraumes Θ beliebig oft partiell differenzierbar, wobei Differentiation und Integration in der Reihenfolge vertauscht werden dürfen.

Für einen Beweis verweisen wir auf PRUSCHA (2000), Abschnitt 2.3 und die dort zitierte Literatur.

Das folgende Resultat ist eine direkte Folgerung aus Lemma 42.

V. Verallgemeinerte lineare Modelle

Satz 14. Unter den Voraussetzungen von Lemma 42 gilt: Die Abbildung $b(\vartheta)$ aus der Darstellung der Dichte f_ϑ in Lemma 42 ist auf dem Inneren $\overset{\circ}{\Theta}$ des Parameterraumes Θ beliebig oft partiell differenzierbar. Ist ferner \mathbf{X} ein Zufallsvektor mit der Dichte f_ϑ , so gilt für alle $\vartheta \in \overset{\circ}{\Theta}$

$$E_\vartheta(\mathbf{t}(\mathbf{X})) = \nabla b(\vartheta), \quad V_\vartheta(\mathbf{t}(\mathbf{X})) = \nabla^2 b(\vartheta).$$

Hierbei bezeichnet ∇ den Gradienten der partiellen Ableitungen und ∇^2 die Hesse-Matrix der zweiten partiellen Ableitungen von $b(\cdot)$. $V_\vartheta(\mathbf{t}(\mathbf{X}))$ bezeichnet die Varianz-Kovarianz-Matrix des Zufallsvektors $\mathbf{t}(\mathbf{X})$.

Beweis: Der erste Teil folgt mit der Setzung $G \equiv 1$ unmittelbar aus Lemma 42. Der zweite Teil ergibt sich so:

$$\begin{aligned} E_\vartheta(\mathbf{t}(\mathbf{X})) &= e^{-b(\vartheta)} \int \mathbf{t}(x) \cdot \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) = e^{-b(\vartheta)} \int \nabla \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) \\ &= e^{-b(\vartheta)} \nabla \int \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) = e^{-b(\vartheta)} \nabla e^{b(\vartheta)} = e^{-b(\vartheta)} e^{b(\vartheta)} \nabla b(\vartheta) = \nabla b(\vartheta), \end{aligned}$$

$$\begin{aligned} E_\vartheta(\mathbf{t}(\mathbf{X}) \cdot (\mathbf{t}(\mathbf{X}))^T) &= e^{-b(\vartheta)} \int \mathbf{t}(x) \cdot \mathbf{t}(x)^T \cdot \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) \\ &= e^{-b(\vartheta)} \int \mathbf{t}(x) \cdot \exp(\vartheta^T \mathbf{t}(x)) \cdot \mathbf{t}(x)^T \cdot h(x) \mu(dx) \\ &= e^{-b(\vartheta)} \int \nabla^2 \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) = e^{-b(\vartheta)} \nabla^2 \int \exp(\vartheta^T \mathbf{t}(x)) \cdot h(x) \mu(dx) \\ &= e^{-b(\vartheta)} \nabla^2 e^{b(\vartheta)} = e^{-b(\vartheta)} \left[e^{b(\vartheta)} \nabla^2 b(\vartheta) + e^{b(\vartheta)} \nabla b(\vartheta) \{ \nabla b(\vartheta) \}^T \right] \\ &= \nabla^2 b(\vartheta) + \nabla b(\vartheta) \{ \nabla b(\vartheta) \}^T = \nabla^2 b(\vartheta) + E_\vartheta(\mathbf{t}(\mathbf{X})) \{ E_\vartheta(\mathbf{t}(\mathbf{X})) \}^T, \\ V_\vartheta(\mathbf{t}(\mathbf{X})) &= E_\vartheta(\mathbf{t}(\mathbf{X}) \cdot (\mathbf{t}(\mathbf{X}))^T) - E_\vartheta(\mathbf{t}(\mathbf{X})) \{ E_\vartheta(\mathbf{t}(\mathbf{X})) \}^T = \nabla^2 b(\vartheta). \end{aligned}$$

Zur Veranschaulichung von Satz 14 dienen die folgenden Beispiele.

Beispiel 35 (Fortsetzung von Lemma 37). Für die folgenden diskreten Verteilungen (der Panjer-Klasse) gilt nach Satz 14, unter Beachtung von $t(x) = x$:

P^X	ϑ	γ	$b(\vartheta)$	$b'(\vartheta)$	$b''(\vartheta)$	$E_\vartheta(X)$	$Var_\vartheta(X)$
$B(n, \gamma)$	$\ln\left(\frac{\gamma}{1-\gamma}\right)$	$\frac{e^\vartheta}{1+e^\vartheta}$	$n \ln(1+e^\vartheta)$	$n \frac{e^\vartheta}{1+e^\vartheta}$	$n \frac{e^\vartheta}{(1+e^\vartheta)^2}$	$b'(\vartheta) = n\gamma$	$b''(\vartheta) = n\gamma(1-\gamma)$
$NB(\beta, \gamma)$	$\ln(1-\gamma)$	$1-e^\vartheta$	$-\beta \ln(1-e^\vartheta)$	$\beta \frac{e^\vartheta}{1-e^\vartheta}$	$\beta \frac{e^\vartheta}{(1-e^\vartheta)^2}$	$b'(\vartheta) = \beta \frac{1-\gamma}{\gamma}$	$b''(\vartheta) = \beta \frac{1-\gamma}{\gamma^2}$
$\mathcal{P}(\gamma)$	$\ln \gamma$	e^ϑ	e^ϑ	e^ϑ	e^ϑ	$b'(\vartheta) = \gamma$	$b''(\vartheta) = \gamma$

mit festen $n \in \mathbb{N}$ und $\beta > 0$.

V. Verallgemeinerte lineare Modelle

Beispiel 36 (Fortsetzung von Lemma 38). Für die folgenden stetigen Verteilungen gilt nach Satz 14:

P^X	ϑ	γ	$b(\vartheta)$	$b'(\vartheta)$	$b''(\vartheta)$	$t(x)$	$E_{\vartheta}(t(X))$	$Var_{\vartheta}(t(X))$
$\mathcal{E}(\gamma)$	$-\gamma$	$-\vartheta$	$-\ln(-\vartheta)$	$-\frac{1}{\vartheta}$	$\frac{1}{\vartheta^2}$	x	$b'(\vartheta) = \frac{1}{\gamma}$	$b''(\vartheta) = \frac{1}{\gamma^2}$
$Pa(\gamma)$	$-\gamma$	$-\vartheta$	$-\ln(-\vartheta)$	$-\frac{1}{\vartheta}$	$\frac{1}{\vartheta^2}$	$\ln(1+x)$	$b'(\vartheta) = \frac{1}{\gamma}$	$b''(\vartheta) = \frac{1}{\gamma^2}$

Beispiel 37 (Fortsetzung von Lemma 41). Für die folgenden stetigen Verteilungen gilt nach Satz 14:

P^X	ϑ	$t(x)$	$b(\vartheta)$	$\nabla b(\vartheta) = E_{\vartheta}(t(\mathbf{X}))$	$\nabla^2 b(\vartheta) = V_{\vartheta}(t(\mathbf{X}))$
$\mathcal{N}(\gamma_1, \gamma_2^2)$	$\begin{pmatrix} \frac{\gamma_1}{\gamma_2^2} \\ -\frac{1}{2\gamma_2^2} \end{pmatrix}$	$\begin{pmatrix} x \\ x^2 \end{pmatrix}$	$-\frac{\vartheta_1^2}{4\vartheta_2} - \frac{1}{2} \ln\left(-\frac{\vartheta_2}{\pi}\right)$	$\begin{pmatrix} -\frac{\vartheta_1}{2\vartheta_2} \\ \frac{\vartheta_1^2}{4\vartheta_2^2} - \frac{1}{2\vartheta_2} \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_1^2 + \gamma_2^2 \end{pmatrix}$	$\begin{pmatrix} -\frac{1}{2\vartheta_2} & \frac{\vartheta_1}{2\vartheta_2^2} \\ \frac{\vartheta_1}{2\vartheta_2^2} & \frac{\vartheta_2 - \vartheta_1^2}{2\vartheta_2^3} \end{pmatrix} = \begin{pmatrix} \gamma_2^2 & 2\gamma_1\gamma_2^2 \\ 2\gamma_1\gamma_2^2 & 2\gamma_2^2(\gamma_1^2 + \gamma_2^2) \end{pmatrix}$
$\Gamma(\gamma_1, \gamma_2)$	$\begin{pmatrix} -\gamma_1 \\ \gamma_2 \end{pmatrix}$	$\begin{pmatrix} x \\ \ln x \end{pmatrix}$	$-\vartheta_2 \ln(-\vartheta_1) + \ln \Gamma(\vartheta_2)$	$\begin{pmatrix} -\frac{\vartheta_2}{\vartheta_1} \\ -\ln(-\vartheta_1) + \Psi(\vartheta_2) \end{pmatrix} = \begin{pmatrix} \frac{\gamma_2}{\gamma_1} \\ -\ln \gamma_1 + \Psi(\gamma_2) \end{pmatrix}$	$\begin{pmatrix} \frac{\vartheta_2}{\vartheta_1^2} & -\frac{1}{\vartheta_1} \\ -\frac{1}{\vartheta_1} & \Psi'(\vartheta_2) \end{pmatrix} = \begin{pmatrix} \frac{\gamma_2}{\gamma_1^2} & \frac{1}{\gamma_1} \\ \frac{1}{\gamma_1} & \Psi'(\gamma_2) \end{pmatrix}$

mit der Digamma-Funktion $\Psi(z) = \frac{d}{dz} \ln \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ und $\Psi'(z) = \frac{\Gamma''(z) \cdot \Gamma(z) - \Gamma'(z)^2}{\Gamma(z)^2}$, $z > 0$.

Bemerkung: Gelegentlich - insbesondere bei den verallgemeinerten linearen Modellen - wird der Fall betrachtet, dass in einer k -parametrischen Exponentialfamilie eine gewisse Anzahl $\tilde{k} < k$ von Parametern als fest angesehen wird. Es entsteht dann eine neue $\bar{k} = k - \tilde{k}$ -parametrische Exponentialfamilie, wie man Definition 16 sofort entnehmen kann.

V. 2. Lineare statistische Modelle

Wir orientieren uns hier zunächst an Kapitel III.3 aus dem Skript zur STOCHASTIK und beginnen mit dem "Methode der kleinsten Quadrate" genannten grundlegenden Verfahren der linearen Regression, das bereits auf Carl Friedrich Gauß (1777 – 1855) zurück geht.

Gegeben seien n Datenpaare (x_i, y_i) , $i = 1, \dots, n$ mit $n \in \mathbb{N}$, $n \geq 2$. Ferner wird hier zunächst vorausgesetzt, dass alle x_i paarweise verschieden sind. Gesucht wird eine *lineare Ausgleichsfunktion* $y = ax + b$ mit reellen Koeffizienten a, b , die die Datenpaare "möglichst gut" approximiert (so genannte *Regressionsgerade*). Als Gütemaß (Fehlermaß) legen wir die quadratische Abweichung zwischen den "Ist-Werten" y_i und den "Soll-Werten" $ax_i + b$ zu Grunde und erhalten damit folgendes Optimierungsproblem:

$$\min_{a,b} F(a,b) = \sum_{i=1}^n (ax_i + b - y_i)^2$$

Dieses Problem kann mit den klassischen Methoden der Analysis gelöst werden; dazu betrachten wir die partiellen Ableitungen und setzen diese Null (notwendige Bedingung für – relative – Extrema):

$$\begin{aligned} \frac{\partial F}{\partial a} &= 2 \sum_{i=1}^n (ax_i + b - y_i) \cdot x_i = 0 \\ \frac{\partial F}{\partial b} &= 2 \sum_{i=1}^n (ax_i + b - y_i) = 0 \end{aligned}$$

mit den Lösungen

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \left(\frac{1}{n} \sum_{i=1}^n x_i \right) \left(\frac{1}{n} \sum_{i=1}^n y_i \right)}{\frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2} = \frac{s_{xy}}{s_{xx}}, \quad \hat{b} = \frac{1}{n} \sum_{i=1}^n y_i - \hat{a} \cdot \frac{1}{n} \sum_{i=1}^n x_i = \bar{y} - \hat{a} \cdot \bar{x}.$$

Die Größen s_{xy} und s_{xx} können dabei als empirische Kovarianz bzw. empirische Varianz der Datenreihen aufgefasst werden, die Größen \bar{x} und \bar{y} entsprechen den üblichen arithmetischen Mittelwerten. Die Abbildung F ist nun aber global konvex wegen

$$\nabla^2 F = \begin{bmatrix} \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial a \partial b} \\ \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{bmatrix} = 2 \begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \quad (\text{Hesse-Matrix})$$

mit

$$\det(\nabla^2 F) = \frac{\partial^2 F}{\partial a^2} \cdot \frac{\partial^2 F}{\partial b^2} - \left(\frac{\partial^2 F}{\partial a \partial b} \right)^2 = 2n^2 \left(\frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 \right) = 2n^2 \text{Var}(X) > 0$$

und

V. Verallgemeinerte lineare Modelle

$$\frac{\partial^2 F}{\partial b^2} = n > 0,$$

wobei formal X eine diskret gleichverteilte Zufallsvariable über der Menge $\{x_1, \dots, x_n\}$ bezeichne. (Man beachte, dass wegen $n \geq 2$ der Ausdruck $\text{Var}(X) > 0$ und somit die Hesse-Matrix positiv definit ist!) Die Lösungen (Schätzwerte) \hat{a} und \hat{b} sind somit die eindeutig bestimmten globalen Minimumstellen der Funktion F .

Die obigen Rechnungen zeigen, dass das Verfahren auch dann noch funktioniert, wenn die x_i mehrfach auftreten (d.h. multiple Zuordnungen vorkommen), solange die Größe

$$s_{xx} = \frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 > 0$$

bleibt, was genau dann eintritt, wenn wenigstens zwei der x_i voneinander verschieden sind. Eine solche Situation liegt beispielsweise vor, wenn es sich bei den y_i um Wiederholte Messwerte aus mehreren Experimenten (z.B. physikalische Körperausdehnung) bei teilweise gleichen x_i (z.B. eingestellte Temperaturen) handelt.

