Risiko- und Extremwerttheorie

Vorlesungsskript

Prof. Dr. Benedikt Jahnel

Institut für Mathematische Stochastik
Technische Universität Braunschweig

Grundlage: J. Blath, M. Ortgiese, M. Scheutzow,
Stochastische Modelle der Versicherungsmathematik,
Mathematik Kompakt, Birkhäuser/Springer 2025.

Stand: 27. Oktober 2026

Vorbemerkungen

Gegenstand. Die Vorlesung behandelt stochastische Modelle für Versicherungsrisiken. Im ersten, kürzeren Teil geht es um das Risiko eines einzelnen Lebens: Zahlungsströme werden bewertet, Prämien nach dem Äquivalenzprinzip bestimmt, das Deckungskapital über die Thielesche Differentialgleichung beschrieben und die Schwankung des Verlusts mit dem Satz von Hattendorff quantifiziert. Im zweiten, längeren Teil steht die Schadenversicherung im Vordergrund: der Gesamtschaden eines Portfolios, seine Abweichungen vom Erwartungswert, die Ruinwahrscheinlichkeit eines Versicherers in der Zeit, Prämienprinzipien und Risikomaße. Ein roter Faden ist dabei die Unterscheidung zwischen *leichten* und *schweren* Verteilungsenden: Sie entscheidet, ob große Abweichungen durch viele kleine Schäden oder durch einen einzigen extremen Schaden entstehen, und sie bildet die Brücke zur Extremwerttheorie.

Grundlage und Bezeichnungen. Das Skript folgt dem Aufbau des Lehrbuchs [BOS25] von Blath, Ortgiese und Scheutzow. Verweise der Form [BOS25, Satz 5.58] beziehen sich auf die Nummerierung dort; im Buch finden sich ausführlichere Beweise, zusätzliche Beispiele und weitere Übungsaufgaben. Die Notation stimmt weitgehend mit dem Buch überein; insbesondere bezeichnet $\overline{F} = 1 - F$ den *Tail* einer Verteilungsfunktion F und „wachsend“ bedeutet stets „nicht fallend“.

Vorkenntnisse. Vorausgesetzt werden die Stochastik-Grundvorlesungen sowie Maß- und Integrationstheorie (insbesondere das Lebesgue–Stieltjes-Integral, siehe Anhang A). Bedingte Erwartungen und Martingale werden an den Stellen wiederholt, an denen sie gebraucht werden.

Zur Extremwerttheorie. Das Buch klammert die klassische Extremwerttheorie bewusst aus und verweist dafür auf Embrechts, Klüppelberg und Mikosch [EKM97]. Die Grundlagen – schwere Verteilungsenden, reguläre Variation, subexponentielle Verteilungen und das Prinzip des einzelnen großen Sprungs – werden aber in Kapitel 5 und 6 sowie in Abschnitt 8.3 ausführlich behandelt. Darauf baut Kapitel 9 auf, das über das Buch hinausgeht: Grenzverteilungen von Maxima, die verallgemeinerte Pareto-Verteilung und die statistische Schätzung von Tails, Quantilen und Expected Shortfall. Sie folgen [EKM97], [dHF06] und [MFE15]. Um dafür Platz zu schaffen, ist die Lebensversicherungsmathematik (Teil I) auf zwei Termine verdichtet.

Inhaltsverzeichnis

Vorbemerkungen **1**

I Lebensversicherungsmathematik 5

1 Finanzmathematische Grundlagen **6**
1.1 Kapitalfunktionen und Zinsintensität . . . 6
1.2 Zahlungsströme und Barwert . . . 7
1.3 Äquivalenzprinzip und Deckungskapital . . . 8

2 Lebensversicherungsverträge und Nettoprämien **10**
2.1 Beschreibung des Todesfallrisikos . . . 10
2.2 Das Modell eines Lebensversicherungsvertrags . . . 11
2.3 Nettoprämien . . . 13

3 Deckungskapital und Thielesche Gleichungen **15**
3.1 Prospektives und retrospektives Nettodeckungskapital . . . 15
3.2 Die Thielesche Differentialgleichung . . . 16
3.3 Die Thielesche Integralgleichung . . . 17
3.4 Erweiterungen des Modells . . . 17

4 Der Satz von Hattendorff **19**
4.1 Varianz bei der Nettoeinmalprämie . . . 19
4.2 Martingale und der kompensierte Einheitsleistungsstrom . . . 19
4.3 Der Verlustprozess . . . 21
4.4 Der Satz . . . 21

II Schadenversicherungsmathematik 23

5 Statische Risikomodelle **24**
5.1 Individuelles und kollektives Modell . . . 24
5.2 Schadenhöhen und Schadenanzahlen . . . 24
5.3 Die Verteilung des Gesamtschadens . . . 26
5.4 Die Panjer-Rekursion . . . 26

6 Abweichungen des Gesamtschadens vom Erwartungswert **29**
6.1 Die Waldschen Gleichungen . . . 29
6.2 Polynomielle Schranken . . . 30
6.3 Große Abweichungen . . . 30

7 Erneuerungstheorie **33**
7.1 Erneuerungsprozesse . . . 33
7.2 Erneuerungsfunktion und Erneuerungsgleichung . . . 33
7.2.1 Die Erneuerungsfunktion . . . 33
7.2.2 Die Erneuerungsgleichung . . . 34
7.2.3 Restwartezeit und Wartezeitparadoxon . . . 34

7.2.4 Der elementare Erneuerungssatz . . . 35
7.3 Verzögerte und stationäre Erneuerungsprozesse . . . 35
7.4 Blackwell und der fundamentale Erneuerungssatz . . . 36
7.5 Grenzverteilungen . . . 37

8 Ruintheorie **39**
8.1 Dynamische Modelle der kollektiven Risikotheorie . . . 39
8.2 Ruin unter der Lundberg-Bedingung . . . 40
8.2.1 Die Lundberg-Ungleichung . . . 40
8.2.2 Eine Integralgleichung für die Ruinfunktion . . . 41
8.2.3 Die Cramér–Lundberg-Approximation . . . 42
8.3 Subexponentielle Verteilungen und Ruin durch Großschäden . . . 43
8.3.1 Subexponentielle Verteilungen . . . 43
8.3.2 Reguläre Variation . . . 44
8.3.3 Asymptotik der Ruinfunktion . . . 45

III Extremwerttheorie **48**

9 Extremwerttheorie **49**
9.1 Maxima und Poisson-Approximation . . . 49
9.2 Grenzverteilungen für Maxima . . . 49
9.3 Beispiele und Anziehungsbereiche . . . 50
9.4 Exzessverteilung und mittlere Exzessfunktion . . . 52
9.5 Die verallgemeinerte Pareto-Verteilung . . . 53
9.6 Die Peaks-over-Threshold-Methode . . . 53
9.7 Der Hill-Schätzer . . . 54

IV Bewertung von Risiken **56**

10 Prämienprinzipien und Risikomaße **57**
10.1 Quantile . . . 57
10.2 Prämienprinzipien . . . 57
10.3 Risikomaße . . . 58

11 Rückversicherung und Erfahrungstarifierung **61**
11.1 Risikoteilung und Rückversicherung . . . 61
11.2 Das heterogene Modell . . . 61
11.3 Credibility nach Bühlmann . . . 62

Rückblick **64**

A Hilfsmittel aus der Maß- und Wahrscheinlichkeitstheorie **65**
A.1 Lebesgue–Stieltjes-Integrale . . . 65
A.2 Bedingte Erwartungen . . . 65

Literaturverzeichnis **67**

Teil I

Lebensversicherungsmathematik

Kapitel 1

Finanzmathematische Grundlagen

In der Lebensversicherung sind die Verträge langfristig. Eine Prämie, die heute gezahlt wird, und eine Leistung, die in dreißig Jahren fällig wird, lassen sich erst vergleichen, wenn die Verzinsung berücksichtigt ist. Dieses Kapitel ist noch rein deterministisch: Wir bewerten feste Zahlungen. Der maßtheoretische Zugang über Lebesgue–Stieltjes-Integrale erlaubt es dabei, zeitdiskrete und zeitstetige Zahlungen einheitlich zu behandeln; im nächsten Kapitel werden die Zahlungszeitpunkte zufällig.

1.1 Kapitalfunktionen und Zinsintensität

Definition 1.1 (Kapitalfunktion)**.** Eine wachsende, rechtsstetige Funktion $K\colon [0,\infty) \to [1,\infty)$ mit $K(0) = 1$ heißt *Kapitalfunktion*. Die Zahl $r := K(1)$ heißt *Aufzinsungsfaktor*, $i := r - 1$ der zugehörige (effektive Jahres-)*Zinssatz*.

Die Zahl $K(t)$ ist der Wert, auf den eine zur Zeit 0 angelegte Geldeinheit bis zur Zeit t anwächst. Ein Betrag A, der zur Zeit s gezahlt wird, hat zur Zeit t den Wert $A\,K(t)/K(s)$; insbesondere ist sein Wert zur Zeit 0 gleich $A/K(s)$ („Abzinsen“).

Beispiel 1.2. (a) *Jährliche Verzinsung mit Zinssatz $i \geq 0$:* $K(t) = (1+i)^{\lfloor t \rfloor}$. Diese Kapitalfunktion springt zu ganzzahligen Zeiten.

(b) *Unterjährige Verzinsung:* Wird das Jahr in m Perioden geteilt und am Ende jeder Periode der Zins $i^{(m)}/m$ gutgeschrieben, so ist $K(t) = (1+i^{(m)}/m)^{\lfloor mt \rfloor}$. Der effektive Jahreszins $(1+i^{(m)}/m)^m - 1$ ist für $m \geq 2$ und $i^{(m)} > 0$ nach dem binomischen Lehrsatz größer als der nominelle Zins $i^{(m)}$, denn $(1+i^{(m)}/m)^m - 1 = i^{(m)} + \sum_{k=2}^{m} \binom{m}{k} (i^{(m)}/m)^k$.

(c) *Stetige Verzinsung mit Zinsrate $\delta \geq 0$:* $K(t) = \mathrm{e}^{\delta t}$. Sie entsteht aus (b) mit $i^{(m)} = \delta$ im Limes $m \to \infty$. Der effektive Jahreszins ist $\mathrm{e}^{\delta} - 1$, für $\delta > 0$ also größer als δ; für $\delta = 0{,}05$ etwa 5,13 %.

Zeitlich schwankende Zinsen beschreibt man über den momentanen relativen Zuwachs der Kapitalfunktion.

Definition 1.3 (Zinsintensität)**.** Besitzt K eine Darstellung $K(t) = 1 + \int_{(0,t]} k(s)\,\mathrm{d}s$ mit messbarem $k \geq 0$, so heißt $\varphi(t) := k(t)/K(t)$ die *Zinsintensität* von K.

Eine solche Darstellung existiert genau dann, wenn K absolut stetig ist (Hauptsatz für absolut stetige Funktionen, Anhang A). Dann ist K fast überall differenzierbar mit $K' = k$ und $\varphi(t) \approx (K(t+h) - K(t))/(hK(t))$ für kleine h. Für $K(t) = \mathrm{e}^{\delta t}$ ist $\varphi \equiv \delta$.

Proposition 1.4 ([BOS25, Prop. 1.8])**.** *Besitzt K eine Zinsintensität φ, so gilt*

$$K(t) = \exp\Big(\int_0^t \varphi(s)\,\mathrm{d}s\Big), \qquad t \geq 0.$$

Beweis. Ist k stetig, so ist K stetig differenzierbar und $(\log K)' = K'/K = \varphi$. Integration mit $\log K(0) = 0$ liefert die Behauptung. Im allgemeinen Fall ist $\log K$ absolut stetig (denn log ist auf kompakten Teilintervallen von $[1,\infty)$ Lipschitz-stetig), die Rechnung gilt fast überall, und der Hauptsatz für absolut stetige Funktionen liefert wieder $\log K(t) = \int_0^t \varphi(s)\,\mathrm{d}s$. □

Bemerkung 1.5 (Kumulierte Zinsintensität). Im absolut stetigen Fall setzt man $\Phi(t) := \int_0^t \varphi(s)\,\mathrm{d}s$, so dass $K = \mathrm{e}^{\Phi}$. Ohne Dichte definiert man

$$\Phi(t) := \int_{(0,t]} \frac{1}{K(s-)}\,\mathrm{d}K(s), \qquad K(s-) := \lim_{u\uparrow s} K(u).$$

Dann gilt im Allgemeinen *nicht* $K = \mathrm{e}^{\Phi}$: Für $K(t) = (1+i)^{\lfloor t\rfloor}$ ist $\Phi(1) = K(1) - K(0) = i$, aber $K(1) = 1 + i \neq \mathrm{e}^{i}$. Die Wahl des linksseitigen Grenzwerts $K(s-)$ im Integranden wird in Definition 2.1 bei der Sterblichkeitsintensität verständlich.

Abbildung 1.1 vergleicht die drei Verzinsungsarten aus Beispiel 1.2 bei gleichem effektivem Jahreszins.

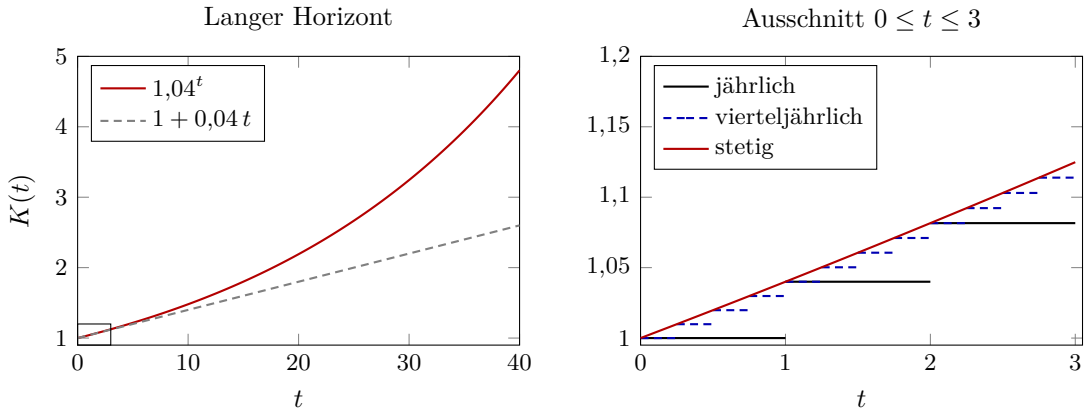

Abbildung 1.1: Kapitalfunktionen mit effektivem Jahreszins 4 %. Links: Über 40 Jahre wird das exponentielle Wachstum $K(t) = 1{,}04^t = \mathrm{e}^{\delta t}$, $\delta = \log 1{,}04$, deutlich; zum Vergleich einfache (lineare) Verzinsung $1 + 0{,}04\,t$, die nach 40 Jahren nur 2,6 statt 4,8 erreicht. Rechts: Auf kurzen Zeiträumen ist die Krümmung praktisch unsichtbar (auf $[0,3]$ weicht $1{,}04^t$ höchstens um etwa 0,002 von der Sehne ab); dafür sieht man den Unterschied zwischen jährlicher Verzinsung (Treppe $1{,}04^{\lfloor t\rfloor}$), vierteljährlicher Verzinsung mit Quartalsfaktor $1{,}04^{1/4}$ und stetiger Verzinsung. Zu ganzzahligen Zeiten stimmen alle überein, dazwischen gilt $1{,}04^{\lfloor t\rfloor} \le K_{\text{viertelj.}}(t) \le 1{,}04^t$; eine unterjährige Zahlung wird daher bei jährlicher Verzinsung am schwächsten abgezinst.

1.2 Zahlungsströme und Barwert

Definition 1.6 (Zahlungsströme)**.** Ein *gerichteter Zahlungsstrom* ist eine rechtsstetige, wachsende Funktion $Z\colon [0,\infty) \to [0,\infty)$; die Menge dieser Funktionen sei $\mathcal{Z}_g$. Der Wert $Z(t)$ ist die Summe aller bis einschließlich t geleisteten Zahlungen. Differenzen $Z = Z_1 - Z_2$ mit $Z_1, Z_2 \in \mathcal{Z}_g$ heißen *ungerichtete Zahlungsströme*.

Wir arbeiten fast ausschließlich mit Paaren gerichteter Ströme; das entspricht auch der Praxis, in der jeder Vertragspartner seinen eigenen Zahlungsstrom leistet, und erspart uns signierte Maße.

Beispiel 1.7 (Diskrete Zeitrente)**.** Zahlungen $z_j \ge 0$ zu Zeiten $0 \le t_0 < t_1 < \dots$ mit $t_j \to \infty$ ergeben $Z(t) = \sum_j z_j \mathbb{1}\{t \ge t_j\}$. Ihr Wert zur Zeit 0 ist

$$b(Z) = \sum_j \frac{z_j}{K(t_j)} = \sum_j \frac{\Delta Z(t_j)}{K(t_j)}, \qquad \Delta Z(t) := Z(t) - Z(t-).$$

Die rechte Seite kann als Integral der Funktion $1/K$ bezüglich des Maßes mit Atomen z_j in t_j interpretiert werden. Das legt die allgemeine Definition nahe.

Definition 1.8 (Barwert)**.** Der *Barwert* von $Z \in \mathcal{Z}_g$ bezüglich der Kapitalfunktion K ist das Lebesgue–Stieltjes-Integral

$$b(Z) := \int_{[0,\infty)} \frac{1}{K(s)}\,\mathrm{d}Z(s) \in [0,\infty].$$

Für $Z = Z_1 - Z_2$ setzt man $b(Z) \coloneqq b(Z_1) - b(Z_2)$, falls einer der beiden Terme endlich ist. Für $t \geq 0$ heißt $b(Z)(t) \coloneqq \int_{[0,t]} K(s)^{-1}\,\mathrm{d}Z(s)$ der Barwert bis t und $s(Z)(t) \coloneqq K(t)\,b(Z)(t)$ der *Zeitwert* zur Zeit t.

In der Praxis zerfällt ein Zahlungsstrom meist in einen diskreten und einen absolut stetigen Anteil, $Z = Z^{(d)} + Z^{(s)}$ mit $Z^{(s)}(t) = Z^{(s)}(0) + \int_0^t z(s)\,\mathrm{d}s$. Dann ist

$$\int g\,\mathrm{d}Z = \sum_j g(t_j)z_j + g(0)Z^{(s)}(0) + \int_0^\infty g(s)z(s)\,\mathrm{d}s$$

für messbares $g \geq 0$, und man kommt ohne die allgemeine Integrationstheorie aus.

1.3 Äquivalenzprinzip und Deckungskapital

Definition 1.9 (Äquivalenz)**.** Zwei Zahlungsströme heißen *äquivalent* (bezüglich K), wenn sie denselben endlichen Barwert haben.

Ein Finanz- oder Versicherungsprodukt beschreiben wir durch ein Paar (Z_L, Z_P) gerichteter Ströme: den *Leistungsstrom* Z_L (Zahlungen des Unternehmens an den Kunden) und den *Prämienstrom* Z_P (Zahlungen des Kunden). Bei Vertragsabschluss soll (Z_L, Z_P) äquivalent sein; im Laufe der Zeit geht diese Balance aber verloren, und das Unternehmen muss Kapital zurückstellen.

Definition 1.10 (Deckungskapital)**.** Sei mindestens einer der Barwerte $b(Z_L)$, $b(Z_P)$ endlich. Das *prospektive Deckungskapital* zur Zeit $t \geq 0$ ist

$$V(t) \coloneqq K(t)\Big(\int_{[t,\infty)} \frac{\mathrm{d}Z_L(s)}{K(s)} - \int_{[t,\infty)} \frac{\mathrm{d}Z_P(s)}{K(s)}\Big),$$

das *retrospektive Deckungskapital* ist

$${}^{(r)}V(t) \coloneqq K(t)\Big(\int_{[0,t)} \frac{\mathrm{d}Z_P(s)}{K(s)} - \int_{[0,t)} \frac{\mathrm{d}Z_L(s)}{K(s)}\Big).$$

Sind Z_L, Z_P äquivalent, spricht man vom *Nettodeckungskapital*.

Das Deckungskapital $V(t)$ ist der Betrag, den das Unternehmen zur Zeit t vorhalten muss, um die noch ausstehenden Leistungen abzüglich der noch ausstehenden Prämien zu decken. Ist $V \geq 0$, spricht man von einem *Sparplan*; ist $V \leq 0$, von einem *Kreditvertrag* mit Restschuld $-V(t)$.

Lemma 1.11 ([BOS25, Lemma 1.22])**.** *Sind Z_L, Z_P äquivalent, so gilt $V(0) = 0$ und $V(t) = {}^{(r)}V(t)$ für alle $t \geq 0$.*

Beweis. Es gilt $V(0) = b(Z_L) - b(Z_P) = 0$. Für $t \geq 0$ zerlegen wir $\int_{[t,\infty)} = \int_{[0,\infty)} - \int_{[0,t)}$ in beiden Termen; die Gesamtbarwerte heben sich wegen der Äquivalenz weg, und es bleibt ${}^{(r)}V(t)$ stehen. □

Die beiden Darstellungen haben unterschiedliche Stärken: Die prospektive Form zeigt, wofür das Kapital benötigt wird; die retrospektive Form zeigt, wie es aus den bisherigen Zahlungen angespart wurde.

Beispiel 1.12 (Sparvertrag)**.** Eine Kundin legt zur Zeit 0 den Betrag $A_1 = 20\,000$ und zur Zeit 3 weitere $A_2 = 10\,000$ an; nach sechs Jahren erhält sie die Einmalzahlung B. Bei stetiger Verzinsung mit $\delta = 0{,}04$ verlangt die Äquivalenz

$$\frac{B}{K(6)} = A_1 + \frac{A_2}{K(3)} \quad\Longrightarrow\quad B = A_1\mathrm{e}^{0{,}24} + A_2\mathrm{e}^{0{,}12} \approx 36\,699{,}95.$$

Das prospektive Deckungskapital ist $V(t) = K(t)\big(B/K(6) - A_2/K(3)\big) = A_1\mathrm{e}^{\delta t}$ für $t \in (0,3]$ und $V(t) = B\mathrm{e}^{-\delta(6-t)}$ für $t \in (3,6]$, sonst 0. Es ist überall ≥ 0: ein Sparplan. An den Zahlungszeitpunkten springt V. Abbildung 1.2 zeigt den Verlauf des Deckungskapitals.

Definition 1.13 (Rendite)**.** Sei $K(t) = (1+i)^t$ und seien $b(Z_P), b(Z_L)$ endlich. Der Zinssatz $i \geq 0$ heißt *Rendite* (Effektivzins) von (Z_L, Z_P), wenn $b(Z_P) = b(Z_L)$.

Eine Rendite muss weder existieren noch eindeutig sein. Für eine Prämie π zur Zeit t_P und eine Leistung A zur Zeit $t_L \neq t_P$ ergibt die Äquivalenz $i = (A/\pi)^{1/(t_L - t_P)} - 1$; das ist nur dann ≥ 0, wenn $A \geq \pi$ für $t_L > t_P$ (Geldanlage) bzw. $A \leq \pi$ für $t_L < t_P$ (Kredit).

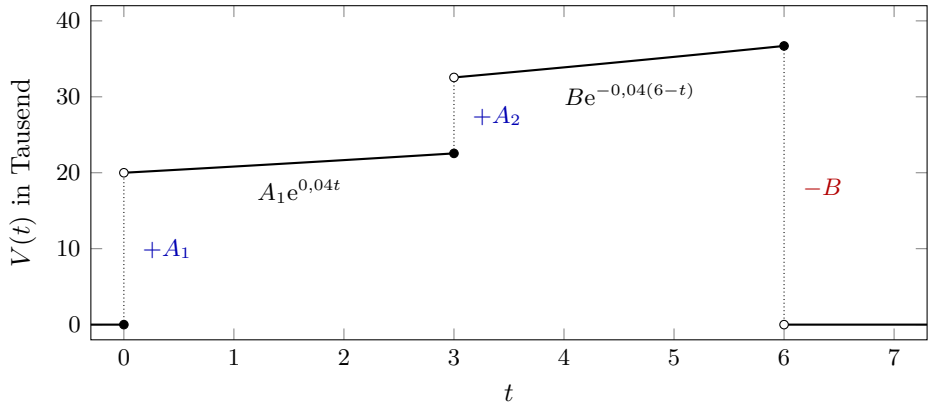

Abbildung 1.2: Prospektives Deckungskapital des Sparvertrags aus Beispiel 1.12 (in Tausend). Zwischen den Zahlungen wächst V mit der Zinsrate $\delta = 0{,}04$; bei jeder Einzahlung springt es um deren Betrag nach oben ($A_1 = 20$, $A_2 = 10$), bei der Auszahlung von $B \approx 36{,}70$ fällt es auf 0. Weil in Definition 1.10 über $[t, \infty)$ integriert wird, zählt eine Zahlung zur Zeit t noch als künftig: V ist linksstetig ($V(3) = A_1 e^{0,12} \approx 22{,}55$ vor, $V(3+) \approx 32{,}55$ nach der zweiten Einzahlung). Gefüllte Punkte markieren den Funktionswert, offene den rechtsseitigen Grenzwert. Wegen $V \geq 0$ ist der Vertrag ein Sparplan.

Aufgaben

Aufgabe 1.1. Zeigen Sie: Für eine absolut stetige Kapitalfunktion stimmen die beiden Definitionen der kumulierten Zinsintensität in Bemerkung 1.5 überein. Berechnen Sie für $K(t) = (1+i)^{\lfloor t \rfloor}$ zusätzlich $\int_{(0,1]} K(s)^{-1}\, \mathrm{d}K(s)$ und vergleichen Sie mit i und $\log(1+i)$.

Aufgabe 1.2 (vgl. [BOS25, Übung 1.25])**.** Konstruieren Sie Zahlungsströme (Z_L, Z_P) mit mehreren Renditen. Gibt es Situationen, in denen die Menge der Renditen kein Minimum besitzt?

Aufgabe 1.3. Ein Kredit über 100 000 wird mit zehn gleichen jährlichen Raten nachschüssig bei $i = 3\,\%$ getilgt. Bestimmen Sie die Rate und die Restschuld $-V(t)$ und skizzieren Sie $t \mapsto V(t)$.

Kapitel 2

Lebensversicherungsverträge und Nettoprämien

2.1 Beschreibung des Todesfallrisikos

Einer Person im Alter x wird zur Zeit 0 ein Vertrag angeboten. Ihre restliche Lebensdauer modellieren wir als Zufallsvariable $T = T_x$ mit Werten in $(0, \infty)$ und Verteilungsfunktion $F = F_{T_x}$, $F(0) = 0$, $F(t) \to 1$. Wir schreiben $\overline{F} = 1 - F$ für die Überlebensfunktion und

$$t_{\max} := \sup\{t \geq 0 \colon F(t) < 1\} \in (0, \infty]$$

für die maximale Restlebensdauer.

Definition 2.1 (Sterblichkeitsintensität)**.** Hat F eine Dichte f, so heißt $\lambda(t) := f(t)/(1 - F(t))$, $t \in (0, t_{\max})$, die *Sterblichkeitsintensität* (hazard rate). Allgemein heißt

$$\Lambda(t) := \int_{[0,t]} \frac{1}{1 - F(u-)} \, \mathrm{d}F(u), \qquad t \geq 0,$$

die *kumulierte Sterblichkeitsintensität*.

Für kleines h ist $\lambda(t)h \approx \mathbb{P}(T \leq t + h | T > t)$. Der linksseitige Grenzwert $F(u-)$ im Nenner sorgt dafür, dass ein Atom von F in u durch die Überlebenswahrscheinlichkeit *bis unmittelbar vor* u geteilt wird: Der Sprung $\Delta\Lambda(u) = \mathbb{P}(T = u | T \geq u)$ ist eine bedingte Sterbewahrscheinlichkeit.

Ist λ stetig, so löst $\overline{F}$ die lineare Differentialgleichung $\overline{F}' = -\lambda \overline{F}$ mit $\overline{F}(0) = 1$, also

$$\overline{F}(t) = \exp\Big(-\int_0^t \lambda(s) \, \mathrm{d}s\Big) = \mathrm{e}^{-\Lambda(t)}.$$

Auch allgemein ist F durch Λ eindeutig bestimmt (Produktintegral).

Beispiel 2.2 (Klassische Sterblichkeitsgesetze)**.** Für einen x-Jährigen:

Gesetz	$\lambda(t)$
de Moivre	$1/(\omega - x - t)$, $t < \omega - x$
Gompertz	Bc^{x+t}
Makeham	$A + Bc^{x+t}$
Weibull	$k(x + t)^n$

Gompertz und Makeham beschreiben das exponentielle Anwachsen der Sterblichkeit im Erwachsenenalter; in der logarithmischen Darstellung einer Sterbetafel sieht man dies als annähernd lineare Kurve. Abbildung 2.1 zeigt, wie eine exponentiell wachsende Intensität die Gestalt der Lebensdauerverteilung bestimmt.

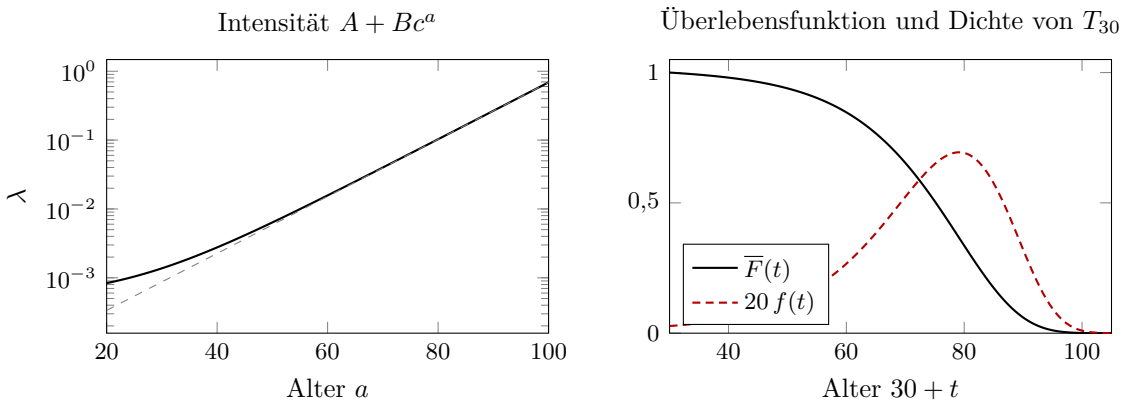

Abbildung 2.1: Gompertz–Makeham-Sterblichkeit mit illustrativen Parametern $A = 5 \cdot 10^{-4}$, $B = 5 \cdot 10^{-5}$, $c = 1{,}1$. Links: Unterhalb von etwa 24 Jahren dominiert die Konstante A, darüber der Gompertz-Anteil Bc^a (gestrichelt); ab etwa 50 Jahren ist A vernachlässigbar, und die Intensität wächst auf logarithmischer Skala linear. Rechts: Die exponentiell wachsende Intensität konzentriert die Todesfälle um ein Alter von etwa 80 Jahren (Modus der Dichte bei 79); die Dichte ist zur besseren Lesbarkeit mit 20 skaliert.

