



TECHNISCHE
UNIVERSITÄT
WIEN

Vienna University of Technology

DIPLOMARBEIT

**Prämienkalkulationsprinzipien
für heavy-tailed Risiken**

ausgeführt am Institut für
Wirtschaftsmathematik
der Technischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Peter Grandits

durch

Lie-Ying Chen
Althanstraße 27/15
1090 Wien

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Wien, im August 2015

Danksagung

Ich möchte mich an dieser Stelle gerne bei all jenen Personen bedanken, die mir während meiner Studienzeit und bei der Erstellung der Diplomarbeit zur Seite standen.

Vor allem gilt mein Dank an Herrn Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Peter Grandits, der mich als Betreuer meiner Diplomarbeit mit seinen Anregungen und seinen wegweisenden und konstruktiven Vorschlägen sehr hilfreich unterstützte.

Ein großes Dankeschön möchte ich außerdem meiner Familie aussprechen, die mir immer schon bei jeder Situation meines bisherigen Werdegangs beistand. Ganz besonderer Dank gebührt meinen Eltern und meinem Freund Matthias, die stets an mich geglaubt haben und mich immer wieder dazu motiviert haben, die Diplomarbeit voranzutreiben.

Nicht zu guter Letzt möchte ich mich vor allem bei meinen Studienkollegen Clemens, Harald, Lukas und Thomas bedanken, mit denen ich neben der ausgezeichneten Zusammenarbeit während des Studiums auch sehr viele lustige Stunden verbracht habe.

Vielen Dank!

Inhaltsverzeichnis

1	Einleitung	4
2	Bekannte Prämienkalkulationsprinzipien	7
2.1	Gütekriterien	7
2.2	Prämienkalkulationsprinzipien	10
2.2.1	Nettoprämienprinzip	10
2.2.2	Erwartungswertprinzip	11
2.2.3	Varianzprinzip	11
2.2.4	Standardabweichungsprinzip	11
2.2.5	Quantilsprinzip	12
2.2.6	Nullnutzenprinzip	13
2.2.7	Exponentialprinzip	14
2.2.8	Verlustfunktionenprinzip	14
2.3	Eigenschaften der Prämienkalkulationsprinzipien	15
3	Extremwerttheorie	17
3.1	Maxima	18
3.1.1	Verallgemeinerte Extremwertverteilung	18
3.1.2	Maximaler Anziehungsbereich	20

3.1.3	Maxima von streng stationären Zeitreihen	22
3.1.4	Die Block Maxima Methode	25
3.1.5	Return-Level und Return-Periode	26
3.2	Schranken-Überschreitungen	30
3.2.1	Verallgemeinerte Paretoverteilung	30
3.2.2	Modellierung von Tails und Messgrößen von Tail-Risiken	35
3.2.3	Die Hill-Methode	38
3.3	Tails aus speziellen Modellen	43
3.3.1	Fréchet-Anziehungsbereich	43
3.3.2	Gumbel-Anziehungsbereich	44
3.3.3	Misch-Modelle	46
4	Prämienkalkulation für heavy-tailed Risiken	50
4.1	Thin-tailed Risiken und das Exponentialprinzip	51
4.2	Heavy-tailed Risiken und das Potenzprinzip	53
4.2.1	Pareto- und Verallgemeinerte Pareto-Verteilung	53
4.2.2	Typ I Extremwert - Fréchet-Verteilung	54
4.2.3	Relation zwischen Verallgemeinerte Pareto- und Fréchet-Prämien	54
4.2.4	Das Potenzprinzip	54
4.3	Risikoaverse Prämien	56
4.3.1	Prämien für große heavy-tailed Schäden	56
4.3.2	Total erwartete Schäden und größt-erwartete Schäden	59
4.4	Risikoaverse Prämien beim ungewissen Tail-Index	62
4.4.1	Heavy-tailed Schadensmodelle	62

4.4.2	Ein Modell bei Ungewissheit von Tail-Heaviness	62
4.4.3	Eigenschaft von unsicheren Prämien	63
4.4.4	Relations-Struktur der unsicheren Prämie $\mathcal{P}(v, \beta)$	65
4.5	Berechnung der Prämie	67
4.5.1	Schätzungen des Tail-Index von Extrem-Relationen	69
4.5.2	Maximum Likelihood auf Basis der k -größten Extremas	70
4.5.3	Ergebnisse des Bedingten Maximum Likelihood-Schätzers (CMLE)	71
4.5.4	Vergleich zwischen Relationen und MLE-Ergebnisse	72
4.5.5	Prämien für Extrem-Schäden	73
4.6	Vergleich mit klassischen Prämienkalkulationsprinzipien	75
Anhang		81
	Berechnung von CMLE	81

Kapitel 1

Einleitung

Eine Prämie ist als jene Zahlung definiert, welche im Versicherungsvertrag vereinbart ist und welche der Versicherungsnehmer nach Abschluss des Versicherungsvertrages dazu verpflichtet diese in der vereinbarten Prämienzahlungsdauer zur Deckung des vereinbarten versicherten Risikos zu tilgen.

Dieses oft in der Praxis auch genannte **versicherungstechnische Risiko** kann in mehrere Aspekte aufgegliedert werden:

Zufallsrisiko: Die Schadenshöhe S ist zufällig und damit risikobehaftet.

Schätzrisiko: Hier ist die Verteilungsfunktion bzw. die Momente von S nicht bekannt, sondern müssen aus Daten geschätzt werden.

Prognoserisiko: Die vorhandenen Daten werden aus vergangenen Perioden entnommen. Würde man S aus den Daten exakt schätzen können, dann hätte man nur das S aus der letzten Periode bestimmt. Zufallsmäßigkeiten können sich aber jederzeit aufgrund von Inflation, Gesetzesänderungen oder technischen Errungenschaften ändern

In der Versicherungsmathematik ist die Berechnung einer gerechten Prämie für die Versicherung eines Risikos einer der wichtigsten Aufgaben.

Sachversicherungsverträge werden im Gegensatz zu Lebensversicherungsverträgen oft nur über kürzere Laufzeiten abgeschlossen, weswegen hier in erster Näherung die Zinsproblematik vernachlässigt werden kann. Des Weiteren werden die in der Praxis ebenso anfallenden Prämienanteile betreffend der kalkulatorischen Kosten nicht berücksichtigt.

Ein naheliegender Ausgangspunkt für die Festlegung der Risikoprämie ist das sogenannte *Nettoprämienprinzip* (oder auch *Äquivalenzprinzip* genannt), bei dem die Prämie durch den Erwartungswert des Verlustes bestimmt wird. Das Äquivalenzprinzip verlangt die faire Prämie.