Manchmal wird auch nur eine lineare Ausgleichsfunktion der Form $y = ax$ gesucht (man spricht dann von einer *Regression durch den Nullpunkt*). In diesem Fall vereinfachen sich die obigen Rechnungen erheblich, mit der Lösung

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i}{\frac{1}{n} \sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.$$

Die Methode der kleinsten Quadrate ist auch anwendbar, wenn nicht-lineare Zusammenhänge zwischen den x - und y -Werten unterstellt werden. Dazu bedient man sich zweckmäßigerweise geeigneter *linear unabhängiger Funktionensysteme* $\{f_k\}_{k \in \mathbb{N}}$, z.B. Polynome unterschiedlichen Grades. Das abstrakte Modell ist hier

$$y = \sum_{k=1}^K \beta_k \cdot f_k(x) \quad \text{mit } K \in \mathbb{N} \text{ und reellen Koeffizienten } \beta_1, \dots, \beta_K.$$

Das zugehörige quadratische Optimierungsproblem lautet dann sinngemäß

$$\min_{\beta_1, \dots, \beta_K} F(\beta_1, \dots, \beta_K) = \sum_{i=1}^n \left(y_i - \sum_{k=1}^K \beta_k f_k(x_i) \right)^2 = \min_{\boldsymbol{\beta}} F(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{D}\boldsymbol{\beta}\|^2$$

V. Verallgemeinerte lineare Modelle

mit dem Koeffizientenvektor $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_K \end{pmatrix}$, dem Wertevektor $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ und der $n \times K$ - **Design-Matrix**

$\mathbf{D} = [f_k(x_i)]_{i=1\dots n, k=1\dots K}$. Wegen

$$F(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{D}\boldsymbol{\beta}\|^2 = (\mathbf{y} - \mathbf{D}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{D}\boldsymbol{\beta}) = \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\beta}^T \mathbf{D}^T \mathbf{D}\boldsymbol{\beta}$$

ergibt sich für den Gradienten von F :

$$\nabla F = -2\mathbf{D}^T \mathbf{y} + 2\mathbf{D}^T \mathbf{D}\boldsymbol{\beta}.$$

Durch Nullsetzen der partiellen Ableitungen ergibt sich also das Gleichungssystem

$$\mathbf{D}^T \mathbf{D}\boldsymbol{\beta} = \mathbf{D}^T \mathbf{y}$$

mit der Lösung

$$\hat{\boldsymbol{\beta}} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{y}.$$

Man beachte, dass die Matrix $\mathbf{D}^T \mathbf{D}$ stets positiv semidefinit ist (vgl. Abschnitt II.9 im Skript zur STOCHASTIK). Die Inverse von $\mathbf{D}^T \mathbf{D}$ existiert also genau dann, wenn $\mathbf{D}^T \mathbf{D}$ positiv definit ist (d.h. alle Eigenwerte von $\mathbf{D}^T \mathbf{D}$ sind positiv). Dieser Fall liegt vor, wenn die Matrix \mathbf{D} vollen Rang hat. Ferner gilt für die Matrix der zweiten partiellen Ableitungen (Hesse-Matrix)

$$\nabla^2 F = 2\mathbf{D}^T \mathbf{D},$$

was im Fall der positiven Definitheit bedeutet, dass bei der gefundenen Lösung ein Minimum vorliegt.

Im Fall der klassischen linearen Regression ist etwa $\boldsymbol{\beta} = \begin{pmatrix} b \\ a \end{pmatrix}$ mit $\mathbf{D} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$, also

$$\mathbf{D}^T \mathbf{D} = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}$$

und

$$\mathbf{D}^T \mathbf{y} = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix},$$

woraus sich wieder die oben angegebene Lösung ergibt.

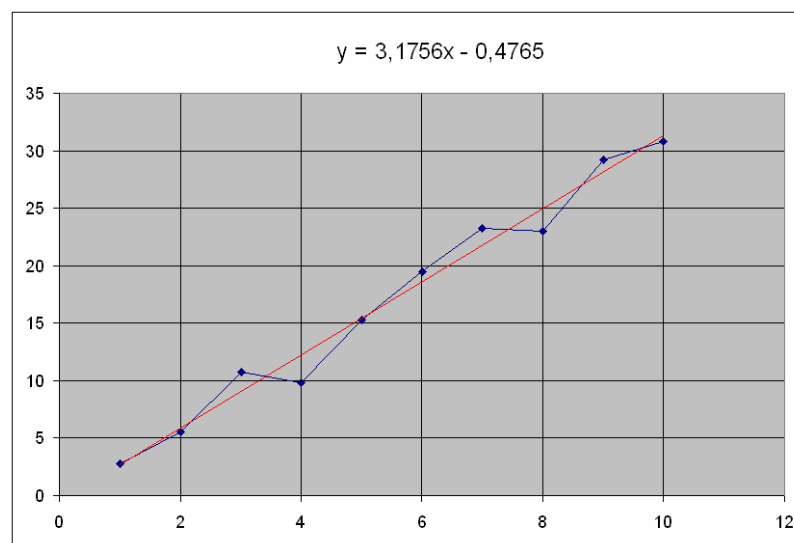
V. Verallgemeinerte lineare Modelle

Beispiel 38. An den folgenden Datensatz soll einmal eine lineare, zum anderen eine quadratische Funktion angepasst werden:

x_i	1	2	3	4	5	6	7	8	9	10
y_i	2,771	5,522	10,753	9,828	15,262	19,480	23,218	23,030	29,231	30,797

1. *Lineare Regression:* Es ist

$$s_{xy} = 26,199 \quad s_{xx} = 8,25 \quad \hat{a} = \frac{s_{xy}}{s_{xx}} = 3,17558 \quad \hat{b} = \bar{y} - \hat{a} \cdot \bar{x} = -0,47647$$



lineare Regression

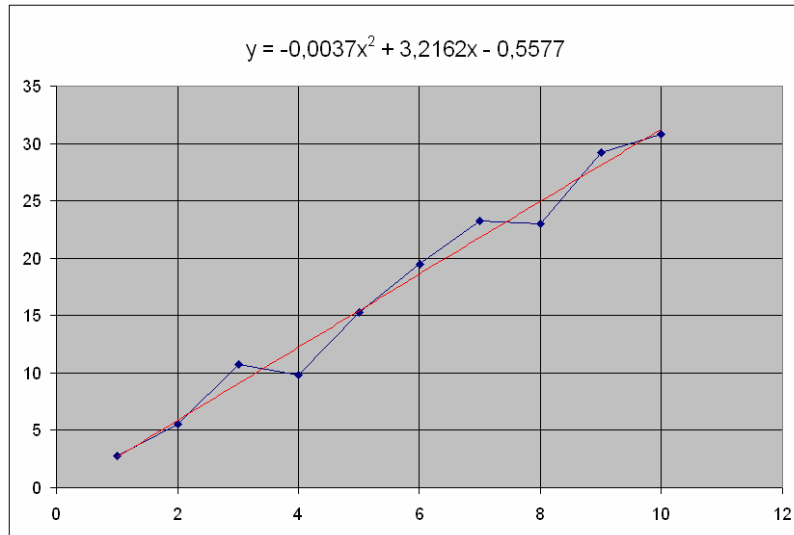
2. *Quadratische Regression:* Hier ist

$$\mathbf{D} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{10} & x_{10}^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \\ 1 & 8 & 64 \\ 1 & 9 & 81 \\ 1 & 10 & 100 \end{bmatrix}, \quad \mathbf{D}^T \mathbf{D} = \begin{bmatrix} 10 & 55 & 385 \\ 55 & 385 & 3025 \\ 385 & 3025 & 25333 \end{bmatrix},$$

V. Verallgemeinerte lineare Modelle

$$(\mathbf{D}^T \mathbf{D})^{-1} = \frac{1}{2640} \begin{bmatrix} 3651 & -1386 & 110 \\ -1386 & 637 & -55 \\ 110 & -55 & 5 \end{bmatrix}, \quad \mathbf{D}^T \mathbf{y} = \begin{pmatrix} 169,892 \\ 1196,391 \\ 9420,727 \end{pmatrix}$$

$$\boldsymbol{\beta} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{y} = \begin{pmatrix} -0,5577 \\ 3,2162 \\ -0,0037 \end{pmatrix}$$



quadratische Regression

Bisher haben wir in dem obigen Ansatz keine *zufälligen* Einflüsse explizit modelliert. Man kann dies aber leicht erreichen, wenn wir etwa annehmen, dass die Messergebnisse \mathbf{y} Realisationen eines Zufallsvektors der Form

$$\mathbf{Y} = \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

mit einem Fehlervektor (Zufallsvektor) $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$ sind, der paarweise unkorrelierte Koordinaten besitzt mit Erwartungswerten $E(\varepsilon_i) = 0$ und (gleichen) Varianzen $\text{Var}(\varepsilon_i) = \sigma^2 > 0$ für $i = 1, \dots, n$. Ist der Fehlervektor beispielsweise multivariat normalverteilt, so erhält man das Verteilungsmodell

$$P^{\mathbf{Y}} = \mathcal{N}(\mathbf{D}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$$

mit der Einheitsmatrix \mathbf{I} . In diesem Fall stimmt die angegebene Lösung des quadratischen Minimierungsproblems mit dem Maximum-Likelihood-Schätzer für den Parametervektor $\boldsymbol{\beta}$ überein; dies ist sofort aus der Form der zugehörigen Dichte ersichtlich (vgl. Abschnitt II.9 im Skript zur STOCHASTIK):

$$f(\mathbf{y}) = \frac{1}{\sqrt{2\pi\sigma^2}^n} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{D}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{D}\boldsymbol{\beta})\right) = \frac{1}{\sqrt{2\pi\sigma^2}^n} \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{y} - \mathbf{D}\boldsymbol{\beta}\|^2\right), \quad \mathbf{y} \in \mathbb{R}^n.$$

V. Verallgemeinerte lineare Modelle

Dieser Schätzer ist auch im allgemeinen Fall erwartungstreu, wie das folgende Resultat zeigt.

Lemma 43. Der Zufallsvektor \mathbf{Y} sei von der Form

$$\mathbf{Y} = \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

mit einem Fehlervektor (Zufallsvektor) $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$, dessen Koordinaten die Erwartungswerten $E(\varepsilon_i) = 0$ für $i = 1, \dots, n$ besitzen mögen, und einer Design-Matrix \mathbf{D} mit vollem Rang. Dann ist

$$\hat{\boldsymbol{\beta}} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{Y}$$

ein erwartungstreuer Schätzer für $\boldsymbol{\beta}$.

Beweis: Durch Nachrechnen erhält man

$$E\hat{\boldsymbol{\beta}} = E\left((\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{Y}\right) = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T E(\mathbf{Y}) = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{D} \boldsymbol{\beta} = \boldsymbol{\beta}.$$

Die bisherigen Betrachtungen zur linearen Regression waren im Wesentlichen an *stetig skalierten* (metrischen) Merkmalen x_i orientiert. Die Designmatrix ergab sich deshalb aus den linear unabhängigen Funktionensystemen $\{f_k\}_{k \in \mathbb{N}}$, ausgewertet an den Stellen x_i . Im *diskreten* Fall gibt es im Unterschied dazu keine metrischen erklärenden Größen. Statt dessen werden die betrachteten Zufallsvektoren \mathbf{Y} nach Kombinationen (meist) mehrerer Gruppen differenziert. Die Merkmale der Gruppierung können rein kategoriell sein (z.B. Geschlecht: *m*(ännlich), *w*(eiblich)) oder einer gewissen plausiblen Ordnung unterliegen (z.B. Altersgruppen 18 – 40, 41 – 60, über 60 Jahre).

Ziel ist es, für jede Merkmalskombination aus den Gruppen einen *Regressor* zu finden, so dass die systematische (d.h. Erwartungswert-)Komponente jeder Beobachtung die Summe der Regressoren aller Merkmalskombinationen ist. Die Design-Matrix enthält dann nur Einsen und Nullen so, dass eine Inzidenz zwischen den Beobachtungen und den Merkmalskombinationen gegeben ist.

Die Merkmalskombinationen aus den verschiedenen Gruppen werden häufig auch als „Zellen“ bezeichnet.

Zur Vereinheitlichung der Notation nehmen wir an, dass es insgesamt d unterschiedliche Gruppen gibt, und dass jede Gruppe j durch k_j Merkmale charakterisiert ist. Ferner nehmen wir zur Vereinfachung an, dass pro Merkmalskombination genau eine Beobachtung (Zufallsvariable) vorliegt.

Die Gesamtzahl n der Beobachtungen (= Anzahl der Zellen) ist dann gegeben durch $n = \prod_{i=1}^d k_i$.

Jeder betrachteten Zufallsvariablen kann unter diesen Bedingungen **in linearen Modellen ohne Wechselwirkung** ein Multi-Index (i_1, \dots, i_d) zugeordnet werden, der angibt, auf welches Merkmal i_j in der j -ten Gruppe sich die Beobachtung bezieht. Die Regressoren $\boldsymbol{\beta}$ lassen sich dabei in der Form

$$\boldsymbol{\beta} = (\beta_{11}, \dots, \beta_{1k_1}, \beta_{21}, \dots, \beta_{2k_2}, \dots, \beta_{d1}, \dots, \beta_{dk_d})^T$$

darstellen.

V. Verallgemeinerte lineare Modelle

Der unterstellte statistische lineare Zusammenhang ist in Modellen *ohne Wechselwirkung* dann gegeben durch

$$Y_{i_1, \dots, i_d} = \beta_{1i_1} + \beta_{2i_2} + \dots + \beta_{di_d} + \varepsilon_{i_1, \dots, i_d} \quad (*)$$

[für unkorrelierte Fehlervariablen $\varepsilon_{i_1, \dots, i_d}$ mit $E(\varepsilon_{i_1, \dots, i_d}) = 0$, $\text{Var}(\varepsilon_{i_1, \dots, i_d}) = \sigma^2$].

Die zugehörige Designmatrix \mathbf{D} enthält hier je Zeile genau d mal eine Eins und sonst lauter Nullen.

Modelle *mit Wechselwirkung* würden zusätzliche Regressoren beinhalten, die einen gleichzeitigen Einfluss bestimmter Zellen abbilden, z.B. mit einem (Multi-) Index in der Form β_{rki, i_k} o.ä.

In der Regel liegen pro Merkmalskombination mehrere (stochastisch unabhängige) Beobachtungen $Y_{i_1, \dots, i_d; 1}, \dots, Y_{i_1, \dots, i_d; m}$ vor. Diese Situation wird aber meist nicht separat (z.B. in einem erweiterten Kontext) modelliert.

Modelle ohne Wechselwirkung können empirisch z.B. durch *Parallelitäten* in *Mittelwertplots* identifiziert werden.

Beispiel 39. Erfasst werden sollen Krankheitskosten abhängig von der Altersgruppe (18 – 40, 41 – 60, ab 61 Jahre) und dem Geschlecht (m/w). Dies impliziert eine tabellarische Darstellung von 3 Zeilen und 2 Spalten; für $\boldsymbol{\beta}$ kann man hier zur besseren Veranschaulichung und Reduzierung des Indizierungsaufwands z.B. $\boldsymbol{\beta} = (\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2)^T$ schreiben. Dabei steht α_i für die jeweilige Altersgruppe und β_j für das Geschlecht. Mit dieser Umindizierung ergibt sich aus (*) die Darstellung

$$Y_{ij} = \alpha_i + \beta_j + \varepsilon_{ij}, \quad i = 1, \dots, 3, \quad j = 1, 2.$$

Diesem Ansatz liegt die Vorstellung zu Grunde, dass sich die systematischen Einflüsse auf die Kosten additiv aus den unterschiedlichen Alters- und Geschlechtsmerkmalen ergeben.