Sterbetafeln. Eine Sterbetafel gibt für ganze Alter x die einjährigen Sterbewahrscheinlichkeiten $q_x = \mathbb{P}(T_x \le 1)$ an. Um daraus die Verteilung von T_x zu gewinnen, nehmen wir *Stationarität* an:

$$\mathbb{P}(T_{x+s} > t) = \mathbb{P}(T_x > s + t | T_x > s), \qquad s, t, x \ge 0. \tag{2.1}$$

Die Lebensdauer eines $(x + s)$-Jährigen verhält sich also wie die eines x-Jährigen, der s weitere Jahre überlebt hat. In einer Bevölkerung mit sich langsam ändernder Sterblichkeit ist das näherungsweise erfüllt.

Lemma 2.3 ([BOS25, Lemma 2.4])**.** *Unter Gleichung* (2.1) *gilt für* $k \in \mathbb{N}$

$$\mathbb{P}(T_x > k) = \prod_{j=0}^{k-1} (1 - q_{x+j}).$$

Beweis. Mit Gleichung (2.1) ist $\mathbb{P}(T_x > k+1) = \mathbb{P}(T_x > k)\mathbb{P}(T_x > k+1 | T_x > k) = \mathbb{P}(T_x > k)\mathbb{P}(T_{x+k} > 1) = \mathbb{P}(T_x > k)(1 - q_{x+k})$; Induktion über k. □

Zwischen ganzzahligen Zeiten interpoliert man. Die übliche Annahme ist, dass der Todeszeitpunkt innerhalb des Sterbejahrs gleichverteilt ist.

Lemma 2.4 ([BOS25, Lemma 2.5])**.** *Sei* $T_x = N_x + U_x$ *mit* $N_x \in \mathbb{N}_0$ *und davon unabhängigem* $U_x \sim \mathrm{U}[0,1]$*. Dann ist* F_{T_x} *auf jedem Intervall* $[k, k+1]$ *linear, und unter Gleichung* (2.1) *gilt für die Sterblichkeitsintensität*

$$\lambda_x(t) = \frac{q_{x+k}}{1 - (t-k)q_{x+k}}, \qquad t \in (k, k+1],\ k \in \mathbb{N}_0.$$

Beweis. Für $t \in (k, k+1]$ ist $\overline{F}_{T_x}(t) = \mathbb{P}(N_x = k)\,(1 - (t-k)) + \mathbb{P}(N_x \ge k+1)$, also linear in t mit Ableitung $-\mathbb{P}(N_x = k)$. Folglich ist $\lambda_x(t) = \mathbb{P}(N_x = k)/\overline{F}_{T_x}(t)$. Division von Zähler und Nenner durch $\mathbb{P}(N_x \ge k) = \mathbb{P}(T_x > k)$ und $\mathbb{P}(N_x = k | N_x \ge k) = \mathbb{P}(T_x \le k+1 | T_x > k) = q_{x+k}$ liefert die Formel. □

2.2 Das Modell eines Lebensversicherungsvertrags

Da die Leistung vom Todeszeitpunkt abhängt, werden Zahlungsströme jetzt zufällig.

Definition 2.5. Ein reellwertiger stochastischer Prozess $(X_t)_{t \ge 0}$ heißt *càdlàg*, wenn seine Pfade rechtsstetig sind und linksseitige Grenzwerte besitzen. Ein *gerichteter zufälliger Zahlungsstrom* ist ein rechtsstetiger Prozess mit wachsenden Pfaden, d. h. $t \mapsto X_t(\omega) \in \mathcal{Z}_g$ für jedes ω.

Integrale gegen zufällige Zahlungsströme werden pfadweise erklärt.

Definition 2.6 (Vertragsmodell, [BOS25, Def. 2.10–2.11])**.** Die Kenngrößen eines Vertrags für ein einzelnes Leben sind:

- die Verteilungsfunktion F der Restlebensdauer T;
- der Endzeitpunkt $\tau \in (0, t_{\max}]$ und der *Leistungszeitpunkt* $Y := \min\{T, \tau\}$;
- das *Auszahlungsspektrum* $A\colon [0,\infty) \to [0,\infty)$ messbar (ist $\tau = \infty$, so ist $Y = T < \infty$ fast sicher, und Terme mit $A(\tau)$ entfallen);
- eine Kapitalfunktion K;
- eine *Prämienfunktion* Π: wachsend, rechtsstetig, $\Pi(t)$ ist die Summe der bis t fälligen Prämien.

Das Modell ist das Tripel (K, Z_L, Z_P) mit

$$Z_L(t) := A(Y)\mathbb{1}\{t \ge Y\}, \qquad Z_P(t) := \begin{cases} \Pi(t), & t < Y, \\ \Pi(Y-), & t \ge Y. \end{cases}$$

Zum Zeitpunkt Y wird $A(Y)$ ausgezahlt; Prämien fließen bis unmittelbar vor Y. Spezialfälle:

- *lebenslange Todesfallversicherung*: $\tau = \infty$;
- *temporäre Todesfallversicherung* (Risikoversicherung): $\tau < \infty$, $A(\tau) = 0$;
- *reine Erlebensfallversicherung*: $\tau < \infty$, $A(t) = 0$ für $t < \tau$;
- *gemischte Versicherung* (Kapitallebensversicherung): $\tau < \infty$, $A > 0$ auf $[0, \tau]$.

Leibrenten oder Zahlungen am Ende des Sterbejahres passen nicht direkt in das Modell; man behandelt sie, indem man Z_L in einen Todesfall- und einen Erlebensfallanteil zerlegt [BOS25, Bem. 2.13].

Definition 2.7 (Barwert, Äquivalenzprinzip)**.** Der zufällige Barwert des Vertrags ist

$$B := b(Z_L) - b(Z_P) = \frac{A(Y)}{K(Y)} - \int_{[0,Y)} \frac{1}{K(s)}\,\mathrm{d}\Pi(s).$$

Die Prämienfunktion Π heißt *Nettoprämienfunktion*, wenn der erwartete Leistungsbarwert und der erwartete Prämienbarwert endlich und gleich sind, d. h. $\mathbb{E}[B] = 0$.

Die Nettoprämie berücksichtigt weder Kosten noch Sicherheitszuschläge; mit diesen spricht man von der *ausreichenden Prämie* (Kapitel 3). Im Folgenden sei der erwartete Leistungsbarwert stets endlich.

Lemma 2.8 ([BOS25, Lemma 2.17])**.** *Sind die Integrale endlich, so gilt*

$$\mathbb{E}[B] = \int_{[0,\tau)} \frac{A(s)}{K(s)}\,\mathrm{d}F(s) + \frac{A(\tau)}{K(\tau)}\big(1 - F(\tau-)\big) - \int_{[0,\tau)} \frac{1-F(s)}{K(s)}\,\mathrm{d}\Pi(s).$$

Die drei Terme sind der erwartete Leistungsbarwert im Todesfall, im Erlebensfall und der erwartete Prämienbarwert.

Beweis. Es ist $\mathbb{P}(Y \le t) = F(t)$ für $t < \tau$ und $\mathbb{P}(Y = \tau) = 1 - F(\tau-)$, also $\mathbb{P}_Y(\mathrm{d}s) = \mathbb{1}\{s < \tau\}F(\mathrm{d}s) + (1 - F(\tau-))\delta_\tau(\mathrm{d}s)$. Das liefert die ersten beiden Terme. Für den Prämienbarwert vertauschen wir mit Fubini:

$$\mathbb{E}\Big[\int_{[0,\infty)} \frac{\mathbb{1}\{s < Y\}}{K(s)}\,\mathrm{d}\Pi(s)\Big] = \int_{[0,\infty)} \frac{\mathbb{P}(Y > s)}{K(s)}\,\mathrm{d}\Pi(s),$$

und $\mathbb{P}(Y > s) = 1 - F(s)$ für $s < \tau$, $\mathbb{P}(Y > s) = 0$ für $s \ge \tau$. □

2.3 Nettoprämien

Das Äquivalenzprinzip legt nur eine Gleichung fest; die Form des Prämienstroms ist frei wählbar. Drei Wahlen sind besonders wichtig.

Nettoeinmalprämie. Eine einzige Zahlung $\tilde{\Pi}$ zur Zeit 0 ist genau dann eine Nettoprämie, wenn

$$\tilde{\Pi} = \mathbb{E}\Big[\frac{A(Y)}{K(Y)}\Big].$$

Beispiel 2.9. Für $\tau = \infty$, $A \equiv A_0$ und $K(t) = \mathrm{e}^{\delta t}$ ist $\tilde{\Pi} = A_0\,\mathbb{E}[\mathrm{e}^{-\delta T}]$, also die Laplace-Transformierte der Lebensdauer an der Stelle δ. Ist $T \sim \mathrm{Exp}(\lambda)$, so folgt $\tilde{\Pi} = A_0\lambda/(\lambda+\delta)$: Die Prämie steigt mit der Sterblichkeit und fällt mit dem Zins. Für $\lambda = 0{,}02$, $\delta = 0{,}03$ zahlt man 40 % der Versicherungssumme.

Laufende konstante Prämie. Wird die Prämie π vorschüssig zu Zeiten $0 = t_0 < \cdots < t_{N-1} < \tau$ gezahlt, d. h. $\Pi(t) = \sum_{k<N} \pi \mathbb{1}\{t \ge t_k\}$, so liefert Lemma 2.8

$$\pi = \mathbb{E}\Big[\frac{A(Y)}{K(Y)}\Big] \Big/ \sum_{k=0}^{N-1} \frac{1-F(t_k)}{K(t_k)}.$$

Der Nenner ist der erwartete Barwert einer temporären Leibrente, die zu den Zeiten $t_0, \dots, t_{N-1}$ eine Einheit zahlt, solange die Person lebt.

Natürliche Prämie. Hier wird zu Beginn jeder Periode genau das Risiko dieser Periode bezahlt.

Definition 2.10 (Natürliche Prämie)**.** Seien $0 = t_0 < \cdots < t_N = \tau < \infty$. Die *natürliche Prämie* zahlt zur Zeit t_k (bei Erleben) den Betrag

$$\pi_k := K(t_k)\,\mathbb{E}\Big[\int_{(t_k,t_{k+1}]} \frac{\mathrm{d}Z_L(s)}{K(s)} \;\Big|\; T > t_k\Big], \qquad k = 0, \dots, N-1.$$

Proposition 2.11 ([BOS25, Prop. 2.27])**.** *Es gilt*

$$\pi_k = \frac{K(t_k)}{1-F(t_k)} \int_{(t_k,t_{k+1}]} \frac{A(s)}{K(s)}\,\mathrm{d}F_Y(s),$$

und die natürliche Prämie ist eine Nettoprämie.

Beweis. Die Formel folgt aus $\int_{(t_k,t_{k+1}]} K^{-1}\,\mathrm{d}Z_L = \mathbb{1}\{t_k < Y \le t_{k+1}\}A(Y)/K(Y)$ und $\{Y > t_k\} = \{T > t_k\}$ (da $t_k < \tau$). Der erwartete Prämienbarwert ist

$$\sum_k \frac{\pi_k}{K(t_k)}\mathbb{P}(T > t_k) = \sum_k \int_{(t_k,t_{k+1}]} \frac{A}{K}\,\mathrm{d}F_Y = \mathbb{E}\Big[\frac{A(Y)}{K(Y)}\Big],$$

denn $Y \in (0, \tau]$ fast sicher. □

Die natürliche Prämie steigt typischerweise mit dem Alter stark an; die konstante Prämie ist zu Beginn höher als das Risiko, und die Differenz wird angespart. Das führt auf das Deckungskapital in Kapitel 3.

Aufgaben

Aufgabe 2.1 (vgl. [BOS25, Übung 2.6])**.** Zeigen Sie unter den Annahmen von Lemma 2.4, dass $F_{T_x}(t) = F_{T_x}(k) + (t-k)(F_{T_x}(k+1) - F_{T_x}(k))$ für $t \in [k, k+1)$.

Aufgabe 2.2 (vgl. [BOS25, Übung 2.24])**.** Leiten Sie die Formel für die laufende konstante Prämie bei vorschüssiger und bei nachschüssiger Zahlungsweise her.

Aufgabe 2.3 (vgl. [BOS25, Übung 2.25])**.** Sei $\Pi(t) = ct$ (kontinuierliche Prämienzahlung). Zeigen Sie, dass c so gewählt werden kann, dass Π eine Nettoprämie ist, sofern $\mathbb{E}[Y]$ und der Leistungsbarwert endlich sind. Berechnen Sie c für $T \sim \mathrm{Exp}(\lambda)$, $\tau = \infty$, $A \equiv 1$, $K(t) = \mathrm{e}^{\delta t}$.

Aufgabe 2.4. Berechnen Sie für eine temporäre Todesfallversicherung ($A = 1$ auf $[0, \tau)$, $A(\tau) = 0$, $\tau \in \mathbb{N}$) bei jährlicher Verzinsung die natürlichen Prämien π_k unter der Interpolationsannahme aus Lemma 2.4.

Kapitel 3

Deckungskapital und Thielesche Gleichungen

3.1 Prospektives und retrospektives Nettodeckungskapital

Bei einer konstanten Prämie zahlt der Versicherungsnehmer anfangs mehr, als sein momentanes Risiko kostet. Der Versicherer muss diesen Überschuss zurückstellen, weil er ihn später zur Finanzierung des steigenden Risikos braucht. Das Nettodeckungskapital misst diese Rückstellung.

Definition 3.1 ([BOS25, Def. 2.29])**.** Sei Π eine Nettoprämienfunktion. Das (erwartete) *prospektive Nettodeckungskapital* zur Zeit $t < t_{\max}$ ist

$$V(t) \coloneqq K(t)\,\mathbb{E}\Big[\frac{A(Y)}{K(Y)}\mathbb{1}\{t \le Y\} - \int_{[t,Y)} \frac{\mathrm{d}\Pi(s)}{K(s)} \;\Big|\; T > t\Big];$$

für $t \ge t_{\max}$ sei $V(t) \coloneqq 0$.

Das Nettodeckungskapital $V(t)$ ist der erwartete Zeitwert aller künftigen Leistungen minus aller künftigen Prämien für einen Vertrag, der zur Zeit t noch aktiv ist. Wegen $F(0) = 0$ und des Äquivalenzprinzips ist $V(0) = \mathbb{E}[B] = 0$. Aus Sicht des Kunden ist $V(t)$ der Nettorückkaufwert, also sein Guthaben beim Versicherer.

Lemma 3.2 ([BOS25, Lemma 2.32, 2.36])**.** *Für $0 \le t < \tau$ gilt die* prospektive Darstellung

$$V(t) = \frac{K(t)}{1-F(t)}\Big(\int_{(t,\tau)} \frac{A(s)}{K(s)}\,\mathrm{d}F(s) + \frac{A(\tau)}{K(\tau)}\big(1-F(\tau-)\big) - \int_{[t,\tau)} \frac{1-F(s)}{K(s)}\,\mathrm{d}\Pi(s)\Big)$$

und die retrospektive Darstellung

$$V(t) = \frac{K(t)}{1-F(t)}\Big(\int_{[0,t)} \frac{1-F(s)}{K(s)}\,\mathrm{d}\Pi(s) - \int_{[0,t]} \frac{A(s)}{K(s)}\,\mathrm{d}F(s)\Big).$$

Beweis. Auf $\{T > t\}$ ist $Y > t$. Wie im Beweis von Lemma 2.8 ist $\mathbb{E}[A(Y)K(Y)^{-1}\mathbb{1}\{T > t\}] = \int_{(t,\tau)} A/K\,\mathrm{d}F + A(\tau)(1-F(\tau-))/K(\tau)$ und, mit Fubini, $\mathbb{E}[\mathbb{1}\{T > t\}\int_{[t,Y)} K^{-1}\,\mathrm{d}\Pi] = \int_{[t,\tau)}(1-F(s))K(s)^{-1}\,\mathrm{d}\Pi(s)$. Division durch $\mathbb{P}(T > t) = 1 - F(t)$ ergibt die prospektive Form. Für die retrospektive Form ziehen wir die Identität $\mathbb{E}[B] = 0$ aus Lemma 2.8 ab: Die Integrale über $[0,\tau)$ werden zu Integralen über $[0,t]$ (Leistungen) bzw. $[0,t)$ (Prämien), mit umgekehrtem Vorzeichen. □

Wie in Lemma 1.11 ist das retrospektive Deckungskapital der aufgezinste Überschuss der bisherigen Prämien über die bisherigen Leistungen – nun aber pro überlebender Person gerechnet (Faktor $1/(1-F(t))$): Die Prämien der Verstorbenen kommen den Überlebenden zugute.

3.2 Die Thielesche Differentialgleichung

Wir nehmen jetzt glatte Kenngrößen an:

$$K(t) = 1 + \int_0^t k, \qquad F(t) = \int_0^t f, \qquad \Pi(t) = \Pi(0) + \int_0^t \pi \tag{3.1}$$

mit stetigen $k, f, \pi \geq 0$; dabei ist π die *Prämienintensität*. Es seien $\varphi = k/K$ die Zinsintensität und $\lambda = f/(1-F)$ die Sterblichkeitsintensität.

Satz 3.3 (Thielesche Differentialgleichung, [BOS25, Satz 2.37])**.** *Unter Gleichung* (3.1) *und für A stetig auf $[0,\tau)$ erfüllt das Nettodeckungskapital auf $(0,\tau)$*

$$V'(t) = \varphi(t)V(t) + \pi(t) - \lambda(t)\big(A(t) - V(t)\big).$$

Beweis. Nach Lemma 3.2 ist $V(t) = (K(t)/\overline{F}(t))\, R(t)$ mit

$$R(t) = \int_t^\tau \frac{A(s)f(s)}{K(s)}\,\mathrm{d}s + \frac{A(\tau)\overline{F}(\tau)}{K(\tau)} - \int_t^\tau \frac{\overline{F}(s)\pi(s)}{K(s)}\,\mathrm{d}s$$

(für $t > 0$ trägt das Atom $\Pi(0)$ nicht bei). Nach dem Hauptsatz ist $R'(t) = -A(t)f(t)/K(t) + \overline{F}(t)\pi(t)/K(t)$, und $\big(K/\overline{F}\big)' = (k\overline{F} + Kf)/\overline{F}^2 = (K/\overline{F})(\varphi + \lambda)$. Die Produktregel ergibt

$$V' = (\varphi + \lambda)V - \lambda A + \pi. \qquad \square$$

Interpretation. Im Zeitintervall $[t, t+\mathrm{d}t]$ wächst das Deckungskapital durch Zinsen ($\varphi V\,\mathrm{d}t$) und durch die eingehende Prämie ($\pi\,\mathrm{d}t$). Mit Wahrscheinlichkeit $\lambda\,\mathrm{d}t$ stirbt die versicherte Person; dann muss statt der Reserve V die Leistung A gezahlt werden, der Versicherer verliert also das *riskierte Kapital* $A - V$.

Spar- und Risikoprämie. Schreibt man die Gleichung nach π um, so erhält man die Zerlegung

$$\pi(t) = \underbrace{V'(t) - \varphi(t)V(t)}_{=:\pi^s(t)\ \text{Sparprämie}} + \underbrace{\lambda(t)\big(A(t) - V(t)\big)}_{=:\pi^r(t)\ \text{Risikoprämie}}.$$

Die Sparprämie baut das Deckungskapital auf; es gilt $V(t) = K(t)\big(\Pi(0) + \int_0^t \pi^s(s)/K(s)\,\mathrm{d}s\big)$ für $t \in (0,\tau)$ (Aufgabe 3.1). Eine Anfangsprämie $\Pi(0)$ geht also vollständig in das Deckungskapital ein: V springt in 0 von $V(0) = 0$ auf $V(0+) = \Pi(0)$. Die Risikoprämie bezahlt das Risiko der laufenden Periode.

Bemerkung 3.4 (Stetige natürliche Prämie). Ist $A(\tau) = 0$, so ist $\pi^{\mathrm{nat}}(t) \coloneqq \lambda(t)A(t)$ eine Nettoprämienintensität mit $V \equiv 0$: Nach Lemma 3.2 ist der Integrand $Af/K - \overline{F}\lambda A/K$ identisch null. Die ganze Prämie ist Risikoprämie, es wird nichts angespart.

Beispiel 3.5. Sei $T \sim \mathrm{Exp}(1/50)$ und eine gemischte Versicherung mit $A \equiv 1$, $\tau = 30$ und konstanter Prämienrate. Man zerlegt sie gedanklich in eine temporäre Todesfallversicherung und eine Erlebensfallversicherung. Bei der ersten ist (Gedächtnislosigkeit!) die konstante Prämie gleich der natürlichen, also $\pi^r_{\mathrm{tT}} \equiv 1/50$. Bei der zweiten ist $\pi^r_{\mathrm{Erl}} = -\lambda V_{\mathrm{Erl}} \leq 0$: Das angesparte Kapital der Verstorbenen fällt an das Kollektiv zurück und wirkt als negative Risikoprämie [BOS25, Bsp. 2.41].

Abbildung 3.1 zeigt Deckungskapital und Prämienzerlegung für eine gemischte Versicherung.

3.3 Die Thielesche Integralgleichung

In der Praxis ist K oft stückweise konstant (jährliche Verzinsung) und F aus einer Sterbetafel interpoliert; dann ist Satz 3.3 nicht anwendbar. Die folgende Integralform gilt allgemein.

Satz 3.6 (Thielesche Integralgleichung, [BOS25, Satz 2.43])**.** *Mit der kumulierten Sterblichkeitsintensität Λ aus Definition 2.1 gilt für alle $t < \tau$*

$$\frac{V(t)}{K(t)} = \int_{[0,t)} \frac{1}{K(s)}\,\mathrm{d}\Pi(s) - \int_{(0,t]} \frac{A(u) - V(u)}{K(u)}\,\mathrm{d}\Lambda(u).$$

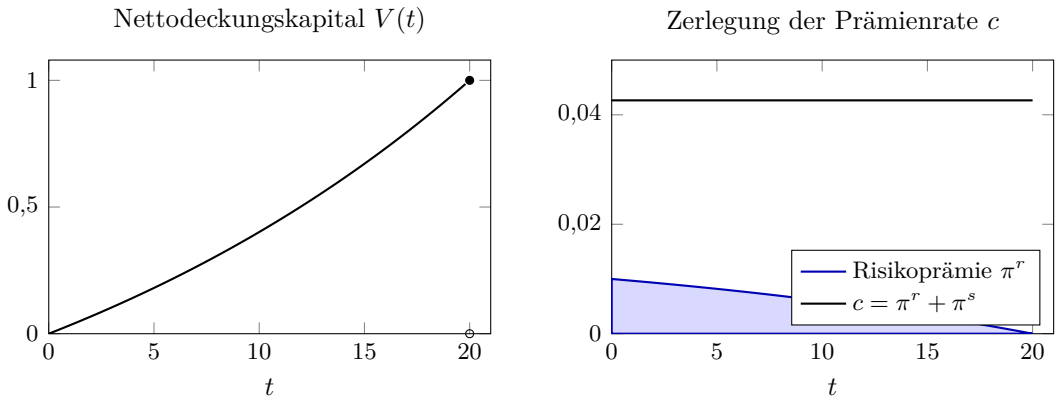

Abbildung 3.1: Gemischte Versicherung mit $A \equiv 1$, $\tau = 20$, $T \sim \mathrm{Exp}(0{,}01)$, $\delta = 0{,}03$ und konstanter Nettoprämienrate $c \approx 0{,}0426$. Links: V wächst gemäß der Thieleschen Gleichung bis zur Erlebensfallleistung $V(20) = A(20) = 1$ an und fällt nach deren Auszahlung auf 0; wie in Abbildung 1.2 ist V linksstetig (gefüllter Punkt: Funktionswert). Rechts: Die Risikoprämie $\pi^r = \lambda(A - V)$ (Fläche) schrumpft, weil das riskierte Kapital $A - V$ schrumpft; der Rest $\pi^s = c - \pi^r$ ist Sparprämie.

Unter Gleichung (3.1) folgt die Integralgleichung durch Integration aus Satz 3.3, angewendet auf V/K. Der allgemeine Beweis verwendet eine Kettenregel für Funktionen mit Sprüngen:

Lemma 3.7 ([BOS25, Lemma 2.44])**.** *Sei G monoton und rechtsstetig mit Werten in einem offenen Intervall J und $h \in C^1(J)$. Dann gilt für $s < t$*

$$h(G(t)) - h(G(s)) = \int_{(s,t]} h'(G(u-))\,\mathrm{d}G(u) + \sum_{s<u\le t} \Big(h(G(u)) - h(G(u-)) - h'(G(u-))\Delta G(u) \Big).$$

Man wendet diese Formel auf $h(x) = 1/x$ und $G = 1 - F$ bzw. auf $1/K$ an und kombiniert sie mit der partiellen Integration für Lebesgue–Stieltjes-Integrale (Anhang A); die Details finden sich in [BOS25, S. 43–47].

3.4 Erweiterungen des Modells

Kosten und ausreichende Prämie. Bezeichnet $C(t)$ den Barwert der Kosten, die anfallen, wenn $Y = t$ ist, so heißt Π *ausreichend*, wenn der erwartete Prämienbarwert gleich $\mathbb{E}[A(Y)/K(Y)] + \mathbb{E}[C(Y)]$ ist. Man kann $\Pi = \Pi_{\text{Netto}} + \Pi_{\text{Kosten}}$ zerlegen und zwei Konten führen.

Mehrzustandsmodelle. Berufsunfähigkeits- oder Pflegeversicherungen benötigen mehr als die Zustände „lebend“ und „tot“. Man modelliert den Zustand des Versicherten dann als zeitinhomogene Markov-Kette mit Übergangsintensitäten $\mu_{ij}(t)$. Für das Deckungskapital V_i im Zustand i gilt das System

$$V_i'(t) = \varphi(t) V_i(t) + \pi_i(t) - \sum_{j \neq i} \mu_{ij}(t)\big(A_{ij}(t) + V_j(t) - V_i(t)\big),$$

wobei A_{ij} die Leistung beim Übergang $i \to j$ ist. Für zwei Zustände und $V_{\text{tot}} = 0$ ist das wieder Satz 3.3.

Aufgaben

Aufgabe 3.1 (vgl. [BOS25, Übung 2.42])**.** Zeigen Sie unter den Voraussetzungen von Satz 3.3: $V(t) = K(t)\big(\Pi(0) + \int_0^t \pi^s(s)/K(s)\,\mathrm{d}s\big)$ für $t \in (0, \tau)$.

Aufgabe 3.2 (vgl. [BOS25, Übung 2.40])**.** Sei A konstant auf $[0, \tau)$, $A(\tau) = 0$ und $T \sim \mathrm{Exp}(\lambda)$. Zeigen Sie, dass die stetige natürliche Prämie gleich der konstanten Nettoprämienrate ist.

Aufgabe 3.3 (vgl. [BOS25, Übung 2.35])**.** Wird zur Zeit t_0 eine Prämie $\pi > 0$ fällig und ist F in t_0 stetig, so gilt $V(t_0) = V(t_0+) - \pi$. Interpretieren Sie den Sprung.

Aufgabe 3.4. Lösen Sie die Thielesche Differentialgleichung für eine lebenslange Todesfallversicherung mit $A \equiv 1$, $T \sim \text{Exp}(\lambda)$, $K(t) = \text{e}^{\delta t}$ und konstanter Nettoprämienrate explizit. Was fällt auf?

Kapitel 4

Der Satz von Hattendorff

Das Äquivalenzprinzip legt den Erwartungswert des Barwerts fest, $\mathbb{E}[B] = 0$, sagt aber nichts über dessen Streuung. Wie stark weicht der realisierte Verlust eines Vertrags von null ab, und in welchen Versicherungsjahren entsteht dieses Risiko? Der Satz von Hattendorff (1868) beantwortet beides.

4.1 Varianz bei der Nettoeinmalprämie

Lemma 4.1 (Korrelationsungleichung, [BOS25, Lemma 3.3])**.** *Sei X reell und seien $f, g\colon \mathbb{R} \to \mathbb{R}$ wachsend mit $f(X), g(X) \in L^2$. Dann gilt $\mathbb{E}[f(X)g(X)] \geq \mathbb{E}[f(X)]\,\mathbb{E}[g(X)]$.*

Beweis. Sei X' eine unabhängige Kopie von X. Da f und g wachsen, ist $(f(X)-f(X'))(g(X)-g(X')) \geq 0$. Erwartungswertbildung ergibt $2\mathbb{E}[f(X)g(X)] - 2\mathbb{E}[f(X)]\mathbb{E}[g(X)] \geq 0$. □

Proposition 4.2 ([BOS25, Prop. 3.1])**.** *Ist $t \mapsto A(t)/K(t)$ fallend, so hat unter allen Nettoprämienfunktionen die Nettoeinmalprämie die kleinste Varianz des Barwerts B.*

Beweis. Schreibe $B = a(Y) - p(Y)$ mit $a(y) = A(y)/K(y)$ fallend und $p(y) = \int_{[0,y)} K^{-1}\,\mathrm{d}\Pi$ wachsend. Dann ist $\operatorname{Var}[B] = \operatorname{Var}[a(Y)] + \operatorname{Var}[p(Y)] - 2\operatorname{Cov}(a(Y), p(Y))$. Nach Lemma 4.1 (angewendet auf $-a$ und p) ist $\operatorname{Cov}(a(Y), p(Y)) \leq 0$, also $\operatorname{Var}[B] \geq \operatorname{Var}[a(Y)]$. Bei der Einmalprämie ist p konstant und es gilt Gleichheit. □

Für eine reine Erlebensfallversicherung ist A/K nicht fallend, und die Aussage wird falsch (Aufgabe 4.1).

4.2 Martingale und der kompensierte Einheitsleistungsstrom

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit Filtration $(\mathcal{F}_t)_{t\geq 0}$. Ein adaptierter Prozess (X_t) mit $\mathbb{E}|X_t| < \infty$ heißt *Martingal*, wenn $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ f. s. für $s \leq t$ (Sub-/Supermartingal mit $\geq$ bzw. $\leq$). Dann ist $\mathbb{E}[X_t] = \mathbb{E}[X_0]$. Ist $X \in L^1$, so ist $M_t \coloneqq \mathbb{E}[X|\mathcal{F}_t]$ ein Martingal (*Doob-Martingal*, Turmeigenschaft).

Lemma 4.3 (Unkorrelierte Zuwächse, Pythagoras)**.** *Sei (M_t) ein quadratintegrierbares Martingal und $u \leq v \leq s \leq t$. Dann ist $\operatorname{Cov}(M_v - M_u, M_t - M_s) = 0$ und $\mathbb{E}[(M_t - M_s)^2] = \mathbb{E}[M_t^2] - \mathbb{E}[M_s^2]$.*

Beweis. Es gilt $\mathbb{E}[(M_v - M_u)(M_t - M_s)] = \mathbb{E}[(M_v - M_u)\mathbb{E}[M_t - M_s|\mathcal{F}_s]] = 0$. Die zweite Aussage folgt aus $\mathbb{E}[M_s(M_t - M_s)] = 0$. □

Der Einheitsleistungsstrom. Sei $Y = T \wedge \tau$ der Leistungszeitpunkt, $N_t \coloneqq \mathbb{1}\{Y \leq t\}$ und $(\mathcal{F}_t)$ die kanonische Filtration von (N_t): $\mathcal{F}_t$ enthält genau die Information, ob und wann Y bis t eingetreten ist. Mit der kumulierten Intensität von Y,

$$\Lambda_Y(t) = \int_{(0,t]} \frac{\mathrm{d}F_Y(u)}{1 - F_Y(u-)},$$

definieren wir den *kompensierten Einheitsleistungsstrom*

$$M_t \coloneqq N_t - \Lambda_Y(t \wedge Y), \qquad t \geq 0.$$

(Ist $\tau < \infty$ und $F(\tau-) = 1$, so ist Λ_Y in τ unendlich; Integrale gegen Λ_Y werden dann nur über $[0, \tau)$ gebildet, siehe [BOS25, Bem. 3.17].)