Zum Einstieg werden im Kapitel 2 einige bisher bekannte theoretische Prämienkalkulationsprinzipien und deren Gütekriterien aus dem 3. Kapitel des Skriptums „Sachversicherungsmathematik“ von A. Zwirchmayr und P. Grandits [1] vorgestellt.

Im dritten Kapitel werden einige wichtige Eigenschaften, Definitionen und Sätze der Extremwerttheorie zusammengefasst. Der Aufbau und der Inhalt dieses Kapitels basiert auf dem siebten Kapitel des Buches „Quantitative Risk Management: Concepts, Techniques and Tools“ von Alexander J. McNeil et al. [2].

Die Verallgemeinerte Extremwertverteilung GEV spielt in der Theorie der Extreme eine sehr wichtige Rolle. Die klassische Extremwerttheorie beschäftigt sich mit Grenzverteilungen für normierte Maxima. Die einzig möglichen nicht-entarteten Grenzverteilungen für normierte Maxima sind in der GEV-Familie zu finden. Ferner wird anhand des Maximalen Anziehungsbereichs MDA gezeigt, dass die Fréchet-, Gumbel- und Weibull Verteilung die möglichen Typen von Extremwertverteilungen sind.

Viele Techniken der Extremwertstatistik, die für unabhängige Daten entwickelt wurden, lassen sich auf Zeitreihen übertragen. Um dabei den Extremalindex zu schätzen, gibt es mehrere Methoden. Wir beschreiben in diesem Kapitel die Grundzüge der bekannten theoretische Schätzmethode - die Block-Maximum Methode.

Jedoch erweist sich die Block-Maximum-Methode als eher „verschwenderisch“ im Umgang mit Datensätzen. In der Praxis wird diese Schätzmethode durch Methoden unter Verwendung von Schranken-Überschreitungen abgelöst.

Das haupt-verbreitete Verteilungsmodell für Schranken-Überschreitungen ist die Verallgemeinerte Paretoverteilung GPD. Das Pickands-Balkema-de Haan-Theorem besagt, dass die Verteilung von Exzessen über eine ausreichend hohe Schranke im Fall von Wahrscheinlichkeitsverteilungen, die im Anziehungsbereich von GEV liegen, durch eine GPD approximiert werden kann.

Als Übergang zum vierten Kapitel werden Methoden zur Schätzung von Tails und Modellierung von Tails sowie deren Messgrößen vorgestellt. Neben der GPD-Methode gibt es noch andere Möglichkeiten die Tails von Verteilungen, vor allem den Tail-Index zu schätzen. Die Grundzüge der Hill-Methode zur Modellierung von Tails von heavy-tailed Verteilungen werden im dritten Kapitel beschrieben. Ebenso werden Tails aus speziellen Modellen im Ende des dritten Kapitels vorgestellt.

Das vierte Kapitel umfasst das eigentliche Thema dieser Diplomarbeit. Dafür werden die Inhalte des Papers „Premium Calculation for Fat-tailed Risk“ von Roger Gay [6] zusammengefasst.

Wenn Versicherungsschäden einer heavy-tailed Verteilung zugrundeliegen, dann verursacht dies erhebliche Fehler über den Wert des Tail-Indexes. Analog zum Exponentialprinzip für thin-tailed Schäden wird das Potenzprinzip für heavy-tailed Schäden vorgestellt.

Ferner wird beschrieben, dass die Prämie bei Schäden, die einer Verallgemeinerten Pareto-Verteilung zugrundeliegen, als das Verhältnis der zwei größten erwarteten Schäden definiert ist. Während thin-tailed Schadens-Prämien weitgehend durch die ersten zwei Momente der Schadensverteilung bestimmt sind, werden heavy-tailed Schadens-Prämien durch die ersten zwei Extremas bestimmt.

Kapitel 2

Bekannte Prämienkalkulationsprinzipien

Zunächst beschäftigen wir uns mit bisher bekannten theoretischen Prämienkalkulationsprinzipien zur Bestimmung der Risikoprämie, wobei die wünschenswerten Eigenschaften dieser Prinzipien auch nicht zu vergessen sind.

Das nachstehende Kapitel basiert auf dem dritten Kapitel des Skriptums „Sachversicherungsmathematik“ von Alexander Zwirchmayer und Peter Grandits [1, Kapitel 3.1].

Zunächst sei \mathcal{H} ein **Prämienkalkulationsprinzip** als Funktional auf dem Raum der Schadensvariablen. Jedem Risiko S wird eine Prämie

$$\mathcal{P} = \mathcal{H}(S) \tag{2.1}$$

zugeordnet.

Zunächst zeigen wir einige Gütekriterien für $\mathcal{H}(S)$, bevor wir einzelne Beispiele für \mathcal{H} und somit Prämienkalkulationsmethoden vorstellen.

2.1 Gütekriterien

Welche Eigenschaften eines Prämienkalkulationsprinzips sind notwendig bzw. wünschenswert? Sowohl in der Theorie als auch in der Praxis sollen Prämienberechnungsprinzipien bestimmte Eigenschaften aufweisen. Natürlich gibt es keine Berechnungsmethode, welche in allen Situationen die optimale Lösung verspricht. Man kann jedoch für ein konkretes Risiko spezielle Eigenschaften an die Prämie verlangen und damit die Auswahl der

Möglichkeiten einschränken.

Zunächst muss definiert sein, dass das zu versichernde Risiko „versicherbar“ ist. Somit sind Risiken gemeint, die vor allem nicht konstant sind, da ansonsten der Abschluss des Versicherungsvertrages nicht sinnvoll ist. Ist nun ein Risiko X nicht konstant, dann muss ferner gelten, dass $\mathbb{E}(X) > 0$ und $\text{Var}(X) > 0$ sind. Dazu würden naheliegend auch die folgenden Bedingungen gelten: $\mathbb{E}(X) < \infty$ und $\text{Var}(X) < \infty$.

Definition 2.1. *Ein Risiko X heißt versicherbar, wenn $0 < \mathbb{E}(X) < \infty$ und $0 < \text{Var}(X) < \infty$ gelten.*

Definition 2.2. *Sei \mathcal{H} ein Prämienkalkulationsprinzip und seien S, S_1, S_2, \dots jeweils beliebige und von einander unabhängige Risiken.*

(a) \mathcal{H} ist maximalschadenbegrenzt, wenn gilt:

$$\mathcal{H}(S) \leq \inf \{x \in \mathbb{R} : P(X \geq x) = 1\} \quad (2.2)$$

Bei einem maximalschadenbegrenzten Prämienkalkulationsprinzip ist die Prämie stets kleiner oder gleich der größtmöglichen Entschädigungsleistung des Versicherers. Würde die Maximalschadenbegrenztheit nicht gelten, dann würde es für den Versicherungsnehmer keinen Sinn ergeben, eine derartige Versicherung abzuschließen. Es wird kein Versicherungsnehmer dazu bereit sein, eine Prämie zu zahlen, welche die maximale Schadensleistung übersteigt.