Die Design-Matrix hat hier demnach die Gestalt

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Dieses Modell ist ersichtlich *wechselwirkungsfrei*. Bei einem Modell *mit Wechselwirkungen* wäre folgender Ansatz zu wählen:

$$Y_{ij} = \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, 3, \quad j = 1, 2.$$

Der Vektor der Regressoren wäre hier gegeben durch

$$\boldsymbol{\beta} = (\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3, \gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23})^T$$

V. Verallgemeinerte lineare Modelle

mit einer erheblich größeren Designmatrix \mathbf{D} (die aber auch wieder nur Nullen und Einsen enthält). Im wechselwirkungsfreien Modell gilt nach obigem

$$E(Y_{ij}) = \alpha_i + \beta_j, \quad i = 1, \dots, 3, \quad j = 1, 2.$$

Liegen pro Zelle ij mehrere (unabhängige) Beobachtungen $Y_{ijk}, k = 1, \dots, m$ vor, so kann man zur Überprüfung der Annahme der Wechselwirkungsfreiheit die empirischen Mittelwerte

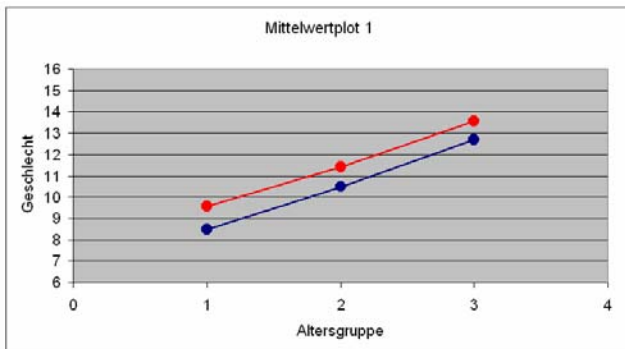
$$\bar{Y}_{ij} = \frac{1}{m} \sum_{k=1}^m Y_{ijk}, \quad i = 1, \dots, 3, \quad j = 1, 2$$

betrachten. Wegen $\bar{Y}_{ij} \approx E(Y_{ij}) = \alpha_i + \beta_j$ (nach dem Gesetz der großen Zahlen, für große Werte von m) sollten die nach Alter bzw. Geschlecht gruppierten Mittelwertplots dann näherungsweise „parallel“ verlaufen.

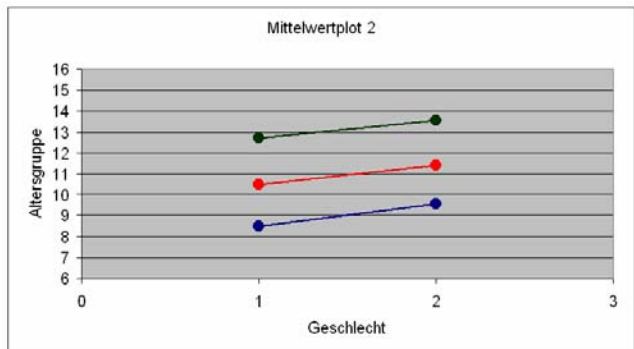
Beispiel 40. Zur Konkretisierung dieses Sachverhalts betrachten wir folgende fiktive Datenmatrix aus einem Modell *ohne* Wechselwirkungen (Kosten in Mio. Euro, AG = Altersgruppe):

Zelle \ Kosten	k	1	2	3	4	5	MW
AG 1, m	Y_{11k}	8,195	8,481	8,773	8,616	8,247	8,463
AG 2, m	Y_{21k}	10,439	10,569	10,497	10,372	10,557	10,487
AG 3, m	Y_{31k}	12,876	12,830	12,782	12,683	12,371	12,708
AG 1, w	Y_{12k}	9,935	9,134	9,988	9,220	9,495	9,554
AG 2, w	Y_{22k}	11,180	11,884	11,550	11,466	11,113	11,439
AG 3, w	Y_{32k}	13,481	13,438	13,768	13,425	13,674	13,557

Als Mittelwertplots erhält man hier:



links: rot: w , blau: m



rechts: grün: AG 3, rot: AG 2, blau: AG 1

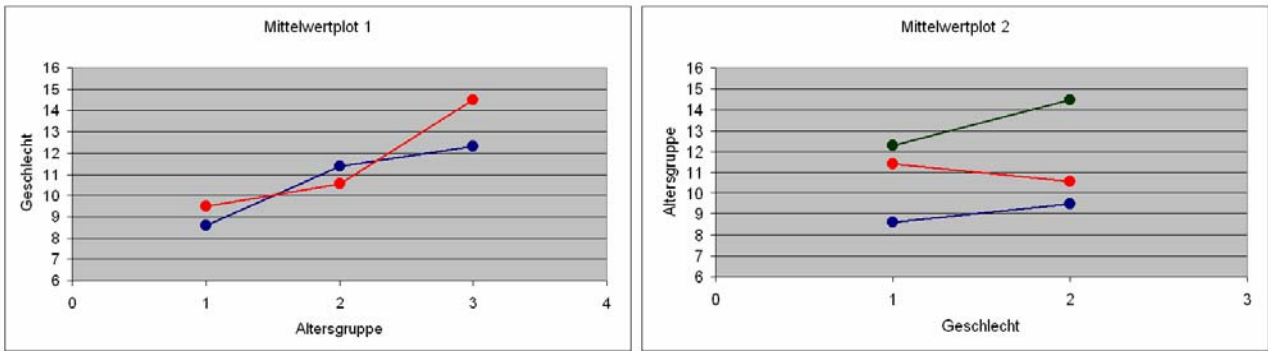
Die Parallelität in den Mittelwertplots ist empirisch gut erkennbar.

Das folgende Datenbeispiel stammt aus einem Modell mit Wechselwirkungen:

Zelle \ Kosten	k	1	2	3	4	5	MW
AG 1, m	Y_{11k}	8,871	8,090	8,809	8,752	8,473	8,599
AG 2, m	Y_{21k}	11,209	11,654	11,105	11,344	11,715	11,405
AG 3, m	Y_{31k}	12,315	12,014	12,583	12,476	12,182	12,314
AG 1, w	Y_{12k}	9,481	9,203	9,354	9,755	9,602	9,479
AG 2, w	Y_{22k}	10,143	10,955	10,749	10,931	10,056	10,567
AG 3, w	Y_{32k}	14,950	14,241	14,896	14,172	14,175	14,487

V. Verallgemeinerte lineare Modelle

Als Mittelwertplots erhält man hier:



links: rot: w , blau: m

rechts: grün: AG 3, rot: AG 2, blau: AG 1

Eine durchgehende Parallelität in den Mittelwertplots ist nicht erkennbar, die Linien überschneiden sich zum Teil.

In der Praxis treten häufig auch Mischfälle aus diskreten und metrischen Merkmalen auf, wie das folgende Beispiel zeigt.

Beispiel 41 (Kraftfahrzeug-Teilkasko-Versicherung). Betrachtet werden die diskreten Merkmale *Fahrzeugwert* (Ausprägungen niedrig / mittel / hoch) und das stetige Merkmal *Jahresfahrleistung*. Abhängige Variable ist der Jahresgesamtschaden einer Kfz-Teilkasko-Police, für den eine proportionale Abhängigkeit von der Jahresfahrleistung mit einem additiven Einfluss des Fahrzeugwerts unterstellt wird. Beobachtungen liegen für folgende Merkmalskombinationen (Zellen) vor:

Zelle	Fahrzeugwert	Jahresfahrleistung
1	niedrig	12000 km
2	niedrig	17000 km
3	niedrig	23000 km
4	mittel	15000 km
5	mittel	21000 km
6	mittel	27000 km
7	mittel	31000 km
8	hoch	15000 km
9	hoch	21000 km
10	hoch	35000 km
11	hoch	42000 km
12	hoch	51000 km

Eine mögliche Designmatrix ist hier gegeben durch

V. Verallgemeinerte lineare Modelle

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 & 12000 \\ 1 & 0 & 0 & 17000 \\ 1 & 0 & 0 & 23000 \\ 0 & 1 & 0 & 15000 \\ 0 & 1 & 0 & 21000 \\ 0 & 1 & 0 & 27000 \\ 0 & 1 & 0 & 31000 \\ 0 & 0 & 1 & 15000 \\ 0 & 0 & 1 & 21000 \\ 0 & 0 & 1 & 35000 \\ 0 & 0 & 1 & 42000 \\ 0 & 0 & 1 & 51000 \end{pmatrix}.$$

Die ersten drei Spalten beschreiben dabei den Einfluß des Fahrzeugwerts (Achsenabschnitt für die lineare Regression), die vierte Spalte beschreibt den Einfluss der Jahresfahrleistung (Steigung für die lineare Regression).

Als lineares statistisches Modell erhält man somit (abgeleitet aus der tabellarischen Form)

$$Y_{i,j_i} = \beta_i + \beta_4 x_{j_i} + \varepsilon_{i,j_i}, \quad i = 1, 2, 3, \quad j_1 = 1, 2, 3, \quad j_2 = 1, \dots, 4, \quad j_3 = 1, \dots, 5.$$

Die Regressionsparameter können wie folgt interpretiert werden:

- Die β_i , $i = 1, 2, 3$ bilden den Unterschied zwischen den verschiedenen Fahrzeugwerten ab,
- β_4 bildet die Steigung des linearen Prädiktors pro Kilometer Jahresfahrleistung ab.

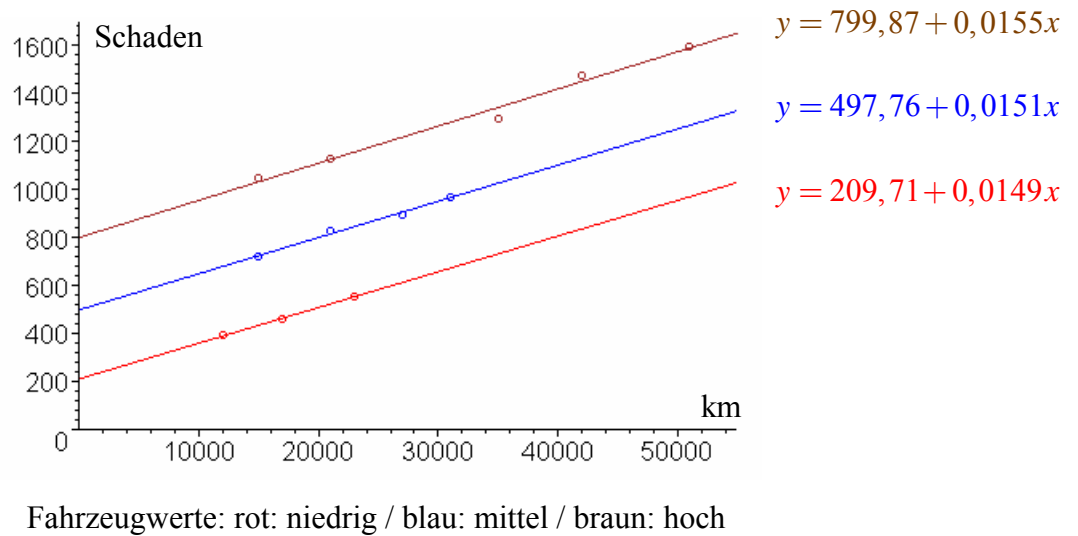
Die Abwesenheit von Wechselwirkungen äußert sich im Parallel-Verlauf der drei aus dem Erwartungswert der Y_{i,j_i} gebildeten Regressionsgeraden $y_i = \beta_i + \beta_4 x$ (gemeinsames β_4 für alle Fahrzeugwerte).

Zur Konkretisierung dieses Modells betrachten wir folgende fiktive Datenmatrix (Kosten in Euro):

Zelle	Kosten	
1	Y_{11}	391
2	Y_{12}	460
3	Y_{13}	555
4	Y_{21}	718
5	Y_{22}	828
6	Y_{23}	897
7	Y_{24}	967
8	Y_{31}	1046
9	Y_{32}	1128
10	Y_{33}	1295
11	Y_{34}	1477
12	Y_{35}	1599

V. Verallgemeinerte lineare Modelle

Als einzelne Regressionsgeraden erhält man mit den (individuellen) Schätzverfahren von oben:



Die Parallelität der Regressionsgeraden ist empirisch gut erkennbar und wird durch die numerisch fast gleichen Steigungswerte bestätigt. Als Gesamt-Schätzer für β erhält man dagegen:

$$\beta = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{Y} = \begin{pmatrix} 201,18 \\ 489,85 \\ 802,84 \\ 0,1543 \end{pmatrix}.$$

Man beachte, dass die aus allen Daten gemeinsam geschätzten Parameter β_1, \dots, β_4 geringfügig von den einzelnen Regressionsschätzern in der Graphik oben abweichen.

Bemerkung: Dieses Modell ist wegen der Parallelität der Regressionsgeraden ein lineares statistisches Modell ohne Wechselwirkungen. Wechselwirkungen lägen z.B. vor im Fall, dass die Steigung des linearen Prädiktors pro Kilometer Jahresfahrleistung für verschiedene Fahrzeugwerte unterschiedlich sind.

V. 3. Verallgemeinerte lineare Modelle

In vielen Anwendungen ist ein strikt linearer Zusammenhang zwischen den Erwartungswerten der betrachteten Zufallsvektoren \mathbf{Y} und dem Parametervektor β vermöge der Beziehung $E(\mathbf{Y}) = \mathbf{D}\beta$ nicht gegeben bzw. nicht gerechtfertigt, z.B. wenn man funktionale Abhängigkeiten von Schaden-eintrittswahrscheinlichkeiten bezüglich kategorialer Merkmale wie Geschlecht oder Altersgruppen modellieren will. In verallgemeinerten linearen Modellen (engl.: generalized linear model, GLM) transformiert man daher den linearen Prädiktor $\mathbf{D}\beta$ in geeigneter Weise. Unterstellt man für die Verteilung der Komponenten des Zufallsvektors \mathbf{Y} geeignete Exponentialfamilien, so lassen sich überraschend einfache Gleichungen zur Schätzung des Parametervektors β mit der Maximum-Likelihood-Methode herleiten, die meist explizit - zumindest auf numerische Weise - lösbar sind.