Proposition 4.4 ([BOS25, Prop. 3.18])**.** *Der Prozess (M_t) ist ein $(\mathcal{F}_t)$-Martingal mit rechtsstetigen Pfaden und $\mathbb{E}[M_t] = 0$.*

Beweis. Vor Y enthält $\mathcal{F}_s$ nur die Information $\{Y > s\}$; daher gilt für Teilintervalle $I \subset (s, \infty)$

$$\mathbb{P}(Y \in I | \mathcal{F}_s) = \frac{\mathbb{P}(Y \in I)}{1 - F_Y(s)} \mathbb{1}\{Y > s\}$$

[BOS25, Lemma 3.21]. Auf $\{Y \leq s\}$ ist $M_t - M_s = 0$. Auf $\{Y > s\}$ ist $N_t - N_s = \mathbb{1}\{s < Y \leq t\}$ und $\Lambda_Y(t \wedge Y) - \Lambda_Y(s) = \int_{(s,t]} \mathbb{1}\{u \leq Y\} \, \mathrm{d}\Lambda_Y(u)$. Mit Fubini ist

$$\mathbb{E}\Big[\int_{(s,t]} \mathbb{1}\{u \leq Y\} \, \mathrm{d}\Lambda_Y(u)\Big] = \int_{(s,t]} \big(1 - F_Y(u-)\big) \frac{\mathrm{d}F_Y(u)}{1 - F_Y(u-)} = F_Y(t) - F_Y(s) = \mathbb{P}(s < Y \leq t),$$

und dasselbe gilt nach Bedingen auf $\{Y > s\}$ (beide Seiten werden durch $1 - F_Y(s)$ geteilt). Also $\mathbb{E}[M_t - M_s | \mathcal{F}_s] = 0$. □

Hier wird die Rolle von $F(u-)$ in der Definition von Λ sichtbar: Der Kompensator integriert gegen $\mathbb{P}(Y \geq u) = 1 - F_Y(u-)$.

Proposition 4.5 (Varianz der Zuwächse und Isometrie, [BOS25, Prop. 3.24, 3.25])**.** *Für $0 \leq s \leq t$ gilt*

$$\mathbb{E}[(M_t - M_s)^2] = \int_{(s,t]} \big(1 - \Delta\Lambda_Y(u)\big) \, \mathrm{d}F_Y(u), \qquad \Delta\Lambda_Y(u) = \Lambda_Y(u) - \Lambda_Y(u-).$$

Für jedes messbare f mit $\int f^2(1 - \Delta\Lambda_Y) \, \mathrm{d}F_Y < \infty$ ist $\int_{(0,\infty)} f \, \mathrm{d}M$ wohldefiniert und

$$\mathbb{E}\Big[\Big(\int_{(0,\infty)} f(u) \, \mathrm{d}M_u\Big)^2\Big] = \int_{(0,\infty)} f(u)^2 \big(1 - \Delta\Lambda_Y(u)\big) \, \mathrm{d}F_Y(u).$$

Beweisskizze. Die erste Formel rechnet man wie in Proposition 4.4 direkt nach; der Faktor $1 - \Delta\Lambda_Y$ korrigiert dafür, dass bei Atomen von F_Y ein Teil des Sprungs von N bereits vorhersehbar ist. Für Treppenfunktionen $f(u) = \sum_j e_j \mathbb{1}\{t_j < u \leq t_{j+1}\}$ ist $\int f \, \mathrm{d}M = \sum_j e_j (M_{t_{j+1}} - M_{t_j})$, und wegen Lemma 4.3 verschwinden die gemischten Terme. Das liefert die Isometrie für Treppenfunktionen; die Fortsetzung auf L^2 ist ein Dichtheitsargument [BOS25, S. 63–66]. □

4.3 Der Verlustprozess

Definition 4.6 (Verlust, [BOS25, Def. 3.27])**.** Der *Verlust* zur Zeit t ist

$$L_t \coloneqq \underbrace{\frac{A(Y)}{K(Y)} \mathbb{1}\{Y \leq t\} + \frac{V(t)}{K(t)} \mathbb{1}\{Y > t\}}_{\text{Leistung oder Reserve}} - \underbrace{\int_{[0, t \wedge Y)} \frac{\mathrm{d}\Pi(s)}{K(s)}}_{\text{bisherige Prämien}}.$$

Für Zeitpunkte $0 = t_0 < t_1 < \dots$ ist $\Delta L_i \coloneqq L_{t_i} - L_{t_{i-1}}$ der Verlust in der i-ten Versicherungsperiode.

Bis zum Leistungsfall ist L_t die abgezinste Reserve minus die bisher eingenommenen Prämien; danach ist $L_t = B$.

Proposition 4.7 ([BOS25, Prop. 3.28, 3.29])**.** *Es gilt $L_t = \mathbb{E}[B | \mathcal{F}_t]$ f. s., also ist (L_t) ein Martingal, und*

$$L_t = \int_{(0,t]} \frac{A(u) - V(u)}{K(u)} \, \mathrm{d}M_u.$$

Beweisskizze. Auf $\{Y \le t\}$ ist B bereits $\mathcal{F}_t$-messbar und gleich L_t. Auf $\{Y > t\}$ ist nach dem Argument aus Proposition 4.4 $\mathbb{E}[B|\mathcal{F}_t]$ die bedingte Erwartung gegeben $T > t$, also nach Definition des Deckungskapitals gleich $V(t)/K(t) - \int_{[0,t)} K^{-1}\,\mathrm{d}\Pi$. Die Integraldarstellung folgt aus der Thieleschen Integralgleichung (Satz 3.6): Vor Y stimmen beide Seiten mit $-\int_{(0,t]}(A-V)/K\,\mathrm{d}\Lambda$ überein, und im Zeitpunkt Y springen beide um $\big(A(Y)-V(Y)\big)\big(1-\Delta\Lambda_Y(Y)\big)/K(Y)$: links, weil V/K nach Satz 3.6 in Y um $-(A(Y)-V(Y))\Delta\Lambda_Y(Y)/K(Y)$ springt, rechts wegen $\Delta M_Y = 1-\Delta\Lambda_Y(Y)$. Insbesondere springt L im Erlebensfall $Y=\tau$ nicht, da dort $\Delta\Lambda_Y(\tau)=1$ ist. □

Damit haben wir drei Beschreibungen des Verlusts: die anschauliche Definition, die Martingaldarstellung als Doob-Martingal und die Integraldarstellung gegen M, auf die sich die Isometrie anwenden lässt.

4.4 Der Satz

Satz 4.8 (Hattendorff, [BOS25, Satz 3.32]). *(i)* $\mathbb{E}[L_t]=0$, $\mathbb{E}[\Delta L_i]=0$ *und* $\mathbb{E}[\Delta L_i|\mathcal{F}_{t_{i-1}}]=0$. *Ist zusätzlich* $\int_{[0,t]}\big((A-V)/K\big)^2(1-\Delta\Lambda_Y)\,\mathrm{d}F_Y<\infty$ *für alle* t, *so gilt:*

(ii) Die Verluste $\Delta L_j, \Delta L_k$ *in verschiedenen Perioden* $j\neq k$ *sind unkorreliert, auch bedingt auf* $\mathcal{F}_{t_i}$ *für* $i<j,k$.

(iii) $$\mathrm{Var}[L_t]=\int_{[0,t]}\Big((A(u)-V(u))/K(u)\Big)^2\big(1-\Delta\Lambda_Y(u)\big)\,\mathrm{d}F_Y(u).$$

(iv) $$\mathrm{Var}[B]=\int_{[0,\infty)}\Big((A(u)-V(u))/K(u)\Big)^2\big(1-\Delta\Lambda_Y(u)\big)\,\mathrm{d}F_Y(u)\in[0,\infty].$$

Beweis. (i) folgt aus der Martingaleigenschaft (Proposition 4.7) und $L_0=V(0)=0$. (ii) ist Lemma 4.3 für das Martingal (L_t), in der bedingten Version mit $\mathbb{E}[\,\cdot|\mathcal{F}_{t_i}]$ statt $\mathbb{E}$. (iii) folgt aus der Integraldarstellung in Proposition 4.7 und der Isometrie (Proposition 4.5) mit $f(u)=((A(u)-V(u))/K(u))\mathbb{1}\{0<u\le t\}$. (iv) folgt aus (iii) mit $t\to\infty$, da $L_t=B$ für $t\ge Y$. □

Abbildung 4.1 zeigt typische Pfade des Verlustprozesses.

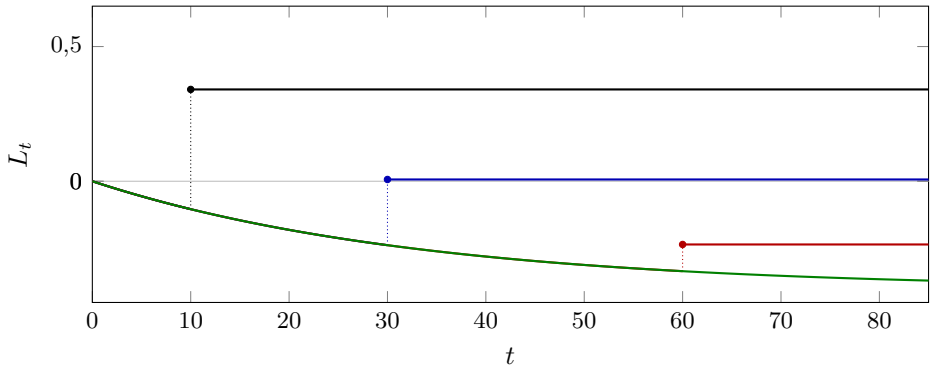

Abbildung 4.1: Vier Pfade des Verlustprozesses (L_t) bei Nettoeinmalprämie $\tilde\Pi=\lambda/(\lambda+\delta)=0{,}4$ für $T\sim\mathrm{Exp}(0{,}02)$, $\delta=0{,}03$, $A\equiv 1$; Todeszeitpunkte $T=10,30,60$ und ein Leben, das länger als 85 Jahre dauert (illustrativ gewählt). Solange die Person lebt, sinkt $L_t=\tilde\Pi(\mathrm{e}^{-\delta t}-1)$ deterministisch; beim Tod springt L um das abgezinste riskierte Kapital $(A-V(T))/K(T)=0{,}6\,\mathrm{e}^{-\delta T}$ und bleibt danach konstant gleich B. Im Mittel heben sich Drift und Sprünge auf: $\mathbb{E}[L_t]=0$.

Interpretation. Das Risiko eines Vertrags zerfällt additiv in die Risiken der einzelnen Perioden: $\mathrm{Var}[B]=\sum_i\mathrm{Var}[\Delta L_i]$. Das Risiko einer Periode ist umso größer, je größer das riskierte Kapital $A-V$ in dieser Periode und je wahrscheinlicher dort der Leistungsfall ist. Am Atom $Y=\tau$ (Erleben des Vertragsendes) ist $\Delta\Lambda_Y(\tau)=1$; dort entsteht kein Risiko, weil der Ausgang bereits feststeht.

Beispiel 4.9 (Natürliche Prämie gegen Einmalprämie)**.** Sei F stetig und $A(\tau) = 0$. Für die stetige natürliche Prämie ist $V \equiv 0$ (Bemerkung 3.4), also

$$\mathrm{Var}[B^{\mathrm{nat}}] = \int \frac{A^2}{K^2}\,\mathrm{d}F_Y = \mathbb{E}\Big[\frac{A(Y)^2}{K(Y)^2}\Big].$$

Für die Einmalprämie ist $\mathrm{Var}[B] = \mathrm{Var}[A(Y)/K(Y)]$. Die natürliche Prämie hat also eine um $\big(\mathbb{E}[A(Y)/K(Y)]\big)^2$ größere Varianz.

Aufgaben

Aufgabe 4.1 (vgl. [BOS25, S. 54])**.** Zeigen Sie an einer reinen Erlebensfallversicherung, dass die Einmalprämie ohne die Monotonievoraussetzung an A/K nicht varianzminimal sein muss.

Aufgabe 4.2 (vgl. [BOS25, Übung 3.23])**.** Die Verteilung von Y habe Atome der Masse $1/2$ in 1 und 2. Bestimmen Sie den Kompensator $\Lambda_Y(t \wedge Y)$ und überprüfen Sie die Varianzformel aus Proposition 4.5.

Aufgabe 4.3. Sei $T \sim \mathrm{Exp}(\lambda)$, $\tau = \infty$, $A \equiv 1$, $K(t) = \mathrm{e}^{\delta t}$ und Einmalprämie. Berechnen Sie $V(t)$ und $\mathrm{Var}[B]$ einmal direkt und einmal mit dem Satz von Hattendorff.

Teil II

Schadenversicherungsmathematik

Kapitel 5

Statische Risikomodelle

In der Schadenversicherung (Kfz, Feuer, Haftpflicht, Naturgefahren) ist nicht nur der Zeitpunkt, sondern auch die *Höhe* der Leistung zufällig. Die Verträge sind meist kurz (ein Jahr), deshalb vernachlässigen wir in Teil II die Verzinsung. Wir betrachten zunächst *statische* Modelle: den Gesamtschaden eines Portfolios in einer festen Periode.

5.1 Individuelles und kollektives Modell

Definition 5.1 ([BOS25, Def. 4.1, 4.2, 4.7])**.** (a) *Individuelles Modell:* $n \in \mathbb{N}$ unabhängige Schäden $Y_1, \dots, Y_n \geq 0$, einer pro Vertrag (oft mit $\mathbb{P}(Y_i = 0) > 0$), Gesamtschaden $S_{\text{ind}} = \sum_{i=1}^{n} Y_i$.

(b) *Kollektives Modell:* eine Schadenanzahl $N \in \mathbb{N}_0$ und identisch verteilte Schadenhöhen $X_1, X_2, \dots > 0$, Gesamtschaden $S_{\text{koll}} = \sum_{j=1}^{N} X_j$.

(c) Sind im kollektiven Modell $N, X_1, X_2, \dots$ unabhängig, so spricht man vom *Standardmodell der kollektiven Risikotheorie*; der Gesamtschaden sei dann $S = \sum_{j=1}^{N} X_j$.

Das individuelle Modell zählt Verträge, das kollektive Modell zählt Schäden. Man kann das erste in das zweite überführen: Setze $N := \#\{i \colon Y_i > 0\}$ und nehme als Schadenhöhen die positiven Y_i. Im allgemeinen Fall sind die X_j dann Mischungen der bedingten Verteilungen $\mathbb{P}(Y_i \in \cdot | Y_i > 0)$ und nicht unabhängig von N [BOS25, Prop. 4.4]. Im *homogenen* Fall (alle Y_i identisch verteilt, $p = \mathbb{P}(Y_1 > 0)$) erhält man dagegen ein Standardmodell mit $N \sim \text{Bin}(n, p)$ und $X_j \sim \mathbb{P}(Y_1 \in \cdot | Y_1 > 0)$, das *Binomialmodell*.

5.2 Schadenhöhen und Schadenanzahlen

Verteilung	Dichte bzw. Tail	Parameter	Tail
Exponential $\text{Exp}(\lambda)$	$\overline{F}(x) = \mathrm{e}^{-\lambda x}$	$\lambda > 0$	leicht
Gamma $\Gamma(\alpha, \beta)$	$f(x) = (\beta^\alpha / \Gamma(\alpha)) x^{\alpha-1} \mathrm{e}^{-\beta x}$	$\alpha, \beta > 0$	leicht
Weibull $W(\lambda, \alpha)$	$\overline{F}(x) = \mathrm{e}^{-(\lambda x)^\alpha}$	$\lambda, \alpha > 0$	schwer gdw. $\alpha < 1$
Pareto $P(\kappa, \alpha)$	$\overline{F}(x) = (\kappa/x)^\alpha$, $x \geq \kappa$	$\kappa, \alpha > 0$	schwer
Lomax $P^*(\kappa, \alpha)$	$\overline{F}(x) = (1 + x/\kappa)^{-\alpha}$	$\kappa, \alpha > 0$	schwer
Log-Normal $LN(\mu, \sigma^2)$	$\log X \sim N(\mu, \sigma^2)$	$\mu \in \mathbb{R}, \sigma > 0$	schwer
Log-Gamma	$\log X \sim \Gamma(\alpha, \beta)$	$\alpha, \beta > 0$	schwer

Die wichtigste qualitative Eigenschaft einer Schadenhöhenverteilung ist, wie schnell ihr Tail abfällt. Großschäden (Feuer, Sturm, Haftpflicht) können den Gesamtschaden eines ganzen Portfolios dominieren.

Definition 5.2 (Leichte und schwere Tails, [BOS25, Def. 4.8])**.** Eine Zufallsvariable $Z \geq 0$ mit Verteilungsfunktion F heißt *light-tailed*, wenn $\mathbb{E}[\mathrm{e}^{\lambda Z}] < \infty$ für ein $\lambda > 0$, und sonst *heavy-tailed*. Die Klasse der heavy-tailed Verteilungen bezeichnen wir mit $\mathcal{K}$.

Proposition 5.3 ([BOS25, Prop. 4.10]). *Für F mit $F(0-)=0$ sind äquivalent:*

(i) *F ist light-tailed;*

(ii) *es gibt $\lambda>0$ mit $\mathrm{e}^{\lambda x}\overline{F}(x)\to 0$ für $x\to\infty$;*

(iii) *es gibt $\lambda>0$ mit $\limsup_{x\to\infty}\mathrm{e}^{\lambda x}\overline{F}(x)<\infty$.*

Beweis. (i)⇒(ii): Ist $\mathbb{E}[\mathrm{e}^{\lambda Z}]<\infty$, so gilt nach Markov $\overline{F}(x)\le \mathrm{e}^{-\lambda x}\mathbb{E}[\mathrm{e}^{\lambda Z}]$, also $\mathrm{e}^{\lambda x/2}\overline{F}(x)\to 0$. (ii)⇒(iii) gilt, da konvergente Folgen beschränkt sind. (iii)⇒(i): Ist $\overline{F}(x)\le C\mathrm{e}^{-\lambda x}$ für große x, so ist für $0<s<\lambda$ nach Fubini

$$\mathbb{E}[\mathrm{e}^{sZ}] = 1+\int_0^\infty s\mathrm{e}^{sx}\overline{F}(x)\,\mathrm{d}x<\infty. \qquad \square$$

Schwere Tails haben also einen Tail, der langsamer als jede Exponentialfunktion abfällt. Ist $\mathbb{E}[Z^n]=\infty$ für ein n, so ist Z heavy-tailed; die Umkehrung gilt nicht (Log-Normal hat alle Momente).

Abbildung 5.1 macht die Unterschiede zwischen den Tails sichtbar.

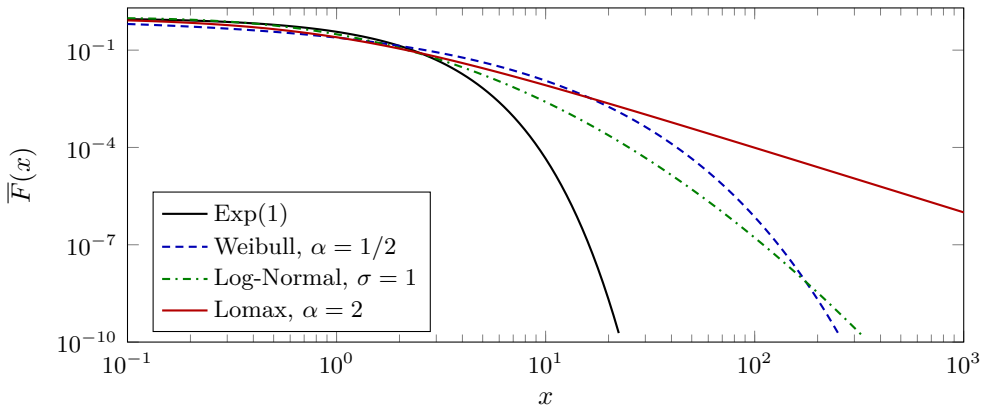

Abbildung 5.1: Tails $\overline{F}(x)$ von vier Schadenhöhenverteilungen mit gleichem Erwartungswert 1 (doppelt logarithmisch). Bei $x=100$ liegen die Werte zwischen $\mathrm{e}^{-100}\approx 4\cdot 10^{-44}$ (exponentiell) und $(101)^{-2}\approx 10^{-4}$ (Lomax). Regulär variierende Tails erscheinen asymptotisch als Geraden mit Steigung $-\alpha$; Weibull mit $\alpha<1$ und Log-Normal sind heavy-tailed, fallen aber schneller als jede Potenz.

Schadenanzahlen.

- *Binomial* $\mathrm{Bin}(n,\theta)$ entsteht im Binomialmodell; für große Portfolios unhandlich.
- *Poisson* $\mathrm{Poi}(\lambda)$ approximiert $\mathrm{Bin}(n,\theta)$ für großes n und kleines θ mit $n\theta\approx\lambda$; Erwartungswert gleich Varianz.
- *Negativ binomial* $\mathrm{NB}(\beta,\theta)$ mit $\mathbb{P}(N=k)=\binom{\beta+k-1}{k}\theta^\beta(1-\theta)^k$: Varianz größer als Erwartungswert („overdispersion“), wie in vielen Daten beobachtet.

Beispiel 5.4 (Negativ binomial als Poisson–Gamma-Mischung). Ist $\Theta\sim\Gamma(\alpha,\beta)$ und N gegeben $\Theta=\theta$ Poisson(θ)-verteilt, so ist

$$\mathbb{P}(N=k)=\int_0^\infty \mathrm{e}^{-\theta}\frac{\theta^k}{k!}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\mathrm{e}^{-\beta\theta}\,\mathrm{d}\theta=\binom{\alpha+k-1}{k}\Big(\frac{\beta}{\beta+1}\Big)^\alpha\Big(\frac{1}{\beta+1}\Big)^k,$$

also $N\sim\mathrm{NB}(\alpha,\beta/(\beta+1))$. Interpretation: Die Versicherten haben unterschiedliche, zufällige Schadenfrequenzen Θ. Wir kommen in Kapitel 11 darauf zurück.

5.3 Die Verteilung des Gesamtschadens

Definition 5.5 (Faltung). Für Verteilungsfunktionen F,G ist $(F*G)(x):=\int_{\mathbb{R}}F(x-t)\,\mathrm{d}G(t)$; ferner $F^{*0}(x)=\mathbb{1}\{x\ge 0\}$ und $F^{*n}=F^{*(n-1)}*F$.

Sind X,Y unabhängig, so ist $F_{X+Y}=F_X*F_Y$. Hat F eine Dichte f, so hat $F*G$ die Dichte $\int f(x-t)\,\mathrm{d}G(t)$.

Proposition 5.6 ([BOS25, Prop. 4.18])**.** *Im Standardmodell mit Schadenhöhenverteilung F gilt*

$$\mathbb{P}(S \le x) = \sum_{n=0}^{\infty} \mathbb{P}(N = n)\, F^{*n}(x).$$

Beweis. Bedinge auf $\{N = n\}$ und nutze die Unabhängigkeit von N und (X_j). □

Die Verteilung von S heißt *zusammengesetzte Verteilung*; für $N \sim \mathrm{Poi}(\lambda)$ spricht man von der *zusammengesetzten Poisson-Verteilung* $\mathrm{CPoi}(\lambda, F)$. Explizit wird die Formel nur selten, etwa bei Gamma-verteilten Schäden, denn $\Gamma(\alpha,\beta) * \Gamma(\bar\alpha,\beta) = \Gamma(\alpha+\bar\alpha,\beta)$.

Erzeugende Funktionen. Für eine Zufallsvariable Z sei

- $\psi_Z(t) = \mathbb{E}[\mathrm{e}^{tZ}]$ die momentenerzeugende Funktion auf $D_Z = \{t\colon \psi_Z(t) < \infty\}$,
- $\varphi_Z(s) = \mathbb{E}[s^Z]$ die wahrscheinlichkeitserzeugende Funktion $(Z \ge 0)$,
- $\chi_Z(t) = \mathbb{E}[\mathrm{e}^{itZ}]$ die charakteristische Funktion.

Für $\mathrm{Exp}(\lambda)$ ist $\psi(t) = \lambda/(\lambda - t)$ auf $(-\infty, \lambda)$, für $\mathrm{Poi}(\lambda)$ ist $\psi(t) = \exp(\lambda(\mathrm{e}^t - 1))$. Ist $Z \ge 0$ heavy-tailed, so ist $D_Z = (-\infty, 0]$. Für $\mathbb{N}_0$-wertiges Z bestimmt φ_Z die Verteilung über $\mathbb{P}(Z = n) = \varphi_Z^{(n)}(0)/n!$; allgemein bestimmt χ_Z die Verteilung, ebenso ψ_Z, falls D_Z eine Umgebung der 0 enthält [BOS25, Satz 4.27]. Für unabhängige Summanden multiplizieren sich die erzeugenden Funktionen.

Proposition 5.7 ([BOS25, Prop. 4.29])**.** *Im Standardmodell gilt $\psi_S(t) = \varphi_N\big(\psi_{X_1}(t)\big)$ für alle t mit $t \in D_{X_1}$ und $\psi_{X_1}(t)$ im Definitionsbereich von φ_N.*

Beweis. Es gilt $\mathbb{E}[\mathrm{e}^{tS}] = \sum_n \mathbb{P}(N=n)\mathbb{E}[\mathrm{e}^{t(X_1+\cdots+X_n)}] = \sum_n \mathbb{P}(N=n)\psi_{X_1}(t)^n$. □

Korollar 5.8. *Für $S \sim \mathrm{CPoi}(\lambda, F)$ ist $\psi_S(t) = \exp\big(\lambda(\psi_{X_1}(t) - 1)\big)$. Sind $S_1, \dots, S_k$ unabhängig mit $S_i \sim \mathrm{CPoi}(\lambda_i, F_i)$, so ist $\sum_i S_i \sim \mathrm{CPoi}(\lambda, F)$ mit $\lambda = \sum_i \lambda_i$ und $F = \sum_i (\lambda_i/\lambda) F_i$.*

Das Zusammenlegen unabhängiger Portfolios bleibt also in der Klasse der zusammengesetzten Poisson-Verteilungen; die Schadenhöhe ist eine Mischung. Die Faltungsstabilität der Schadenanzahlverteilungen ($\mathrm{Bin}(n,\theta) * \mathrm{Bin}(\bar n,\theta) = \mathrm{Bin}(n+\bar n,\theta)$, analog für Poisson und für $\mathrm{NB}(\cdot,\theta)$) beweist man ebenso mit φ_N.

Abbildung 5.2 zeigt zwei zusammengesetzte Poisson-Verteilungen.

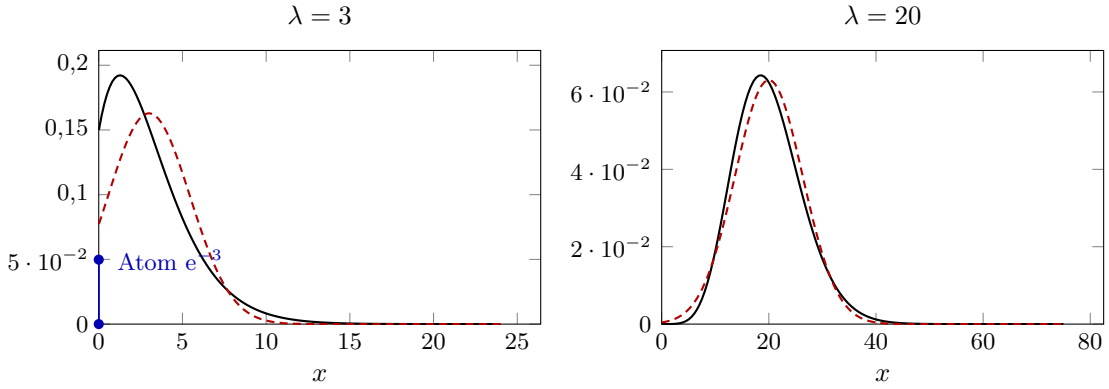

Abbildung 5.2: Zusammengesetzte Poisson-Verteilung $\mathrm{CPoi}(\lambda, \mathrm{Exp}(1))$: Dichte des absolut stetigen Teils (durchgezogen) und Normalverteilung mit gleichem Erwartungswert λ und gleicher Varianz 2λ (gestrichelt). Für $\lambda = 3$ trägt $\{S = 0\}$ die Masse $\mathrm{e}^{-3} \approx 0{,}05$ (Punkt, keine Dichte), und die Verteilung ist deutlich rechtsschief. Für $\lambda = 20$ ist die Normalapproximation im Zentrum gut, unterschätzt aber den rechten Tail.

5.4 Die Panjer-Rekursion

Die Faltungsformel aus Proposition 5.6 ist numerisch teuer. Sind die Schadenhöhen ganzzahlig (in Geldeinheiten diskretisiert), so lassen sich die Gewichte des Gesamtschadens für eine wichtige Klasse von Schadenanzahlverteilungen rekursiv berechnen. In diesem Abschnitt seien die X_j $\mathbb{N}_0$-wertig mit $f_k := \mathbb{P}(X_1 = k)$ (wir erlauben $f_0 > 0$), sowie $p_n := \mathbb{P}(N = n)$ und $g_s := \mathbb{P}(S = s)$.