(b) \mathcal{H} ist erwartungswertübersteigend, wenn gilt:

$$\mathcal{H}(S) \geq \mathbb{E}S \quad (2.3)$$

Die Eigenschaft eines Prämienkalkulationsprinzips erwartungswertübersteigend zu sein, ist eine offensichtliche Forderung. Würde die Prämie unter dem zu erwarteten Schaden liegen, dann ist für den Versicherer der Ruin auf längere Sicht nur schwer vermeidbar.

(c) \mathcal{H} ist translationsinvariant, wenn für alle $c \in \mathbb{R}_+$ gilt:

$$\mathcal{H}(S + c) = \mathcal{H}(S) + c \quad (2.4)$$

Erhöht sich das Risiko um eine konstante Leistung der Höhe c , dann steigt zugleich die Prämie um denselben Betrag.

(d) \mathcal{H} ist positiv homogen, wenn für alle $c \in \mathbb{R}_+$ gilt:

$$\mathcal{H}(cS) = c\mathcal{H}(S) \quad (2.5)$$

Die Homogenität wird oft für bestimmte Werte von c verlangt. Als Beispiel könnte man die Quotenversicherung bei der Rückversicherung betrachten, für die $c \in [0; 1]$ als ausreichend erscheint.

Diese Eigenschaft verhindert eine Arbitragemöglichkeit, denn wenn die Prämie von cS größer als das c -fache von S wäre, dann würde man S c -mal versichern und dadurch einen Gewinn erzielen. Dies gilt ebenso für den gegenteiligen Fall wenn $\mathcal{H}(cS) \leq c\mathcal{H}(S)$ ist. Solche Risiken cS mit $c \gg 1$ würden relativ gesehen im Schadensfall hohe Verluste für den Versicherer verursachen.

(e) \mathcal{H} ist additiv, wenn gilt:

$$\mathcal{H}(S_1 + S_2) = \mathcal{H}(S_1) + \mathcal{H}(S_2) \quad (2.6)$$

Die Gesamtprämie für unabhängige Risiken soll gleich der Summe aus den Einzelprämien entsprechen. Wie bei der Eigenschaft der Homogenität liegt bei nichtzutreffender Additivität eine theoretische Arbitragemöglichkeit vor.

(f) \mathcal{H} ist subadditiv, wenn gilt:

$$\mathcal{H}(S_1 + S_2) \leq \mathcal{H}(S_1) + \mathcal{H}(S_2) \quad (2.7)$$

Die Subadditivität ist auch eine offensichtliche Konsequenz des Ausgleichs im Kollektiv. Durch das Zusammenfassen mehrerer Risiken zu einem Gesamtbestand soll nicht mehr Prämie gezahlt werden. Liegt die Subadditivität vor, dann kann ein hohes Risiko nicht für eine geringere Prämie versichert werden, indem dieses als zwei oder mehreren unterteilten Risiken versichert wird.

(g) \mathcal{H} ist iterativ, wenn gilt:

$$\mathcal{H}(S) = \mathcal{H}(\mathcal{H}(S|X)) \quad (2.8)$$

Diese Eigenschaft wird durch die Iterativität des Erwartungswertes motiviert. $\mathcal{H}(\mathcal{H}(S|X))$ ist eine Kurzform dafür, dass zuerst für die bedingte Schadensvariable $S|(X = x)$ die Prämie $\mathcal{H}(S|(X = x)) =: g(X)$ bestimmt wird. Danach wird noch einmal das Prämienkalkulationsprinzip auf die Zufallsvariable $g(X)$ angewendet.

Nun betrachten wir im Folgenden einige Beispiele für \mathcal{H} :

2.2 Prämienkalkulationsprinzipien

2.2.1 Nettoprämienprinzip

Definition 2.3. Man bezeichnet das Nettoprämienprinzip auch als Äquivalenzprinzip, da hier als Prämie die erwartete Schadenshöhe verlangt wird:

$$\mathcal{P} = \mathbb{E}[S] \quad (2.9)$$

Motiviert wird dieser Zugang durch die theoretischen Überlegungen betreffend dem Ausgleich im Kollektiv (Zusammenfassung mehrerer Risiken zu einem Bestand. Dies führt zu einer Schadensverteilung, die im Allgemeinen gegenüber jeder Einzelverteilung vorteilhafter ist, da sich günstige und ungünstige Einzelrisikoverläufe ausgleichen können. - Näheres siehe [1, Seite 31]).

Mit dem Erwartungswert allein erzielt man jedoch nicht die passende Risikoprämie. Auch die Ruinwahrscheinlichkeit beträgt in diesem Fall fast sicher 1.

Dies erfordert nun zusätzlich **Schwankungs-** bzw. **Sicherheitszuschläge** zur Nettoprämie:

2.2.2 Erwartungswertprinzip

Definition 2.4. *Wir wählen den Schwankungszuschlag ähnlich wie in der Ruintheorie proportional zu $\mathbb{E}[S]$. $\forall \theta > 0$ gilt:*

$$\mathcal{P} = (1 + \theta)\mathbb{E}[S] = \mathbb{E}[S] + \theta\mathbb{E}[S] \quad (2.10)$$

Derartige Prämienkalkulationsprinzipien haben natürlich die Einfachheit, dass nur die Kenntnis über $\mathbb{E}[S]$ ausreicht um die Risikoprämie zu berechnen. Diese Einfachheit ist jedoch zugleich auch ein Nachteil, da $\mathbb{E}[S]$ kein Risikomaß darstellt und daher die alleinige Abhängigkeit aus langfristiger Sicht zu „gefährlich“ wird, zumal die Variabilität des Risikos hier nicht berücksichtigt wird.