Zur Vereinfachung der Notation führen wir folgende Bezeichnungen zur Charakterisierung eines verallgemeinerten linearen Modells ein, das aus folgenden Modellkomponenten besteht:

V. Verallgemeinerte lineare Modelle

- Einem n -dimensionaler Zufallsvektor $\mathbf{Y} = (Y_1, \dots, Y_n)^{tr}$ mit $n \in \mathbb{N}$; die Y_i werden auch *Response-Variablen* genannt. Die Verteilung jedes Y_i gehört zu einer *einparametrischen* Exponentialfamilie mit Dichten der Form

$$f_{Y_i}(y; \vartheta_i) = \exp \left\{ \frac{y\vartheta_i + a(y, \tau_i) - b(\vartheta_i)}{\tau_i} \right\}, \quad \vartheta_i \in \Theta_{\tau_i}, \quad i = 1, \dots, n, \quad y \in \mathbb{R}$$

bezüglich eines dominierenden Maßes μ , wobei die $\tau_i > 0$ sogenannte *Störparameter* darstellen (in der Versicherungstechnik entspricht τ_i meist einer Varianz und damit $1/\tau_i$ einem *Volumenmaß*), mit $b''(\vartheta_i) > 0$ für alle $\vartheta_i \in \Theta_{\tau_i}$, den natürlichen Parameterräumen

$$\Theta_{\tau_i} := \left\{ \vartheta \in \mathbb{R} \mid \int \exp \left(\frac{y\vartheta + a(y, \tau_i)}{\tau_i} \right) \mu(dy) < \infty \right\}, \quad i = 1, \dots, n$$

- dem zugehörigen Erwartungswert-Vektor $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^{tr}$ mit $\mu_i = E(Y_i)$

- der Design-Matrix $\mathbf{D} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}$ der (bekannten) Kontroll- bzw. Einflussgrößen (mit Zeilenvektoren \mathbf{x}_i)

- einer sogenannten *Linkfunktion* $g: \mathbb{R} \supseteq G \rightarrow \mathbb{R}$ (mit einer offenen Teilmenge G), die zweimal stetig differenzierbar ist mit $g' \neq 0$ auf G ; die (dann existierende) Umkehrfunktion $h = g^{-1}$ heißt auch *Response Funktion*
- der komponentenweisen Beziehung

$$\boldsymbol{\eta} := g(\boldsymbol{\mu}) = \mathbf{D}\boldsymbol{\beta}$$

bzw.

$$\eta_i := g(\mu_i) = \mathbf{x}_i \boldsymbol{\beta} \quad \text{für } i = 1, \dots, n.$$

Bemerkung: Die letzte Beziehung läßt sich auch mit der Umkehrfunktion h schreiben:

$$\boldsymbol{\mu} = h(\mathbf{D}\boldsymbol{\beta})$$

bzw.

$$\mu_i = h(\mathbf{x}_i \boldsymbol{\beta}) \quad \text{für } i = 1, \dots, n.$$

Lemma 44. Unter den obigen Bedingungen an ein GLM gilt

$$\mu_i = E_{\vartheta_i}(Y_i) = b'(\vartheta_i), \quad \text{Var}_{\vartheta_i}(Y_i) = \tau_i \cdot b''(\vartheta_i) \quad \text{für } i = 1, \dots, n.$$

V. Verallgemeinerte lineare Modelle

Beweis: Durch die Umparametrisierung $\tilde{\vartheta}_i := \frac{\vartheta_i}{\tau_i}$, $\tilde{b}(\tilde{\vartheta}_i) := \frac{b(\tau_i \cdot \tilde{\vartheta}_i)}{\tau_i}$ erhält man die kanonische Darstellung

$$f_{Y_i}(y; \tilde{\vartheta}_i) = \exp\{y\tilde{\vartheta}_i + \tilde{a}(y) - \tilde{b}(\tilde{\vartheta}_i)\}$$

mit einer geeigneten Abbildung $\tilde{a}(\cdot)$ und daraus nach Satz 14

$$E(Y_i) = \tilde{b}'(\tilde{\vartheta}_i) = \frac{d}{d\tilde{\vartheta}_i} \tilde{b}(\tilde{\vartheta}_i) = \frac{d}{d\tilde{\vartheta}_i} \frac{b(\tau_i \cdot \tilde{\vartheta}_i)}{\tau_i} = b'(\tau_i \cdot \tilde{\vartheta}_i) = b'(\vartheta_i)$$

sowie

$$\text{Var}(Y_i) = \tilde{b}''(\tilde{\vartheta}_i) = \frac{d^2}{d\tilde{\vartheta}_i^2} \tilde{b}(\tilde{\vartheta}_i) = \frac{d^2}{d\tilde{\vartheta}_i^2} \frac{b(\tau_i \cdot \tilde{\vartheta}_i)}{\tau_i} = \frac{\tau_i^2}{\tau_i} b''(\tau_i \cdot \tilde{\vartheta}_i) = \tau_i \cdot b''(\vartheta_i).$$

Es soll jetzt gezeigt werden, dass die beiden unterschiedlichen Parametrisierungen der Verteilung des Zufallsvektors \mathbf{Y} - zum einen über den Parametervektor $\boldsymbol{\vartheta} = (\vartheta_1, \dots, \vartheta_n)^T$, zum anderen über die Regressoren $\boldsymbol{\beta} = (\beta_1, \dots, \beta_d)^T$ - in einem sinnvollen funktionalen Zusammenhang stehen, der es erlaubt, die Regressoren ohne separate Bestimmung von $\boldsymbol{\vartheta}$ zu schätzen.

Aus der Voraussetzung $b''(\cdot) > 0$ folgt zunächst die Monotonie von $b'(\cdot)$, so dass die Gleichungen

$$\mu_i = E_{\vartheta_i}(Y_i) = b'(\vartheta_i), \quad i = 1, \dots, n$$

eindeutig nach ϑ_i aufgelöst werden können, in der Form

$$\vartheta_i = \psi(\mu_i), \quad i = 1, \dots, n \quad \text{mit der Umkehrfunktion } \psi = (b')^{-1}.$$

Damit ergibt sich aber auch ein direkter Zusammenhang zwischen $\boldsymbol{\vartheta}$ und $\boldsymbol{\beta}$ vermöge der Beziehung

$$\boldsymbol{\vartheta} = \psi(\boldsymbol{\mu}) = \psi \circ h(\mathbf{D}\boldsymbol{\beta}) \quad \text{bzw.} \quad \vartheta_i = \psi(\mu_i) = \psi \circ h(\mathbf{x}_i\boldsymbol{\beta}), \quad i = 1, \dots, n.$$

Ferner erhält man noch für die Varianzen die Beziehungen

$$\sigma_i^2 = \tau_i \cdot b''(\vartheta_i) = \tau_i \cdot b''(\psi(\mu_i)), \quad i = 1, \dots, n,$$

die für eine pragmatische Identifikation eines geeigneten verallgemeinerten linearen Modells wichtig sind. Die Abbildung $b'' \circ \psi$ heißt in diesem Zusammenhang auch *Varianzfunktion*; sie wird üblicherweise mit dem Symbol V bezeichnet.

Einen besonders wichtigen Spezialfall beschreibt

Definition 17. Die Linkfunktion g heißt *natürlich* (auch *kanonisch*), wenn $g = \psi = (b')^{-1}$ gilt.

Bei natürlichen Linkfunktionen erhält man die stark vereinfachte Beziehung

$$g(\boldsymbol{\mu}) = \boldsymbol{\vartheta} = \boldsymbol{\eta} = \mathbf{D}\boldsymbol{\beta}.$$

V. Verallgemeinerte lineare Modelle

Im Folgenden werden wir nur noch verallgemeinerte lineare Modelle mit natürlicher Linkfunktion betrachten. Für die *Score-Funktionen* (d.h. die partiellen Ableitungen der log-Likelihood-Funktion) erhält man dann folgendes Resultat.

Lemma 45. Gegeben sei ein verallgemeinertes lineares Modell mit natürlicher Linkfunktion g . Der Vektor $U_n(\boldsymbol{\beta}) = (U_{n1}(\boldsymbol{\beta}), \dots, U_{nd}(\boldsymbol{\beta}))$ der Score-Funktionen bzgl. des Parametervektors $\boldsymbol{\beta}$ ist dann gegeben durch

$$U_{nj}(\boldsymbol{\beta}) = \sum_{i=1}^n \frac{x_{ij}}{\tau_i} (Y_i - h(\mathbf{x}_i \boldsymbol{\beta})), \quad j = 1, \dots, d.$$

Beweis: Für die log-likelihood-Funktion erhält man

$$\ell_n(\boldsymbol{\beta}) = \sum_{i=1}^n \ln f_{Y_i}(Y_i; \vartheta_i) = \sum_{i=1}^n \frac{1}{\tau_i} \{Y_i \vartheta_i - b(\vartheta_i)\} + \sum_{i=1}^n \frac{a(Y_i, \tau_i)}{\tau_i}$$

und daraus durch Bildung der partiellen Ableitungen

$$\begin{aligned} U_{nj}(\boldsymbol{\beta}) &= \frac{\partial}{\partial \beta_j} \ell_n(\boldsymbol{\beta}) = \frac{\partial}{\partial \beta_j} \sum_{i=1}^n \frac{1}{\tau_i} \{Y_i \vartheta_i - b(\vartheta_i)\} = \sum_{i=1}^n \frac{1}{\tau_i} \left\{ Y_i \frac{\partial}{\partial \beta_j} \mathbf{x}_i \boldsymbol{\beta} - \frac{\partial}{\partial \beta_j} b(\vartheta_i) \right\} \\ &= \sum_{i=1}^n \frac{x_{ij}}{\tau_i} Y_i - \sum_{i=1}^n \frac{1}{\tau_i} \cdot \frac{\partial}{\partial \vartheta_i} b(\vartheta_i) \cdot \frac{\partial \vartheta_i}{\partial \beta_j} = \sum_{i=1}^n \frac{x_{ij}}{\tau_i} Y_i - \sum_{i=1}^n \frac{1}{\tau_i} \cdot \underbrace{b'(\vartheta_i)}_{=\mu_i} \cdot \underbrace{\frac{\partial}{\partial \beta_j} \mathbf{x}_i \boldsymbol{\beta}}_{=x_{ij}} \\ &= \sum_{i=1}^n \frac{x_{ij}}{\tau_i} (Y_i - \mu_i) = \sum_{i=1}^n \frac{x_{ij}}{\tau_i} (Y_i - h(\mathbf{x}_i \boldsymbol{\beta})), \quad j = 1, \dots, d. \end{aligned}$$

Bemerkungen:

i) Wenn die Störvariablen τ_i alle identisch sind, also $\tau_i \equiv \tau > 0$ gilt, vereinfacht sich die Darstellung des Score-Vektors wie folgt:

$$U_n(\boldsymbol{\beta}) = \frac{1}{\tau} \mathbf{D}^r (\mathbf{Y} - h(\mathbf{D}\boldsymbol{\beta})).$$

ii) Einen Maximum-Likelihood-Schätzer (ML-Schätzer) für $\boldsymbol{\beta}$ erhält man über $U_n(\boldsymbol{\beta}) = \mathbf{0}$. Die Lösung ist eindeutig bestimmt, sofern die Design-Matrix \mathbf{D} vollen Rang hat; vgl. hierzu etwa die Argumentation in PRUSCHA (2000), Kapitel VII, Abschnitt 2.3. Für den Fall, dass die Design-Matrix vom Inzidenz-Typ ist (z.B. bei kategoriellen Merkmalen), gibt es in der Regel mehrere Lösungen. Eine Eindeutigkeit der Lösung kann jedoch durch lineare Zusatzbedingungen erreicht werden; vgl. hierzu etwa PRUSCHA (2000), Kapitel III. Für Anwendungen im Versicherungsbereich spielt das aber keine wesentliche Rolle, weil der lineare Prediktor $\boldsymbol{\vartheta} = \mathbf{D}\boldsymbol{\beta}$ für alle Lösungen identisch ist und nur dieser in die Verteilung der Response-Variablen eingeht, welche z.B. Grundlage einer Tarifierung ist. Alternativ kann man durch Streichen passender Spalten von \mathbf{D} (und damit auch zugehöriger Parameterkomponenten) die Designmatrix in eine mit vollem Rang überführen.

iii) Im Falle eines einfachen linearen Modells gilt $g(\mu) = \mu = h(\mu)$, so dass sich - bei vollem Rang von \mathbf{D} - der ML-Schätzer für $\boldsymbol{\beta}$ aus der Gleichung $\mathbf{D}^r (\mathbf{Y} - h(\mathbf{D}\boldsymbol{\beta})) = \mathbf{0}$ zu der bekannten Lösung $\hat{\boldsymbol{\beta}} = (\mathbf{D}^r \mathbf{D})^{-1} \mathbf{D}^r \mathbf{Y}$ ergibt.

V. Verallgemeinerte lineare Modelle

iv) Liegen für den Zufallsvektor \mathbf{Y} mehrere stochastisch unabhängige Kopien $\mathbf{Y}_1, \dots, \mathbf{Y}_r$ mit $r \in \mathbb{N}$ vor, so zeigt der Beweis zu Lemma 45 unter Beachtung von Lemma 40, dass die resultierenden Score-Funktionen durch

$$U_{nj}(\boldsymbol{\beta}) = \sum_{m=1}^r \sum_{i=1}^n \frac{x_{ij}}{\tau_i} (Y_{im} - h(\mathbf{x}_i, \boldsymbol{\beta})), \quad j = 1, \dots, d$$

gegeben sind. Den ML-Schätzer für $\boldsymbol{\beta}$ erhält man dann aus den Gleichungen

$$\sum_{i=1}^n \frac{x_{ij}}{\tau_i} (\bar{Y}_i - h(\mathbf{x}_i, \boldsymbol{\beta})) = 0, \quad j = 1, \dots, d,$$

wobei $\bar{Y}_i = \frac{1}{r} \sum_{m=1}^r Y_{im}$ das arithmetische Mittel der Kopien von Y_i bezeichnet.

Die folgende Tabelle enthält für einige wichtige verallgemeinerte lineare Modelle mit natürlicher Linkfunktion g die wesentlichen Informationen. Vgl. hierzu auch MCCULLAGH UND NELDER, (1989). Hierbei bezeichnet abweichend von der üblichen Konvention $B^*(m, p)$ mit $m \in \mathbb{N}$ und $p \in (0, 1)$ eine *skalierte Binomialverteilung* über der Menge $\left\{ \frac{1}{m}, \frac{2}{m}, \dots, 1 \right\}$, die sich als Verteilung von $\frac{X}{m}$ ergibt, wenn X einer $B(m, p)$ -Verteilung genügt.

Lemma 46. Es gilt (mit der üblichen Parametrisierung der Verteilungen):

Verteilung	$\mathcal{N}(\mu, s^2)$	$\mathcal{P}(\lambda)$	$B^*(m, p)$	$\Gamma(\alpha, \lambda)$
Wertebereich	$(-\infty, \infty)$	$\{0, 1, 2, \dots\}$	$\left\{ 0, \frac{1}{m}, \frac{2}{m}, \dots, 1 \right\}$	$(0, \infty)$
ϑ	μ	$\ln \lambda$	$\ln \left(\frac{p}{1-p} \right)$	$-\frac{\lambda}{\alpha}$
τ	s^2	1	$\frac{1}{m}$	$\frac{1}{\alpha}$
$b(\vartheta)$	$\frac{\vartheta^2}{2}$	e^ϑ	$\ln(1 + e^\vartheta)$	$-\ln(-\vartheta)$
$a(y, \tau)$	$-\frac{1}{2}(y^2 + \tau \ln(2\pi\tau))$	$-\ln(y!)$	$\frac{1}{\tau} \ln \left(\frac{1}{\tau} \frac{y}{\tau} \right)$	$(\tau - 1) \ln y - \ln \tau - \tau \ln \Gamma \left(\frac{1}{\tau} \right)$
$\mu(\vartheta)$	ϑ	e^ϑ	$\frac{e^\vartheta}{1 + e^\vartheta}$	$-\frac{1}{\vartheta}$
$g(\mu)$	μ	$\ln \mu$	$\ln \left(\frac{\mu}{1-\mu} \right)$	$-\frac{1}{\mu}$
$V(\mu)$	1	μ	$\mu(1-\mu)$	μ^2

V. Verallgemeinerte lineare Modelle

Beweis: Bei der Normal- und Poisson-Verteilung sowie der Gammaverteilung ergeben sich die natürliche Parametrisierung sowie die daraus abgeleiteten Größen unmittelbar aus den Lemmata 37 und 41. Die Angaben zur skalierten Binomialverteilung ergeben sich aus der Darstellung

$$P(Y = y) = \binom{m}{ym} p^{my} (1-p)^{m-my} \text{ für } y \in \left\{0, \frac{1}{m}, \frac{2}{m}, \dots, 1\right\}.$$

Zum Abschluß dieses Abschnitts wollen wir noch zwei typische Anwendungsbeispiele aus der Versicherungspraxis vorstellen.