Definition 5.9 (Panjer-Klasse, [BOS25, Def. 4.64])**.** Eine Verteilung (p_n) auf $\mathbb{N}_0$ gehört zur *Panjer-Klasse*, wenn es $a, b \in \mathbb{R}$ mit $a + b > 0$ gibt, so dass

$$p_n = \Big(a + \frac{b}{n}\Big) p_{n-1}, \qquad n \geq 1.$$

	a	b	$g_0 = \varphi_N(f_0)$
$\mathrm{Bin}(m, \theta)$	$-\theta/(1-\theta)$	$(m+1)\theta/(1-\theta)$	$(1-\theta+\theta f_0)^m$
$\mathrm{Poi}(\lambda)$	0	λ	$\mathrm{e}^{-\lambda(1-f_0)}$
$\mathrm{NB}(\beta, \theta)$	$1-\theta$	$(\beta-1)(1-\theta)$	$\big(\theta/(1-(1-\theta)f_0)\big)^\beta$

Satz 5.10 (Sundt–Jewell 1981, [BOS25, Satz 4.68])**.** *Die Panjer-Klasse besteht genau aus den Binomial-, Poisson- und negativen Binomialverteilungen.*

Beweisskizze. Ist $a < 0$, so wird $a + b/n$ für große n negativ; damit alle $p_n \geq 0$ sind, muss die Folge vorher abbrechen, d. h. $b/a = -(m+1)$ mit $m \in \mathbb{N}$: binomial. Für $a = 0$ ist $p_n = p_0 b^n / n!$: Poisson. Für $0 < a < 1$ ist $p_n = p_0 \binom{\beta+n-1}{n} a^n$ mit $\beta = 1 + b/a$: negativ binomial mit $\theta = 1 - a$. Für $a \geq 1$ ist $\sum p_n = \infty$. □

Satz 5.11 (Panjer-Rekursion 1981, [BOS25, Satz 4.70])**.** *Ist N aus der Panjer-Klasse und sind $N, X_1, X_2, \ldots$ unabhängig, so gilt $g_0 = \varphi_N(f_0)$ und*

$$g_s = \frac{1}{1 - a f_0} \sum_{k=1}^{s} \Big(a + \frac{bk}{s}\Big) f_k \, g_{s-k}, \qquad s \geq 1.$$

Beweis. Es gilt $g_0 = \sum_n p_n f_0^n = \varphi_N(f_0)$. Sei $s \geq 1$ und $S_n = X_1 + \cdots + X_n$. Aus Symmetriegründen ist $\mathbb{E}[X_j | S_n = s]$ für alle $j \leq n$ gleich, also gleich s/n. Daher gilt für $n \geq 1$

$$\Big(a + \frac{b}{n}\Big) \mathbb{P}(S_n = s) = \mathbb{E}\Big[\Big(a + \frac{bX_1}{s}\Big) \mathbb{1}\{S_n = s\}\Big] = \sum_{k=0}^{s} \Big(a + \frac{bk}{s}\Big) f_k \, \mathbb{P}(S_{n-1} = s - k).$$

Multipliziert man mit p_{n-1} und summiert über $n \geq 1$ (der Term $n = 0$ verschwindet wegen $s \geq 1$), so erhält man

$$g_s = \sum_{n \geq 1} p_n \mathbb{P}(S_n = s) = \sum_{k=0}^{s} \Big(a + \frac{bk}{s}\Big) f_k \, g_{s-k}.$$

Bringt man den Summanden $k = 0$, also $a f_0 g_s$, auf die linke Seite, folgt die Behauptung. □

Für $s \leq s_{\max}$ benötigt die Rekursion $O(s_{\max}^2)$ Operationen, statt alle Faltungspotenzen f^{*n} zu berechnen. In der Praxis konkurriert sie mit Fourier-Methoden (schnelle Fourier-Transformation von $\varphi_N(\chi_{X_1})$).

Abbildung 5.3 zeigt das Ergebnis der Rekursion für ein konkretes Beispiel.

Aufgaben

Aufgabe 5.1 (vgl. [BOS25, Übung 4.13])**.** Zeigen Sie: Pareto-, Lomax-, Log-Gamma- und Log-Normalverteilung sind heavy-tailed; die Weibull-Verteilung ist genau für $\alpha < 1$ heavy-tailed.

Aufgabe 5.2 (vgl. [BOS25, Übung 4.11])**.** Finden Sie eine Verteilung mit $\liminf_x \mathrm{e}^{\lambda x} \overline{F}(x) = 0$ und $\limsup_x \mathrm{e}^{\lambda x} \overline{F}(x) = \infty$ für alle $\lambda > 0$.

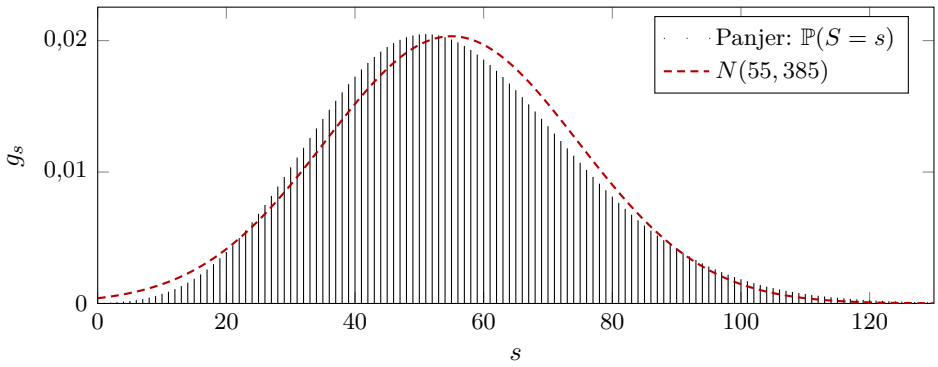

Abbildung 5.3: Exakte Verteilung des Gesamtschadens für $N \sim \text{Poi}(10)$ und auf $\{1, \dots, 10\}$ gleichverteilte Schadenhöhen, berechnet mit der Panjer-Rekursion ($\mathbb{P}(S = 0) = \mathrm{e}^{-10}$ ist nicht sichtbar). Die Normalverteilung mit gleichem Mittel und gleicher Varianz trifft das Zentrum, aber nicht den Tail: $\mathbb{P}(S > 100) \approx 0{,}0177$ exakt gegenüber 0,0102 in der Normalapproximation mit Stetigkeitskorrektur.

Aufgabe 5.3. Berechnen Sie $\mathbb{E}[S]$ und $\text{Var}[S]$ für $S \sim \text{CPoi}(\lambda, F)$ durch Ableiten von $\log \psi_S$ in 0. Was sind die höheren Kumulanten?

Aufgabe 5.4 (vgl. [BOS25, Übung 4.33])**.** Beweisen Sie das Korollar über das Zusammenlegen unabhängiger zusammengesetzter Poisson-Portfolios.

Aufgabe 5.5 (vgl. [BOS25, Übung 4.63])**.** Bestimmen Sie a, b und p_0 für die Binomial- und die negative Binomialverteilung und verifizieren Sie die Tabelle.

Aufgabe 5.6. Implementieren Sie die Panjer-Rekursion für $N \sim \text{Poi}(10)$ und X_1 gleichverteilt auf $\{1, \dots, 10\}$. Vergleichen Sie $\mathbb{P}(S > 100)$ mit einer Normal- und einer Gamma-Approximation.

Kapitel 6

Abweichungen des Gesamtschadens vom Erwartungswert

Wie groß muss eine Prämie sein, damit der Gesamtschaden sie nur mit kleiner Wahrscheinlichkeit übersteigt? Dazu brauchen wir Schranken für $\mathbb{P}(S \geq \mathbb{E}[S] + c)$. Wir sehen zwei Arten: polynomielle aus Momenten und exponentielle aus der momentenerzeugenden Funktion. Welche zur Verfügung steht, hängt vom Tail der Schadenhöhe ab.

6.1 Die Waldschen Gleichungen

Satz 6.1 (Erste Waldsche Gleichung, [BOS25, Satz 4.34])**.** *Seien $X_1, X_2, \dots$ u. i. v. mit $\mathbb{E}|X_1| < \infty$, $N \in \mathbb{N}_0$ mit $\mathbb{E}[N] < \infty$, und für jedes n sei $\mathbb{1}\{N = n\}$ unabhängig von $(X_{n+1}, X_{n+2}, \dots)$. Dann gilt $\mathbb{E}\big[\sum_{k=1}^{N} X_k\big] = \mathbb{E}[N]\,\mathbb{E}[X_1]$.*

Beweis. Es ist $\sum_{k=1}^{N} X_k = \sum_{k\geq 1} X_k \mathbb{1}\{N \geq k\}$, und $\{N \geq k\} = \{N \leq k-1\}^{\mathrm{c}}$ ist unabhängig von X_k. Also $\mathbb{E}\sum_k |X_k|\mathbb{1}\{N \geq k\} = \mathbb{E}|X_1| \sum_k \mathbb{P}(N \geq k) = \mathbb{E}|X_1|\mathbb{E}[N] < \infty$, und dieselbe Rechnung ohne Beträge liefert die Behauptung. □

Die Voraussetzung ist insbesondere erfüllt, wenn N eine *Stoppzeit* ist, d. h. $\{N = n\} \in \sigma(X_1, \dots, X_n)$ für alle $n \in \mathbb{N}_0$; die Anzahl N darf also von den bereits beobachteten $X_1, \dots, X_N$ abhängen. Diese Allgemeinheit nutzen wir in der Erneuerungstheorie (Kapitel 7).

Satz 6.2 (Zweite Waldsche Gleichung, [BOS25, Satz 4.37, 4.38])**.** *Gelte zusätzlich $\mathbb{E}[X_1^2] < \infty$; sei $\mu = \mathbb{E}[X_1]$, $\sigma^2 = \operatorname{Var}[X_1]$. Ist N eine Stoppzeit bezüglich einer Filtration $(\mathcal{F}_n)_{n\geq 0}$, für die X_n $\mathcal{F}_n$-messbar und X_{n+1} unabhängig von $\mathcal{F}_n$ ist, so gilt $\mathbb{E}\big[(S_N - \mu N)^2\big] = \sigma^2 \mathbb{E}[N]$. Sind $N, X_1, X_2, \dots$ unabhängig und $\mathbb{E}[N^2] < \infty$, so gilt*

$$\operatorname{Var}[S_N] = \sigma^2\,\mathbb{E}[N] + \mu^2\,\operatorname{Var}[N].$$

Beweis. Wir zeigen nur die zweite Formel. Gegeben $N = n$ hat S_N Erwartungswert $n\mu$ und Varianz $n\sigma^2$. Die Formel der totalen Varianz $\operatorname{Var}[S_N] = \mathbb{E}[\operatorname{Var}[S_N|N]] + \operatorname{Var}[\mathbb{E}[S_N|N]]$ liefert $\mathbb{E}[N]\sigma^2 + \operatorname{Var}[N\mu]$. Die erste Formel beweist man mit dem Martingal $M_n = \sum_{i\leq n}(X_i - \mu)\mathbb{1}\{N \geq i\}$ [BOS25, S. 98–100]. Die Stoppzeiteigenschaft ist dabei wesentlich: Die Bedingung aus der ersten Waldschen Gleichung genügt nicht (Aufgabe 6.3). □

Die zweite Formel zerlegt die Varianz in einen Anteil aus den Schadenhöhen und einen aus der Schadenanzahl. Im zusammengesetzten Poisson-Modell ist $\operatorname{Var}[S] = \lambda \mathbb{E}[X_1^2]$.

6.2 Polynomielle Schranken

Proposition 6.3 (Markov, Tschebyschow, Cantelli)**.** *Sei X reell.*

(a) Für wachsendes $h \geq 0$ und $c > 0$ mit $h(c) > 0$: $\mathbb{P}(|X| \geq c) \leq \mathbb{E}[h(|X|)]/h(c)$.

(b) Ist $\operatorname{Var}[X] < \infty$, so gilt $\mathbb{P}(|X - \mathbb{E}X| \geq c) \leq \operatorname{Var}[X]/c^2$.

(c) (Cantelli) Ist $\mathrm{Var}[X] = \sigma^2 < \infty$*, so gilt* $\mathbb{P}(X - \mathbb{E}X \geq c) \leq \sigma^2/(c^2 + \sigma^2)$ *für* $c > 0$.

Beweis. (a) folgt aus $h(c)\mathbb{1}\{|X| \geq c\} \leq h(|X|)$; (b) ist (a) mit $h(x) = x^2$. (c) Sei o. E. $\mathbb{E}X = 0$. Für jedes $x \geq 0$ ist nach (a)

$$\mathbb{P}(X \geq c) \leq \mathbb{P}\big((X+x)^2 \geq (c+x)^2\big) \leq \frac{\sigma^2 + x^2}{(c+x)^2}.$$

Die Wahl $x = \sigma^2/c$ minimiert die rechte Seite und liefert $\sigma^2/(c^2+\sigma^2)$. □

Die Cantelli-Ungleichung ist scharf: Für $\mathbb{P}(X = c) = 1/(1+c^2)$, $\mathbb{P}(X = -1/c) = c^2/(1+c^2)$ gilt Gleichheit. Für den Gesamtschaden im Standardmodell erhält man mit Satz 6.2

$$\mathbb{P}\big(S \geq \mathbb{E}[N]\mu + c\big) \leq \frac{\mathbb{E}[N]\sigma^2 + \mathrm{Var}[N]\mu^2}{c^2 + \mathbb{E}[N]\sigma^2 + \mathrm{Var}[N]\mu^2}.$$

Anwendung: Prämie mit Sicherheitszuschlag. Soll die Prämie $P = \mathbb{E}[S] + c$ nur mit Wahrscheinlichkeit $\leq \varepsilon$ überschritten werden, so genügt nach Cantelli

$$c \geq \sqrt{\mathrm{Var}[S]}\,\sqrt{(1-\varepsilon)/\varepsilon}.$$

Das ist ein Sicherheitszuschlag proportional zur Standardabweichung; er benötigt nur zwei Momente, ist dafür aber grob.

6.3 Große Abweichungen

Definition 6.4 (Ratenfunktion, [BOS25, Def. 4.48])**.** Für eine reelle Zufallsvariable X sei $\eta_X(s) \coloneqq \log \mathbb{E}[\mathrm{e}^{sX}] \in (-\infty, \infty]$ die *kumulantenerzeugende Funktion* und

$$I_X(b) \coloneqq \sup_{s \geq 0}\big(sb - \eta_X(s)\big), \qquad b \in \mathbb{R},$$

die (einseitige) *Ratenfunktion.*

Die Funktionen η_X und I_X sind konvex (Hölder), $I_X \geq 0$ (wähle $s = 0$), I_X ist wachsend und $I_X(\mathbb{E}X) = 0$ (Jensen). Ist $\mathbb{P}(X \geq x) = 0$, so ist $I_X(x') = \infty$ für $x' > x$.

Proposition 6.5 (Chernoff-Schranke, [BOS25, Prop. 4.55])**.** *Für u. i. v.* $X_1, \dots, X_n$ *und* $S_n = X_1 + \cdots + X_n$ *gilt* $\mathbb{P}(S_n \geq bn) \leq \mathrm{e}^{-nI_{X_1}(b)}$ *für alle* $b \in \mathbb{R}$.

Beweis. Für $s \geq 0$ ist nach Markov $\mathbb{P}(S_n \geq bn) \leq \mathrm{e}^{-sbn}\mathbb{E}[\mathrm{e}^{sS_n}] = \exp\big(-n(sb - \eta_{X_1}(s))\big)$. Nimm das Infimum über $s \geq 0$. □

Im homogenen individuellen Modell folgt $\mathbb{P}(S_{\mathrm{ind}} \geq a) \leq \mathrm{e}^{-nI_{Y_1}(a/n)}$; im Standardmodell durch Bedingen auf N $\mathbb{P}(S \geq a) \leq \sum_n \mathbb{P}(N = n)\mathrm{e}^{-nI_{X_1}(a/n)}$. Die Schranke ist im exponentiellen Sinne optimal:

Satz 6.6 (Cramér 1938, [BOS25, Satz 4.58])**.** *Seien* $X_1, X_2, \dots$ *u. i. v. und sei* η_{X_1} *in einer rechtsseitigen Umgebung der* 0 *endlich. Dann ist* $\mu \coloneqq \mathbb{E}[X_1] \in [-\infty, \infty)$ *wohldefiniert, und für alle* $b \geq \mu$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq bn) = -I_{X_1}(b).$$

Beweisskizze. Die obere Schranke ist Proposition 6.5. Für die untere Schranke nehmen wir an, dass $b = \eta'(s)$ für ein $s > 0$ im Inneren des Definitionsbereichs. Unter dem *exponentiell gekippten Maß* $\mathrm{d}Q = \mathrm{e}^{sX_1 - \eta(s)}\,\mathrm{d}\mathbb{P}$ (für jeden Summanden) hat X_1 den Erwartungswert $\eta'(s) = b$. Für $\varepsilon > 0$ ist

$$\mathbb{P}(S_n \geq bn) \geq \mathbb{E}_Q\big[\mathrm{e}^{-sS_n + n\eta(s)}\mathbb{1}\{bn \leq S_n \leq (b+\varepsilon)n\}\big] \geq \mathrm{e}^{-n(sb-\eta(s)) - sn\varepsilon}\, Q\big(bn \leq S_n \leq (b+\varepsilon)n\big),$$

und die letzte Wahrscheinlichkeit konvergiert nach dem zentralen Grenzwertsatz gegen $1/2$. Da $sb - \eta(s) = I(b)$ für dieses s, folgt $\liminf (1/n)\log \mathbb{P}(S_n \geq bn) \geq -I(b) - s\varepsilon$. Den allgemeinen Fall behandelt [BOS25, S. 106–109]. □

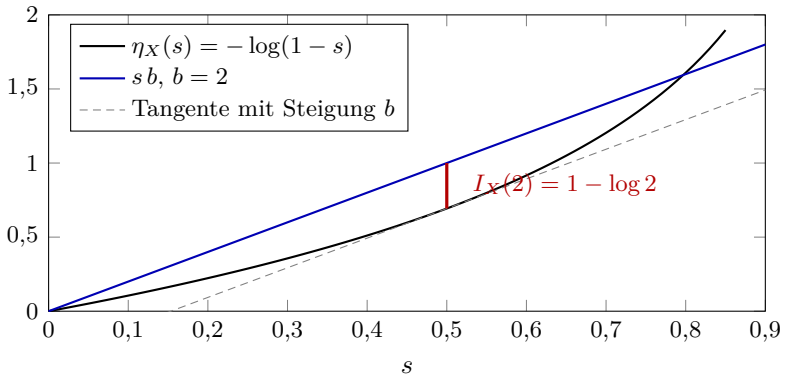

Abbildung 6.1: Die Ratenfunktion als Legendre-Transformierte, für $X \sim \mathrm{Exp}(1)$ und $b = 2$: $I_X(b)$ ist der größte Abstand der Geraden $s \mapsto sb$ über der konvexen Funktion η_X. Er wird dort angenommen, wo $\eta_X'(s^*) = b$, hier bei $s^* = 1/2$; das ist zugleich der Parameter des gekippten Maßes im Beweis des Satzes von Cramér, unter dem X den Erwartungswert b hat.

Die Idee des Maßwechsels ist zentral: Das unwahrscheinliche Ereignis wird *typisch* realisiert, indem alle Summanden ein wenig größer ausfallen. Bei schweren Tails ist das anders (Abschnitt 8.3).

Abbildung 6.1 veranschaulicht die Definition der Ratenfunktion.

Beispiel 6.7 (Ratenfunktionen, [BOS25, Bsp. 4.59])**.** Für $b \geq \mathbb{E}X$ gilt:

- $X \sim \mathrm{Exp}(\lambda)$: $I(b) = \lambda b - 1 - \log(\lambda b)$;
- $X \sim \mathrm{Poi}(\lambda)$: $I(b) = b\log(b/\lambda) - b + \lambda$;
- $X \sim \mathrm{Ber}(\theta)$: $I(b) = b\log(b/\theta) + (1-b)\log((1-b)/(1-\theta))$, $b \leq 1$.

Beispiel 6.8 (Vergleich der Schranken)**.** Wir betrachten $n = 1000$ Verträge mit Schadenwahrscheinlichkeit $\theta = 0{,}01$ und Schadenhöhe 1. Gefragt ist $\mathbb{P}(S \geq 20)$, also doppelt so viele Schäden wie erwartet. Cantelli mit $c = 10$, $\sigma^2 = 9{,}9$ liefert 0,090; Chernoff mit $b = 0{,}02$ liefert $\mathrm{e}^{-1000 \cdot I(0{,}02)} \approx 0{,}020$. Der exakte Wert ist etwa 0,0033.

Abbildung 6.2 vergleicht die Schranken über einen ganzen Bereich von Schwellen.

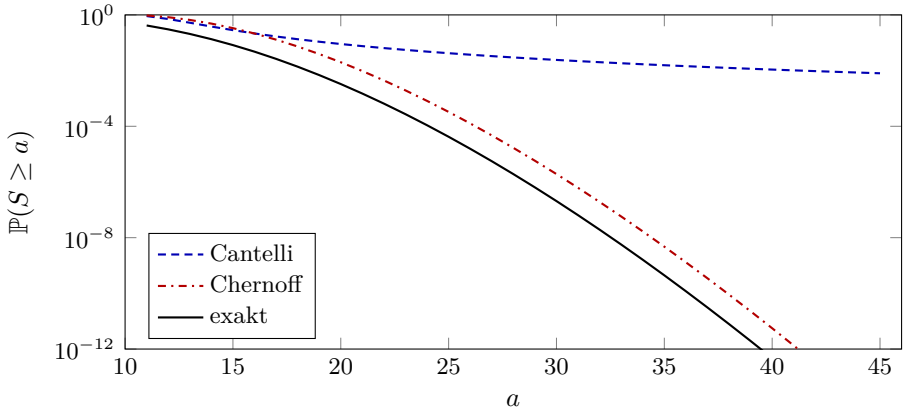

Abbildung 6.2: Schranken für $\mathbb{P}(S \geq a)$ mit $S \sim \mathrm{Bin}(1000; 0{,}01)$, $\mathbb{E}S = 10$. Cantelli fällt nur wie a^{-2}; die Chernoff-Schranke fällt exponentiell mit der richtigen Rate, liegt aber um einen subexponentiellen Faktor über dem exakten Wert (bei $a = 40$ etwa um den Faktor 12). Logarithmische Skala.

Proposition 6.9 ([BOS25, Prop. 4.61])**.** *Sei $X \geq 0$ mit Ratenfunktion I. Ist X heavy-tailed, so ist $I \equiv 0$; andernfalls gibt es x mit $I(x) > 0$.*

Beweis. Ist X heavy-tailed, so ist $\eta_X(s) = \infty$ für alle $s > 0$ und das Supremum wird bei $s = 0$ angenommen. Ist $\mathbb{E}[\mathrm{e}^{\lambda X}] < \infty$, so ist $I(x) \geq \lambda x - \eta_X(\lambda) > 0$ für großes x. □

Für schwere Tails liefert die Chernoff-Methode also nichts: Die Wahrscheinlichkeit großer Abweichungen fällt dann nicht exponentiell in n.

Aufgaben

Aufgabe 6.1 (vgl. [BOS25, Übung 4.60])**.** Leiten Sie die Ratenfunktionen aus dem Beispiel her und bestimmen Sie zusätzlich η_X und I_X für die Gammaverteilung.

Aufgabe 6.2. Zeigen Sie an einem Beispiel, dass die Varianzformel in Satz 6.2 ohne Unabhängigkeit von N und (X_j) falsch sein kann [BOS25, Bem. 4.39].

Aufgabe 6.3. Seien $X_1, X_2, \dots$ u. i. v. mit $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$, und sei $N = 2$ für $X_1 X_2 = 1$, sonst $N = 1$. Zeigen Sie, dass $\mathbb{1}\{N = n\}$ für jedes n von $(X_{n+1}, X_{n+2}, \dots)$ unabhängig ist, dass aber $\mathbb{E}[S_N^2] \neq \sigma^2 \mathbb{E}[N]$ gilt. Ist N eine Stoppzeit?

Aufgabe 6.4. Sei $S \sim \mathrm{CPoi}(\lambda, \mathrm{Exp}(1))$. Bestimmen Sie die Chernoff-Schranke für $\mathbb{P}(S \geq a)$ explizit und vergleichen Sie sie für $\lambda = 100$ und $a = 150$ numerisch mit Cantelli und mit einer Normalapproximation.

Kapitel 7

Erneuerungstheorie

Die Ruinwahrscheinlichkeit wird sich als Lösung einer *Erneuerungsgleichung* herausstellen (Abschnitt 8.2). Wir stellen daher die nötige Erneuerungstheorie bereit. Sie ist auch darüber hinaus nützlich, etwa in der Warteschlangen- und Zuverlässigkeitstheorie.

7.1 Erneuerungsprozesse

Definition 7.1 (Erneuerungsprozess, [BOS25, Def. 5.4])**.** Seien $W_1, W_2, \cdots > 0$ u. i. v. (*Wartezeiten*) mit Verteilungsfunktion G und $\mathbb{E}[W_1] = 1/\lambda \in (0, \infty]$. Mit $T_0 = 0$, $T_n = W_1 + \cdots + W_n$ heißt $N_t := \#\{n \geq 1 \colon T_n \leq t\}$ *Erneuerungsprozess* mit Intensität λ. Für $W_i \sim \mathrm{Exp}(\lambda)$ heißt er *Poisson-Prozess* mit Rate λ.

Proposition 7.2 (Starkes Gesetz, [BOS25, Prop. 5.5])**.** *Für einen Erneuerungsprozess gilt $N_t/t \to \lambda$ fast sicher.*

Beweis. Es ist $T_{N_t} \leq t < T_{N_t+1}$ und $N_t \to \infty$ f. s. Also $T_{N_t}/N_t \leq t/N_t < T_{N_t+1}/(N_t+1) \cdot (N_t+1)/N_t$, und beide Schranken konvergieren nach dem starken Gesetz der großen Zahlen gegen $\mathbb{E}[W_1] = 1/\lambda$ (auch für $\mathbb{E}W_1 = \infty$ durch Abschneiden). □

Der Poisson-Prozess hat unabhängige, stationäre Zuwächse mit $N_{t+s} - N_t \sim \mathrm{Poi}(\lambda s)$; diese Eigenschaften legen seine Verteilung fest. Der Prozess $(N_t - \lambda t)$ ist ein Martingal: λt ist der Kompensator wie in Kapitel 4.

7.2 Erneuerungsfunktion und Erneuerungsgleichung

Sei (N_t) ein Erneuerungsprozess mit Wartezeitverteilung G, $G(0) = 0$, Sprungzeiten T_n und $m := \mathbb{E}[W_1] \in (0, \infty]$.

7.2.1 Die Erneuerungsfunktion

Wegen $\{N_t \geq k\} = \{T_k \leq t\}$ gilt

$$\mathbb{P}(N_t \geq k) = G^{*k}(t), \qquad \mathbb{P}(N_t = k) = G^{*k}(t) - G^{*(k+1)}(t).$$

Proposition 7.3 ([BOS25, Prop. 5.21])**.** *Für alle $t \geq 0$ und $\gamma > 0$ ist $\mathbb{E}[\mathrm{e}^{\gamma N_t}] < \infty$.*

Beweis. Für $\theta > 0$ ist $L(\theta) := \mathbb{E}[\mathrm{e}^{-\theta W_1}] < 1$ wegen $W_1 > 0$, und $L(\theta) \to 0$ für $\theta \to \infty$. Nach Markov ist $\mathbb{P}(N_t \geq k) = \mathbb{P}(\mathrm{e}^{-\theta T_k} \geq \mathrm{e}^{-\theta t}) \leq \mathrm{e}^{\theta t} L(\theta)^k$. Wähle θ so groß, dass $\mathrm{e}^{\gamma} L(\theta) < 1$; dann ist $\sum_k \mathrm{e}^{\gamma k} \mathbb{P}(N_t \geq k) < \infty$. □

Definition 7.4 (Erneuerungsfunktion)**.** Die Funktion $M(t) := \mathbb{E}[N_t]$, $t \geq 0$, heißt *Erneuerungsfunktion.*

Es ist $M(t) = \sum_{k \geq 1} \mathbb{P}(N_t \geq k) = \sum_{k \geq 1} G^{*k}(t) < \infty$, $M(0) = 0$, und M ist wachsend und rechtsstetig. Für den Poisson-Prozess ist $M(t) = \lambda t$.

Definition 7.5 (Faltung)**.** Für lokal beschränktes messbares A und B von lokal beschränkter Variation (Differenz zweier wachsender, rechtsstetiger Funktionen) sei

$$(A * B)(t) := \int_{(0,t]} A(t-s)\,\mathrm{d}B(s), \qquad t \geq 0.$$

Für $A(0) = B(0) = 0$ ist die Faltung kommutativ (partielle Integration), und sie ist assoziativ, sofern alle Terme definiert sind.

Proposition 7.6 ([BOS25, Prop. 5.26])**.** *Es gilt $M = G + M * G$.*

Beweis. Es gilt $M = G + \sum_{k\geq 2} G^{*(k-1)} * G = G + \big(\sum_{j\geq 1} G^{*j}\big) * G$. □

7.2.2 Die Erneuerungsgleichung

Erneuerungsargument. Um eine Größe $H(t)$, die vom Prozess abhängt, zu berechnen, unterscheidet man nach dem ersten Sprung $T_1 = W_1$: Ist $W_1 > t$, so kennt man die Situation explizit; ist $W_1 = y \leq t$, so beginnt der Prozess in y von neuem, und die gesuchte Größe ist $H(t-y)$. Das führt auf

$$H(t) = a(t) + \int_{(0,t]} H(t-y)\,\mathrm{d}G(y), \qquad \text{kurz } H = a + H * G. \tag{7.1}$$

Satz 7.7 (Erneuerungsgleichung, [BOS25, Satz 5.28])**.** *Sei a lokal beschränkt und messbar. Dann besitzt Gleichung* (7.1) *genau eine lokal beschränkte messbare Lösung, nämlich*

$$H = a + a * M, \qquad H(t) = a(t) + \int_{(0,t]} a(t-x)\,\mathrm{d}M(x).$$

Beweis. Existenz: Mit $M * G = M - G$ ist $(a + a*M)*G = a*G + a*(M*G) = a*G + a*M - a*G = a*M$, also $a + (a + a*M)*G = a + a*M$. Die Lösung H ist lokal beschränkt, da M endlich ist. *Eindeutigkeit:* Die Differenz D zweier Lösungen erfüllt $D = D*G$, also $D = D*G^{*n}$ für alle n und $|D(t)| \leq \sup_{[0,t]} |D|\, G^{*n}(t)$. Da $\sum_n G^{*n}(t) = M(t) < \infty$, folgt $G^{*n}(t) \to 0$ und $D = 0$. □

Für $a = G$ ist $H = M$.

7.2.3 Restwartezeit und Wartezeitparadoxon

Definition 7.8. Zur Zeit t sei $R_t := T_{N_t+1} - t$ die *Restwartezeit* bis zum nächsten Sprung und $A_t := t - T_{N_t}$ das *Alter* (die aktuelle Wartezeit).

Abbildung 7.1 illustriert die Größen A_t und R_t.

Lemma 7.9 ([BOS25, Lemma 5.31])**.** *Für $z \geq 0$ gilt*

$$\mathbb{P}(R_t > z) = \overline{G}(t+z) + \int_{(0,t]} \overline{G}(t+z-x)\,\mathrm{d}M(x).$$

Ist $m < \infty$, so gilt außerdem $\mathbb{E}[R_t] = m\big(1 + M(t)\big) - t$.