2.2.3 Varianzprinzip

Als Ergänzung zum Erwartungswertprinzip haben wir das Varianzprinzip. Hier wird die Variabilität des Risikos anhand der Varianz mitberücksichtigt:

Definition 2.5. *Für den Parameter $\alpha > 0$ erhalten wir:*

$$\mathcal{P} = \mathbb{E}[S] + \alpha\text{Var}[S] \quad (2.11)$$

2.2.4 Standardabweichungsprinzip

Würden wir im Gegensatz zum Varianzprinzip den Schwankungszuschlag proportional zur Standardabweichung wählen, so erhalten wir wie beim Nettoprämienprinzip eine Kalkulation für die Prämie mit der Einheit von S :

Definition 2.6. *Für den Parameter $\beta > 0$:*

$$\mathcal{P} = \mathbb{E}[S] + \beta\sqrt{\text{Var}[S]} \quad (2.12)$$

2.2.5 Quantilsprinzip

Bemerkung: Bevor wir zur eigentlichen Prämienformel des Quantilsprinzips kommen, müssen wir zunächst die Verteilung betrachten in welcher das Quantil bestimmt wird. Dazu trifft man eine pessimistische Verteilungsannahme, ähnlich wie in der Lebensversicherung im Vergleich der Rechnungsgrundlagen erster Ordnung mit jenen der zweiten Ordnung. Wir ersetzen zunächst das Risiko S durch ein Risiko \tilde{S} mit:

$$F_{\tilde{S}}(s) = \mathbb{P}(\tilde{S} \leq s) \leq \mathbb{P}(S \leq s) = F_S(s) \quad (2.13)$$

Zur Bestimmung der Prämie verlangen wir den Erwartungswert von \tilde{S} . Durch partielles Integrieren erhält man folgende Beziehung

$$\mathbb{E}[S] = \int_0^\infty 1 - F_S(s) ds \quad (2.14)$$

womit die Prämie wie folgt bestimmt wird:

$$\begin{aligned} \mathcal{P} &= \mathbb{E}[\tilde{S}] = \int_0^\infty 1 - F_{\tilde{S}}(s) ds \frac{\mathbb{E}[S]}{\mathbb{E}[S]} \\ &= \int_0^\infty 1 - F_S(s) + F_S(s) - F_{\tilde{S}}(s) ds \frac{\mathbb{E}[S]}{\int_0^\infty 1 - F_S(s) ds} \\ &= \mathbb{E}[S] \left(1 + \frac{\int_0^\infty F_S(s) - F_{\tilde{S}}(s) ds}{\int_0^\infty 1 - F_S(s) ds} \right) \end{aligned}$$

Somit haben wir ein Erwartungswertprinzip erhalten mit von F_S und $F_{\tilde{S}}$ abhängigen θ .

Definition 2.7. Um nun die Prämie mittels des Quantilsprinzips zu ermitteln, nehmen wir zunächst eine obere Schranke $\epsilon > 0$ für die Illiquiditätswahrscheinlichkeit $\mathbb{P}(P < S)$ an und erhalten dazu folgende Prämie:

$$\mathcal{P} = \inf \{P : \mathbb{P}(P < S) \leq \epsilon\} \quad (2.15)$$

Existiert nun das $(1 - \epsilon)$ -Quantil der Verteilung $F_S(s)$, so ist dies ident mit der gesuchten Prämie.

2.2.6 Nullnutzenprinzip

Beim Nullnutzenprinzip wählt das Versicherungsunternehmen seine Prämie gemäß einer **Nutzenfunktion** $u(\cdot)$, welcher den Nutzen für das Versicherungsunternehmen bei jeder Geldmenge $x \geq 0$ $u(x)$ darstellt mit:

- $u' > 0$: Je größer die Geldmenge ist, desto größer ist ihr Nutzen.
- $u'' \leq 0$: Bei wachsendem Vermögen wird dem Zuwachs der Geldeinheit ein immer geringer werdender Nutzen beigemessen.

Definition 2.8. Die dazugehörige **Risikoaversion** $r(x)$ ist wie folgt definiert:

$$r(x) := -\frac{u''(x)}{u'(x)} \quad (2.16)$$

Wird nun ausgehend von der Geldeinheit x , diese um einen kleinen Betrag dx erhöht, so wird die dazugehörige Risikoaversion umso größer, je kleiner der absolute Nutzenzuwachs mit $u(x + dx) - u(x) \approx u'(x)dx$ ausmacht. $r(x)$ wird ebenso umso größer, je stärker das Wachstum mit $|u'(x + dx) - u'(x)| \approx u''(x)dx$ abnimmt.

Definition 2.9. Sei nun x der Überschuß des Versicherungsunternehmens vor der Übernahme des Risikos und $x + \mathcal{P} - S$ jener danach.

Und betrachtet man die Nutzenfunktionen $u(x)$ und $u(x + \mathcal{P} - S)$, so stellt zweiteres im Gegensatz zu $u(x)$ eine Zufallsvariable dar, welche von S abhängt.

Nach der Theorie des Nullnutzenprinzips, errechnet sich hier die Prämie unter der Annahme, dass der Nutzen vor Versicherungsübernahme gleich dem erwarteten Nutzen nach der Versicherungsübernahme bleibt.

Somit lässt sich die Prämie nach dem Nullnutzenprinzip aus folgender Abhängigkeit ableiten:

$$u(x) = \mathbb{E}(u(x + \mathcal{P} - S)) \quad (2.17)$$

2.2.7 Exponentialprinzip

Man erhält hier die Prämie indem man die Nutzenfunktion wie folgt wählt:

$$u(x) = \frac{1 - e^{-ax}}{a} \quad (2.18)$$

mit $a > 0$ als die Risikoaversion von \mathcal{P} .

Definition 2.10. Die dazugehörige Prämie nach dem Exponentialprinzip ist gegeben durch:

$$\mathcal{P} = \frac{1}{a} \ln M_S(a) \quad (2.19)$$

mit:

$$\lim_{a \rightarrow 0^+} \mathcal{P} = E[S] \quad \text{und} \quad \lim_{a \rightarrow \infty} \mathcal{P} = \max S$$

Jede beliebige Prämie, die einen Wert zwischen $\mathbb{E}[S]$ und $F^{-1}(1) = \max S$ annimmt, kann gemäß Definition 2.10 bei geeigneter Wahl der Risikoaversion a nach dem Exponentialprinzip ermittelt werden. Die Prämienzahlung erhöht sich bei wachsender Risikoaversion a . Insbesondere erhält man für $a \rightarrow 0$ das *Nettoprämienprinzip*.

Falls $F^{-1}(1)$ endlich ist, dann beziffert dieser Wert den maximalen Gesamtschaden. Das Prämienprinzip

$$\mathcal{P} = F^{-1}(1)$$

wird als *Maximalschadenprinzip* bezeichnet.

2.2.8 Verlustfunktionenprinzip

Das Verlustfunktionenprinzip basiert auf der Minimierung des erwarteten Verlustes. Dazu führen wir eine **Verlustfunktion** $L(s, p)$ ein, als Verlust des Versicherungsunternehmens im Falle $S = s$ und $\mathcal{P} = p$. Man beachte auch, dass der Verlust in dem Fall ein (negatives) Maß für die Güte der Situation $S = s$ und $\mathcal{P} = p$ angibt und nicht notwendigerweise den Geldverlust des Versicherungsunternehmens im Falle von Ruin.