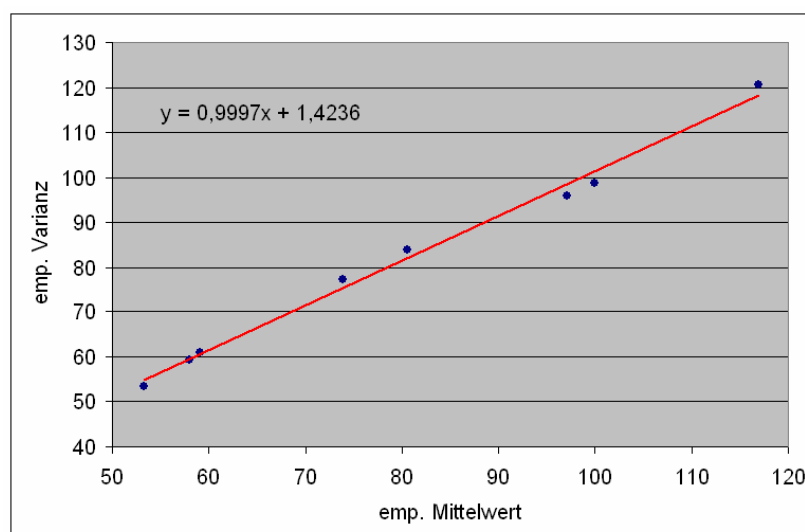
Beispiel 42. Die folgende Tabelle enthält die Anzahl von Versicherungsleistungen aus einem Kfz-Geschäft der letzten 10 Jahre, gruppiert nach Geschlecht G (*m* / *w*), Fahrerfahrung FE (*kurz* / *lang*) und jährlicher Fahrleistung FL (*hoch* / *niedrig*).

G / FE / FL	Jahr 1	Jahr 2	Jahr 3	Jahr 4	Jahr 5	Jahr 6	Jahr 7	Jahr 8	Jahr 9	Jahr 10	MW	emp. Varianz
m / lang / niedrig	52	55	51	46	66	71	65	61	62	62	59,1	60,99
m / lang / hoch	91	100	96	108	109	117	86	106	89	98	100,0	98,67
m / kurz / niedrig	71	72	74	76	61	74	66	85	91	68	73,8	77,29
m / kurz / hoch	101	119	109	124	118	106	115	116	141	120	116,9	120,54
w / lang / niedrig	42	67	61	53	49	46	52	50	57	56	53,3	53,34
w / lang / hoch	88	79	58	80	79	83	88	82	77	91	80,5	83,83
w / kurz / niedrig	61	65	67	62	56	55	63	60	49	42	58,0	59,33
w / kurz / hoch	97	85	119	84	101	93	97	100	94	101	97,1	95,88

Die Daten sollen an ein verallgemeinertes lineares Modell angepaßt werden.

Eine pragmatische Vorgehensweise ist die folgende:

1. Bestimmung eines geeigneten Verteilungsmodells anhand eines empirischen Varianz-Mittelwertplots:



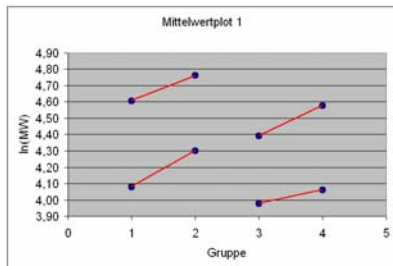
Die eingezeichnete Regressionsgerade hat eine Steigung von näherungsweise 1, dies legt nach Lemma 46 ein Poisson-Modell nahe.

V. Verallgemeinerte lineare Modelle

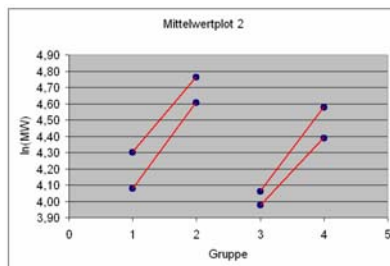
2. Ermittlung einer geeigneten Design-Matrix:

Im Poisson-Modell ist der Logarithmus die natürliche Linkfunktion. Zur empirischen Prüfung einer Wechselwirkungsfreiheit kann man also die logarithmierten Mittelwerte heranziehen, die in der folgenden Tabelle wiedergegeben sind:

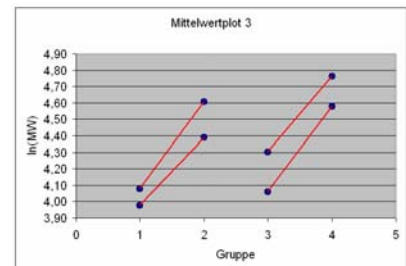
G / FE / FL	Gruppen Plot 1	Gruppen Plot 2	Gruppen Plot 3	ln(MW)
m / lang / niedrig	1	1	1	4,08
m / lang / hoch	1	2	2	4,61
m / kurz / niedrig	2	1	3	4,30
m / kurz / hoch	2	2	4	4,76
w / lang / niedrig	3	3	1	3,98
w / lang / hoch	3	4	2	4,39
w / kurz / niedrig	4	3	3	4,06
w / kurz / hoch	4	4	4	4,58



Mittelwertplot, gruppiert nach Geschlecht und Fahrerfahrung



Mittelwertplot, gruppiert nach Geschlecht und Fahrleistung



Mittelwertplot, gruppiert nach Fahrerfahrung und Fahrleistung

Alle Plots zeigen eine hinreichend sichtbare Parallelität, so dass die Annahme einer Wechselwirkungsfreiheit plausibel erscheint. Dies legt folgenden Modellansatz nahe:

$$\ln \mu_{ijk} = \alpha_i + \beta_j + \gamma_k \quad \text{für } i, j, k = 1, 2,$$

wobei α das Geschlecht, β die Fahrerfahrung und γ die Fahrleistung repräsentiert. Als Parametervektor β ergibt sich damit

$$\beta = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2)^T$$

sowie als Designmatrix

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

V. Verallgemeinerte lineare Modelle

Diese Matrix besitzt allerdings nicht den (maximalen) Rang 6, sondern lediglich den Rang 4, wie man z.B. mit Hilfe des Gauß'schen Algorithmus feststellen kann.

3. Schätzung von $\boldsymbol{\beta}$ mit der Maximum-Likelihood-Methode:

Da 10 unterstellt unabhängige Kopien der zugehörigen Zufallsvariablen $Y_i, i = 1, \dots, 8 = n$ (nach Umindizierung) vorliegen, kann wegen $\tau = 1$ der Parameter $\boldsymbol{\beta}$ durch Lösen von

$$\sum_{i=1}^8 x_{ij} (\bar{Y}_i - \exp(\mathbf{x}_i \boldsymbol{\beta})) = \sum_{i=1}^8 x_{ij} \left(\bar{Y}_i - \exp \left(\sum_{k=1}^6 x_{ik} \boldsymbol{\beta}_k \right) \right) = 0, \quad j = 1, \dots, 6 = d$$

bestimmt werden. Im Poisson-Modell mit Designmatrizen vom Inzidenztyp wie oben ergibt sich dabei eine praktische Vereinfachung, wenn man $\xi_k = \exp(\boldsymbol{\beta}_k) > 0$ für $k = 1, \dots, d$ setzt. Man erhält dann das (nicht-lineare) Gleichungssystem

$$\sum_{i=1}^n x_{ij} \prod_{k=1}^d \xi_k^{x_{ik}} = \sum_{i=1}^n x_{ij} \bar{Y}_i, \quad j = 1, \dots, d,$$

welches nach den Unbekannten ξ_k aufzulösen ist. Im letzten Schritt ergibt sich dann $\boldsymbol{\beta}_k = \ln \xi_k$ für $k = 1, \dots, d$. Im betrachteten Beispiel lautet das Gleichungssystem

$$\begin{aligned} \xi_1 \xi_3 \xi_5 + \xi_1 \xi_3 \xi_6 + \xi_1 \xi_4 \xi_5 + \xi_1 \xi_4 \xi_6 &= 349,8 \\ \xi_2 \xi_3 \xi_5 + \xi_2 \xi_3 \xi_6 + \xi_2 \xi_4 \xi_5 + \xi_2 \xi_4 \xi_6 &= 288,9 \\ \xi_1 \xi_3 \xi_5 + \xi_1 \xi_3 \xi_6 + \xi_2 \xi_3 \xi_5 + \xi_2 \xi_3 \xi_6 &= 292,9 \\ \xi_1 \xi_4 \xi_5 + \xi_1 \xi_4 \xi_6 + \xi_2 \xi_4 \xi_5 + \xi_2 \xi_4 \xi_6 &= 345,8 \\ \xi_1 \xi_3 \xi_5 + \xi_1 \xi_4 \xi_5 + \xi_2 \xi_3 \xi_5 + \xi_2 \xi_4 \xi_5 &= 244,2 \\ \xi_1 \xi_3 \xi_6 + \xi_1 \xi_4 \xi_6 + \xi_2 \xi_3 \xi_6 + \xi_2 \xi_4 \xi_6 &= 394,5 \end{aligned}$$

Dieses Gleichungssystem ist allerdings nicht eindeutig lösbar. Eine eindeutige Lösung erhält man aber z.B. durch Streichen der vierten und sechsten Spalte der Designmatrix (der resultierende Rang ist dann 4) bzw. durch die Fixierung von $\beta_4 := \beta_6 := 0$ bzw. $\xi_4 := \xi_6 := 1$. Die Lösung ist dann - auf vier Stellen gerundet - gegeben durch

k	1	2	3	4	5	6
$\hat{\xi}_k$	116,9765	96,6108	0,8470	1,0000	0,6190	1,0000
$\hat{\boldsymbol{\beta}}_k$	4,7620	4,5707	-0,1660	0,0000	-0,4796	0,0000

Als Schätzung für den linearen Prediktor ergibt sich damit

$$\hat{\boldsymbol{\vartheta}} = \mathbf{D} \hat{\boldsymbol{\beta}} = (4,1164 \quad 4,5960 \quad 4,2824 \quad 4,7620 \quad 3,9251 \quad 4,4047 \quad 4,0911 \quad 4,5707)^{tr}$$

sowie

$$\hat{\boldsymbol{\mu}} = \exp(\hat{\boldsymbol{\vartheta}}) = (61,338 \quad 99,087 \quad 72,414 \quad 116,980 \quad 50,658 \quad 81,835 \quad 59,806 \quad 96,612)^{tr}.$$

(Man vergleiche dieses Ergebnis mit den Mittelwerten aus der gegebenen Datentabelle.)

Bemerkung: Auch ohne die Fixierung gewisser Parameter ergibt sich aus *jeder* Lösung $\hat{\boldsymbol{\beta}}$ des Gleichungssystems stets der gleiche Prediktorwert $\hat{\boldsymbol{\vartheta}} = \mathbf{D} \hat{\boldsymbol{\beta}}$.

V. Verallgemeinerte lineare Modelle

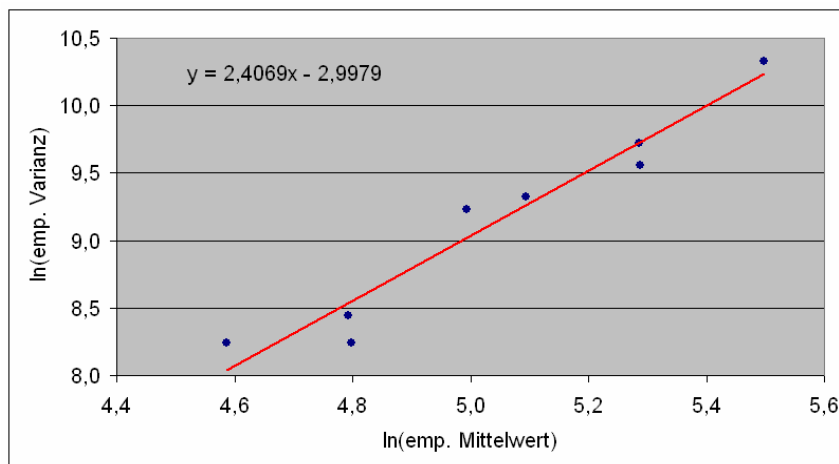
Beispiel 43. Die folgende Tabelle enthält die *Schadenzahlungen* in Mio. € aus einem Kfz-Geschäft der letzten 10 Jahre, gruppiert nach Geschlecht G (*m* / *w*), Fahrerfahrung FE (*kurz* / *lang*) und jährlicher Fahrleistung FL (*hoch* / *niedrig*).

G / FE / FL	Jahr 1	Jahr 2	Jahr 3	Jahr 4	Jahr 5	Jahr 6	Jahr 7	Jahr 8	Jahr 9	Jahr 10	MW	Varianz
m / lang / niedrig	213,69	53,08	73,45	95,63	170,59	66,81	196,46	209,83	73,77	53,80	120,71	4648,22
m / lang / hoch	208,15	275,56	210,95	54,08	172,13	440,83	300,68	158,55	88,62	68,23	197,78	14171,25
m / kurz / niedrig	114,17	157,10	18,77	127,78	385,86	61,07	71,32	199,87	167,92	171,29	147,52	10219,67
m / kurz / hoch	140,40	642,65	112,94	441,62	235,68	90,95	231,98	205,15	88,80	249,92	244,01	30572,77
w / lang / niedrig	151,87	39,68	100,47	52,69	114,14	84,16	38,63	61,79	96,52	241,55	98,15	3806,35
w / lang / hoch	350,08	96,87	32,60	93,45	312,47	149,23	80,15	186,12	93,74	234,69	162,94	11231,28
w / kurz / niedrig	113,20	164,61	131,93	139,06	44,83	114,79	52,45	216,52	39,59	195,28	121,23	3801,89
w / kurz / hoch	456,81	189,89	103,23	176,73	32,11	148,29	155,21	85,60	347,58	281,65	197,71	16701,27

Die Daten sollen an ein verallgemeinertes lineares Modell angepaßt werden.

Eine entsprechende pragmatische Vorgehensweise ist wieder die folgende:

1. Bestimmung eines geeigneten Verteilungsmodells anhand eines empirischen Varianz-Mittelwertplots: da die empirische Varianz offensichtlich überproportional mit dem empirischen Mittelwert wächst, werden die logarithmierten Werte gegeneinander aufgetragen:



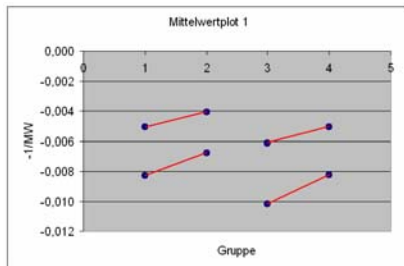
Die eingezeichnete Regressionsgerade hat eine Steigung von etwas oberhalb von 2, was einen quadratischen Zusammenhang nahelegt. Ein passendes Modell ist daher nach Lemma 43 das Gamma-Modell.