Beweis. Sei $H(t) := \mathbb{P}(R_t > z)$. Ist $W_1 > t$, so ist $R_t = W_1 - t$, also $R_t > z$ genau dann, wenn $W_1 > t+z$. Ist $W_1 = y \leq t$, so verhält sich der Prozess ab y wie ein neuer, unabhängiger Erneuerungsprozess. Formal:

$$H(t) = \mathbb{P}(W_1 > t+z) + \int_{(0,t]} \mathbb{P}(R'_{t-y} > z)\,\mathrm{d}G(y) = \overline{G}(t+z) + (H*G)(t),$$

und Satz 7.7 mit $a(t) = \overline{G}(t+z)$ liefert die Formel. Für den Erwartungswert: $N_t + 1$ ist eine Stoppzeit bezüglich (W_n), denn $\{N_t + 1 = n\} = \{T_{n-1} \leq t < T_n\}$ hängt nur von $W_1, \ldots, W_n$ ab. Die erste Waldsche Gleichung ergibt $\mathbb{E}[T_{N_t+1}] = m\,\mathbb{E}[N_t + 1]$. □

Beispiel 7.10 (Wartezeitparadoxon)**.** Busse kommen gemäß einem Poisson-Prozess mit Rate 1, also im Mittel alle $m = 1$ Minute. Wer zur Zeit t an der Haltestelle ankommt, wartet im Mittel $\mathbb{E}[R_t] = 1\cdot(1+t) - t = 1$ Minute, nicht $1/2$. Die Auflösung: Ein fester Zeitpunkt t fällt mit größerer Wahrscheinlichkeit in ein *langes* Wartezeitintervall. Die Länge des Intervalls, das t überdeckt, ist größenverzerrt (Abschnitt 7.5); für exponentielle Wartezeiten hat es im Limes die erwartete Länge 2.

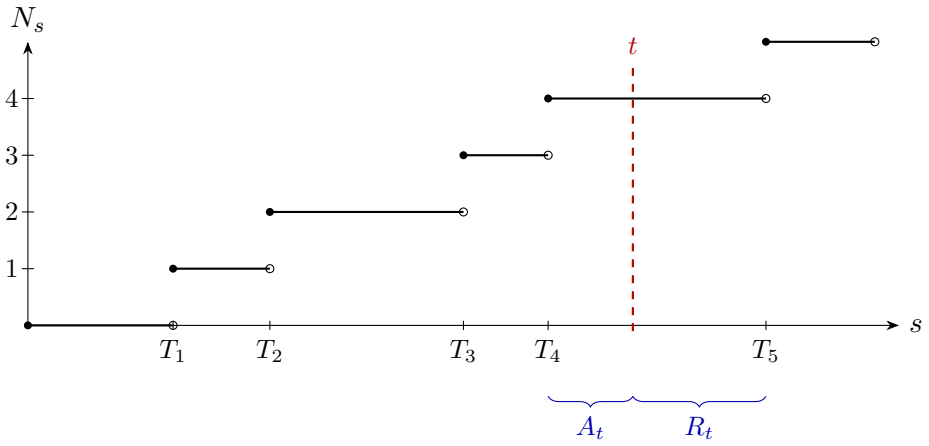

Abbildung 7.1: Schema eines Erneuerungsprozesses: N_s springt in den Erneuerungszeitpunkten $T_n = W_1 + \cdots + W_n$ um 1 und ist rechtsstetig. Zur festen Zeit t (hier $N_t = 4$) misst das Alter $A_t = t - T_{N_t}$ die seit dem letzten Sprung vergangene Zeit, die Restwartezeit $R_t = T_{N_t+1} - t$ die Zeit bis zum nächsten. Ihre Summe ist die Länge des Intervalls, das t enthält.

7.2.4 Der elementare Erneuerungssatz

Satz 7.11 (Elementarer Erneuerungssatz, [BOS25, Satz 5.35])**.** *Es gilt* $\lim_{t\to\infty} M(t)/t = 1/m \in [0, \infty)$.

Beweis. Untere Schranke ($m < \infty$): Nach Lemma 7.9 ist $m(1 + M(t)) = \mathbb{E}[T_{N_t+1}] > t$, also $\liminf_t M(t)/t \geq 1/m$. Für $m = \infty$ ist nichts zu zeigen.

Obere Schranke: Für $c > 0$ sei $W_i^c := W_i \wedge c$ mit zugehörigem Erneuerungsprozess N^c und Erneuerungsfunktion M^c. Da die gestutzten Wartezeiten kleiner sind, ist $N_t^c \geq N_t$, also $M^c \geq M$. Der Überschuss über t ist jetzt durch c beschränkt: $T^c_{N_t^c+1} \leq t + c$. Mit Wald folgt $m_c(1 + M^c(t)) \leq t + c$, $m_c := \mathbb{E}[W_1 \wedge c]$. Also $\limsup_t M(t)/t \leq 1/m_c$, und $m_c \uparrow m$ für $c \to \infty$. □

Die fast sichere Aussage $N_t/t \to 1/m$ (Proposition 7.2) impliziert die Konvergenz der Erwartungswerte nicht ohne Weiteres; der Beweis oben umgeht die fehlende gleichgradige Integrierbarkeit mit Hilfe der Waldschen Gleichung.

Abbildung 7.2 zeigt den elementaren Erneuerungssatz an einem explizit lösbaren Beispiel.

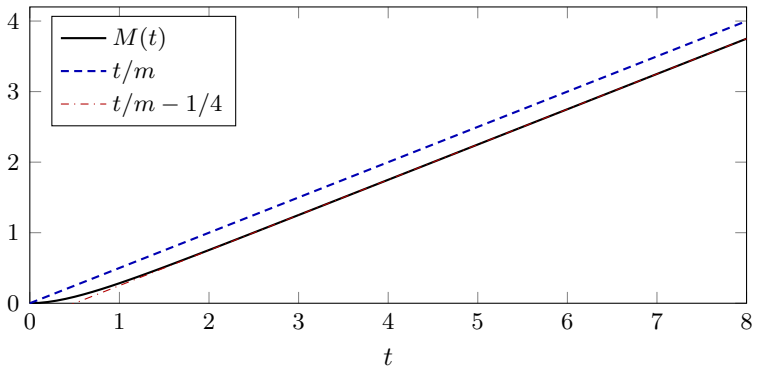

Abbildung 7.2: Erneuerungsfunktion für $\Gamma(2,1)$-verteilte Wartezeiten ($m = 2, \sigma^2 = 2$), explizit $M(t) = t/2 - (1/4)(1 - e^{-2t})$. Der elementare Erneuerungssatz liefert $M(t)/t \to 1/m$; hier gilt sogar $M(t) - t/m \to (\sigma^2 - m^2)/(2m^2) = -1/4$ (vgl. die Aufgabe in Abschnitt 7.3).

7.3 Verzögerte und stationäre Erneuerungsprozesse

Definition 7.12 ([BOS25, Def. 5.36, 5.37])**.** Seien $W_1, W_2, \cdots > 0$ unabhängig, $W_2, W_3, \ldots$ mit Verteilungsfunktion G und W_1 mit Verteilungsfunktion H. Der zugehörige Zählprozess N^H heißt *verzögerter*

Erneuerungsprozess. Ist $m = \mathbb{E}[W_2] < \infty$ und

$$H(t) = G_I(t) \coloneqq \frac{1}{m} \int_0^t \big(1 - G(y)\big)\,\mathrm{d}y,$$

so heißt N^H *stationärer* (Gleichgewichts-)Erneuerungsprozess. Die Verteilung G_I heißt *integrierte Tail-Verteilung* von G.

Proposition 7.13. *Für den stationären Erneuerungsprozess gilt $M^H(t) \coloneqq \mathbb{E}[N_t^H] = t/m$ für alle $t \geq 0$. Seine Zuwächse sind stationär: $N_{t+s}^H - N_s^H \overset{d}{=} N_t^H$.*

Beweis. Wie in Abschnitt 7.2 ist $M^H = \sum_{k\geq 1} H * G^{*(k-1)} = H + M^H * G$; nach Satz 7.7 besitzt diese Gleichung genau eine lokal beschränkte Lösung. Die Funktion t/m ist eine Lösung, denn

$$\frac{1}{m}\int_{(0,t]} (t-y)\,\mathrm{d}G(y) = \frac{1}{m}\int_0^t G(y)\,\mathrm{d}y, \qquad H(t) + \frac{1}{m}\int_0^t G(y)\,\mathrm{d}y = \frac{t}{m}.$$

Für die Stationarität der Zuwächse verweisen wir auf [BOS25, Prop. 5.39]: Die Restwartezeit zur Zeit s hat wieder die Verteilung G_I (vgl. Korollar 7.19). □

Der Poisson-Prozess ist stationär, denn für $G = \mathrm{Exp}(\lambda)$ ist $G_I = G$.

7.4 Blackwell und der fundamentale Erneuerungssatz

Der elementare Erneuerungssatz beschreibt $M(t)$ im Mittel über lange Zeiträume. Feinere Aussagen über $M(t+s) - M(t)$ erfordern, dass G nicht auf einem Gitter konzentriert ist.

Definition 7.14. Die Verteilungsfunktion G heißt *arithmetisch*, wenn es $d > 0$ gibt mit $\mathbb{P}(W_1 \in d\mathbb{N}) = 1$; das größte solche d heißt Spanne.

Satz 7.15 (Blackwell 1948, [BOS25, Satz 5.42])**.** *Ist G nicht arithmetisch, so gilt $M(t+s) - M(t) \to s/m$ für jedes $s > 0$. Ist G arithmetisch mit Spanne d, so gilt dies für $s, t \in d\mathbb{N}$.*

Definition 7.16 ([BOS25, Def. 5.43])**.** Eine Funktion $h\colon [0,\infty) \to [0,\infty)$ heißt *direkt Riemann-integrierbar* (d. R. i.), wenn die Ober- und Untersummen $\delta \sum_{k\geq 0} \sup_{[k\delta,(k+1)\delta]} h$ und $\delta \sum_{k\geq 0} \inf_{[k\delta,(k+1)\delta]} h$ für alle $\delta > 0$ endlich sind und für $\delta \downarrow 0$ gegen denselben Grenzwert konvergieren.

Kompakt getragene Riemann-integrierbare Funktionen sind d. R. i., ebenso fallende Funktionen mit endlichem uneigentlichem Integral. Uneigentliche Riemann-Integrierbarkeit allein genügt nicht (schmale, hohe Zacken). Für Anwendungen genügt meist:

Lemma 7.17 ([BOS25, Lemma 5.48])**.** *Ist $a \geq 0$ auf jedem $[0,x]$ Riemann-integrierbar und gilt $a \leq A$ für eine fallende, uneigentlich integrierbare Funktion A, so ist a d. R. i.*

Satz 7.18 (Fundamentaler Erneuerungssatz, [BOS25, Satz 5.45])**.** *Sei G nicht arithmetisch und $h \geq 0$ d. R. i. Dann gilt*

$$\lim_{t\to\infty} \int_{(0,t]} h(t-s)\,\mathrm{d}M(s) = \frac{1}{m}\int_0^\infty h(s)\,\mathrm{d}s.$$

Insbesondere konvergiert die Lösung $H = h + h * M$ der Erneuerungsgleichung gegen $\frac{1}{m}\int_0^\infty h$.

Der Satz ist äquivalent zum Satz von Blackwell. Für $h(u) = \mathbb{1}\{0 \leq u < s\}$ erhält man diesen sofort; die Umkehrung folgt durch Approximation mit Treppenfunktionen, was genau die direkte Riemann-Integrierbarkeit erfordert. Beweise findet man in [Fel71, XI] oder [RSST99, Kap. 6].

7.5 Grenzverteilungen

Korollar 7.19 ([BOS25, Kor. 5.50, 5.52])**.** *Sei G nicht arithmetisch und $m < \infty$. Dann gilt für $z \geq 0$*

$$\lim_{t\to\infty} \mathbb{P}(R_t \leq z) = \lim_{t\to\infty} \mathbb{P}(A_t \leq z) = G_I(z).$$

Beweis. Nach Lemma 7.9 ist $\mathbb{P}(R_t > z) = h(t) + (h * M)(t)$ mit $h(t) = \overline{G}(t+z)$. Diese Funktion ist fallend mit $\int_0^\infty h = \int_z^\infty \overline{G}(y)\,\mathrm{d}y \leq m < \infty$, also d. R. i. Mit Satz 7.18 und $h(t) \to 0$ folgt $\mathbb{P}(R_t > z) \to \frac{1}{m}\int_z^\infty \overline{G}(y)\,\mathrm{d}y = 1 - G_I(z)$. Für A_t argumentiert man analog mit $h(t) = \overline{G}(t)\mathbb{1}\{t > z\}$. □

Abbildung 7.3 zeigt die Konvergenz der Restwartezeit und die Größenverzerrung.

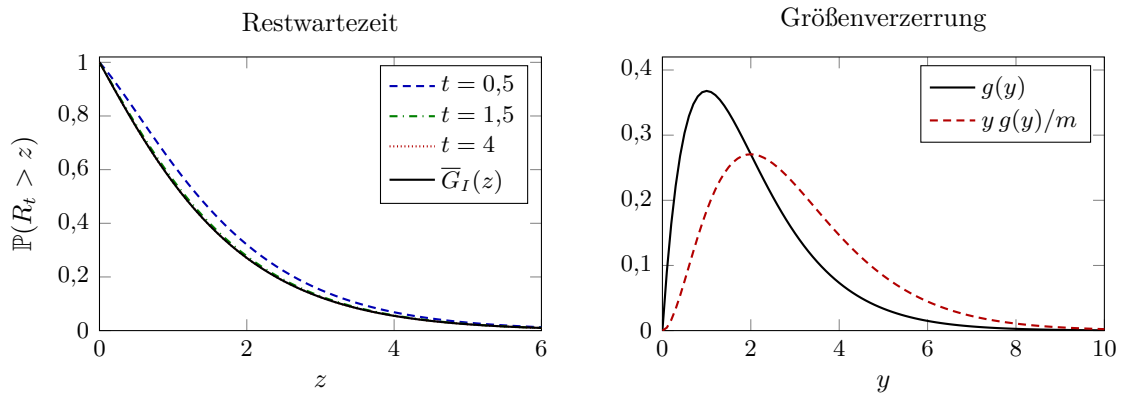

Abbildung 7.3: Wartezeiten $W_i \sim \Gamma(2,1)$ mit Dichte $g(y) = y\mathrm{e}^{-y}$, $m = 2$. Links: $\mathbb{P}(R_t > z)$ nach Lemma 7.9, numerisch ausgewertet für $t = 0{,}5;\ 1{,}5;\ 4$, konvergiert gegen $\overline{G}_I(z) = (1/2)(2+z)\mathrm{e}^{-z}$. Rechts: Das Intervall, das einen festen späten Zeitpunkt überdeckt, hat im Limes die größenverzerrte Dichte $yg(y)/m$ mit Erwartungswert $\mathbb{E}[W^2]/m = 3$ statt $m = 2$ – die Erklärung des Wartezeitparadoxons.

Größenverzerrung. Im Limes hat die Restwartezeit den Erwartungswert

$$\int_0^\infty \big(1 - G_I(z)\big)\,\mathrm{d}z = \frac{1}{m}\int_0^\infty y\,\overline{G}(y)\,\mathrm{d}y = \frac{\mathbb{E}[W_1^2]}{2m} = \frac{m^2+\sigma^2}{2m}.$$

Die Länge $A_t + R_t$ des Intervalls, das t überdeckt, hat im Limes die Verteilung $\frac{1}{m}\int_{(0,z]} y\,\mathrm{d}G(y)$: Ein Intervall wird proportional zu seiner Länge „getroffen“. Ist $\sigma^2 = \infty$ bei $m < \infty$ (etwa Pareto mit $\alpha \in (1,2]$), so ist die erwartete Restwartezeit im Limes unendlich – ein erstes Beispiel dafür, wie schwere Tails Mittelwerte dominieren.

Aufgaben

Aufgabe 7.1. Zeigen Sie, dass für einen Poisson-Prozess $(N_t - \lambda t)$ und $\big((N_t - \lambda t)^2 - \lambda t\big)$ Martingale bezüglich der kanonischen Filtration sind.

Aufgabe 7.2 (vgl. [BOS25, Übung 5.34])**.** Leiten Sie die Formel für $\mathbb{E}[R_t]$ ohne die Waldsche Gleichung mit einem Erneuerungsargument her.

Aufgabe 7.3. Bestimmen Sie die Erneuerungsfunktion für $G = \Gamma(2,1)$ mit Laplace-Transformierten und berechnen Sie $M(t) - t/m$ für $t \to \infty$.

Aufgabe 7.4. Stellen Sie für $\mathbb{P}(A_t > z)$ eine Erneuerungsgleichung auf und lösen Sie sie.

Aufgabe 7.5 (vgl. [BOS25, Übung 5.53])**.** Zeigen Sie, dass die Intervalllänge $A_t + R_t$ im nicht arithmetischen Fall mit $m < \infty$ gegen die Verteilung $\frac{1}{m}\int_{(0,z]} y\,\mathrm{d}G(y)$ konvergiert.

Aufgabe 7.6. Berechnen Sie G_I für die Gamma-, die Pareto- und die Lomax-Verteilung. Welche Tails haben die integrierten Verteilungen?

Aufgabe 7.7. Zeigen Sie mit dem fundamentalen Erneuerungssatz: Ist $\sigma^2 < \infty$ und G nicht arithmetisch, so gilt $M(t) - t/m \to (\sigma^2 - m^2)/(2m^2)$.

Kapitel 8

Ruintheorie

In diesem Kapitel verfolgen wir das Kapital eines Versicherers in der Zeit. Zentral ist die Wahrscheinlichkeit, dass es jemals negativ wird. Ob sie exponentiell oder nur polynomiell im Startkapital fällt, hängt vom Tail der Schadenhöhe ab.

8.1 Dynamische Modelle der kollektiven Risikotheorie

Nun verfolgen wir den Gesamtschaden in der Zeit und vergleichen ihn mit den laufend eingehenden Prämien.

Definition 8.1 ([BOS25, Def. 5.1, 5.3])**.** Ein *Schadeneintrittsprozess* ist ein $\mathbb{N}_0$-wertiger Prozess (N_t) mit $N_0 = 0$, stückweise konstanten, rechtsstetigen Pfaden und Sprüngen der Höhe 1. Sei (X_i) eine davon unabhängige u. i. v. Folge positiver Schadenhöhen mit Verteilungsfunktion F, und sei Π eine Prämienfunktion mit $\Pi(0) = 0$. Dann heißen

$$S_t \coloneqq \sum_{i=1}^{N_t} X_i \quad \text{(Gesamtschadenprozess)}, \qquad U_t \coloneqq u + \Pi(t) - S_t \quad \text{(Risikoprozess)}$$

mit Startkapital $u \geq 0$.

Als Schadeneintrittsprozess verwenden wir im Folgenden einen Erneuerungsprozess (Definition 7.1).

Definition 8.2 (Sparre-Andersen- und Cramér–Lundberg-Modell)**.** Ist (N_t) ein Erneuerungsprozess mit Intensität $\lambda > 0$, gilt $\mu \coloneqq \mathbb{E}[X_1] < \infty$ und $\Pi(t) = ct$ mit Prämienrate $c > 0$, so heißt das Modell *Sparre-Andersen-Modell*. Ist (N_t) ein Poisson-Prozess, so heißt es *Cramér–Lundberg-Modell*.

Im Cramér–Lundberg-Modell ist (S_t) ein *zusammengesetzter Poisson-Prozess* und (U_t) ein Lévy-Prozess (unabhängige und stationäre Zuwächse, càdlàg-Pfade).

Definition 8.3 (Ruinfunktion, [BOS25, Def. 5.14])**.** Die *Ruinwahrscheinlichkeit* bei Startkapital $u \geq 0$ ist

$$\Psi(u) \coloneqq \mathbb{P}\big(U_t < 0 \text{ für ein } t \geq 0 \;\big|\; U_0 = u\big);$$

für $u < 0$ sei $\Psi(u) \coloneqq 1$. Ferner sei $\overline{\Psi}(u) \coloneqq 1 - \Psi(u)$ die Überlebenswahrscheinlichkeit.

Satz 8.4 ([BOS25, Satz 5.15])**.** *Im Sparre-Andersen-Modell gilt $S_t/t \to \lambda\mu$ und $U_t/t \to c - \lambda\mu$ fast sicher.*

Beweis. Es gilt $S_t/t = \frac{1}{N_t} \sum_{i \leq N_t} X_i \cdot N_t/t \to \mu\lambda$ nach dem starken Gesetz und Proposition 7.2. □

Definition 8.5 (Sicherheitszuschlag, NPC)**.** Die Größe $\rho \coloneqq c/(\lambda\mu) - 1$ heißt *relativer Sicherheitszuschlag*. Gilt $\rho > 0$, d. h. $c > \lambda\mu$, so sagt man, die *Nettogewinnbedingung* (net profit condition, NPC) sei erfüllt.

Für $\rho < 0$ strebt U_t linear gegen $-\infty$, also $\Psi \equiv 1$. Für $\rho > 0$ ist $\inf_t U_t > -\infty$ f. s., und da eine Erhöhung von u den Pfad nur verschiebt, folgt $\Psi(u) \to 0$ für $u \to \infty$. Auch im kritischen Fall $\rho = 0$ tritt Ruin fast sicher ein (Korollar 8.11 für das Cramér–Lundberg-Modell). Der Fall $\rho = 0$ ist das Äquivalenzprinzip der Schadenversicherung; er reicht also *nicht* aus.

Abbildung 8.1 zeigt einen typischen Pfad bis zum Ruin.

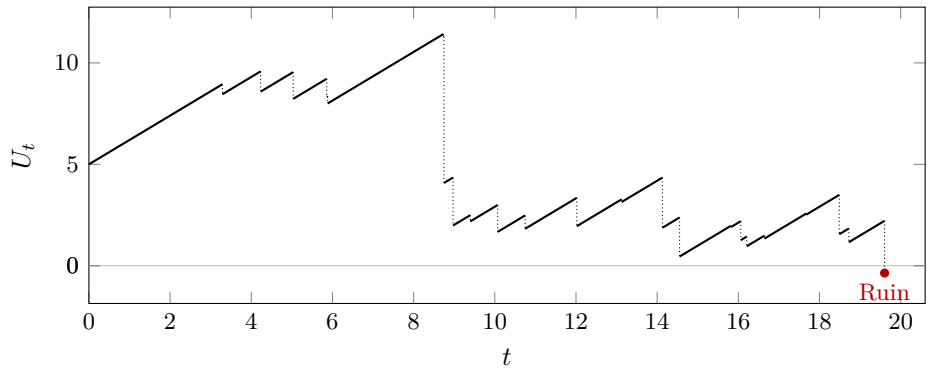

Abbildung 8.1: Ein simulierter Pfad des Risikoprozesses im Cramér–Lundberg-Modell mit $u = 5$, $c = 1{,}2$, $\lambda = 1$ und Exp(1)-Schäden ($\rho = 0{,}2$), bis zum Ruin. Zwischen den Schäden wächst U linear mit Steigung c; jeder Schaden ist ein Sprung nach unten (gepunktet). Ruin kann daher nur in einem Schadenzeitpunkt eintreten, was die Irrfahrtsdarstellung Gleichung (8.1) begründet.

Irrfahrtsdarstellung. Zwischen zwei Schäden wächst U linear. Ruin kann daher nur in einem Schadenzeitpunkt T_n eintreten, und dort ist $U_{T_n} = u - \sum_{i=1}^{n}(X_i - cW_i)$. Mit den *Pseudoschadenhöhen* $\bar{X}_i := X_i - cW_i$ und der Irrfahrt $\bar{S}_n := \bar{X}_1 + \cdots + \bar{X}_n$ gilt

$$\Psi(u) = \mathbb{P}\Big(\sup_{n\ge 1} \bar{S}_n > u\Big). \tag{8.1}$$

Die NPC ist äquivalent zu $\mathbb{E}[\bar{X}_1] = \mu - c/\lambda < 0$: Die Irrfahrt hat negative Drift.

8.2 Ruin unter der Lundberg-Bedingung

8.2.1 Die Lundberg-Ungleichung

Wir arbeiten im Sparre-Andersen-Modell mit NPC und den Pseudoschadenhöhen $\bar{X}_i = X_i - cW_i$. Nach Gleichung (8.1) ist $\Psi(u)$ die Wahrscheinlichkeit, dass die Irrfahrt $\bar{S}_n$ mit negativer Drift jemals über u steigt. Eine exponentielle Schranke erwarten wir, wenn $\bar{X}_1$ exponentielle Momente hat.

Definition 8.6 (Lundberg-Bedingung, [BOS25, Def. 5.55])**.** Sei

$$f(r) := \mathbb{E}[\mathrm{e}^{r\bar{X}_1}] = \psi_{X_1}(r)\,\mathbb{E}[\mathrm{e}^{-rcW_1}], \qquad r \ge 0.$$

Gibt es $r_0 > 0$ mit $f(r_0) = 1$, so heißt r_0 *Lundberg-Koeffizient* (Anpassungskoeffizient).

Die Funktion f ist konvex mit $f(0) = 1$ und $f'(0+) = \mathbb{E}[\bar{X}_1] < 0$ (NPC). Ist X_1 heavy-tailed, so ist $f(r) = \infty$ für alle $r > 0$, und es gibt keinen Lundberg-Koeffizienten. Ist X_1 light-tailed, so existiert r_0, sofern f vor dem Rand seines Definitionsbereichs über 1 wächst. Im Cramér–Lundberg-Modell ist $\mathbb{E}[\mathrm{e}^{-rcW_1}] = \lambda/(\lambda + cr)$, und die Bedingung $f(r) = 1$ wird zur *Lundberg-Gleichung*

$$\lambda\big(\psi_{X_1}(r) - 1\big) = cr \quad\Longleftrightarrow\quad \frac{\lambda}{c}\int_0^\infty \mathrm{e}^{rx}\overline{F}(x)\,\mathrm{d}x = 1. \tag{8.2}$$

Abbildung 8.2 zeigt die Lundberg-Gleichung für exponentielle Schäden.

Satz 8.7 (Lundberg-Ungleichung, [BOS25, Satz 5.58])**.** *Im Sparre-Andersen-Modell gilt* $\Psi(u) \le \mathrm{e}^{-ru}$ *für alle* $u \ge 0$ *und jedes* $r > 0$ *mit* $f(r) \le 1$*; insbesondere* $\Psi(u) \le \mathrm{e}^{-r_0 u}$*.*

Beweis. Sei $\Psi_n(u) := \mathbb{P}(\max_{k\le n} \bar{S}_k > u)$; dann gilt $\Psi_n \uparrow \Psi$. Wir zeigen $\Psi_n(u) \le \mathrm{e}^{-ru}$ für alle $u \in \mathbb{R}$ durch Induktion. Für $n = 1$ ist nach Markov $\Psi_1(u) = \mathbb{P}(\bar{X}_1 > u) \le \mathrm{e}^{-ru} f(r) \le \mathrm{e}^{-ru}$. Bedingen auf $\bar{X}_1 = x$ liefert mit der Verteilung $P_{\bar{X}}$ von $\bar{X}_1$

$$\begin{aligned}\Psi_{n+1}(u) &= \mathbb{P}(\bar{X}_1 > u) + \int_{(-\infty,u]} \Psi_n(u-x)\,\mathrm{d}P_{\bar{X}}(x)\\ &\le \int_{(u,\infty)} \mathrm{e}^{r(x-u)}\,\mathrm{d}P_{\bar{X}}(x) + \int_{(-\infty,u]} \mathrm{e}^{-r(u-x)}\,\mathrm{d}P_{\bar{X}}(x),\end{aligned}$$

und die rechte Seite ist $\mathrm{e}^{-ru} f(r) \le \mathrm{e}^{-ru}$. □

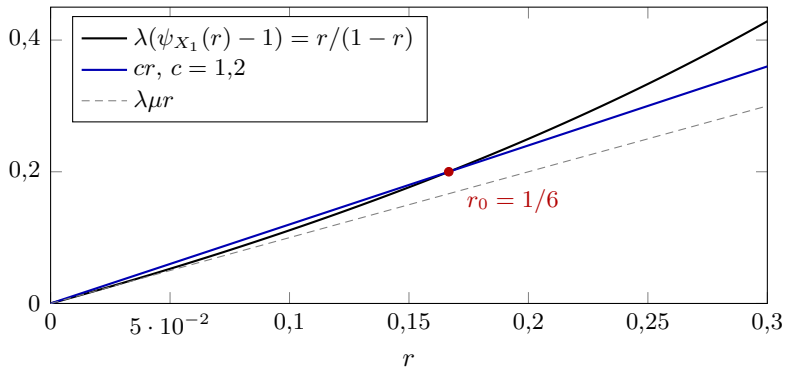

Abbildung 8.2: Die Lundberg-Gleichung Gleichung (8.2) im Cramér–Lundberg-Modell mit $\lambda = 1$, $c = 1{,}2$ und Exp(1)-Schäden. Die linke Seite ist konvex mit Steigung $\lambda\mu = 1$ in 0; wegen der NPC ($c > \lambda\mu$) liegt die Gerade cr zunächst darüber, und es gibt genau einen positiven Schnittpunkt $r_0 = (1/\mu)\rho/(1+\rho) = 1/6$. Bei heavy-tailed Schäden ist die linke Seite für $r > 0$ unendlich, und kein Schnittpunkt existiert.

Bemerkung 8.8 (Martingalbeweis). Ist $f(r_0) = 1$, so ist $\mathrm{e}^{r_0 \bar{S}_n}$ ein nichtnegatives Martingal. Für die Ruinzeit $\nu = \inf\{n\colon \bar{S}_n > u\}$ liefert optionales Stoppen $1 = \mathbb{E}[\mathrm{e}^{r_0 \bar{S}_{\nu\wedge n}}] \geq \mathrm{e}^{r_0 u}\mathbb{P}(\nu \leq n)$, also wieder $\Psi(u) \leq \mathrm{e}^{-r_0 u}$.

Beispiel 8.9 (Exponentielle Schäden, [BOS25, Bsp. 5.60])**.** Im Cramér–Lundberg-Modell mit $X_1 \sim \mathrm{Exp}(1/\mu)$ ist $\psi_{X_1}(r) = 1/(1-\mu r)$, und Gleichung (8.2) hat die Lösung

$$r_0 = \frac{1}{\mu} - \frac{\lambda}{c} = \frac{1}{\mu}\cdot\frac{\rho}{1+\rho}.$$

8.2.2 Eine Integralgleichung für die Ruinfunktion

Ab jetzt betrachten wir das Cramér–Lundberg-Modell. Sei F_I die integrierte Tail-Verteilung der Schadenhöhe, $F_I(t) = \frac{1}{\mu}\int_0^t \overline{F}(x)\,\mathrm{d}x$.