Die dazugehörige Risikofunktion erhalten wir durch den Erwartungswert von $L(S, p)$:

$$R(p) = \mathbb{E}L(S, p) \quad (2.20)$$

Definition 2.11. Die Prämie aus dem Verlustfunktionenprinzip erhält man somit durch das Minimieren der Risikofunktion:

$$\mathcal{P} = \min R(p) \tag{2.21}$$

2.3 Eigenschaften der Prämienkalkulationsprinzipien

Zum Abschluss dieses Kapitels fassen wir die im Abschnitt 2.1 genannten Gütekriterien zu den so eben vorgestellten Prämienkalkulationsprinzipien zusammen.

Prinzip	maximal-schaden-begrenzt	erwartungswertübersteigend	translationsinvariant	positiv-homogen	additiv	subadditiv	iterativ
Nettoprämienprinzip	ja	ja	ja	ja	ja	ja	ja
Erwartungswertprinzip	nein	ja	nein	ja	ja	ja	nein
Varianzprinzip	nein	ja	ja	nein	ja	nein	nein
Standardabweichungsprinzip	nein	ja	ja	ja	nein	ja	nein
Quantilsprinzip	ja	nein	ja	ja	nein	nein	nein
Nullnutzenprinzip	ja	ja	ja	nein	nein	nein	nein
Exponentialprinzip	ja	ja	ja	nein	ja	ja	ja

Tabelle 2.1: Zusammenfassung der Gütekriterien zu den Prämienkalkulationsprinzipien

Das Verlustfunktionenprinzip, welches etliche andere Prämienprinzipien als Spezialfälle umfasst, kann in Abhängigkeit von der konkreten Verlustfunktion sehr unterschiedliche Prämienbemessungen implizieren. Ohne Kenntnis der Verlustfunktion lassen sich bei diesem Prinzip keine allgemeinen Aussagen über die Gütekriterien dieses Prinzips treffen (vgl. [10, Seite 124]).

Die konkrete Auswahl eines Prämienkalkulationsprinzips muss nun durch einen Kompromiss zwischen den gewünschten Eigenschaften und obiger Tabelle erzielt werden. Von allen genannten Prämienkalkulationsprinzipien erfüllt nur das Nettoprämienprinzip alle

KAPITEL 2. BEKANNTE PRÄMIENKALKULATIONSPRINZIPIEN

genannten Gütekriterien. Dieses Prinzip hat aber den Nachteil, dass hierbei der Ruin relativ sicher eintritt und daher in der Praxis nicht von Gebrauch ist.

Sehr oft eignet sich das Exponentialprinzip zur Prämienberechnung, da hier nur die positive Homogenität nicht zutrifft. Spätestens aber im Kapitel 4 erweist sich auch dieses Prinzip als nicht geeignet für heavy-tailed Verteilungen, sondern lediglich für thin-tailed Risiken.

Kapitel 3

Extremwerttheorie

Im täglichen Leben sprechen wir immer wieder von Rekorden. Sie sind ein Phänomen, das uns schon seit etlicher Zeit verfolgt und interessiert. Seit die Menschheit existiert, sind wir ständig auf der Suche nach Rekorden - dem höchsten Berg, dem längsten Fluss, dem heißesten Tag oder der höchsten Flutwelle.

Rekorde sind aber auch von großer Wichtigkeit für praktische Belange. In der Wahrscheinlichkeitstheorie haben wir kennen gelernt, wie sich Summen von Zufallsvariablen verhalten. Dies hilft uns, etliche praktische Problemstellungen zu analysieren bzw. zu lösen.

Ebenso ist man am Extremalverhalten interessiert, wenn man die notwendige Sicherheitsreserve einer Bank oder eines Versicherungsunternehmens festlegen möchte. Denkt man an die jährlichen Schäden aus Naturkatastrophen wie Sturm bzw. Hurrikane, Tornados, Taifune, Hagel, Überschwemmungen, Erdbeben, Erdrutsch oder Lawinen, welche weltweit die Portfolios der Versicherungswirtschaft in erheblichem Maße belasten, so ist man gezwungen auch solche Rekorde zu analysieren bzw. diese für die weitere Bestimmung der notwendigen Sicherheitsreserve eines Versicherungsunternehmens zu berücksichtigen. Finanziell am Stärksten von Naturkatastrophen betroffen, sind vor allem die international operierenden großen Rückversicherer, die ihre anteiligen Schadenslasten vor allem durch die globale Diversifikation des Geschäfts kontrollierbar halten.

Eine möglichst genaue versicherungsmathematische Analyse und Modellierung von Katastrophenrisiken ist daher aus der Perspektive eines effizienten Risikomanagements sowohl für Erst- als auch für Rückversicherer von großem Interesse.

In der Praxis berechnen wir die maximalen Forderungen an ein Versicherungsunternehmen mit einem stochastischen Ansatz, die im Rahmen der Extremwertstatistik auf ökonomisch geeignet bereinigten aggregierten Jahresschadendaten beruhen. Wir wollen daher das Extremwertverhalten von stochastischen Prozessen untersuchen.

Das nachstehende Kapitel basiert auf dem siebten Kapitel des Buches „Quantitative Risk Management: Concepts, Techniques and Tools“ von Alexander J. McNeil et al. [2].

3.1 Maxima

Zunächst betrachten wir eine Folge von iid Zufallsvariablen X_1, \dots, X_n , welche die finanziellen Verluste repräsentieren. Diese können interpretiert werden als operative Verluste, Verluste aus Versicherungsschäden oder Verluste aus einem Kreditportfolio über fixe Zeitintervalle.

Später werden wir die Annahme der Unabhängigkeit belassen und betrachten die Zufallsvariablen als einen Teil von streng stationären Zeitreihen $(X_t)_{t \in \mathbb{Z}}$ der abhängigen Verluste. Sie könnten (negative) Renditen einer Investition in ein einziges Wertpapier oder ein Portfolio von Wertpapieren sein.

3.1.1 Verallgemeinerte Extremwertverteilung - GEV

Die Rolle der verallgemeinerten Extremwertverteilung in der Theorie der Extreme ist analog zu der Normalverteilung im Sinne des zentralen Grenzwertsatzes für Summen von Zufallsvariablen.