2. Ermittlung einer geeigneten Design-Matrix:

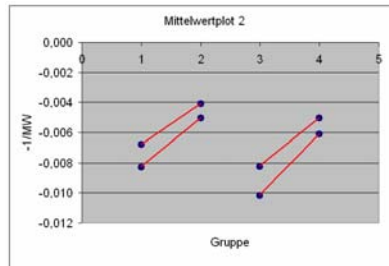
Im Gamma-Modell ist das negative Inverse die natürliche Linkfunktion. Zur empirischen Prüfung einer Wechselwirkungsfreiheit kann man also die entsprechend transformierten Mittelwerte heranziehen, die in der folgenden Tabelle wiedergegeben sind:

V. Verallgemeinerte lineare Modelle

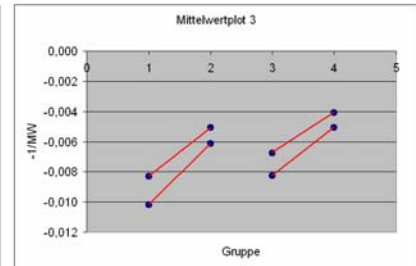
G / FE / FL	Gruppen Plot 1	Gruppen Plot 2	Gruppen Plot 3	- 1/MW
m / lang / niedrig	1	1	1	- 0,0083
m / lang / hoch	1	2	2	- 0,0051
m / kurz / niedrig	2	1	3	- 0,0068
m / kurz / hoch	2	2	4	- 0,0041
w / lang / niedrig	3	3	1	- 0,0102
w / lang / hoch	3	4	2	- 0,0061
w / kurz / niedrig	4	3	3	- 0,0082
w / kurz / hoch	4	4	4	- 0,0051



Mittelwertplot, gruppiert nach Geschlecht und Fahrerfahrung



Mittelwertplot, gruppiert nach Geschlecht und Fahrleistung



Mittelwertplot, gruppiert nach Fahrerfahrung und Fahrleistung

Alle Plots zeigen auch hier eine hinreichend sichtbare Parallelität, so dass die Annahme einer Wechselwirkungsfreiheit wieder plausibel erscheint. Dies legt folgenden Modellansatz nahe:

$$-\frac{1}{\mu_{ijk}} = \alpha_i + \beta_j + \gamma_k \quad \text{für } i, j, k = 1, 2,$$

wobei α das Geschlecht, β die Fahrerfahrung und γ die Fahrleistung repräsentiert. Als Parametervektor $\boldsymbol{\beta}$ ergibt sich damit wieder

$$\boldsymbol{\beta} = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2)^T$$

sowie als Designmatrix

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Diese Designmatrix ist identisch mit der aus Beispiel 42 und besitzt lediglich den Rang 4.

3. Schätzung von $\boldsymbol{\beta}$ mit der Maximum-Likelihood-Methode:

Da wieder 10 unabhängige Kopien der zugehörigen Zufallsvariablen $Y_i, i = 1, \dots, 8 = n$ (nach Umindizierung) vorliegen, kann bei einem unterstellten konstanten Störparameter τ der Parameter $\boldsymbol{\beta}$ durch Lösen von

V. Verallgemeinerte lineare Modelle

$$\sum_{i=1}^8 x_{ij} \left(\bar{Y}_i + \frac{1}{\mathbf{x}_i \boldsymbol{\beta}} \right) = 0, \quad j = 1, \dots, 6 = d$$

bestimmt werden. Man erhält dann das (nicht-lineare) Gleichungssystem

$$\sum_{i=1}^n \frac{x_{ij}}{\mathbf{x}_i \boldsymbol{\beta}} = -\sum_{i=1}^n x_{ij} \bar{Y}_i, \quad j = 1, \dots, d,$$

welches nach den Unbekannten $\boldsymbol{\beta}_k$ aufzulösen ist. Im betrachteten Beispiel lautet das Gleichungssystem

$$\begin{aligned} \frac{1}{\beta_1 + \beta_3 + \beta_5} + \frac{1}{\beta_1 + \beta_3 + \beta_6} + \frac{1}{\beta_1 + \beta_4 + \beta_5} + \frac{1}{\beta_1 + \beta_4 + \beta_6} &= -710,02 \\ \frac{1}{\beta_2 + \beta_3 + \beta_5} + \frac{1}{\beta_2 + \beta_3 + \beta_6} + \frac{1}{\beta_2 + \beta_4 + \beta_5} + \frac{1}{\beta_2 + \beta_4 + \beta_6} &= -580,03 \\ \frac{1}{\beta_1 + \beta_3 + \beta_5} + \frac{1}{\beta_1 + \beta_3 + \beta_6} + \frac{1}{\beta_2 + \beta_3 + \beta_5} + \frac{1}{\beta_2 + \beta_3 + \beta_6} &= -579,58 \\ \frac{1}{\beta_1 + \beta_4 + \beta_5} + \frac{1}{\beta_1 + \beta_4 + \beta_6} + \frac{1}{\beta_4 + \beta_4 + \beta_5} + \frac{1}{\beta_4 + \beta_4 + \beta_6} &= -710,47 \\ \frac{1}{\beta_1 + \beta_3 + \beta_5} + \frac{1}{\beta_1 + \beta_4 + \beta_5} + \frac{1}{\beta_2 + \beta_3 + \beta_5} + \frac{1}{\beta_2 + \beta_4 + \beta_5} &= -487,61 \\ \frac{1}{\beta_1 + \beta_3 + \beta_6} + \frac{1}{\beta_1 + \beta_4 + \beta_6} + \frac{1}{\beta_2 + \beta_3 + \beta_6} + \frac{1}{\beta_2 + \beta_4 + \beta_6} &= -802,44 \end{aligned}$$

Um eine eindeutige Lösung zu erhalten, kann wieder $\beta_4 := \beta_6 := 0$ fixiert werden. Die damit eindeutige Lösung ergibt sich - auf 5 Stellen gerundet - zu

k	1	2	3	4	5	6
$\boldsymbol{\beta}_k$	-0,00394	-0,00512	-0,00119	0,00000	-0,00316	0,00000

mit

$$\hat{\boldsymbol{\vartheta}} = \mathbf{D}\hat{\boldsymbol{\beta}} = -(0,00829 \quad 0,00513 \quad 0,00710 \quad 0,00394 \quad 0,00947 \quad 0,00631 \quad 0,00828 \quad 0,00512)^T$$

sowie

$$\hat{\boldsymbol{\mu}} = h(\hat{\boldsymbol{\vartheta}}) = (120,58 \quad 194,95 \quad 140,76 \quad 253,74 \quad 105,56 \quad 158,49 \quad 120,71 \quad 195,27)^T.$$

(Man vergleiche dieses Ergebnis mit den Mittelwerten aus der gegebenen Datentabelle.)

Bemerkung: Ohne die Fixierung gewisser Parameter ergibt sich auch hier aus *jeder* Lösung $\hat{\boldsymbol{\beta}}$ des Gleichungssystems stets der gleiche Prediktorwert $\hat{\boldsymbol{\vartheta}} = \mathbf{D}\hat{\boldsymbol{\beta}}$.

Bemerkung: Für eine Bestimmung von Lösungen der ML-Gleichungen werden in der Literatur oft das *Newton-Raphson-Verfahren* oder Varianten davon vorgeschlagen (vgl. etwa OHLSSON AND JOHANSSON (2010), Kapitel 3.2.3 oder DE JONG AND HELLER (2008), Kapitel 5.5). Dieses Verfahren setzt geeignete Differenzierbarkeitsannahmen voraus (die im Kontext von GLM's praktisch immer erfüllt sind). Ein stochastisches Suchverfahren, das ohne solche Voraussetzungen auskommt und in sehr viel allgemeineren Situationen angewendet werden kann, wird im Folgenden beschrieben.

Satz 15 (Stochastische Maximum-Suche). Es sei $D \subseteq \mathbb{R}^k$ für $k \in \mathbb{N}$ eine nicht-leere offene Borel-Menge von positivem Lebesgue-Maß, d.h. $m^k(D) > 0$, und $f : D \rightarrow \mathbb{R}$ eine beliebige stetige nicht-konstante Abbildung mit $M := \sup \{f(\mathbf{x}) | \mathbf{x} \in D\} = \max \{f(\mathbf{x}) | \mathbf{x} \in D\} < \infty$, d.h. f nimmt auf D das Supremum in mindestens einem Punkt an. Ferner sei $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$ eine Folge stochastisch unabhängiger, identisch wie \mathbf{X} verteilter Zufallsvektoren auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit Werten in D und der Eigenschaft, dass die Verteilung $P^{\mathbf{X}}$ von \mathbf{X} absolut stetig bezüglich m^k mit auf D positiver Lebesgue-Dichte sei. Mit der Setzung $M_n := \max \{f(\mathbf{X}_1), \dots, f(\mathbf{X}_n)\}$ für $n \in \mathbb{N}$ gilt dann:

$$\lim_{n \rightarrow \infty} M_n = M \text{ } P\text{-fast sicher.}$$

Beweis: Es sei $Y_n := f(\mathbf{X}_n)$ für $n \in \mathbb{N}$. Dann ist $\{Y_n\}_{n \in \mathbb{N}}$ eine Folge stochastisch unabhängiger D -wertiger identisch wie Y verteilter Zufallsvariablen auf (Ω, \mathcal{A}, P) mit $M_n = \max \{Y_1, \dots, Y_n\}$ für alle $n \in \mathbb{N}$. Die Folge $\{M_n\}_{n \in \mathbb{N}}$ ist trivialerweise (schwach) monoton wachsend und nach oben durch M beschränkt, also überall und damit insbesondere P -fast sicher konvergent gegen eine reelle Zufallsvariable Z . Wir zeigen, dass die Folge $\{M_n\}_{n \in \mathbb{N}}$ schwach gegen M konvergiert. Dann folgt nämlich $Z = M$ P -fast sicher (vgl. die Sätze 36 und 40 im Skript zur STOCHASTIK). Wir betrachten dazu die Verteilungsfunktion von M_n , die gegeben ist durch

$$F_{M_n}(z) = P\left(\bigcap_{i=1}^n \{Y_i \leq z\}\right) = \prod_{i=1}^n P(Y_i \leq z) = F_Y(z)^n \text{ für } z \in \mathbb{R}.$$

Hierbei bezeichnet F_Y die Verteilungsfunktion von Y . Dabei gilt nach Voraussetzung:

$$F_Y(z) = 1 - P(f(\mathbf{X}) > z) = 1 - P(\mathbf{X} \in f^{-1}((z, M])) \leq 1 - P(\mathbf{X} \in f^{-1}((z, M))) < 1 \text{ für alle } z < M,$$

da $f^{-1}((z, M)) \subseteq D$ wegen der Stetigkeit und Nicht-Konstanz von f sowie der Existenz des Maximums von f auf D ein offenes Intervall positiven Lebesgue-Maßes enthält. Damit folgt aber $\lim_{n \rightarrow \infty} F_{M_n}(z) = 0$ für alle $z < M$. Wegen $F_{M_n}(z) = 1$ für alle $z \geq M$ entspricht dies aber gerade der schwachen Konvergenz der Folge $\{M_n\}_{n \in \mathbb{N}}$ gegen M .

Bemerkungen:

1. Bei diesem Verfahren ist die Verteilung der Anzahl A_N der Updates (d.h. der gefundenen strikt größeren Werte in der Folge M_1, \dots, M_N) unabhängig von der Funktion f , mit dem Erwartungswert $E(A_N) = \sum_{k=2}^N \frac{1}{k} \approx \ln N$.

V. Verallgemeinerte lineare Modelle

2. Mit diesem Verfahren läßt sich auch - zumindest approximativ - ein Wert $\mathbf{x}_{\max} \in D$ bestimmen, für den $f(\mathbf{x}_{\max}) = M$ gilt. Dazu definiert man rekursiv die Folge $\{\mathbf{Z}_n\}_{n \in \mathbb{N}}$ in D durch

$$\mathbf{Z}_1 := \mathbf{X}_1, \quad \mathbf{Z}_{n+1} := \begin{cases} \mathbf{X}_n, & \text{wenn } f(\mathbf{X}_{n+1}) \leq f(\mathbf{X}_n) \\ \mathbf{X}_{n+1}, & \text{wenn } f(\mathbf{X}_{n+1}) > f(\mathbf{X}_n) \end{cases} \quad \text{für } n \in \mathbb{N}.$$

Dann ist $M_n = f(\mathbf{Z}_n)$ für alle $n \in \mathbb{N}$. Ist der Punkt $\mathbf{x}_{\max} \in D$ eindeutig bestimmt, so konvergiert die Folge $\{\mathbf{Z}_n\}_{n \in \mathbb{N}}$ fast sicher gegen \mathbf{x}_{\max} . Ist der Punkt $\mathbf{x}_{\max} \in D$ nicht eindeutig bestimmt, so enthält die Folge $\{\mathbf{Z}_n\}_{n \in \mathbb{N}}$ eine fast sicher konvergente Teilfolge.

3. Eine Beschleunigung des Verfahrens kann im Falle isolierter Maximumstellen von f dadurch erreicht werden, dass man im Laufe der Suche die Menge D geeignet verkleinert, bis schließlich nur eine einzige Maximumstelle von f in D übrig bleibt, oder indem man die "Suchverteilung" nach jedem größeren gefunden Funktionswert stärker um die letzte Näherungslösung herum konzentriert. Gilt beispielsweise $D = \mathbb{R}^k$ und ist $\{\mathbf{W}_n\}_{n \in \mathbb{N}}$ eine Folge stochastisch unabhängiger, identisch bivariat $\mathcal{N}(\mathbf{0}, \mathbf{I})$ -verteilter k -dimensionaler Zufallsvektoren (d.h. mit unabhängigen Komponenten), so kann man die folgende rekursiv definierte Zufallsfolge $\{\mathbf{Z}_n\}_{n \in \mathbb{N}}$ zur Maximumsuche verwenden:

$$\mathbf{Z}_0 := \mathbf{z}_0, \quad \mathbf{Z}_{n+1} := \begin{cases} \mathbf{Z}_n, & \text{wenn } f(\mathbf{Z}_n + \sigma \cdot \mathbf{W}_n) \leq f(\mathbf{Z}_n) \\ \mathbf{Z}_n + \sigma \cdot \mathbf{W}_n, & \text{wenn } f(\mathbf{Z}_n + \sigma \cdot \mathbf{W}_n) \geq f(\mathbf{Z}_n) \end{cases} \quad \text{für } n \in \mathbb{Z}^+,$$

wobei $\mathbf{z}_0 \in D$ ein beliebiger Startwert und $\sigma > 0$ ein frei wählbarer Parameter ist, der im Laufe des Verfahrens auch (geringfügig) abnehmen darf. Für allgemeinere Definitionsbereiche D kann man analog verfahren.