Satz 8.10 ([BOS25, Satz 5.61])**.** *Die Überlebenswahrscheinlichkeit $\overline{\Psi} = 1 - \Psi$ erfüllt*

$$\overline{\Psi}(u) = \overline{\Psi}(0) + \frac{\lambda\mu}{c}\int_{(0,u]} \overline{\Psi}(u-x)\,\mathrm{d}F_I(x), \qquad u \geq 0.$$

Beweis. Bedinge auf den ersten Schaden: Zur Zeit $W_1 = t$ ist das Kapital $u+ct$, und der Schaden $X_1 = x$ ruiniert genau dann nicht, wenn $x \leq u+ct$; danach beginnt der Prozess neu mit Kapital $u+ct-x$. Also

$$\overline{\Psi}(u) = \int_0^\infty \lambda\mathrm{e}^{-\lambda t}\int_{[0,u+ct]} \overline{\Psi}(u+ct-x)\,\mathrm{d}F(x)\,\mathrm{d}t = \frac{\lambda}{c}\mathrm{e}^{\lambda u/c}\int_u^\infty \mathrm{e}^{-\lambda s/c}\int_{[0,s]} \overline{\Psi}(s-x)\,\mathrm{d}F(x)\,\mathrm{d}s.$$

Daher ist $\overline{\Psi}$ differenzierbar mit $\overline{\Psi}'(u) = (\lambda/c)\overline{\Psi}(u) - \frac{\lambda}{c}\int_{[0,u]} \overline{\Psi}(u-x)\,\mathrm{d}F(x)$. Integration über $[0,t]$ und Fubini, $\int_0^t\int_{[0,u]} \overline{\Psi}(u-x)\,\mathrm{d}F(x)\,\mathrm{d}u = \int_0^t \overline{\Psi}(v)F(t-v)\,\mathrm{d}v$, ergeben

$$\overline{\Psi}(t) - \overline{\Psi}(0) = \frac{\lambda}{c}\int_0^t \overline{\Psi}(t-y)\overline{F}(y)\,\mathrm{d}y = \frac{\lambda\mu}{c}\int_{(0,t]} \overline{\Psi}(t-y)\,\mathrm{d}F_I(y). \qquad \square$$

Korollar 8.11 ([BOS25, Kor. 5.62])**.** *Ist $\rho > 0$, so gilt $\Psi(0) = \lambda\mu/c = 1/(1+\rho)$. Ist $\rho = 0$, so ist $\Psi \equiv 1$.*

Beweis. Für $\rho > 0$ gilt $\overline{\Psi}(u) \to 1$, und monotone Konvergenz in Satz 8.10 liefert $1 = \overline{\Psi}(0) + \lambda\mu/c$. Für $\rho = 0$ ist $\overline{\Psi} = \overline{\Psi}(0) + \overline{\Psi} * F_I$ eine Erneuerungsgleichung mit der echten Verteilung F_I; ihre Lösung ist $\overline{\Psi}(0)(1 + M_I(u))$ mit der Erneuerungsfunktion $M_I(u) \to \infty$. Da $\overline{\Psi} \leq 1$, folgt $\overline{\Psi}(0) = 0$ und damit $\overline{\Psi} \equiv 0$. $\square$

Bemerkenswert: $\Psi(0)$ hängt nur vom Sicherheitszuschlag ab, nicht von der Schadenhöhenverteilung. Mit $q \coloneqq 1/(1+\rho)$ lautet Satz 8.10

$$\Psi(u) = q\,\overline{F}_I(u) + q\int_{(0,u]} \Psi(u-x)\,\mathrm{d}F_I(x), \tag{8.3}$$

eine *defekte* Erneuerungsgleichung (das Maß qF_I hat Masse $q<1$). Wie in Satz 7.7 hat sie genau eine lokal beschränkte Lösung, denn $q^n F_I^{*n} \to 0$.

Beispiel 8.12 (Exponentielle Schäden, [BOS25, Bsp. 5.63])**.** Für $X_1 \sim \mathrm{Exp}(1/\mu)$ ist $F_I = F$, und man rechnet nach, dass

$$\Psi(u) = \frac{1}{1+\rho}\exp\Big(-\frac{\rho}{1+\rho}\frac{u}{\mu}\Big)$$

die Gleichung Gleichung (8.3) löst. Der Exponent ist genau r_0: Die Lundberg-Schranke ist bis auf den Faktor q scharf.

8.2.3 Die Cramér–Lundberg-Approximation

Satz 8.13 (Lundberg-Approximation, [BOS25, Satz 5.65])**.** *Im Cramér–Lundberg-Modell mit $\rho>0$ existiere r_0, und es sei $\vartheta \coloneqq \int_0^\infty v\mathrm{e}^{r_0 v}\overline{F}(v)\,\mathrm{d}v < \infty$. Dann gilt*

$$\Psi(u) \sim \frac{\rho\mu}{r_0\vartheta}\,\mathrm{e}^{-r_0 u}, \qquad u\to\infty.$$

Beweis. Multipliziere Gleichung (8.3) mit $\mathrm{e}^{r_0 u}$. Für $Z(u) \coloneqq \mathrm{e}^{r_0 u}\Psi(u)$ folgt

$$Z(u) = a(u) + \int_{(0,u]} Z(u-x)\,\mathrm{d}G(x), \qquad a(u) \coloneqq q\,\mathrm{e}^{r_0 u}\overline{F}_I(u), \quad G(\mathrm{d}x) \coloneqq q\,\mathrm{e}^{r_0 x}F_I(\mathrm{d}x).$$

Nach Gleichung (8.2) ist $\int G(\mathrm{d}x) = \frac{q}{\mu}\int \mathrm{e}^{r_0 x}\overline{F}(x)\,\mathrm{d}x = \frac{\lambda}{c}\int \mathrm{e}^{r_0 x}\overline{F}(x)\,\mathrm{d}x = 1$: Der Maßwechsel macht die defekte Gleichung zu einer echten Erneuerungsgleichung. Das Maß G hat eine Dichte, ist also nicht arithmetisch, und den Erwartungswert $m_G = \frac{q}{\mu}\int x\mathrm{e}^{r_0 x}\overline{F}(x)\,\mathrm{d}x = q\vartheta/\mu$. Ferner ist $a(u) \le \frac{q}{\mu}\int_u^\infty \mathrm{e}^{r_0 x}\overline{F}(x)\,\mathrm{d}x$, und diese Schranke ist fallend mit Integral $(q/\mu)\vartheta<\infty$; nach Lemma 7.17 ist a d. R. i. Partielle Integration ergibt

$$\int_0^\infty \mathrm{e}^{r_0 u}\overline{F}_I(u)\,\mathrm{d}u = \frac{1}{r_0}\Big(\int \mathrm{e}^{r_0 x}F_I(\mathrm{d}x) - 1\Big) = \frac{1}{r_0}\Big(\frac{1}{q}-1\Big) = \frac{\rho}{r_0}.$$

Nach Satz 7.18 konvergiert $Z(u)$ gegen $\frac{1}{m_G}\int_0^\infty a = \mu/(q\vartheta)\cdot q\rho/r_0 = \rho\mu/(r_0\vartheta)$. □

Für exponentielle Schäden ergibt sich wieder die exakte Formel. Die Lundberg-Ungleichung zeigt zusätzlich $\rho\mu/(r_0\vartheta)\le 1$.

Abbildung 8.3 vergleicht Schranke, Approximation und numerisch berechnete Ruinwahrscheinlichkeit.

Wie tritt Ruin ein? Unter der Lundberg-Bedingung entsteht Ruin bei großem Startkapital durch eine ungünstige *Häufung vieler mäßiger Schäden*: Gegeben Ruin verhält sich der Prozess wie unter dem gekippten Maß aus Satz 6.6, das die Drift positiv macht. In Abschnitt 8.3 sehen wir das Gegenteil.

8.3 Subexponentielle Verteilungen und Ruin durch Großschäden

Bei schweren Tails gibt es keinen Lundberg-Koeffizienten, und die Ruinfunktion fällt nicht mehr exponentiell. Wir brauchen eine Klasse von Verteilungen, für die sich Summen und Maxima gleich verhalten.

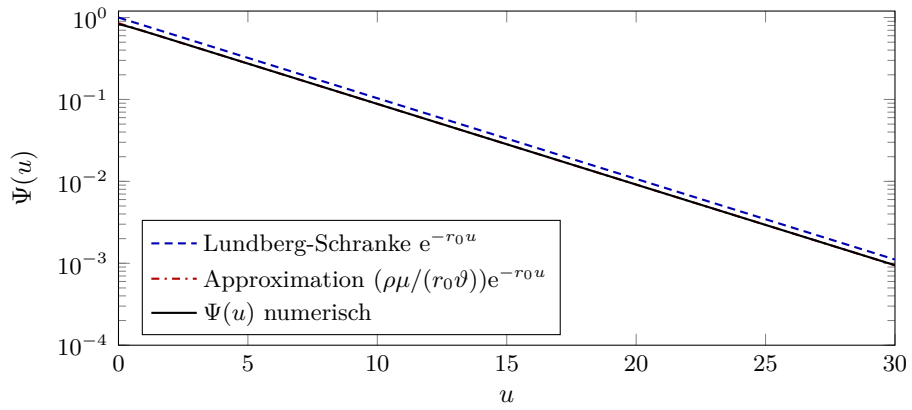

Abbildung 8.3: Ruinwahrscheinlichkeit für $\Gamma(2,2)$-Schäden ($\mu = 1$), $\lambda = 1$, $c = 1{,}2$, logarithmische Skala. Die Funktion Ψ wurde durch Diskretisierung der defekten Erneuerungsgleichung Gleichung (8.3) berechnet (Schrittweite 0,01; Halbierung der Schrittweite ändert die Werte um weniger als 0,05 %, und für exponentielle Schäden reproduziert das Verfahren die exakte Formel auf 0,02 %). Hier ist $r_0 \approx 0{,}2268$ und $\rho\mu/(r_0\vartheta) \approx 0{,}8518$. Die Cramér–Lundberg-Approximation ist schon ab $u \approx 2$ kaum vom exakten Wert zu unterscheiden; die Lundberg-Schranke hat die richtige Rate, aber einen zu großen Vorfaktor.

8.3.1 Subexponentielle Verteilungen

Seien $X_1, X_2, \dots$ u. i. v. mit Verteilungsfunktion F, $F(0) = 0$, $\overline{F}(x) > 0$ für alle x; $S_n = X_1 + \cdots + X_n$, $M_n = \max_{i \le n} X_i$.

Definition 8.14 ([BOS25, Def. 5.69])**.** Die Verteilungsfunktion F heißt *subexponentiell*, $F \in \mathcal{S}$, wenn für alle $n \ge 2$

$$\lim_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} = n.$$

Proposition 8.15 ([BOS25, Prop. 5.70])**.** *Es gilt $F \in \mathcal{S}$ genau dann, wenn $\mathbb{P}(S_n > x) \sim \mathbb{P}(M_n > x)$ für alle $n \ge 2$.*

Beweis. Es gilt $\mathbb{P}(M_n > x) = 1 - F(x)^n = \overline{F}(x) \sum_{k=0}^{n-1} F(x)^k \sim n\overline{F}(x)$. □

Eine Summe ist also nur deshalb groß, weil *ein* Summand groß ist. Es genügt, die Bedingung für $n = 2$ zu fordern [EKM97, Lemma 1.3.4].

Proposition 8.16 ([BOS25, Prop. 5.72])**.** *(i) Ist $F \in \mathcal{S}$, so gilt $\overline{F}(x-y)/\overline{F}(x) \to 1$ für $x \to \infty$, gleichmäßig für y in Kompakta (*Langschwänzigkeit*).*

(ii) Ist F langschwänzig, so gilt $\mathrm{e}^{\varepsilon x}\overline{F}(x) \to \infty$ für jedes $\varepsilon > 0$. Insbesondere ist $\mathcal{S} \subset \mathcal{K}$.

Beweis. (i) Es ist $\overline{F^{*2}}(x) = \overline{F}(x) + \int_{[0,x]} \overline{F}(x-t)\,\mathrm{d}F(t)$. Für $0 < y \le x$ schätzen wir den Integranden für $t \le y$ durch $\overline{F}(x)$ und für $t \in (y, x]$ durch $\overline{F}(x-y)$ ab:

$$\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} - 1 \ge F(y) + \frac{\overline{F}(x-y)}{\overline{F}(x)}\big(F(x) - F(y)\big).$$

Mit $x \to \infty$ folgt $\limsup_x \overline{F}(x-y)/\overline{F}(x) \le (2 - 1 - F(y))/(1 - F(y)) = 1$; die Ungleichung ≥ 1 folgt aus der Monotonie von $\overline{F}$. Gleichmäßigkeit folgt aus der Monotonie in y. (ii) Sei $\varepsilon > 0$. Für $x \ge x_0$ ist $\overline{F}(x) \ge \mathrm{e}^{-\varepsilon}\overline{F}(x-1)$, also $\overline{F}(x_0 + n) \ge \mathrm{e}^{-\varepsilon n}\overline{F}(x_0)$ und damit $\mathrm{e}^{2\varepsilon x}\overline{F}(x) \to \infty$. Daher ist $\mathbb{E}[\mathrm{e}^{sX_1}] = \infty$ für alle $s > 0$. □

8.3.2 Reguläre Variation

Definition 8.17 ([BOS25, Def. 5.75, 5.77])**.** Eine messbare Funktion $\ell > 0$ heißt *langsam variierend*, $\ell \in \mathcal{R}_0$, wenn $\ell(tx)/\ell(x) \to 1$ für alle $t > 0$. Eine Funktion $h > 0$ heißt *regulär variierend mit Index α*, $h \in \mathcal{R}_\alpha$, wenn $h(tx)/h(x) \to t^\alpha$; äquivalent $h(x) = x^\alpha \ell(x)$ mit $\ell \in \mathcal{R}_0$.

Beispiele für $\overline{F} \in \mathcal{R}_{-\alpha}$: Pareto, Lomax ($\overline{F}(x) = (1 + x/\kappa)^{-\alpha}$), Log-Gamma (mit $\overline{F}(x) \sim c(\log x)^{\alpha-1}x^{-\beta}$, also Index $-\beta$), Burr, Student-t. Logarithmen und iterierte Logarithmen sind langsam variierend.

Lemma 8.18 ([BOS25, Lemma 5.78]). *Seien $Y_1, Y_2 \geq 0$ unabhängig mit $\overline{F}_1, \overline{F}_2 \in \mathcal{R}_{-\alpha}$, $\alpha \geq 0$. Dann gilt $\mathbb{P}(Y_1 + Y_2 > x) \sim \overline{F}_1(x) + \overline{F}_2(x)$.*

Beweis. Untere Schranke: $\mathbb{P}(Y_1 + Y_2 > x) \geq \mathbb{P}(Y_1 > x \text{ oder } Y_2 > x) = \overline{F}_1(x) + \overline{F}_2(x) - \overline{F}_1(x)\overline{F}_2(x)$. Obere Schranke: Für $\delta \in (0, 1/2)$ ist

$$\{Y_1 + Y_2 > x\} \subset \{Y_1 > (1-\delta)x\} \cup \{Y_2 > (1-\delta)x\} \cup \{Y_1 > \delta x,\, Y_2 > \delta x\}.$$

Teilt man durch $\overline{F}_1(x) + \overline{F}_2(x)$, so konvergieren die ersten beiden Terme zusammen gegen $(1-\delta)^{-\alpha}$, der letzte gegen 0, da $\overline{F}_1(\delta x)\overline{F}_2(\delta x)/\overline{F}_1(x) \to \delta^{-\alpha} \cdot 0$. Dann $\delta \downarrow 0$. □

Proposition 8.19 ([BOS25, Prop. 5.79]). *Ist $\overline{F} \in \mathcal{R}_{-\alpha}$ mit $\alpha \geq 0$, so ist $F \in \mathcal{S}$.*

Beweis. Induktion mit Lemma 8.18: Ist $\overline{F^{*(n-1)}} \sim (n-1)\overline{F}$, so ist $\overline{F^{*(n-1)}} \in \mathcal{R}_{-\alpha}$ und $\overline{F^{*n}} \sim (n-1)\overline{F} + \overline{F} = n\overline{F}$. □

Auch die Log-Normal- und die Weibull-Verteilung mit $\alpha < 1$ sind subexponentiell, aber nicht regulär variierend [EKM97, Kap. 1.4].

Proposition 8.20 (Kesten, [BOS25, Prop. 5.80]). *Ist $F \in \mathcal{S}$ und $\varepsilon > 0$, so gibt es $K < \infty$ mit $\overline{F^{*n}}(x) \leq K(1+\varepsilon)^n \overline{F}(x)$ für alle $x \geq 0$, $n \geq 2$.*

Der Beweis ist eine geschickte Induktion über n mit der Zerlegung aus dem Beweis von Proposition 8.16 [BOS25, S. 169–171].

8.3.3 Asymptotik der Ruinfunktion

Satz 8.21 (Pollaczek–Chintschin, [BOS25, Satz 5.81]). *Im Cramér–Lundberg-Modell mit $\rho > 0$ und $q = 1/(1+\rho)$ gilt*

$$\Psi(u) = (1-q)\sum_{n=1}^{\infty} q^n\, \overline{F_I^{*n}}(u), \qquad u \geq 0.$$

Beweis. Nach Satz 8.10 und Korollar 8.11 ist $\overline{\Psi} = (1-q) + q\,\overline{\Psi} * F_I$. Die Reihe $\overline{\Psi} = (1-q)\sum_{n\geq 0} q^n F_I^{*n}$ konvergiert und löst diese Gleichung; die Lösung ist eindeutig, weil $q^n F_I^{*n} \to 0$. Wegen $(1-q)\sum_{n\geq 0} q^n = 1$ folgt die Formel für $\Psi = 1 - \overline{\Psi}$. □

Probabilistisch: $\Psi(u) = \mathbb{P}(Y_1 + \cdots + Y_K > u)$ mit K geometrisch, $\mathbb{P}(K = n) = (1-q)q^n$, und $Y_i \sim F_I$ unabhängig. Die Y_i sind die Höhen, um die der Prozess $\sup_{s\leq t}(S_s - cs)$ jeweils einen neuen Rekord setzt (*Leiterhöhen*), und K ist die Anzahl der Rekorde.

Satz 8.22 ([BOS25, Satz 5.83]). *Im Cramér–Lundberg-Modell mit $\rho > 0$ und $F_I \in \mathcal{S}$ gilt*

$$\Psi(u) \sim \frac{1}{\rho}\,\overline{F}_I(u), \qquad u \to \infty.$$

Beweis. Nach Satz 8.21 ist $\Psi(u)/\overline{F}_I(u) = (1-q)\sum_{n\geq 1} q^n\, \overline{F_I^{*n}}(u)/\overline{F}_I(u)$. Jeder Quotient konvergiert gegen n, und nach Proposition 8.20 mit $q(1+\varepsilon) < 1$ sind die Summanden durch $Kq^n(1+\varepsilon)^n$ summierbar majorisiert. Dominierte Konvergenz liefert den Limes $(1-q)\sum_n nq^n = q/(1-q) = 1/\rho$. □

Beispiel 8.23 (Pareto-Schäden, [BOS25, Bsp. 5.84]). Für $\overline{F}(x) = x^{-\alpha}$, $x \geq 1$, $\alpha > 1$, ist $\mu = \alpha/(\alpha-1)$ und $\overline{F}_I(u) = u^{1-\alpha}/\alpha$ für $u \geq 1$. Also

$$\Psi(u) \sim \frac{u^{-(\alpha-1)}}{\alpha\rho}.$$

Die Ruinwahrscheinlichkeit fällt nur polynomiell: Eine Verdopplung des Startkapitals senkt sie lediglich um den Faktor $2^{\alpha-1}$, während sie im Lundberg-Fall quadriert würde.

Abbildung 8.4 zeigt die polynomielle Abklingrate und die langsame Einstellung der Asymptotik.

Satz 8.24 (Embrechts–Veraverbeke 1982, [BOS25, Satz 5.85]). *Im Cramér–Lundberg-Modell mit $\rho > 0$ sind äquivalent: (i) $F_I \in \mathcal{S}$; (ii) $1 - \Psi \in \mathcal{S}$; (iii) $\Psi(u)/\overline{F}_I(u) \to 1/\rho$.*

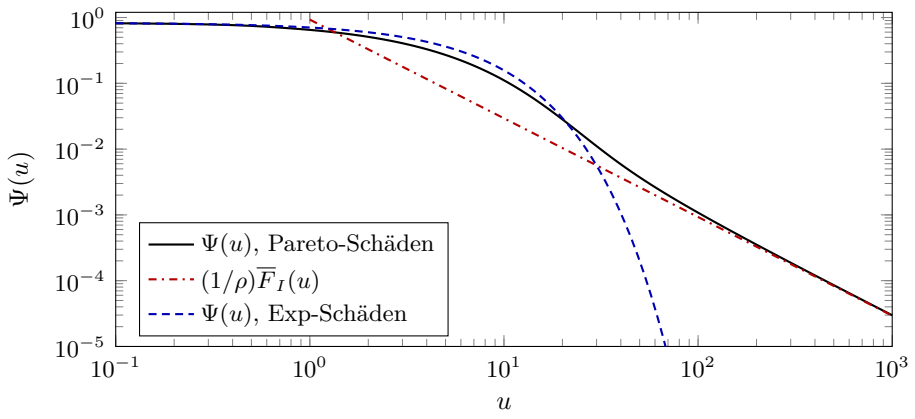

Abbildung 8.4: Ruinwahrscheinlichkeit bei Pareto-Schäden ($\kappa = 0{,}6$, $\alpha = 2{,}5$, $\mu = 1$) und bei exponentiellen Schäden mit gleichem Mittel, jeweils $\rho = 0{,}2$, doppelt logarithmisch. Pareto: numerische Lösung der defekten Erneuerungsgleichung (Schrittweite 0,1; Halbierung ändert die Werte um weniger als 0,5 %). Die Ruinwahrscheinlichkeit Ψ fällt nur polynomiell wie $u^{-3/2}$, also wie $\overline{F}_I$. Die Asymptotik aus Satz 8.22 stellt sich langsam ein: Das Verhältnis $\Psi(u)\big/\big(\overline{F}_I(u)/\rho\big)$ beträgt 3,78 bei $u = 10$, 1,16 bei $u = 100$ und 1,01 bei $u = 1000$.

Das Prinzip des einzelnen großen Sprungs. Im subexponentiellen Fall tritt Ruin bei großem Startkapital typischerweise durch *einen einzigen* Großschaden ein, nicht durch eine Häufung (vgl. Abschnitt 8.2). Eine Heuristik bestätigt die Formel: Bis zur Zeit t ist das Kapital etwa $u + (c - \lambda\mu)t = u + \rho\lambda\mu t$, und ein Schaden, der dieses übersteigt, tritt mit Rate $\lambda\overline{F}(u + \rho\lambda\mu t)$ auf. Also

$$\Psi(u) \approx \int_0^\infty \lambda\overline{F}(u + \rho\lambda\mu t)\,\mathrm{d}t = \frac{1}{\rho\mu}\int_u^\infty \overline{F}(y)\,\mathrm{d}y = \frac{1}{\rho}\overline{F}_I(u).$$

Abbildung 8.5 vergleicht simulierte Ruinpfade bei leichten und schweren Tails.

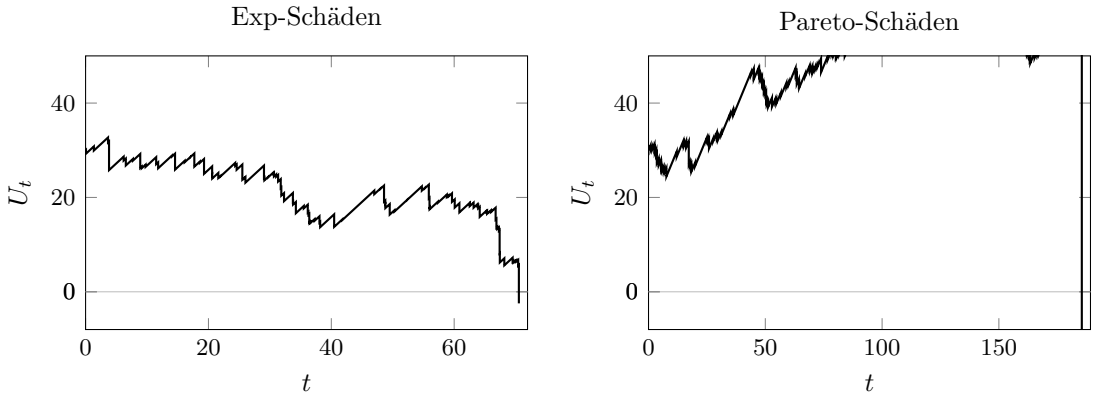

Abbildung 8.5: Je ein simulierter Pfad bis zum Ruin bei $u = 30$, $\rho = 0{,}2$, $\mu = 1$; links Exp(1)-, rechts Pareto-Schäden ($\kappa = 0{,}6$, $\alpha = 2{,}5$). Die Ruinwahrscheinlichkeiten haben dieselbe Größenordnung (Simulation: 0,55 % bzw. 1,11 %; exakt bzw. numerisch aus Abbildung 8.4: 0,56 % bzw. 1,11 %). In 60 000 Läufen pro Modell war der größte Schaden vor dem Ruin bei exponentiellen Schäden im Median nur $0{,}22\,u$ groß (größer als $u/2$ in 0 % der Ruinfälle), bei Pareto-Schäden im Median $1{,}45\,u$ (größer als $u/2$ in 90 %): Ruin durch Häufung gegenüber Ruin durch einen großen Sprung. Die Pfade sind Einzelbeispiele, die Prozentangaben Simulationsergebnisse.

Brücke zur Extremwerttheorie. Bei schweren Tails bestimmt das *Maximum* der Schäden das Verhalten der Summe. Es liegt daher nahe, Maxima und große Überschreitungen direkt zu studieren: Wie groß ist der größte Schaden unter n Schäden typischerweise, und wie schätzt man $\overline{F}(x)$ weit außerhalb der beobachteten Daten? Das ist der Gegenstand von Kapitel 9.

Aufgaben

Aufgabe 8.1. Verifizieren Sie die Formel für Ψ bei exponentiellen Schäden durch Einsetzen in Gleichung (8.3).

Aufgabe 8.2 (vgl. [BOS25, Übung 5.67])**.** Zeigen Sie: Ist $f(r) < \infty$ für ein $r > r_0$, so ist $\vartheta < \infty$.

Aufgabe 8.3. Bestimmen Sie r_0 numerisch für $\lambda = 1$, $c = 1{,}2$ und $X_1 \sim \Gamma(2, 2)$, und vergleichen Sie Lundberg-Schranke und -Approximation für $u = 5, 10, 20$ mit einer Monte-Carlo-Simulation.

Aufgabe 8.4 (vgl. [BOS25, Bsp. 5.74])**.** Zeigen Sie, dass die „Peter-und-Paul-Verteilung“ $F(x) = \sum_{n\in\mathbb{N}:\ 2^n \le x} 2^{-n}$ heavy-tailed, aber nicht langschwänzig und daher nicht subexponentiell ist.

Aufgabe 8.5. Bestimmen Sie die Asymptotik von Ψ für Lomax-verteilte Schäden mit $\alpha > 1$.

Aufgabe 8.6. Simulieren Sie den Risikoprozess mit Pareto-Schäden ($\alpha = 1{,}5$) und mit Gamma-Schäden gleichen Erwartungswerts bei hohem Startkapital. Bedingen Sie auf Ruin und vergleichen Sie die Pfade kurz vor dem Ruinzeitpunkt.

Teil III

Extremwerttheorie

Kapitel 9

Extremwerttheorie

Bisher haben wir Summen von Schäden untersucht. Bei schweren Tails wird die Summe von ihrem größten Summanden dominiert (Abschnitt 8.3). In der Praxis fragt man oft direkt nach dem größten Schaden eines Jahres, nach dem „Jahrhundertsturm“ oder nach einem Quantil weit außerhalb der Daten. Die Extremwerttheorie beantwortet solche Fragen durch Grenzwertsätze für Maxima, die eine ähnliche Rolle spielen wie der zentrale Grenzwertsatz für Summen.

9.1 Maxima und Poisson-Approximation

Seien $X_1, X_2, \dots$ u. i. v. mit Verteilungsfunktion F und rechtem Endpunkt $x_F := \sup\{x\colon F(x) < 1\} \in (-\infty, \infty]$. Für $M_n := \max\{X_1, \dots, X_n\}$ gilt $\mathbb{P}(M_n \le x) = F(x)^n$, und $M_n \to x_F$ fast sicher. Um eine nichtdegenerierte Grenzverteilung zu erhalten, muss man M_n geeignet normieren. Das zentrale Hilfsmittel ist die folgende einfache Beobachtung.

Proposition 9.1 (Poisson-Approximation)**.** *Sei $\zeta \in [0, \infty)$ und (u_n) eine reelle Folge. Dann gilt*

$$n\overline{F}(u_n) \to \zeta \quad \Longleftrightarrow \quad \mathbb{P}(M_n \le u_n) \to \mathrm{e}^{-\zeta}.$$

Beweis. „⇒“: Dann ist $\overline{F}(u_n) \to 0$, und wegen $\log(1-y) = -y(1+o(1))$ für $y \to 0$ gilt $\log \mathbb{P}(M_n \le u_n) = n \log(1 - \overline{F}(u_n)) = -n\overline{F}(u_n)(1+o(1)) \to -\zeta$. „⇐“: Aus $F(u_n)^n \to \mathrm{e}^{-\zeta} > 0$ folgt $F(u_n) \to 1$, also $\overline{F}(u_n) \to 0$, und dieselbe Entwicklung liefert $n\overline{F}(u_n) \to \zeta$. □

Interpretation: Die Anzahl $B_n = \sum_{i \le n} \mathbb{1}\{X_i > u_n\}$ der Überschreitungen von u_n ist $\mathrm{Bin}(n, \overline{F}(u_n))$-verteilt und konvergiert gegen $\mathrm{Poi}(\zeta)$; $\{M_n \le u_n\} = \{B_n = 0\}$.

9.2 Grenzverteilungen für Maxima

Definition 9.2. Zwei Verteilungsfunktionen $H, \tilde{H}$ heißen *vom selben Typ*, wenn $\tilde{H}(x) = H(ax+b)$ für Konstanten $a > 0$, $b \in \mathbb{R}$. Eine Verteilungsfunktion F liegt im *Maximumanziehungsbereich* $\mathrm{MDA}(H)$ einer nichtdegenerierten Verteilungsfunktion H, wenn es $a_n > 0$, $b_n \in \mathbb{R}$ gibt mit

$$\mathbb{P}\Big(\frac{M_n - b_n}{a_n} \le x\Big) = F(a_n x + b_n)^n \longrightarrow H(x)$$

in allen Stetigkeitsstellen von H.

Nach dem *Satz über die Konvergenz zu Typen* ist der Grenzwert bis auf den Typ eindeutig bestimmt [Res87, Prop. 0.2]. Ersetzt man n durch nk, so sieht man, dass jeder Grenzwert *max-stabil* ist: Für jedes k gibt es $\alpha_k > 0, \beta_k$ mit $H(\alpha_k x + \beta_k)^k = H(x)$. Die Lösung dieser Funktionalgleichung liefert:

Satz 9.3 (Fisher–Tippett 1928, Gnedenko 1943)**.** *Ist $F \in \mathrm{MDA}(H)$, so ist H vom Typ einer der drei Standard-Extremwertverteilungen*

$$\begin{aligned}
&\textit{Fréchet:} && \Phi_\alpha(x) = \exp(-x^{-\alpha}),\ x > 0 \quad (\alpha > 0),\\
&\textit{Gumbel:} && \mathcal{G}(x) = \exp(-\mathrm{e}^{-x}),\ x \in \mathbb{R},\\
&\textit{Weibull:} && \mathcal{W}_\alpha(x) = \exp(-(-x)^\alpha),\ x \le 0 \quad (\alpha > 0).
\end{aligned}$$

Umgekehrt ist jede dieser Verteilungen max-stabil und daher Grenzwert.