Die Standard-Version des zentralen Grenzwertsatz zeigt, wie Summen von iid Zufallsvariablen mit endlicher Varianz, wenn sie entsprechend normiert sind, in der Verteilung in eine Standardnormalverteilung konvergieren. Entsprechend normiert bedeutet, wenn man für alle n die Zufallsvariable $\frac{S_n - a_n}{b_n}$ betrachtet, wobei $a_n = n\mathbb{E}(X_1)$ und $b_n = \sqrt{n\text{Var}(X_1)}$ Terme in zwei normierte Folgen sind. Der zentrale Grenzwertsatz besagt dann:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{S_n - a_n}{b_n} \leq x \right) = \Phi(x), \quad x \in \mathbb{R}$$

Die klassische Extremwerttheorie beschäftigt sich mit Grenzverteilungen für normierte *Maxima* von iid $M_n = \max(X_1, \dots, X_n)$ mit der Annahme, dass die X_i 's absolut stetig sind. Die einzig möglichen nicht-entarteten Grenzverteilungen für normierte Maxima sind in der GEV-Familie zu finden.

Definition 3.1. (Verallgemeinerte Extremwertverteilung)

$$H_\xi(x) = \begin{cases} \exp(-(1 + \xi x)^{-1/\xi}) & \text{für } \xi \neq 0, \\ \exp(-e^{-x}) & \text{für } \xi = 0 \end{cases}$$

wobei $\forall x$ und ξ gilt $1 + \xi x > 0$. Die zugehörige drei-parametrig skalierte Familie $H_{\xi, \mu, \sigma}$ aus der verallgemeinerten Extremwertverteilung erhält man durch die Einführung eines Skalenparameters σ und eines Lageparameters μ und Substitution von $x := \frac{x - \mu}{\sigma}$, $\forall \mu \in \mathbb{R}$ und $\sigma > 0$. Als Definitionsbereich gilt nun in diesem Fall für alle x : $1 + \xi \frac{x - \mu}{\sigma} > 0$.

Der Parameter ξ ist der sogenannte Formparameter der verallgemeinerten Extremwertverteilung und H_ξ definiert den Typ der Verteilung. Die allgemein parametrisierte Form klassifiziert abhängig vom Wert von ξ drei Arten von Verteilungen, welche unter anderen Namen bekannt sind.

Somit entspricht die verallgemeinerte Extremwertverteilung offensichtlich für:

$\xi > 0$ dem Typ der Fréchet-Verteilung $\Phi_{\xi^{-1}}$
 $\xi < 0$ dem Typ der Weibull-Verteilung $\Psi_{-\xi^{-1}}$
 $\xi = 0$ dem Typ der Gumbel-Verteilung Λ .

Ferner gilt für:

$\xi > 0 : H_\xi(x) = \Phi_{\xi^{-1}}(1 + \xi x), \forall 1 + \xi x > 0$
 $\xi < 0 : H_\xi(x) = \Psi_{-\xi^{-1}}(-(1 + \xi x)), \forall 1 + \xi x > 0$
 $\xi = 0 : H_\xi(x) = \Lambda(x)$.

Für fixe x erhalten wir $\lim_{\xi \rightarrow \infty} H_\xi(x) = H_0(x)$ (beidseitiger Grenzwert), sodass die Parametrisierung in Definition 3.1. stetig in ξ ist, welches die Verwendung dieser Verteilung in einem statistischen Modell erleichtert.

Angenommen, die Folge der Maxima M_n , welche von iid Zufallsvariablen abgeleitet werden, konvergiert in eine Verteilung unter einer entsprechenden Normierung. Dann existieren Folgen von realen Konstanten d_n und $c_n > 0$, sodass

$$\lim_{n \rightarrow \infty} \mathbb{P}((M_n - d_n)/c_n \leq x) = \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x), \quad (3.1)$$

für eine nicht-entartete Verteilungsfunktion $H(x)$.

Definition 3.2. (Maximaler Anziehungsbereich MDA)

Wenn (3.1) gilt für nicht-entartete Verteilungsfunktionen H , dann befindet sich F im sogenannten Maximalen Anziehungsbereich von H , geschrieben $F \in MDA(H)$.

Der folgende Satz ist eines der Hauptresultate in der Extremwertverteilung und zeigt, dass wir mit der Fréchet-, Weibull- und Gumbel-Verteilung schon alle möglichen Typen von Extremwertverteilungen erhalten haben.

Satz 3.1. (ohne Beweis) (Fisher-Tippett, Gnedenko)

Wenn $F \in MDA(H)$ für nicht-entartete Verteilungsfunktionen H , dann muss H eine Verteilungsfunktion vom Typ H_ξ sein, somit eine GEV-Verteilung.

Das Grenzverhalten für die Konvergenz von Maxima kann ebenso übergeleitet werden in jene der Minima unter der Verwendung von

$$\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n). \quad (3.2)$$

Es ist nicht schwer zu erkennen, dass normierte Minima von iid Zufallszahlen mit Verteilungsfunktion F in der Verteilung konvergieren wird, wenn die Verteilungsfunktion $\tilde{F}(x) = 1 - F(-x)$, welche die Verteilungsfunktion von $-X_1, \dots, -X_n$ ist, enthalten ist im Maximalen Anziehungsbereich einer Extremwertverteilung.

Geschrieben $M_n^* = \max(-X_1, \dots, -X_n)$ und angenommen $\tilde{F} \in MDA(H_\xi)$, erhalten wir

$$\lim_{n \rightarrow \infty} \mathbb{P}((M_n^* - d_n)/c_n \leq x) = H_\xi(x),$$

wobei wir durch Einsetzen von (3.2) folgendes erhalten

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{\min(X_1, \dots, X_n) + d_n}{c_n} \leq x\right) = 1 - H_\xi(-x).$$

Derartig geeignete Grenzen für Minima sind Verteilungen vom Typ $1 - H_\xi(-x)$. Für eine symmetrische Verteilung F haben wir $\tilde{F}(x) = F(x)$, sodass wenn H_ξ die Grenzverteilung für Maxima für bestimmte Werte von ξ ist, dann ist $1 - H_\xi(-x)$ die Grenzverteilung für Minima.

3.1.2 Maximaler Anziehungsbereich - MDA

Nachdem wir die Grenzverteilungen von normierten Maxima beschrieben haben, ist es natürlich auch interessant nachzuforschen, wie allgemein anwendbar das Fisher-Tippett-Theorem ist und welche unterliegende Verteilungen zu welchen Grenzen führen. Für die meiste Anwendung ist es ausreichend zu wissen, dass im Wesentlichen alle der gängigen stetigen Verteilungen der Statistik oder der Versicherungswissenschaft im $MDA(H_\xi)$ für gewisse Werte von ξ liegen.

Der Fréchet-Fall: Die Verteilungen, welche der Fréchet-Grenze $H_\xi, \xi > 0$ unterliegen, haben eine besonders elegante Charakterisierung.