4. Satz 15 bleibt offensichtlich richtig, wenn sup und max durch inf und min ersetzt werden. (Man gehe einfach von f zu $-f$ über.)
5. Satz 15 läßt sich unter den getroffenen Annahmen mit der in 4. beschriebenen Modifikation auch für eine approximative Lösung einer Gleichung $f(\mathbf{x}) = 0$, $\mathbf{x} \in D$ heranziehen. Mit f ist nämlich auch $f^2 \geq 0$ stetig, und die Lösungen der gegebenen Gleichung sind identisch mit den Minimalstellen von f^2 . Gibt man jetzt ein kleines $\varepsilon > 0$ vor und setzt man $N_\varepsilon := \inf \{n \in \mathbb{N} \mid f^2(\mathbf{X}_n) < \varepsilon\}$, so ist N_ε eine P -fast sicher endliche Stoppzeit, und $\mathbf{X}_{N_\varepsilon}$ kann als approximative Lösung von $f(\mathbf{x}) = 0$, $\mathbf{x} \in D$ angesehen werden. Liegt in D nur eine einzige Nullstelle \mathbf{x}_0 von f vor, so gilt sogar $\lim_{\varepsilon \downarrow 0} \mathbf{X}_{N_\varepsilon} = \mathbf{x}_0$ P -fast sicher.
6. Eine Beschleunigung des Verfahrens kann im Falle isolierter Nullstellen von f wie oben dadurch erreicht werden, dass man im Laufe der Suche die Menge D geeignet verkleinert, bis schließlich nur eine einzige Nullstelle von f in D übrig bleibt, oder indem man die "Suchverteilung" nach jedem kleineren gefunden Funktionswert stärker um die letzte Näherungslösung herum konzentriert.
7. Für eine Anwendung zur Lösung eines Gleichungssystems $f_j(\mathbf{x}) = 0$, $\mathbf{x} \in D$, $j = 1, \dots, p$ für $p \in \mathbb{N}$ mit stetigen Funktionen f_1, \dots, f_p , die jeweils mindestens eine Nullstelle in D besitzen,

V. Verallgemeinerte lineare Modelle

kann man die Abbildung $f(\mathbf{x}) := \sum_{j=1}^p f_j^2(\mathbf{x})$, $\mathbf{x} \in D$ betrachten. Zur (approximativen) Bestimmung von Lösungen der ML-Gleichungen bei GLMs,

$$\sum_{i=1}^n \frac{x_{ij}}{\tau_i} (\bar{Y}_i - h(\mathbf{x}_i \boldsymbol{\beta})) = 0, \quad j = 1, \dots, d,$$

siehe oben, setzt man zweckmäßigerweise

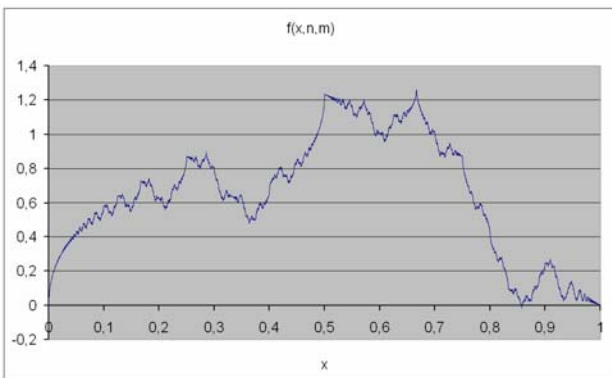
$$f(\boldsymbol{\beta}) = \sum_{j=1}^p \left(\sum_{i=1}^n \frac{x_{ij}}{\tau_i} h(\mathbf{x}_i \boldsymbol{\beta}) - \sum_{i=1}^n \frac{x_{ij}}{\tau_i} \bar{Y}_i \right)^2$$

und betrachtet dann zufällige Wahlen von $\boldsymbol{\beta}$ in einem geeigneten Bereich D . Die Lösungen zu den Beispielen 42 und 43 wurden auf diese Weise mit Hilfe eines Tabellenkalkulationsprogramms ermittelt.

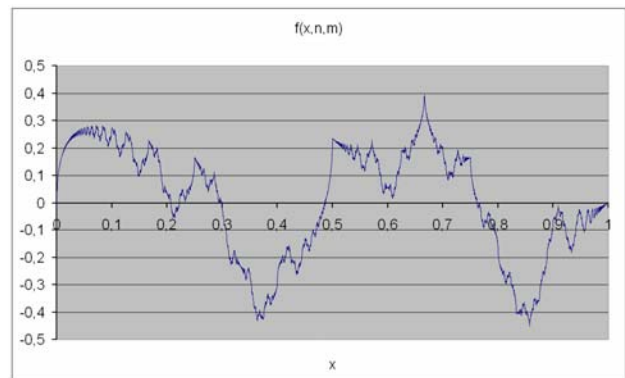
Beispiel 44. Es soll ein Maximum der Testfunktion

$$f(x, n, m) = \sum_{k=n}^m \frac{\sin(k^2 \pi x)}{k^2} \quad \text{für } 0 < x < 1 \text{ und ganzzahlige } 1 \leq n \leq m$$

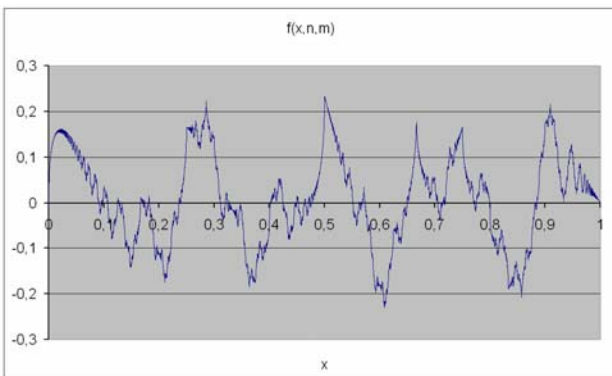
durch stochastische Suche bestimmt werden. Je nach Wahl von n und m besitzt die Testfunktion zahlreiche relative Extrema, wegen $f(0, n, m) = f(1, n, m) = 0$ auch (mindestens) ein absolutes Maximum. Für eine Veranschaulichung des Verfahrens wählen wir $m = 100$ und verschiedene Werte für n . Die folgenden Graphiken zeigen den Verlauf der ausgewählten Testfunktionen.



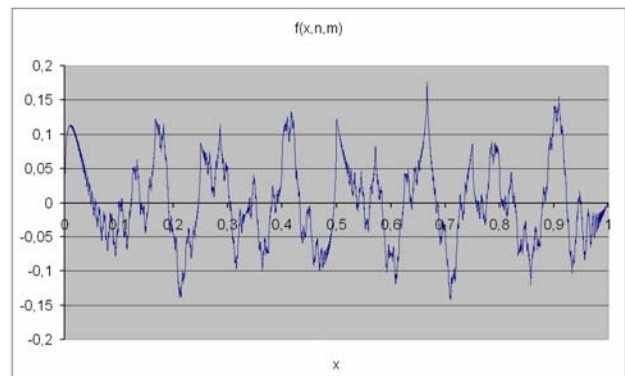
$n = 1$



$n = 2$



$n = 3$



$n = 4$

V. Verallgemeinerte lineare Modelle

Für die stochastische Suche verwenden wir zunächst stochastisch unabhängige Standard-Zufallszahlen X_1, \dots, X_N , die einer stetigen Gleichverteilung über dem Intervall $(0,1)$ entstammen. Wir definieren rekursiv

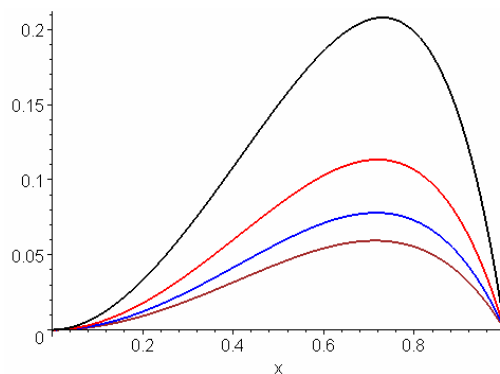
$$Z_1 := X_1, \quad Z_{n+1} := \begin{cases} Z_n, & \text{wenn } f(X_{n+1}) \leq f(Z_n) \\ X_{n+1}, & \text{wenn } f(X_{n+1}) > f(Z_n) \end{cases} \quad \text{für } n \in \mathbb{N}. \quad (*)$$

Dann ist $M_n = f(Z_n)$ für alle $n \in \mathbb{N}$. Die folgende Tabelle zeigt, dass - auch bei einem für übliche lokale Suchverfahren ungünstigem Ausgangspunkt - das absolute Maximum gefunden wird. Es wurden in den vier Fällen jeweils $N = 10000$ Suchschritte durchgeführt. Die Ergebnisse sind auf 4 Dezimalen gerundet. Jeder Fall wurde drei mal wiederholt.

n	Z_1	$f(Z_1, n, m)$	Z_N	$f(Z_N, n, m)$	A_N	x_{\max}	$f(x_{\max}, n, m)$
1	0,8305	0,1675	0,6667	1,2605	8	0,6667	1,2605
1	0,4048	0,7372	0,6667	1,2604	5		
1	0,2492	0,8299	0,6667	1,2596	7		
2	0,1112	0,2006	0,6667	0,3945	7	0,6667	0,3945
2	0,0789	0,2267	0,6667	0,3935	9		
2	0,5943	0,0492	0,6667	0,3944	10		
3	0,8026	-0,0671	0,5001	0,2337	8	0,5001	0,2345
3	0,2012	-0,1120	0,5001	0,2345	11		
3	0,6244	-0,0872	0,5001	0,2335	11		
4	0,1074	-0,0214	0,6667	0,1780	8	0,6667	0,1780
4	0,8452	0,0326	0,6666	0,1771	9		
4	0,7612	-0,0034	0,6667	0,1736	5		

Alternativ benutzen wir jetzt ein beschleunigtes Suchverfahren. Dazu verwenden wir (bedingte) Beta-Verteilungen, die nach jedem Update eine stärkere Konzentration um den zuletzt gewonnenen Suchpunkt herum bewirken. Dazu wählen wir in (*) nach einem Update - d.h. im Fall $f(Z_n) > f(Z_{n-1})$ - die Zufallsvariable Z_{n+1} nach der Verteilung $P^{Z_{n+1}}(\bullet | Z_n = x) = \mathcal{B}\left(\frac{c}{1-x}, \frac{c}{x}\right)$ für $0 < x < 1$, mit einer Konstanten $c \geq 1$. Es ist dann

$$E(Z_{n+1} | Z_n = x) = x \quad \text{und} \quad \text{Var}(Z_{n+1} | Z_n = x) = \frac{x^2(1-x^2)}{c+x(1-x)} \lesssim \frac{5}{20c+4}.$$



Graph von $\text{Var}(Z_{n+1} | Z_n = x)$ für $c = 1$ [oben]...4 [unten]

V. Verallgemeinerte lineare Modelle

Die folgende Tabelle zeigt das Ergebnis in den vier Fällen mit jeweils $N = 10000$ Suchschritten und $c = 3$. Die Ergebnisse sind wieder auf 4 Dezimalen gerundet. Jeder Fall wurde drei mal wiederholt.

n	Z_1	$f(Z_1, n, m)$	Z_N	$f(Z_N, n, m)$	A_N	x_{\max}	$f(x_{\max}, n, m)$
1	0,0287	0,3270	0,6667	1,2605	14	0,6667	1,2605
1	0,4606	0,8400	0,6667	1,2605	12		
1	0,0151	0,2485	0,6667	1,2604	12		
2	0,7886	-0,0934	0,6666	0,3932	10	0,6667	0,3945
2	0,1885	0,1178	0,6667	0,3944	14		
2	0,4474	-0,1539	0,6667	0,3945	9		
3	0,2102	-0,1664	0,5001	0,2342	10	0,5001	0,2345
3	0,6260	-0,0596	0,5001	0,2344	8		
3	0,7600	0,0469	0,5001	0,2339	14		
4	0,1567	0,0352	0,6667	0,1779	13	0,6667	0,1780
4	0,6538	0,0642	0,6667	0,1779	10		
4	0,4832	-0,0693	0,6667	0,1779	11		

Man erkennt - erwartungsgemäß - tendenziell mehr Updates und genauere Endwerte.

Beispiel 45. Es soll eine Lösung des Gleichungssystems

$$\begin{aligned} x^2 y + y^3 &= 1 \\ e^x y^2 - x &= 1 \end{aligned} \quad \text{für } (x, y) \in D = \mathbb{R}^2$$

durch stochastische Suche bestimmt werden. Dazu betrachten wir die Funktion

$$f(x, y) = (x^2 y + y^3 - 1)^2 + (e^x y^2 - x - 1)^2 \quad \text{für } (x, y) \in D = \mathbb{R}^2$$

und suchen ein Minimum. Für die stochastische Suche wird analog zur letzten Bemerkung - ausgehend von einem Startwert $\mathbf{z}_0 \in D$ - und einer Folge stochastisch unabhängiger, identisch bivariat $\mathcal{N}(\mathbf{0}, \mathbf{I})$ -verteilter Zufallsvektoren $\{\mathbf{W}_n\}_{n \in \mathbb{N}}$ (mit unabhängigen Komponenten) die folgende rekursiv definierte Zufallsfolge $\{\mathbf{Z}_n\}_{n \in \mathbb{N}}$ verwendet:

$$\mathbf{Z}_0 := \mathbf{z}_0, \quad \mathbf{Z}_{n+1} := \begin{cases} \mathbf{Z}_n, & \text{wenn } f(\mathbf{Z}_n + \sigma \cdot \mathbf{W}_n) \geq f(\mathbf{Z}_n) \\ \mathbf{Z}_n + \sigma \cdot \mathbf{W}_n, & \text{wenn } f(\mathbf{Z}_n + \sigma \cdot \mathbf{W}_n) < f(\mathbf{Z}_n) \end{cases} \quad \text{für } n \in \mathbb{Z}^+,$$

wobei wieder $\sigma > 0$ ein frei wählbarer Parameter ist.

Die folgenden Tabellen zeigen, dass bei entsprechendem Vorgehen (mindestens) eine Lösung gefunden wird. Es wurden nacheinander jeweils 10000 Suchschritte durchgeführt. Die Ergebnisse sind auf 5 Dezimalen gerundet.

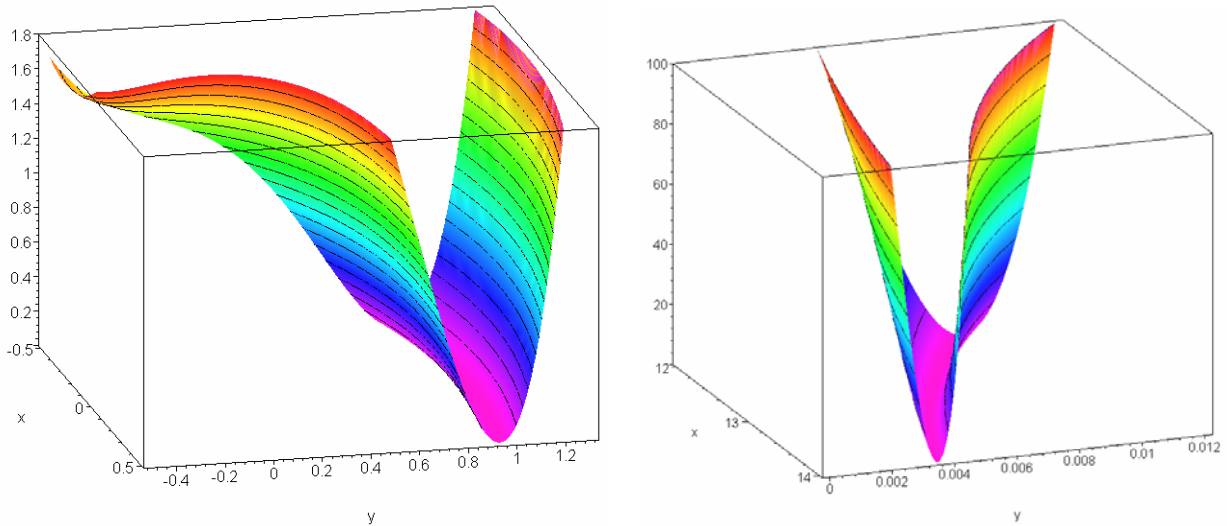
σ	Startwert \mathbf{z}_0	$f(\mathbf{z}_0)$	Endwert \mathbf{z}_N	$f(\mathbf{z}_N)$	Updates
2	(1,00000;1,00000)	1,51593	(-0,16168;0,98657)	0,00030	9
0,01	(-0,16168;0,98657)	0,00030	(-0,00787;0,99987)	0,00000	23
0,001	(-0,00787;0,99987)	0,00000	(-0,00297;0,99999)	0,00000	8

Die zugehörige genaue Lösung wäre hier $x = 0, y = 1$.