Die drei Typen fasst die *verallgemeinerte Extremwertverteilung* (GEV) zusammen:

$$H_\xi(x) = \exp\bigl(-(1+\xi x)^{-1/\xi}\bigr), \quad 1+\xi x > 0, \qquad H_0(x) = \exp(-\mathrm{e}^{-x}).$$

Der *Extremwertindex* ξ entscheidet über den Typ: $\xi > 0$ ist Fréchet mit $\alpha = 1/\xi$, $\xi = 0$ Gumbel, $\xi < 0$ Weibull mit $\alpha = -1/\xi$. In der Statistik passt man meist $H_\xi((x-\mu)/\sigma)$ mit drei Parametern an.

Abbildung 9.1 zeigt die drei Typen der Extremwertverteilung.

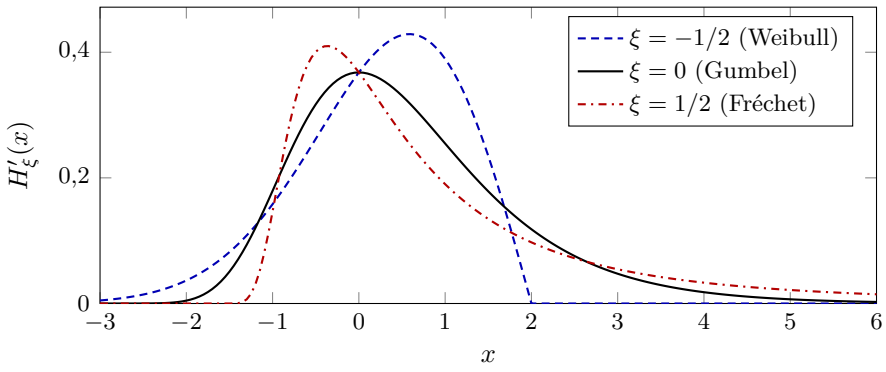

Abbildung 9.1: Dichten der verallgemeinerten Extremwertverteilung H_ξ. Für $\xi < 0$ ist der Träger nach rechts durch $-1/\xi$ beschränkt (hier $x < 2$), für $\xi > 0$ nach links durch $-1/\xi$ (hier $x > -2$), und der rechte Tail ist regulär variierend mit Index $-1/\xi$. Die Gumbel-Dichte hat einen exponentiell fallenden rechten Tail.

9.3 Beispiele und Anziehungsbereiche

Beispiel 9.4 (Exponentialverteilung: Gumbel)**.** Für $F = \mathrm{Exp}(1)$ und $x \in \mathbb{R}$ ist, sobald $x > -\log n$,

$$\mathbb{P}(M_n - \log n \le x) = \bigl(1 - \mathrm{e}^{-x}/n\bigr)^n \to \exp(-\mathrm{e}^{-x}).$$

Also $a_n = 1$, $b_n = \log n$: Das Maximum wächst logarithmisch.

Abbildung 9.2 zeigt die Konvergenz im Exponentialbeispiel.

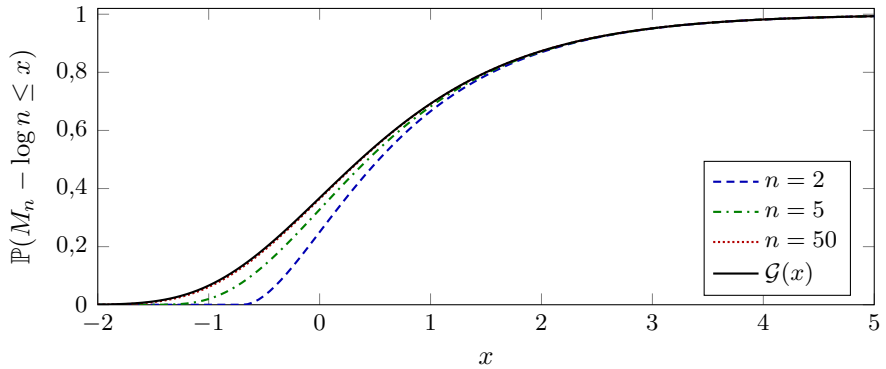

Abbildung 9.2: Verteilungsfunktion des normierten Maximums $M_n - \log n$ von n unabhängigen Exp(1)-Variablen, exakt $(1 - \mathrm{e}^{-x}/n)^n$, und der Gumbel-Grenzwert $\mathcal{G}$. Die Konvergenz ist schnell: Für $n = 5$ beträgt der maximale Abstand zum Grenzwert noch etwa 0,058, für $n = 50$ ist kaum noch ein Unterschied zu sehen.

Satz 9.5 (Gnedenko)**.** *Sei $x_F = \infty$. Ist $\overline{F} \in \mathcal{R}_{-\alpha}$ mit $\alpha > 0$, so gilt $F \in \mathrm{MDA}(\Phi_\alpha)$ mit $b_n = 0$ und $a_n := F^{\leftarrow}(1 - 1/n) = \inf\{x\colon F(x) \ge 1 - 1/n\}$. Es gilt auch die Umkehrung.*

Beweis der Hinrichtung. Es ist $a_n \to \infty$ und $\overline{F}(a_n) \le 1/n \le \overline{F}(a_n-) \le \overline{F}((1-\delta)a_n)$ für $\delta \in (0,1)$. Mit regulärer Variation folgt $(1-\delta)^\alpha \le \liminf n\overline{F}(a_n) \le \limsup n\overline{F}(a_n) \le 1$, also $n\overline{F}(a_n) \to 1$. Für $x > 0$ ist dann

$$n\overline{F}(a_n x) = n\overline{F}(a_n) \cdot \frac{\overline{F}(a_n x)}{\overline{F}(a_n)} \to x^{-\alpha},$$

und Proposition 9.1 liefert $\mathbb{P}(M_n \le a_n x) \to \exp(-x^{-\alpha})$. Für $x \le 0$ ist $\mathbb{P}(M_n \le a_n x) \le F(0)^n \to 0$, sofern $F(0) < 1$. □

Für Pareto-Schäden $\overline{F}(x) = (\kappa/x)^\alpha$ ist $a_n = \kappa n^{1/\alpha}$: Der größte von n Schäden wächst polynomiell in n, und für $\alpha < 1$ ist er von derselben Größenordnung wie die ganze Summe.

Beispiel 9.6 (Gleichverteilung: Weibull)**.** Für $F = \mathrm{U}[0,1]$ und $x \le 0$ ist $\mathbb{P}(n(M_n - 1) \le x) = (1 + x/n)^n \to \mathrm{e}^x = \mathcal{W}_1(x)$. Allgemein gilt $F \in \mathrm{MDA}(\mathcal{W}_\alpha)$ genau dann, wenn $x_F < \infty$ und $\overline{F}(x_F - 1/x) \in \mathcal{R}_{-\alpha}$.

Der Gumbel-Bereich ist komplizierter zu charakterisieren; eine hinreichende Bedingung (von Mises) ist die Darstellung $\overline{F}(x) = c\exp\bigl(-\int_z^x \mathrm{d}t/a(t)\bigr)$ mit einer positiven, differenzierbaren Hilfsfunktion a mit $a'(t) \to 0$; dann kann man $b_n = F^{\leftarrow}(1-1/n)$ und $a_n = a(b_n)$ wählen.

Verteilung	Anziehungsbereich	Tail-Klasse
Pareto, Lomax, Burr, Student-t, Log-Gamma	Fréchet, $\xi = 1/\alpha > 0$	$\mathcal{R}_{-\alpha} \subset \mathcal{S}$
Log-Normal, Weibull mit $\alpha < 1$	Gumbel, $\xi = 0$	$\mathcal{S}$, nicht $\mathcal{R}$
Exponential, Gamma, Normal	Gumbel, $\xi = 0$	light-tailed
Gleich-, Beta-Verteilung	Weibull, $\xi < 0$	beschränkter Träger

Die Klassifikation nach ξ deckt sich also nicht mit der Unterscheidung leicht/schwer: Der Gumbel-Bereich enthält leichte und subexponentielle Verteilungen. Für die Versicherung besonders relevant ist der Fall $\xi > 0$: Dann existieren nur die Momente der Ordnung $< 1/\xi$.

Wiederkehrniveaus. Beschreibt H die Verteilung des Jahresmaximums, so heißt $z_T \coloneqq H^{\leftarrow}(1 - 1/T)$ das *T-Jahres-Wiederkehrniveau*: Es wird im Mittel einmal in T Jahren überschritten (die Wartezeit bis zur ersten Überschreitung ist geometrisch mit Erwartungswert T). Für die GEV mit Parametern (ξ, μ, σ) und $\xi \neq 0$ ist

$$z_T = \mu - \frac{\sigma}{\xi}\Bigl(1 - \bigl(-\log(1-1/T)\bigr)^{-\xi}\Bigr).$$

Die *Block-Maxima-Methode* passt die GEV an beobachtete Jahresmaxima an (Maximum-Likelihood, regulär für $\xi > -1/2$) und liest daraus z_T ab. Ihr Nachteil: Sie nutzt pro Block nur eine Beobachtung. Das motiviert die Überschreitungsmethode in Abschnitt 9.4.

9.4 Exzessverteilung und mittlere Exzessfunktion

Definition 9.7. Sei $X \sim F$ und $u < x_F$. Die *Exzessverteilung* über der Schwelle u ist

$$F_u(x) \coloneqq \mathbb{P}(X - u \le x \mid X > u), \qquad \overline{F}_u(x) = \frac{\overline{F}(u+x)}{\overline{F}(u)},\ x \ge 0.$$

Ist $\mathbb{E}[X^+] < \infty$, so heißt $e(u) \coloneqq \mathbb{E}[X - u \mid X > u]$ die *mittlere Exzessfunktion*.

In der Rückversicherung ist $e(u)$ die erwartete Zahlung eines Schadenexzedentenvertrags mit Priorität u pro betroffenem Schaden; in der Lebensversicherung ist $e(x)$ die Restlebenserwartung.

Lemma 9.8. *Für $X \ge 0$ mit $\mu = \mathbb{E}[X] < \infty$ ist $e(u)\overline{F}(u) = \int_u^\infty \overline{F}(y)\,\mathrm{d}y$, also*

$$\overline{F}_I(u) = \frac{e(u)\overline{F}(u)}{\mu}.$$

Beweis. Es gilt $\mathbb{E}[(X-u)^+] = \int_0^\infty \mathbb{P}(X - u > y)\,\mathrm{d}y = \int_u^\infty \overline{F}(y)\,\mathrm{d}y$, und $\mathbb{E}[(X-u)^+] = e(u)\overline{F}(u)$. □

Zusammen mit Satz 8.22 erhält man im subexponentiellen Fall die gut interpretierbare Form $\Psi(u) \sim e(u)\overline{F}(u)/(\rho\mu)$: Die Ruinwahrscheinlichkeit ist bis auf den Faktor $1/(\rho\mu)$ die Wahrscheinlichkeit eines Schadens über u, multipliziert mit dessen mittlerem Überschuss.

Beispiel 9.9.

- Exp(λ): $e(u) \equiv 1/\lambda$ (Gedächtnislosigkeit).
- Pareto (κ, α), $\alpha > 1$, $u \geq \kappa$: $e(u) = u/(\alpha - 1)$.
- Lomax (κ, α), $\alpha > 1$: $e(u) = (\kappa + u)/(\alpha - 1)$.
- Log-Normal: $e(u) \sim \sigma^2 u / \log u$; Weibull mit $\alpha < 1$: $e(u) \sim u^{1-\alpha}/(\alpha\lambda^\alpha)$.
- Gamma mit Form > 1: $e(u)$ fällt gegen $1/\beta$.

Ein *linear wachsendes* empirisches $e(u)$ („mean excess plot“) ist daher ein Hinweis auf einen Pareto-artigen Tail, ein konstantes auf exponentielles Verhalten.

Abbildung 9.3 vergleicht mittlere Exzessfunktionen.

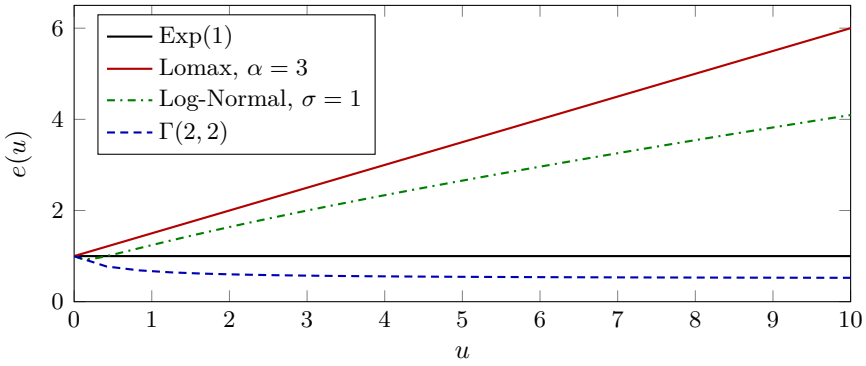

Abbildung 9.3: Mittlere Exzessfunktionen von vier Verteilungen mit Erwartungswert 1 (also $e(0) = 1$). Lomax: linear, $e(u) = (\kappa + u)/(\alpha - 1)$; Log-Normal: wachsend, aber sublinear; exponentiell: konstant; $\Gamma(2, 2)$: fallend gegen $1/\beta = 1/2$. Das Wachstum von e ist ein empirisch gut ablesbares Merkmal schwerer Tails.

9.5 Die verallgemeinerte Pareto-Verteilung

Definition 9.10. Für $\xi \in \mathbb{R}$, $\beta > 0$ ist die *verallgemeinerte Pareto-Verteilung* $\mathrm{GPD}(\xi, \beta)$ gegeben durch

$$\overline{G}_{\xi,\beta}(x) = \left(1 + \frac{\xi x}{\beta}\right)^{-1/\xi}, \qquad \overline{G}_{0,\beta}(x) = \mathrm{e}^{-x/\beta},$$

für $x \geq 0$ (und $x \leq -\beta/\xi$, falls $\xi < 0$).

Für $\xi > 0$ ist das eine Lomax-Verteilung mit $\overline{G} \in \mathcal{R}_{-1/\xi}$, für $\xi = 0$ die Exponentialverteilung, für $\xi < 0$ hat sie beschränkten Träger.

Lemma 9.11 (Schwellenstabilität)**.** *Ist $X \sim \mathrm{GPD}(\xi, \beta)$ und $v \geq 0$ im Träger, so ist der Exzess über v wieder $\mathrm{GPD}(\xi, \beta + \xi v)$-verteilt. Für $\xi < 1$ ist $e(v) = (\beta + \xi v)/(1 - \xi)$, also linear in v.*

Beweis. Es gilt $\overline{G}(v + x)/\overline{G}(v) = \big((\beta + \xi v + \xi x)/(\beta + \xi v)\big)^{-1/\xi} = \big(1 + \xi x/(\beta + \xi v)\big)^{-1/\xi}$. Der Erwartungswert von $\mathrm{GPD}(\xi, \beta)$ ist $\beta/(1 - \xi)$ für $\xi < 1$. □

Der folgende Satz erklärt, warum die GPD die natürliche Approximation für Exzesse über hohe Schwellen ist; er ist das Gegenstück zu Satz 9.3.

Satz 9.12 (Pickands 1975, Balkema–de Haan 1974)**.** *Es gilt $F \in \mathrm{MDA}(H_\xi)$ genau dann, wenn es eine positive Funktion $\beta(u)$ gibt mit*

$$\lim_{u \uparrow x_F} \sup_{0 \leq x < x_F - u} \big|F_u(x) - G_{\xi,\beta(u)}(x)\big| = 0.$$

Beweis der punktweisen Aussage im Fréchet-Fall. Sei $\overline{F} \in \mathcal{R}_{-\alpha}$, $\xi = 1/\alpha$ und $\beta(u) \coloneqq \xi u$. Für $x > 0$ ist

$$\overline{F}_u(ux) = \frac{\overline{F}(u(1+x))}{\overline{F}(u)} \to (1+x)^{-\alpha} = \overline{G}_{\xi,\xi u}(ux).$$

Da beide Seiten in x monoton und der Limes stetig ist, ist die Konvergenz gleichmäßig (Pólya); das ergibt die Aussage für den Fréchet-Fall. □

9.6 Die Peaks-over-Threshold-Methode

Gegeben seien Schäden $X_1, \dots, X_n$. Man wählt eine hohe Schwelle u, sei N_u die Anzahl der Überschreitungen und $Y_j = X_{i_j} - u$ die Exzesse. Nach Satz 9.12 modelliert man $Y_j \sim \mathrm{GPD}(\xi, \beta)$ und schätzt (ξ, β), etwa durch Maximum-Likelihood. Wegen $\overline{F}(u+x) = \overline{F}(u)\overline{F}_u(x)$ und $\overline{F}(u) \approx N_u/n$ erhält man den *Tail-Schätzer*

$$\widehat{\overline{F}}(x) = \frac{N_u}{n}\Big(1 + \hat{\xi}\,\frac{x-u}{\hat{\beta}}\Big)^{-1/\hat{\xi}}, \qquad x > u.$$

Er erlaubt Aussagen *jenseits* der größten Beobachtung – genau das, was die empirische Verteilungsfunktion nicht kann.

Proposition 9.13 (Quantil- und Expected-Shortfall-Schätzer)**.** *Im GPD-Modell oberhalb von u ist das p-Quantil (p nahe 1)*

$$q_p = u + \frac{\beta}{\xi}\Big(\Big(\frac{n}{N_u}(1-p)\Big)^{-\xi} - 1\Big),$$

und für $\xi < 1$ ist der Expected Shortfall

$$\mathrm{ES}_p \coloneqq \mathbb{E}[X \mid X > q_p] = \frac{q_p + \beta - \xi u}{1-\xi}.$$

Beweis. Die erste Formel ist die Lösung von $\widehat{\overline{F}}(q_p) = 1 - p$. Für die zweite ist $\mathrm{ES}_p = q_p + e(q_p)$, und nach Lemma 9.11 mit $v = q_p - u$ ist $e(q_p) = (\beta + \xi(q_p - u))/(1-\xi)$. □

In der Notation von Kapitel 10, in der Risikomaße auf Positionen (Gewinn positiv) wirken, ist $q_p = \mathrm{V@R}_{1-p}(-X)$ und für stetige Verteilungen $\mathrm{ES}_p = \mathrm{AV@R}_{1-p}(-X)$. Für $\xi > 0$ gilt $\mathrm{ES}_p/q_p \to 1/(1-\xi)$: Bei schweren Tails liegt der mittlere Verlust jenseits des V@R deutlich über diesem.

Momentenschätzer. Für $\xi < 1/2$ hat die GPD Mittelwert $m = \beta/(1-\xi)$ und Varianz $s^2 = \beta^2/((1-\xi)^2(1-2\xi))$. Auflösen ergibt $\hat{\xi} = (1/2)(1 - \bar{m}^2/\bar{s}^2)$, $\hat{\beta} = \bar{m}(1-\hat{\xi})$ mit empirischem Mittel $\bar{m}$ und Varianz $\bar{s}^2$ der Exzesse.

9.7 Der Hill-Schätzer

Für $\xi > 0$ gibt es einen besonders einfachen Schätzer. Seien $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(n)}$ die Ordnungsstatistiken.

Definition 9.14. Für $1 \leq k < n$ heißt

$$\hat{\xi}^{\mathrm{H}}_k \coloneqq \frac{1}{k}\sum_{i=1}^{k} \log X_{(i)} - \log X_{(k+1)}$$

Hill-Schätzer (1975) für $\xi = 1/\alpha$.

Motivation. Ist $\overline{F}(x) = (x/u)^{-\alpha}$ für $x \geq u$, so ist $\mathbb{P}(\log(X/u) > t | X > u) = \mathrm{e}^{-\alpha t}$: Die logarithmischen Exzesse sind $\mathrm{Exp}(\alpha)$-verteilt mit Mittelwert ξ. Der Hill-Schätzer ist ihr empirisches Mittel mit der zufälligen Schwelle $u = X_{(k+1)}$; er ist zugleich der Maximum-Likelihood-Schätzer in diesem Modell.

Satz 9.15 (Mason 1982)**.** *Ist $\overline{F} \in \mathcal{R}_{-1/\xi}$ mit $\xi > 0$ und gilt $k = k_n \to \infty$, $k/n \to 0$, so konvergiert $\hat{\xi}^{\mathrm{H}}_k \to \xi$ in Wahrscheinlichkeit.*

Unter einer Bedingung zweiter Ordnung ist $\sqrt{k}(\hat{\xi}^{\mathrm{H}}_k - \xi)$ asymptotisch normal mit Varianz ξ^2 [dHF06, Abschn. 3.2]. Die Wahl von k ist ein Bias-Varianz-Kompromiss: Kleines k bedeutet hohe Varianz, großes k zieht Daten aus dem Zentrum der Verteilung heran, für die das Pareto-Modell nicht gilt. In der Praxis betrachtet man den *Hill-Plot* $k \mapsto \hat{\xi}^{\mathrm{H}}_k$ und sucht einen stabilen Bereich; analog wählt man die POT-Schwelle u dort, wo der mean excess plot linear wird.

Abbildung 9.4 zeigt einen Hill-Plot für simulierte Daten.

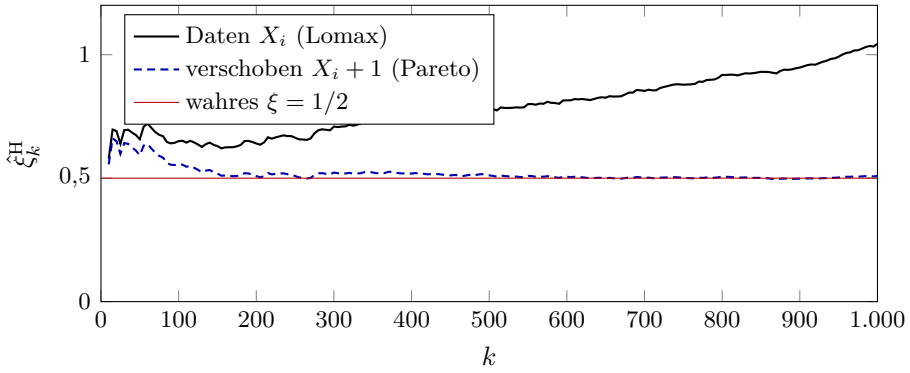

Abbildung 9.4: Hill-Plot für eine simulierte Stichprobe von $n = 2000$ Lomax-Schäden mit $\kappa = 1$, $\alpha = 2$, also $\xi = 1/2$. Für die Daten selbst (durchgezogen) liegt der Schätzer deutlich über ξ und steigt für große k stark an (0,66; 0,64; 1,04 bei $k = 50$; 200; 1000): Der Tail $(1 + x)^{-2}$ sieht erst weit draußen wie x^{-2} aus. Für die um κ verschobenen, exakt Pareto-verteilten Werte (gestrichelt) schwankt er nur um ξ (0,59; 0,51; 0,51). Der Hill-Schätzer ist nicht verschiebungsinvariant, und die Wahl von k ist ein Bias-Varianz-Kompromiss.

Ausblick

Für abhängige Daten (Zeitreihen von Schäden, Clusterbildung bei Stürmen) benötigt man den *Extremalindex*; für mehrere Risiken die multivariate Extremwerttheorie und Tail-Abhängigkeit. Beides ist für Rückversicherung und Solvency II zentral, siehe [EKM97, Kap. 8], [MFE15, Kap. 7].

Aufgaben

Aufgabe 9.1. Zeigen Sie, dass Φ_α, $\mathcal{G}$ und $\mathcal{W}_\alpha$ max-stabil sind, und bestimmen Sie jeweils α_k, β_k.

Aufgabe 9.2. Sei $F = \mathrm{Exp}(\lambda)$ und $N \sim \mathrm{Poi}(\nu)$ unabhängig. Bestimmen Sie die Verteilung des größten Schadens $\max\{X_1, \ldots, X_N\}$ (mit $\max \emptyset = 0$) im zusammengesetzten Poisson-Modell und ihre Asymptotik für $\nu \to \infty$.

Aufgabe 9.3. Zeigen Sie mit der von-Mises-Bedingung, dass die Standardnormalverteilung im Gumbel-Bereich liegt, und bestimmen Sie a_n (Hinweis: Mills' ratio $\overline{F}(x) \sim \mathrm{e}^{-x^2/2}/(\sqrt{2\pi}\, x)$ für $x \to \infty$).

Aufgabe 9.4. Lomax-verteilte Schäden mit $\alpha = 2$, $\kappa = 1$: Berechnen Sie a_n und die Wahrscheinlichkeit, dass der größte von $n = 1000$ Schäden die Summe der übrigen übersteigt, per Simulation.

Aufgabe 9.5. Berechnen Sie $e(u)$ für die Lomax-Verteilung und zeigen Sie: Ist $e(u) = a + bu$ für alle $u \geq 0$ mit $a > 0$, $0 \leq b < 1$, so ist F eine GPD.

Aufgabe 9.6. Zeigen Sie, dass der Hill-Schätzer im exakten Pareto-Modell mit *fester* Schwelle u der ML-Schätzer für $1/\alpha$ ist, und bestimmen Sie seine Verteilung.

Aufgabe 9.7. Simulieren Sie $n = 2000$ Lomax-Schäden mit $\alpha = 2$. Erstellen Sie mean excess plot und Hill-Plot, schätzen Sie $q_{0,999}$ und $\text{ES}_{0,999}$ mit der POT-Methode und vergleichen Sie mit den wahren Werten und mit dem empirischen Quantil.

Teil IV

Bewertung von Risiken

Kapitel 10

Prämienprinzipien und Risikomaße

10.1 Quantile

Definition 10.1 ([BOS25, Def. 6.1, 6.5])**.** Für $\lambda \in (0,1)$ heißt q ein λ-*Quantil* von X, wenn $\mathbb{P}(X < q) \leq \lambda \leq \mathbb{P}(X \leq q)$. Die Menge der λ-Quantile ist ein Intervall $[q_X^-(\lambda), q_X^+(\lambda)]$. Die *Linksinverse* $F_X^{-1}(t) := \inf\{x \colon F_X(x) \geq t\}$ erfüllt $F_X^{-1}(\lambda) = q_X^-(\lambda)$. Eine wachsende Funktion q_X mit $q_X(\lambda) \in [q_X^-(\lambda), q_X^+(\lambda)]$ heißt *Quantilfunktion*.

Ist $U \sim \mathrm{U}[0,1]$, so gilt $q_X(U) \overset{d}{=} X$ für jede Quantilfunktion (*Quantiltransformation*); insbesondere ist $\mathbb{E}[g(X)] = \int_0^1 g(q_X(u))\,\mathrm{d}u$.

10.2 Prämienprinzipien

Die Nettorisikoprämie $\mathbb{E}[X]$ reicht auf Dauer nicht aus: Nach Korollar 8.11 führt sie sicher zum Ruin. Eine *Risikoprämie* enthält deshalb einen Sicherheitszuschlag; mit Kosten und Gewinn entsteht die Bruttoprämie. Im Folgenden sei $\mathcal{X}$ die Menge der nichtnegativen Zufallsvariablen (Risiken).

Definition 10.2 (Prämienprinzip, [BOS25, Def. 6.11])**.** Ein *Prämienprinzip* auf $\mathcal{X}' \subset \mathcal{X}$ ist eine Abbildung $\mathcal{H} \colon \mathcal{X}' \to [0,\infty)$, die nur von der Verteilung abhängt.

Prinzip	$\mathcal{H}(X)$	Parameter
(P1) Netto	$\mathbb{E}X$	
(P2) Erwartungswert	$(1+\delta)\mathbb{E}X$	$\delta > 0$
(P3) Varianz	$\mathbb{E}X + \delta \operatorname{Var} X$	$\delta > 0$
(P4) Standardabweichung	$\mathbb{E}X + \delta\sqrt{\operatorname{Var} X}$	$\delta > 0$
(P5) Maximalschaden	$\operatorname{ess\,sup} X$	
(P6) Perzentil	$F_X^{-1}(1-\varepsilon)$	$\varepsilon \in (0,1)$
(P8) Nullnutzen	$\mathbb{E}[u(\mathcal{H}(X) - X)] = 0$	u wachsend, $u(0) = 0$
(P9) Mittelwert	$v^{-1}(\mathbb{E}[v(X)])$	v wachsend, $v(0) = 0$
(P10) Exponential	$(1/a)\log \mathbb{E}[\mathrm{e}^{aX}]$	$a > 0$
(P12) Esscher	$\mathbb{E}[Xg(X)]/\mathbb{E}[g(X)]$	$g > 0$ wachsend
(P14) Wang	$\int_0^\infty g(\overline{F}_X(t))\,\mathrm{d}t$	g konkav, $g(0) = 0, g(1) = 1$
(P15) absolute Abweichung	$\mathbb{E}X + a\mathbb{E}\lvert X - F_X^{-1}(1/2)\rvert$	$a \in (0,1]$

Definition 10.3 (Eigenschaften, [BOS25, Def. 6.13])**.** Ein Prämienprinzip $\mathcal{H}$ heißt (E1) *erwartungswertübersteigend*, wenn $\mathcal{H}(X) \geq \mathbb{E}X$; (E2) *maximalschadenbegrenzt*, wenn $\mathcal{H}(X) \leq \operatorname{ess\,sup} X$; (E3) *monoton*, wenn $X \leq Y$ f. s. $\Rightarrow \mathcal{H}(X) \leq \mathcal{H}(Y)$; (E4) *translationsinvariant*, wenn $\mathcal{H}(X+c) = \mathcal{H}(X) + c$; (E5) *homogen*, wenn $\mathcal{H}(cX) = c\mathcal{H}(X)$ für $c > 0$; (E6) *subadditiv* bzw. (E6*) *stark subadditiv*, wenn $\mathcal{H}(X+Y) \leq \mathcal{H}(X) + \mathcal{H}(Y)$ für unabhängige bzw. beliebige X, Y; (E7) *additiv*, wenn Gleichheit für unabhängige X, Y gilt; (E8) *konvex*; (E9) *konstantenerhaltend*, wenn $\mathcal{H}(c) = c$.

Monotonie, (E2) und (E9) wird man als unverzichtbar ansehen. Ohne starke Subadditivität lohnt es sich, ein Risiko aufzuspalten und getrennt zu versichern.

Proposition 10.4 ([BOS25, Prop. 6.17])**.** *Das Varianz- und das Standardabweichungsprinzip sind nicht monoton; das Erwartungswertprinzip ist nicht translationsinvariant; das Varianzprinzip ist nicht homogen.*

Beweis. Sei $Y \equiv c$ und $X = c\mathbb{1}_A$ mit $\mathbb{P}(A) = p$, also $X \le Y$ und $\mathcal{H}(Y) = c$. Beim Varianzprinzip ist $\mathcal{H}(X) = cp + \delta c^2 p(1-p) > c$, sobald $\delta cp > 1$; beim Standardabweichungsprinzip ist $\mathcal{H}(X) = cp + \delta c\sqrt{p(1-p)} > c$, sobald $p > 1/(1+\delta^2)$. Ferner ist $(1+\delta)(\mathbb{E}X + c) \neq (1+\delta)\mathbb{E}X + c$ und $\delta \operatorname{Var}[cX] = c^2 \delta \operatorname{Var} X$. □

Das Exponentialprinzip. Für $a > 0$ ist $\mathcal{H}(X) = (1/a)\log \mathbb{E}[\mathrm{e}^{aX}]$ das Nullnutzenprinzip zur exponentiellen Nutzenfunktion $u(x) = 1 - \mathrm{e}^{-ax}$ (konstante Risikoaversion a). Es ist monoton, translationsinvariant, additiv für unabhängige Risiken und nach Jensen erwartungswertübersteigend, aber nicht homogen und nicht stark subadditiv. Für heavy-tailed X ist es unendlich – solche Risiken sind mit diesem Prinzip nicht versicherbar.