Satz 3.2. (ohne Beweis) (**Fréchet MDA, Gnedenko**)

Für $\xi > 0$ gilt:

$$F \in MDA(H_\xi) \iff \bar{F}(x) = x^{-1/\xi} L(x), \quad (3.3)$$

für gewisse Funktionen von langsamer Variation $L(x)$.

Dies bedeutet, dass Verteilungen, welche zu einem Fréchet-Fall führen jene sind, deren Tails grundsätzlich wie bei einer Potenzfunktion abfallen, da Funktionen von langsamer Variationen dazu behandelt werden sich wie Konstanten im Bereich der Tails einer Verteilung zu verhalten. Ein Beispiel für eine derartige Funktion ist der Logarithmus $L(x) = \ln(x)$.

Definition 3.3. (Funktion von langsamer Variation)

Eine Funktion von langsamer Variation ist eine positive und Lebesgue-messbare Funktion auf $(0, \infty)$, sodass $\lim_{x \rightarrow \infty} L(tx)/L(x) = 1$ für $t > 0$ ist.

Verteilungen deren Tails in der Form von $\bar{F}(x) = x^{-1/\xi}L(x)$ sind, werden oft als Verteilungen mit Tails von *regulärer Variation* oder *Power-Tails* bezeichnet und der Parameter $\alpha = 1/\xi$ als *Tail-Index* der Verteilung genannt. Derartige Verteilungen sind die am meisten untersuchten in der Extremwerttheorie. Diese sind vor allem vom besonderen Interesse in Finanz-Anwendungen als solche von heavy-tailed Verteilungen dessen höhere Momente unendlich sind.

Wenn für gewisse kleine $\epsilon > 0$ die Verteilung im $MDA(H_{1/2+\epsilon})$ liegt, dann ist dies eine Verteilung mit unendlicher Varianz und wenn die Verteilung im $MDA(H_{1/4+\epsilon})$ liegt, dann handelt es sich hierbei um eine Verteilung mit unendlichem vierten Moment.

Der Gumbel-Fall: Die Charakterisierung der Verteilungen in dieser Klasse ist wesentlich komplizierter als im Fréchet-Fall. Im Beispiel aus [2, Beispiel 7.5] wurde gezeigt, dass die Exponentialverteilung sich in dieser Klasse befindet. Grob kann man beschreiben, dass die Klasse der Gumbel-Verteilung für Stichproben-Maxima von Verteilungen, deren Tails im wesentlichen exponentiell zerfallen, erreicht wird. In dieser Klasse gibt es jedoch eine Vielzahl an unterschiedlichem Tailverhalten. Dies gilt für Verteilungen mit moderaten Tails wie die Normalverteilung oder für Verteilungen mit schweren Tails wie die Lognormalverteilung. Die Lognormalverteilung hat viel „heavier“ Tails, sodass wir wesentlich mehr Daten sammeln müssen bevor wir das Tail-Verhalten von dem der Verteilungen aus der Fréchet-Klasse unterscheiden können. Eine weitere Verteilung, welche der Gumbel-Klasse angehört, ist beispielsweise auch die Hyperbolische Verteilung.

In der Finanzmodellierung wird oft irrtümlicherweise angenommen, dass die einzig interessanten Modelle für die Berechnung von Financial Returns die power-tailed Verteilungen der Fréchet-Klasse sind. Die Gumbel-Klasse ist umso interessanter, weil sie viele noch mehr heavy-tailed Verteilungen enthält als normal auch wenn diese asymptotisch betrachtet keine regulärer-variierenden Power-Tails sind.

Der Weibull-Fall: Dieser ist möglicherweise der am wenigsten wichtige Fall für die finanzielle Modellierung, zumindest im Bereich der Marktrisiken, da die Verteilungen dieser

Klasse in der Praxis beschränkt sind. Der rechte Endpunkt dieser Verteilungen ist endlich. Als Beispiel für derartige Verteilungen ist die Beta-Verteilung bzw. als Spezialfall davon die Gleichverteilung auf $(0, 1)$.

Satz 3.3. (ohne Beweis) (**Weibull MDA, Gnedenko**)

Es sei $\xi < 0$ und x_F der rechte Endpunkt der Verteilung, dann gilt:

$$F \in MDA(H_\xi) \iff x_F < \infty \text{ und } \bar{F}(x_F - x^{-1}) = x^{-1/\xi} L(x), \quad (3.4)$$

für gewisse Funktionen von langsamer Variation $L(x)$.

3.1.3 Maxima von streng stationären Zeitreihen

Wir wissen nun, dass die GEV eine angemessene Grenzverteilung für Maxima von iid Zufallsvariablen von vielen zugrundeliegenden Wahrscheinlichkeitsverteilungen ist. Jedoch in der Finanzmodellierung insbesondere im Zusammenhang mit Marktrisiken, sind die bestmodellierten Renditen-Werten aus Zeitreihen-Prozessen. In diesem Abschnitt werden wir kurz auf die Theorie eingehen, welche darauf hindeutet, dass normierte Maxima von Zufallsvariablen aus streng stationären Zeitreihen die gleichen Typen von Grenzverteilungen haben.

Wir nehmen an, dass die Zufallsvariablen X_1, \dots, X_n einen Block von einer streng stationären Zeitreihe $(X_t)_{t \in \mathbb{Z}}$ mit stationärer Verteilungsfunktion F bilden. Ebenso berücksichtigen wir einen Block von iid Zufallsvariablen $\widetilde{X}_1, \dots, \widetilde{X}_n$ von einem entsprechenden iid Prozess mit derselben Verteilungsfunktion F .

Wir bezeichnen die Block-Maxima der abhängigen Blöcke mit $M_n = \max(X_1, \dots, X_n)$ und die unabhängigen Blöcke mit $\widetilde{M}_n = \max(\widetilde{X}_1, \dots, \widetilde{X}_n)$.

Für viele Prozesse $(X_t)_{t \in \mathbb{Z}}$ kann man zeigen, dass es ein reelles θ in $(0, 1]$ existiert, für die gilt:

$$\lim_{n \rightarrow \infty} P\{(\widetilde{M}_n - d_n)/c_n \leq x\} = H(x) \quad (3.5)$$

für nicht-entartete Grenzen $H(x)$, nur wenn:

$$\lim_{n \rightarrow \infty} P\{(M_n - d_n)/c_n \leq x\} = H^\theta(x) \quad (3.6)$$

Für derartige Prozesse ist der Wert θ bekannt als der *Extremalindex* des Prozesses, welcher gekennzeichnet ist durch den Tail-Index der Verteilungen in der Fréchet-Klasse.

Für Prozesse mit einem Extremalindex konvergieren normierte Block Maxima in der Verteilung, vorausgesetzt die Maxima des zugeordneten iid Prozesses konvergiert in der Verteilung - sprich, vorausgesetzt die zugrundeliegende Verteilungsfunktion F der X_t liegt im $MDA(H_\xi)$ für gewisse ξ .