V. Verallgemeinerte lineare Modelle

σ	Startwert \mathbf{z}_0	$f(\mathbf{z}_0)$	Endwert \mathbf{z}_N	$f(\mathbf{z}_N)$	Updates
2	(-1,00000;0,00000)	1,00000	(-0,33520;0,96330)	0,00001	11
0,01	(-0,33520;0,96330)	0,00001	(-0,38804;0,94982)	0,00000	14

Die zugehörige genaue (numerische) Lösung wäre hier $x = -0,38636$, $y = 0,95029$.



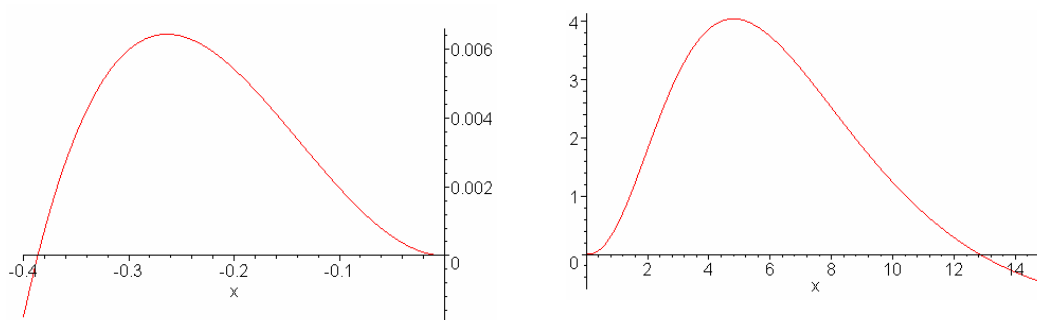
Graph der Funktion f

Bemerkung: Aus der ersten Gleichung folgt für jede Lösung $z = (x, y)$, dass $y \geq 0$ sein muß.

Substitution von $y = \sqrt{(1+x)e^{-x}}$ in die erste Gleichung ergibt

$$h(x) = \sqrt{(1+x)e^{-x}} (x^2 + (1+x)e^{-x}) - 1 = 0.$$

Diese Funktion hat genau drei Nullstellen.



Graph der Funktion h

Verzeichnis der Definitionen, Sätze, Lemmata und Beispiele

Definition	Satz	Lemma	Beispiel	Seite
			1	7
			2	8
			3	9
			4	11
1				12
		1		13
		2		15
			5	16
			6	18
			7	22
		3		24
		4		24
		5		27
			8	30
			9	33
		6		34
			10	36
		7		39
			11	40
			12	42
2				53
		8		53
		9		53
		10		54
3				55
	1			56
	2			60
			13	62
		11		66
		12		66
	3			67
		13		68
	4			72
			14	74
4				75
		14		76
5				76
		15		76
	5			77
	6			78
		16		79
		17		83
			15	84
			16	85
		18		89
		19		91
		20		92
			17	93

Definition	Satz	Lemma	Beispiel	Seite
			18	95
			19	97
			20	99
			21	103
6				104
7				104
8				105
		21		105
			22	109
9				109
10				109
		22		110
			23	110
		23		110
			24	111
			25	112
		24		115
		25		116
	7			119
			26	120
		26		121
	8			123
		27		123
			27	124
	9			126
			28	126
			29	130
		28		131
		29		132
			30	133
		30		134
			31	136
		31		137
		32		137
	10			138
			32	139
		33		140
			33	141
11				141
12				143
	11			143
13				145
		34		145
14				145
		35		146
		36		147
	12			148
	13			148

Verzeichnis der Definitionen, Sätze, Lemmata und Beispiele

Definition	Satz	Lemma	Beispiel	Seite
15				149
		37		150
		38		151
16				151
		39		152
		40		152
			34	153
		41		153
		42		154
	14			155
			35	155
			36	156
			37	156
			38	160

Definition	Satz	Lemma	Beispiel	Seite
		43		162
			39	163
			40	164
			41	165
		44		168
17				169
		45		170
		46		171
			42	172
			43	175
	15			178
			44	180
			45	182

Literatur

- [1] D. ANDERSON ET AL. (2007): *A Practitioner's Guide to Generalized Linear Models*. A foundation for theory, interpretation and application. 3rd ed., Watson Wyatt, London.
- [2] J. BANNISTER (1998): *Alternative Risk Financing*. Changing the Face of Insurance. Jim Bannister Developments Ltd.
- [3] P.J. BOLAND (2007): *Statistical and Probabilistic Methods in Actuarial Science*. Chapman & Hall, Boca Raton.
- [4] H. BÜHLMANN (1970): *Mathematical Methods in Risk Theory*. Springer, Berlin.
- [5] H. BÜHLMANN UND A. GISLER (2005): *A Course in Credibility Theory and its Applications*. Springer, Berlin.
- [6] C. COTTIN UND S. DÖHLER (2013): *Risikoanalyse*. Modellierung, Beurteilung und Management von Risiken mit Praxisbeispielen. 2. Aufl., Vieweg+Teubner, Wiesbaden.
- [7] P. DE JONG AND G.Z. HELLER (2008): *Generalized Linear Models for Insurance Data*. International Series on Actuarial Science, Cambridge Univ. Press, Cambridge.
- [8] D.C.M. DICKSON (2005): *Insurance Risk and Ruin*. International Series on Actuarial Science, Cambridge Univ. Press, Cambridge.
- [9] D.C.M. DICKSON, M.R. HARDY, AND H.R. WATERS (2009): *Actuarial Mathematics for Life Contingent Risks*. International Series on Actuarial Science, Cambridge Univ. Press, Cambridge.
- [10] H.R. DIENST (Hrsg.)(1988): *Mathematische Verfahren der Rückversicherung*. Schriftenreihe Angewandte Versicherungsmathematik, Heft 19. VVW, Karlsruhe.
- [11] W. DONG (2001): *Building a More Profitable Portfolio*. Modern Portfolio Theory with Application to Catastrophe Insurance. Reactions Publishing Group, London.
- [12] G. DRUDE (1988): *Ausgewählte Themen der kollektiven Risikothorie*. Schriftenreihe Angewandte Versicherungsmathematik, Heft 18. VVW, Karlsruhe.
- [13] H.U. GERBER (1986): *Lebensversicherungsmathematik*. Springer, Berlin.
- [14] H.U. GERBER (1997): *Life Insurance Mathematics*. 3rd ed., Springer, Berlin.

- [15] H.U. GERBER (1979): *An Introduction to Mathematical Risk Theory*. Richard D.Irwin, Homewood, IL.
- [16] R.J. GRAY AND S.M. PITTS (2012): *Risk Modelling in General Insurance*. From Principles to Practice. International Series on Actuarial Science, Cambridge Univ. Press, Cambridge.
- [17] P. GROSSI AND H. KUNREUTHER (2005): *Catastrophe Modeling: A New Approach to Managing Risk*. Springer, N.Y.
- [18] W.-R. HEILMANN (1987): *Grundbegriffe der Risikotheorie*. VVW, Karlsruhe.
- [19] D. FARNY (2000): *Versicherungsbetriebslehre*. 3. Aufl., VVW, Karlsruhe.
- [20] CH. FÜHRER UND A. GRIMMER (2006): *Einführung in die Lebensversicherungsmathematik*. VVW, Karlsruhe.
- [21] W. GRUNDMANN (1996): *Finanz- und Versicherungsmathematik*. Teubner, Stuttgart.
- [22] W. GRUNDMANN UND B. LUDERER (2001): *Formelsammlung Finanzmathematik, Versicherungsmathematik, Wertpapieranalyse*. Teubner, Stuttgart.
- [23] M. HELBIG (Hrsg.)(1987): *Beiträge zum versicherungsmathematischen Grundwissen*. Schriftenreihe Angewandte Versicherungsmathematik, Heft 12. VVW, Karlsruhe.
- [24] E. HELTEN (Hrsg.)(1989): *Beiträge zur Credibility-Theorie*. Schriftenreihe Angewandte Versicherungsmathematik, Heft 22. VVW, Karlsruhe.
- [25] CH. HIPPE UND H. MICHEL (1990): *Risikotheorie: Stochastische Modelle und Statistische Methoden*. Schriftenreihe Angewandte Versicherungsmathematik, Heft 24. VVW, Karlsruhe.
- [26] R. KAAS, M. GOOVAERTS, J. DHAENE AND M. DENUIT (2001): *Modern Actuarial Risk Theory*. Kluwer, Dordrecht.
- [27] P. KAKIES ET AL. (1985): *Methodik von Sterblichkeitsuntersuchungen*. Schriftenreihe Angewandte Versicherungsmathematik, Heft 15. VVW, Karlsruhe.
- [28] S.A. KLUGMAN, H.H. PANJER AND G.E. WILLMOT (2008): *Loss Models - From Data to Decisions*. 3rd Ed., Wiley, New York.
- [29] E. KNOBLOCH UND J.-M. GRAF VON DER SCHULENBURG (Hrsg.) (2000): *Gottfried Wilhelm Leibniz. Hauptschriften zur Versicherungs- und Finanzmathematik*. Akademie-Verlag, Berlin.
- [30] P. KOCH (1998): *Geschichte der Versicherungswissenschaft in Deutschland*. VVW, Karlsruhe.
- [31] M. KOLLER (2000): *Stochastische Modelle in der Lebensversicherung*. Springer, Berlin.
- [32] P. LIEBWEIN (2009): *Klassische und moderne Formen der Rückversicherung*. 2. Aufl., VVW, Karlsruhe.
- [33] TH. MACK (2002): *Schadenversicherungsmathematik*. 2. Aufl., Schriftenreihe Angewandte Versicherungsmathematik, Heft 28. VVW, Karlsruhe.
- [34] P. MCCULLAGH AND J.A. NELDER (1989): *Generalized Linear Models*. Chapman and Hall, 2nd Ed., New York.
- [35] T. MIKOSCH (2009): *Non-Life Insurance Mathematics*. An Introduction with the Poisson Process. Springer, 2nd Ed., Berlin.
- [36] H. MILBRODT UND M. HELBIG (1999): *Mathematische Methoden der Personenversicherung*. De Gruyter, Berlin.
- [37] H. MILBRODT (2005): *Aktuarielle Methoden der deutschen Privaten Krankenversicherung*. Schriftenreihe Angewandte Versicherungsmathematik, Heft 34. VVW, Karlsruhe.
- [38] E. OHLSSON AND B. JOHANSSON (2010): *Non-Life Insurance Pricing with Generalized Linear Models*. Springer, Berlin.

- [39] K.M. ORTMANN (2009): *Praktische Lebensversicherungsmathematik*. Mit zahlreichen Beispielen, Abbildungen und Anwendungen. Vieweg+Teubner, Wiesbaden.
- [40] CH. PFEIFFER (1994): *Einführung in die Rückversicherung*. Gabler, Wiesbaden.
- [41] H. PRUSCHA (2000): *Vorlesungen über Mathematische Statistik*. B.G. Teubner Verlag, Stuttgart.
- [42] M. RADTKE UND K.D. SCHMIDT (2012): *Handbuch zur Schadenreservierung*. 2. Auflage, VVW, Karlsruhe.
- [43] R.-D. REISS AND M. THOMAS (2001): *Statistical Analysis of Extreme Values*. With Applications to Insurance, Finance, Hydrology and Other Fields. 2. Aufl., Birkhäuser, Basel.
- [44] T. ROLSKI, H. SCHMIDLI, V. SCHMIDT AND J. TEUGELS (1998): *Stochastic Processes for Insurance and Finance*. Wiley, New York.
- [45] K.D. SCHMIDT (2002): *Versicherungsmathematik*. Springer, Berlin.
- [46] W. SCHMID-GROTJOHANN (1995): *Das Änderungsrisiko in der Privaten Krankenversicherung*. Schriftenreihe "Versicherung und Risikoforschung" des Instituts für betriebswirtschaftliche Risikoforschung und Versicherungswirtschaft der Ludwig-Maximilians-Universität München, Band 16. Gabler Verlag, Wiesbaden.
- [47] A. SCHWEPKE (2001): *Rückversicherung*. Grundlagen und aktuelles Wissen. VVW, Karlsruhe.
- [48] E. STRAUB (1988): *Non-Life Insurance Mathematics*. Springer, Berlin.
- [49] B. SUNDT (1994): *An Introduction to Non-Life Insurance*. VVW, Karlsruhe.
- [50] G. TAYLOR (2000): *Loss Reserving*. An Actuarial Perspective. Kluwer, Dordrecht.
- [51] A. TOSBERG (1957): *Die Technik der Lebensversicherung*. Selbstverlag, Berlin-Dahlem.
- [52] G. TAYLOR (2000): *Loss Reserving*. An Actuarial Perspective. Kluwer Academic Publishers, Boston.
- [53] Y.-K. TSE (2009): *Nonlife Actuarial Models*. Theory, Methods and Evaluation. International Series on Actuarial Science, Cambridge Univ. Press, Cambridge.
- [54] K. WOLFSDORF (1997): *Versicherungsmathematik*. Teil 1: Personenversicherung. 2. Aufl., Teubner, Stuttgart.
- [55] K. WOLFSDORF (1988): *Versicherungsmathematik*. Teil 2: Theoretische Grundlagen, Risikotheorie, Sachversicherung. Teubner, Stuttgart.
- [56] M.V. WÜTHRICH AND M. MERZ (2008): *Stochastic Claims Reserving Methods in Insurance*. Wiley, New York.

Danksagung

Ein erheblicher Teil unserer versicherungs- und finanzmathematischen Arbeit in Forschung und Lehre ist regelmäßig durch Spenden aus der nationalen und internationalen Versicherungs- und Finanzwirtschaft, insbesondere durch den *Verein zur Förderung der Versicherungs- und Finanzmathematik an der Universität Oldenburg e.V.* unterstützt worden. Dafür gebührt allen Institutionen mein herzlicher Dank.

Besonders danken möchte ich in an dieser Stelle auch Frau Dipl.-Math. Mirjam Langmann und Frau Dr. Doreen Straßburger für eine gründliche Durchsicht von früheren Teilen des Manuskripts und die Erstellung elektronischer Lehrinhalte im Rahmen des *estat*-Projektes sowie Herrn Dipl.-Math. Martin Hampel für die kompetente und engagierte Begleitung der Erstellung der aktuellen Fassung des Skripts.