10.3 Risikomaße

Nun betrachten wir *Finanzpositionen* $X \in \mathbb{X}$ (beschränkte Zufallsvariablen, Gewinn positiv, Verlust negativ).

Definition 10.5 ([BOS25, Def. 6.30, 6.32])**.** Eine Abbildung $\varrho\colon \mathbb{X} \to \mathbb{R}$ heißt *monetäres Risikomaß*, wenn

(a) $X \le Y \Rightarrow \varrho(X) \ge \varrho(Y)$ (Monotonie),

(b) $\varrho(X + m) = \varrho(X) - m$ für $m \in \mathbb{R}$ (Translationsinvarianz).

Ein monetäres Risikomaß ϱ heißt *konvex*, wenn $\varrho(cX + (1-c)Y) \le c\varrho(X) + (1-c)\varrho(Y)$, und *kohärent*, wenn es zusätzlich positiv homogen ist: $\varrho(cX) = c\varrho(X)$ für $c \ge 0$.

Der Wert $\varrho(X)$ ist der kleinste Betrag, den man der Position hinzufügen muss, damit sie akzeptabel wird; es gilt $\varrho(X + \varrho(X)) = 0$. Für positiv homogenes ϱ ist Konvexität äquivalent zur Subadditivität $\varrho(X+Y) \le \varrho(X)+\varrho(Y)$: Diversifikation erhöht das Risiko nicht. Ein monotones, translationsinvariantes und konstantenerhaltendes Prämienprinzip entspricht über $\varrho(X) = \mathcal{H}(-X)$ einem normierten Risikomaß und umgekehrt [BOS25, Bem. 6.31].

Definition 10.6 ([BOS25, Def. 6.35])**.** Für $\lambda \in (0,1)$ ist der *Value at Risk* $\mathrm{V@R}_\lambda(X) := \inf\{m \in \mathbb{R}\colon \mathbb{P}(m + X < 0) \le \lambda\}$, und für $\lambda \in (0,1]$ der *Average Value at Risk*

$$\mathrm{AV@R}_\lambda(X) := \frac{1}{\lambda}\int_0^\lambda \mathrm{V@R}_\gamma(X)\,\mathrm{d}\gamma.$$

Es gilt $\mathrm{V@R}_\lambda(X) = -q_X^+(\lambda)$ und $\mathrm{AV@R}_\lambda \ge \mathrm{V@R}_\lambda$. Ist $X \sim N(m,\sigma^2)$, so ist $\mathrm{V@R}_\lambda(X) = -m + \sigma\,\Phi_{0,1}^{-1}(1-\lambda)$ mit der Standardnormal-Verteilungsfunktion $\Phi_{0,1}$. Für den Verlust $-X$ ist $\mathrm{V@R}_\lambda(X)$ das $(1-\lambda)$-Quantil von $-X$, also q_p aus Abschnitt 9.4 mit $p = 1-\lambda$.

Abbildung 10.1 veranschaulicht V@R und AV@R an der Quantilfunktion.

Beispiel 10.7 (V@R ist nicht subadditiv)**.** Zwei unabhängige Kredite verlieren je mit Wahrscheinlichkeit 0,04 den Betrag 1: $X, Y \in \{-1, 0\}$. Für $\lambda = 0{,}05$ ist $\mathrm{V@R}_\lambda(X) = \mathrm{V@R}_\lambda(Y) = 0$, denn $\mathbb{P}(X < 0) = 0{,}04 \le \lambda$. Aber $\mathbb{P}(X + Y < 0) = 1 - 0{,}96^2 = 0{,}0784 > \lambda$, und $\mathbb{P}(X + Y = -2) = 0{,}0016$; also $\mathrm{V@R}_\lambda(X + Y) = 1 > 0$. Der V@R bestraft hier Diversifikation.

Lemma 10.8 ([BOS25, Lemma 6.38, 6.40])**.** *Für jede Quantilfunktion q_X und $\lambda \in (0,1)$ gilt*

$$\mathrm{AV@R}_\lambda(X) = -\frac{1}{\lambda}\int_0^\lambda q_X(u)\,\mathrm{d}u = \frac{1}{\lambda}\mathbb{E}\big[(q_X(\lambda) - X)^+\big] - q_X(\lambda).$$

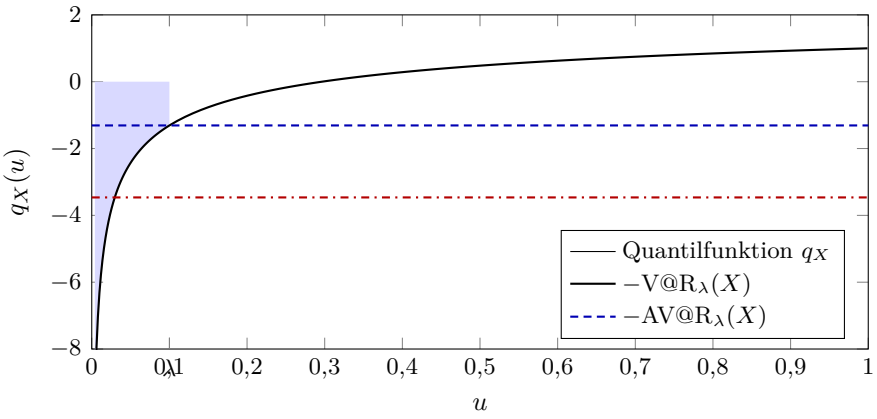

Abbildung 10.1: V@R und AV@R zum Niveau $\lambda = 0{,}1$ für eine Position X, deren Verlust $1-X$ Lomax-verteilt ist ($\kappa = 2$, $\alpha = 3$, also $\mathbb{E}X = 0$). Die Größe $-\mathrm{V@R}_\lambda(X)$ ist der Wert der Quantilfunktion in λ; die Größe $-\mathrm{AV@R}_\lambda(X)$ ist ihr Mittelwert über $(0, \lambda)$ (Lemma 10.8), also die mittlere Höhe der schattierten Fläche. Der V@R ($\approx 1{,}31$) ignoriert, wie tief der Tail links von λ reicht; der AV@R ($\approx 3{,}46$) nicht.

Beweis. Die erste Gleichung folgt aus $\mathrm{V@R}_\gamma = -q_X^+(\gamma)$, da sich q_X^+ und q_X nur auf einer Lebesgue-Nullmenge unterscheiden. Für die zweite ist nach der Quantiltransformation $\mathbb{E}[(q-X)^+] = \int_0^1 (q - q_X(u))^+ \,\mathrm{d}u$ mit $q = q_X(\lambda)$, und $q_X(u) \le q$ für $u < \lambda$, $q_X(u) \ge q$ für $u > \lambda$. Also $\mathbb{E}[(q-X)^+] = \lambda q - \int_0^\lambda q_X(u)\,\mathrm{d}u$. □

Satz 10.9 ([BOS25, Satz 6.41])**.** *Für $\lambda \in (0,1]$ ist $\mathrm{AV@R}_\lambda$ kohärent, und es gilt*

$$\mathrm{AV@R}_\lambda(X) = \max_{Q \in \mathcal{Q}_\lambda} \mathbb{E}_Q[-X], \qquad \mathcal{Q}_\lambda = \Big\{ Q \ll \mathbb{P} \colon \frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} \le \frac{1}{\lambda} \Big\}.$$

Beweis. Sei $\lambda < 1$ ($\lambda = 1$: $\mathcal{Q}_1 = \{\mathbb{P}\}$). Für $Q \in \mathcal{Q}_\lambda$ ist $Z := \lambda \mathrm{d}Q/\mathrm{d}\mathbb{P} \in [0,1]$ mit $\mathbb{E}Z = \lambda$. Mit $q = q_X(\lambda)$ gilt

$$\mathbb{E}_Q[-X] = \frac{1}{\lambda}\mathbb{E}[(q-X)Z] - q \le \frac{1}{\lambda}\mathbb{E}[(q-X)^+] - q = \mathrm{AV@R}_\lambda(X)$$

nach Lemma 10.8. Gleichheit gilt für $Z = \mathbb{1}\{X < q\} + r\mathbb{1}\{X = q\}$, wobei $r \in [0,1]$ so gewählt wird, dass $\mathbb{E}Z = \lambda$; das ist möglich, da $\mathbb{P}(X < q) \le \lambda \le \mathbb{P}(X \le q)$. Als Maximum linearer Funktionale, die jeweils monoton, translationsinvariant und positiv homogen sind, ist $\mathrm{AV@R}_\lambda$ kohärent. □

Interpretation: $\mathrm{AV@R}_\lambda$ ist der erwartete Verlust unter dem ungünstigsten Szenario, das die Wahrscheinlichkeiten höchstens um den Faktor $1/\lambda$ verzerrt; für stetiges X ist $\mathrm{AV@R}_\lambda(X) = \mathbb{E}[-X \mid X \le q_X(\lambda)]$ (Expected Shortfall).

Zurück zu den Prämienprinzipien. Setzt man $\mathcal{H}_\lambda(X) := \mathrm{AV@R}_\lambda(-X) = \frac{1}{\lambda}\int_{1-\lambda}^1 q_X(u)\,\mathrm{d}u$ für Risiken $X \ge 0$, so gilt $\mathcal{H}_\lambda(X) = \int_0^\infty \min\{1, \overline{F}_X(t)/\lambda\}\,\mathrm{d}t$. Allgemeiner ist für jedes Wahrscheinlichkeitsmaß μ auf $(0,1]$

$$\int \mathcal{H}_\lambda(X)\,\mu(\mathrm{d}\lambda) = \int_0^\infty g(\overline{F}_X(t))\,\mathrm{d}t, \qquad g(s) = \int \min\{1, s/\lambda\}\,\mu(\mathrm{d}\lambda),$$

also ein Wangsches Prinzip [BOS25, Satz 6.45]. Das absolute Abweichungsprinzip ist der Spezialfall $\mu = (1-a)\delta_1 + a\delta_{1/2}$, d. h. $\mathbb{E}X + a\mathbb{E}|X-m| = (1-a)\mathbb{E}X + a\,\mathcal{H}_{1/2}(X)$ für den Median m. Daraus folgt, dass es alle Eigenschaften (E1)–(E9) außer der Additivität besitzt [BOS25, Satz 6.29].

Aufgaben

Aufgabe 10.1 (vgl. [BOS25, Übung 6.8])**.** Beweisen Sie die Quantiltransformation $q_X(U) \overset{d}{=} X$.

Aufgabe 10.2 (vgl. [BOS25, Übung 6.24])**.** Zeigen Sie mit der Hölder-Ungleichung, dass das Exponentialprinzip mit $a > 0$ konvex ist.

Aufgabe 10.3. Berechnen Sie $\mathrm{V@R}_\lambda$ und $\mathrm{AV@R}_\lambda$ für einen Pareto-verteilten Verlust ($\alpha > 1$) und das Verhältnis $\mathrm{AV@R}_\lambda/\mathrm{V@R}_\lambda$. Vergleichen Sie mit der Normalverteilung für $\lambda \to 0$.

Aufgabe 10.4 (vgl. [BOS25, Übung 6.28])**.** Bestimmen Sie die Funktion g, für die das absolute Abweichungsprinzip ein Wangsches Prinzip ist.

Kapitel 11

Rückversicherung und Erfahrungstarifierung

11.1 Risikoteilung und Rückversicherung

Ein Risiko $X \geq 0$ soll auf Versicherer $1, \dots, N$ mit Prämienprinzipien $\mathcal{H}_i$ aufgeteilt werden: $X = X_1 + \cdots + X_N$ mit $X_i \geq 0$, so dass $\sum_i \mathcal{H}_i(X_i)$ minimal ist. Verwenden alle dasselbe stark subadditive Prinzip (etwa AV@R oder absolutes Abweichungsprinzip), lohnt sich keine Teilung: Einer übernimmt alles. Eine Teilung kann sich also nur lohnen, wenn die Prinzipien verschieden oder nicht stark subadditiv sind.

Bei der *Rückversicherung* ($N = 2$) teilt der Erstversicherer $X = X_E + X_R$ in Selbstbehalt und Anteil des Rückversicherers:

- *Quotenrückversicherung*: $X_R = (1-p)X$, $X_E = pX$, $p \in [0,1]$;
- *Schadenexzedentenrückversicherung* (Excess-of-Loss; auf den Gesamtschaden angewandt Stop-Loss): $X_R = (X-r)^+$, $X_E = \min\{X, r\}$ mit Priorität $r > 0$.

Beim Schadenexzedentenvertrag ist $\mathbb{E}[X_R] = e(r)\overline{F}(r)$ (Lemma 9.8); seine Bepreisung ist eine Anwendung der POT-Methode.

Beispiel 11.1 (Quote mit Exponentialprinzip, [BOS25, Bsp. 6.48])**.** Sei $X \sim \mathrm{Exp}(\lambda)$, und beide Versicherer verwenden das Exponentialprinzip mit $a > 0$. Für $ab < \lambda$ ist $\mathcal{H}(bX) = (1/a)\log(\lambda/(\lambda - ab))$. Die Gesamtprämie

$$\frac{1}{a}\log\frac{\lambda}{\lambda - ap} + \frac{1}{a}\log\frac{\lambda}{\lambda - a(1-p)}$$

ist bei $p = 1/2$ minimal und dort gleich $(2/a)\log(\lambda/(\lambda - a/2))$, endlich für $a < 2\lambda$. Ohne Teilung ist die Prämie nur für $a < \lambda$ endlich. Für $\lambda \leq a < 2\lambda$ ist das Risiko also *nur* mit Rückversicherung versicherbar – möglich, weil das Exponentialprinzip nicht stark subadditiv ist.

11.2 Das heterogene Modell

Ein Portfolio ist selten homogen: Fahrer fahren unterschiedlich, Gebäude sind unterschiedlich gefährdet. Man beobachtet für Police i die Schäden $X_{i,1}, \dots, X_{i,n}$ der vergangenen Jahre und möchte den Schaden $X_{i,0}$ des kommenden Jahres prognostizieren.

Definition 11.2 (Heterogenes Modell, [BOS25, Def. 6.49])**.** Jede Police i hat einen zufälligen *Risikoparameter* Y_i mit Verteilung ν. Die Paare $(Y_i, (X_{i,j})_{j\geq 0})$ sind u. i. v., und gegeben $Y_i = \theta$ sind $X_{i,0}, X_{i,1}, \dots$ u. i. v. mit Verteilung $\mathcal{P}_\theta$.

Wir lassen den Index i weg. Unbedingt sind die X_j *nicht* unabhängig: Vergangene Schäden enthalten Information über Y und damit über X_0.

Beispiel 11.3 (Mischung erzeugt schwere Tails, [BOS25, Bsp. 6.51])**.** Ist $Y \sim \mathrm{Exp}(\alpha)$ und gegeben $Y = \theta$ die Schadenhöhe $\mathrm{Exp}(\theta)$, so ist

$$\mathbb{P}(X_0 > x) = \int_0^\infty \mathrm{e}^{-\theta x}\alpha\mathrm{e}^{-\alpha\theta}\,\mathrm{d}\theta = \frac{\alpha}{\alpha + x} = \Big(1 + \frac{x}{\alpha}\Big)^{-1} :$$

eine Lomax-Verteilung mit Index 1, also ein extrem schwerer Tail ($\xi = 1$) aus lauter exponentiellen Einzelrisiken. Heterogenität ist eine wichtige Quelle schwerer Tails.

Abbildung 11.1 zeigt diesen Effekt.

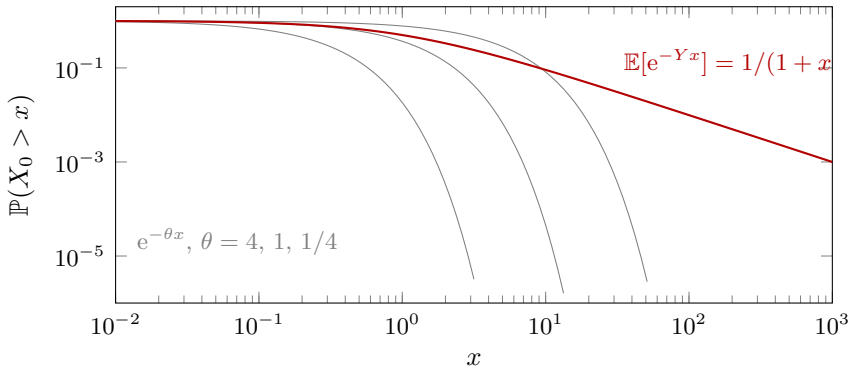

Abbildung 11.1: Heterogenität erzeugt schwere Tails: Gegeben $Y = \theta$ ist der Schaden $\mathrm{Exp}(\theta)$-verteilt mit leichtem Tail (grau, drei Werte von θ). Mischt man über $Y \sim \mathrm{Exp}(1)$, so entsteht der Lomax-Tail $(1+x)^{-1}$ (rot), der in doppelt logarithmischer Darstellung eine Gerade ist – Policen mit kleinem θ, also großen mittleren Schäden, dominieren den Tail.

Beispiel 11.4 (Poisson–Gamma, [BOS25, Bsp. 6.52])**.** Seien $Y \sim \Gamma(\gamma, \beta)$ und die Schadenanzahlen gegeben $Y = \theta$ Poisson(θ)-verteilt. Nach Bayes ist die a-posteriori-Verteilung von Y gegeben $X_1, \dots, X_n$ wieder Gamma mit Parametern $(\gamma + \sum_j X_j,\ \beta + n)$. Die *Bayes-Prämie* ist daher

$$\mathbb{E}[X_0 | X_1, \dots, X_n] = \mathbb{E}[Y | X_1, \dots, X_n] = \frac{\gamma + \sum_j X_j}{\beta + n} = \frac{n}{n+\beta}\,\bar{X} + \frac{\beta}{n+\beta} \cdot \frac{\gamma}{\beta},$$

ein gewichtetes Mittel aus individueller Erfahrung $\bar{X}$ und Kollektivmittel $\mathbb{E}[Y] = \gamma/\beta$. Unbedingt ist X_0 negativ binomial (Beispiel 5.4); Bichsel (1964) zeigte, dass dies Kfz-Daten weit besser beschreibt als ein Poisson-Modell.

11.3 Credibility nach Bühlmann

Meist ist die bedingte Erwartung nicht explizit berechenbar. Man sucht dann die beste *lineare* Prognose: die orthogonale Projektion von X_0 in L^2 auf $\mathrm{span}\{1, X_1, \dots, X_n\}$.

Satz 11.5 (Bühlmann 1967)**.** *Im heterogenen Modell mit* $\mathbb{E}[X_0^2] < \infty$ *seien* $\mu = \mathbb{E}X_0$, $A := \mathrm{Var}\big[\mathbb{E}[X_0|Y]\big] > 0$ *und* $k := \big(\mathrm{Var}\, X_0 - A\big)/A = \mathbb{E}[\mathrm{Var}(X_0|Y)]/A$. *Die beste lineare Prognose von* X_0 *aus* $X_1, \dots, X_n$ *ist*

$$\widetilde{X}_0 = \frac{n}{n+k}\,\bar{X} + \frac{k}{n+k}\,\mu, \qquad \bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j.$$

Beweis. Schritt 1: Kovarianzen. Mit Turmeigenschaft und bedingter Unabhängigkeit ist für $j \neq l$

$$\mathrm{Cov}(X_j, X_l) = \mathbb{E}\big[\mathbb{E}[X_j - \mu|Y]\,\mathbb{E}[X_l - \mu|Y]\big] = \mathrm{Var}\big[\mathbb{E}[X_0|Y]\big] = A,$$

und $\mathrm{Var}\, X_j = \mathrm{Var}\, X_0$. Daher $\mathrm{Cov}(X_0, \bar{X}) = A$ und $\mathrm{Var}\,\bar{X} = (1/n)\,\mathrm{Var}\, X_0 + ((n-1)/n)A$.

Schritt 2: Projektion auf $\mathrm{span}\{1, \bar{X}\}$. Das Minimum von $\mathbb{E}[(X_0 - a - b\bar{X})^2]$ wird bei $b = \mathrm{Cov}(X_0, \bar{X})/\,\mathrm{Var}\,\bar{X}$, $a = (1-b)\mu$ angenommen. Mit Schritt 1 ist

$$b = \frac{A}{\frac{1}{n}\,\mathrm{Var}\, X_0 + \frac{n-1}{n}A} = \frac{nA}{\mathrm{Var}\, X_0 - A + nA} = \frac{n}{n+k}.$$

Schritt 3: Es genügt $\bar{X}$. Für $\check{X}_0 = (1-b)\mu + b\bar{X}$ und $j \geq 1$ ist

$$\mathbb{E}[(X_0 - \check{X}_0)X_j] = \mathrm{Cov}(X_0, X_j) - \frac{b}{n}\sum_{l=1}^{n}\mathrm{Cov}(X_l, X_j) = A - \frac{b}{n}\big((n-1)A + \mathrm{Var}\, X_0\big) = 0,$$

und $\mathbb{E}[X_0 - \check{X}_0] = 0$. Also ist $X_0 - \check{X}_0$ orthogonal zu $\mathrm{span}\{1, X_1, \dots, X_n\}$, und $\check{X}_0 = \widetilde{X}_0$. □

Der *Credibility-Faktor* $z = n/(n+k) \in [0,1]$ wächst mit der Anzahl der Beobachtungen und mit der Heterogenität A des Portfolios, und er fällt mit der Streuung $\mathbb{E}[\operatorname{Var}(X_0|Y)]$ innerhalb einer Police. Im Poisson–Gamma-Modell ist $A = \operatorname{Var} Y = \gamma/\beta^2$ und $\mathbb{E}[\operatorname{Var}(X_0|Y)] = \mathbb{E}Y = \gamma/\beta$, also $k = \beta$: Die lineare Prognose stimmt mit der Bayes-Prämie überein. In der Praxis schätzt man μ, A und k aus dem Portfolio (empirische Bayes-Methode).

Abbildung 11.2 zeigt, wie der Credibility-Faktor von n und k abhängt.

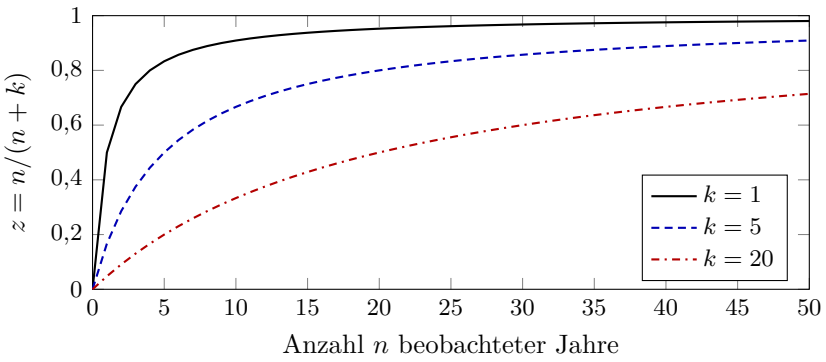

Abbildung 11.2: Credibility-Faktor z als Gewicht der eigenen Schadenerfahrung $\bar{X}$ in Bühlmanns Prognose. Je kleiner $k = \mathbb{E}[\operatorname{Var}(X_0|Y)]/\operatorname{Var}(\mathbb{E}[X_0|Y])$, d. h. je heterogener das Portfolio im Verhältnis zur Streuung innerhalb einer Police, desto schneller vertraut man der individuellen Erfahrung. Bei $n = k$ erhalten Individual- und Kollektivmittel gleiches Gewicht.

Aufgaben

Aufgabe 11.1. Leiten Sie die a-posteriori-Verteilung im Poisson–Gamma-Modell her.

Aufgabe 11.2. Berechnen Sie den Credibility-Faktor im Modell $Y \sim \Gamma(\gamma, \beta)$, $X|Y = \theta \sim \operatorname{Exp}(\theta)$ mit $\gamma > 2$. Ist hier die lineare Prognose gleich der Bayes-Prämie?

Aufgabe 11.3. Bestimmen Sie für $X \sim \operatorname{Lomax}(\kappa, \alpha)$ die Netto-Rückversicherungsprämie eines Schadenexzedentenvertrags mit Priorität r und deren Abhängigkeit von α.

Rückblick

- *Lebensversicherung:* Bewertung über Barwerte, Äquivalenzprinzip, Deckungskapital (Thiele), Verlust als Martingal mit expliziter Varianz (Hattendorff).
- *Statische Modelle:* Gesamtschaden als zusammengesetzte Verteilung; Wald, Cantelli, Cramér; exakte Berechnung mit Panjer.
- *Ruintheorie:* Erneuerungsgleichungen als Werkzeug. Unter der Lundberg-Bedingung: $\Psi(u) \sim C\mathrm{e}^{-r_0 u}$ für $u \to \infty$, Ruin durch viele kleine Schäden. Subexponentielle Tails: $\Psi(u) \sim \overline{F}_I(u)/\rho$ für $u \to \infty$, Ruin durch einen großen Schaden.
- *Extremwerttheorie:* Maxima konvergieren gegen die GEV, Exzesse gegen die GPD; der Extremwertindex ξ steuert den Tail und wird mit POT- oder Hill-Methoden geschätzt.
- *Bewertung von Risiken:* Prämienprinzipien und Risikomaße; AV@R ist kohärent, V@R nicht; Rückversicherung und Credibility.

Der rote Faden ist das Verhalten im Tail: Es entscheidet über die Form der Ruinwahrscheinlichkeit, über die Versicherbarkeit unter bestimmten Prämienprinzipien und über die Wahl geeigneter statistischer Methoden.

Anhang A

Hilfsmittel aus der Maß- und Wahrscheinlichkeitstheorie

A.1 Lebesgue–Stieltjes-Integrale

Zu jeder wachsenden, rechtsstetigen Funktion $G\colon [0,\infty) \to \mathbb{R}$ gibt es genau ein Maß μ_G auf $[0,\infty)$ mit $\mu_G([0,t]) = G(t) - G(0-)$, wobei wir $G(0-) := 0$ setzen; ein Wert $G(0) > 0$ erzeugt also ein Atom in 0. Man schreibt $\int g \,\mathrm{d}G := \int g \,\mathrm{d}\mu_G$. Ist G eine Differenz zweier solcher Funktionen, so integriert man gegen beide getrennt und subtrahiert.

Spezialfälle. Für eine Treppenfunktion $G(t) = \sum_j z_j \mathbb{1}\{t \ge t_j\}$ ist $\int_{[0,t]} g \,\mathrm{d}G = \sum_{t_j \le t} g(t_j) z_j$. Ist G *absolut stetig*, d. h. $G(t) = G(0) + \int_0^t g(s)\,\mathrm{d}s$ mit integrierbarem g, so ist G fast überall differenzierbar mit Ableitung g (Hauptsatz für absolut stetige Funktionen), und $\int_{[0,t]} h \,\mathrm{d}G = h(0)G(0) + \int_0^t h(s)g(s)\,\mathrm{d}s$.

Partielle Integration. Für wachsende, rechtsstetige F, G und $0 \le s < t$ gilt

$$F(t)G(t) - F(s)G(s) = \int_{(s,t]} F(u-)\,\mathrm{d}G(u) + \int_{(s,t]} G(u)\,\mathrm{d}F(u).$$

Die Asymmetrie ($F(u-)$ gegen $G(u)$) erfasst die gemeinsamen Sprungstellen.

Fubini für Stieltjes-Integrale. Für messbares $f \ge 0$ und eine Zufallsvariable Y gilt $\mathbb{E}\big[\int_{[0,Y)} f\,\mathrm{d}G\big] = \int_{[0,\infty)} f(s)\mathbb{P}(Y > s)\,\mathrm{d}G(s)$; das wird in den Beweisen von Lemma 2.8 und Lemma 3.2 verwendet.

A.2 Bedingte Erwartungen

Sei $X \in L^1$ und $\mathcal{G} \subset \mathcal{F}$ eine Teil-σ-Algebra. Die bedingte Erwartung $\mathbb{E}[X|\mathcal{G}]$ ist die (f. s. eindeutige) $\mathcal{G}$-messbare Zufallsvariable Z mit $\mathbb{E}[Z\mathbb{1}_B] = \mathbb{E}[X\mathbb{1}_B]$ für alle $B \in \mathcal{G}$. Es gelten:

(a) Linearität und Monotonie;

(b) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}X$;

(c) Turmeigenschaft: $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$ für $\mathcal{H} \subset \mathcal{G}$;

(d) Herausziehen: $\mathbb{E}[ZX|\mathcal{G}] = Z\mathbb{E}[X|\mathcal{G}]$ für $\mathcal{G}$-messbares Z mit $ZX \in L^1$;

(e) Jensen: $\varphi(\mathbb{E}[X|\mathcal{G}]) \le \mathbb{E}[\varphi(X)|\mathcal{G}]$ für konvexes φ;

(f) Unabhängigkeit: Ist X unabhängig von $\mathcal{G}$, so ist $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}X$;

(g) Varianzzerlegung: $\operatorname{Var} X = \mathbb{E}[\operatorname{Var}(X|\mathcal{G})] + \operatorname{Var}(\mathbb{E}[X|\mathcal{G}])$.

In L^2 ist $\mathbb{E}[X|\mathcal{G}]$ die orthogonale Projektion auf $L^2(\mathcal{G})$, also die beste $\mathcal{G}$-messbare Prognose im quadratischen Mittel – das ist der Ausgangspunkt der Credibility-Theorie (Kapitel 11).

Literaturverzeichnis

[AA10] S. Asmussen and H. Albrecher, *Ruin Probabilities*, 2. ed., World Scientific, 2010.

[BOS25] J. Blath, M. Ortgiese, and M. Scheutzow, *Stochastische Modelle der Versicherungsmathematik*, Mathematik Kompakt, Birkhäuser, 2025.

[Col01] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer, 2001.

[dHF06] L. de Haan and A. Ferreira, *Extreme Value Theory. An Introduction*, Springer, 2006.

[EKM97] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer, 1997.

[Fel71] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2. ed., vol. II, Wiley, 1971.

[FS04] H. Föllmer and A. Schied, *Stochastic Finance*, 2. ed., de Gruyter, 2004.

[Ger97] H. U. Gerber, *Life Insurance Mathematics*, 3. ed., Springer, 1997.

[MFE15] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management*, revised ed., Princeton University Press, 2015.

[MH99] H. Milbrodt and M. Helbig, *Mathematische Methoden der Personenversicherung*, de Gruyter, 1999.

[Mik09] T. Mikosch, *Non-Life Insurance Mathematics*, 2. ed., Springer, 2009.

[Res87] S. I. Resnick, *Extreme Values, Regular Variation, and Point Processes*, Springer, 1987.

[RSST99] T. Rolski, H. Schmidli, V. Schmidt, and J. Teugels, *Stochastic Processes for Insurance and Finance*, Wiley, 1999.

[Sch09] K. D. Schmidt, *Versicherungsmathematik*, 3. ed., Springer, 2009.