Ferner, da $H_\xi^\theta(x)$ leicht als eine Verteilung desselben Typs wie $H_\xi(x)$ verifiziert werden kann, ist die Grenzverteilung für die normierten Block Maxima der abhängigen Reihen eine GEV mit exakt demselben Parameter ξ als Grenze der zugeordneten iid Daten. Einzig der Lage- und der Skalenparameter der Verteilung könnte sich ändern.

Mit $u = c_n x + d_n$ erkennt man, dass für große n (3.5) und (3.6) folgendes impliziert:

$$P(M_n \leq u) \approx P^\theta(\widetilde{M}_n \leq u) = F^{n\theta}(u) \quad (3.7)$$

sodass bei hohem Level die Wahrscheinlichkeitsverteilung des Maximums von n Beobachtungen der Zeitreihe mit dem Extremalindex θ wie die des Maximums von $n\theta$ Beobachtungen der entsprechenden iid Reihe ist.

Nicht jeder streng stationäre Prozess hat einen Extremalindex, aber jene Arten von Zeitreihen-Prozessen, für die wir uns in der Finanz-Modellierung interessieren, existiert im Allgemeinen immer ein Extremalindex. Im Wesentlichen müssen wir uns nur zwischen den Fällen wenn $\theta = 1$ oder wenn $\theta < 1$ ist, unterscheiden.

Im Folgenden eine kurze Auflistung der wichtigsten Informationen mit zusätzlichen Referenzen aus [4]:

- Da die X_t s unabhängig sind, hat striktes weißes Rauschen automatisch den Extremalindex $\theta = 1$.
- Dass alle $ARMA(p, q)$ -Prozesse (autoregressive moving average Prozesse der Ordnung (p, q)) mit normalverteilten Innovationen ebenfalls den Extremalindex $\theta = 1$ haben, ist nicht offensichtlich. Ihre Maxima verhalten sich also wie Maxima von unabhängigen Daten. Einerseits haben ARMA-Prozesse ein exponentiell abfallendes Gedächtnis, d.h. Beobachtungen $X_t, X_{t+\tau}$ sind für genügend große Zeitabstände τ praktisch unabhängig, und andererseits ist die Wahrscheinlichkeit für zwei extrem große Beobachtungen innerhalb desselben, nicht zu langen Zeitintervallen, klein.
- ARMA-Prozesse mit strikten weißen Rausch-Innovationen mit einer heavy-tailed Verteilung in $MDA(H_\xi)$ für $\xi > 0$ haben den Extremalindex $\theta < 1$
- Bei ARCH(1)-Prozessen (autoregressive conditional heteroskedasticity Prozesse der Ordnung 1) mit Parameter ω, α ist der Extremalindex stets $\theta < 1$. Er kann approximativ berechnet werden.

Es lassen sich viele Techniken der Extremwertstatistik, die für unabhängige Daten entwickelt wurden, auf Zeitreihen übertragen. Dabei benötigt man allerdings mehr Daten, da der effektive Stichprobenumfang nur $n\theta$ statt n ist.

Ein Problem besteht auch darin, den Extremalindex zu schätzen.

Wir gehen von einer Schätzmethode aus, die sich ohne große technische Vorbereitungen darstellen lässt, die sogenannte *Block Maximum Methode*. Zuerst werden die Zeitreihendaten Z_1, \dots, Z_n in b Blöcke der Länge jeweils l unterteilt ($n = bl$, b, l groß). $M_l^{(k)}$ sei das Maximum der Beobachtungen im k -ten Block:

$$M_l^{(k)} = \max(Z_{(k-1)l+1}, \dots, Z_{kl}), \quad k = 1, \dots, b.$$

Für einen großen Schwellenwert u sei $N(u) = \#\{t \leq n; Z_t > u\}$ die Zahl der Beobachtungen oberhalb der Schwelle und $B(u) = \#\{k \leq b; M_l^{(k)} > u\}$ die Zahl der Blöcke mit wenigstens einer Beobachtung oberhalb von u .

Der Schätzer für den Extremalindex ist dann

$$\hat{\theta} = \frac{1}{l} \frac{\log \left\{ 1 - \frac{B(u)}{b} \right\}}{\log \left\{ 1 - \frac{N(u)}{n} \right\}}. \quad (3.8)$$

Dieser Schätzer lässt sich heuristisch aus den folgenden drei Bemerkungen herleiten vgl. [4, Kapitel 19.4]:

- (i) Aus der Definition des Extremalindex folgt $P(M_n \leq u) \approx F^{\theta n}(u)$, wenn $n, u \rightarrow \infty$, sodass $n\bar{F}(u) \rightarrow \tau$.

Durch Auflösen nach θ folgt

$$\theta \approx \frac{\log \{P(M_n \leq u)\}}{n \log \{F(u)\}} \quad (3.9)$$

- (ii) F lässt sich durch die empirische Verteilungsfunktion \hat{F}_n schätzen, sodass

$$F(u) = 1 - P(Z_t > u) \approx 1 - \frac{N(u)}{n} \quad (3.10)$$

- (iii) Mit $n = bl$ folgt

$$\begin{aligned} P(M_n \leq u) &\approx \prod_{k=1}^b P(M_l^{(k)} \leq u) \approx \left\{ P(M_l^{(1)} \leq u) \right\}^b \\ &\approx \left\{ \frac{1}{b} \sum_{k=1}^b \mathbf{1}(M_l^{(k)} \leq u) \right\}^b = \left(1 - \frac{B(u)}{b} \right)^b \end{aligned} \quad (3.11)$$

Durch Kombination der drei Beziehungen erhalten wir

$$\theta \approx \frac{b \log \left(1 - \frac{B(u)}{b}\right)}{n \log \left(1 - \frac{N(u)}{n}\right)} = \hat{\theta} \quad (3.12)$$

□

3.1.4 Die Block Maxima Methode

Angenommen wir haben die Daten X_1, X_2, \dots : Die zugrundeliegende Verteilungsfunktion F , welche im Anziehungsbereich einer Extremwertverteilungsfunktion H_ξ für gewisse ξ liegt, ist unbekannt. Wir konstruieren das Maximum M_n eines benachbarten n -Blocks von Beobachtungen. Ob die Daten Ausführungen von iid Variablen oder von Variablen eines Prozesses mit Extremalindex sind, ist nicht bekannt. Die Theorie besagt, dass die wahre Verteilungsfunktion des n -Block Maximums M_n approximiert werden kann für genug große n mittels einer drei-parametrischen GEV-Verteilung $H_{\xi, \mu, \sigma}$.

Wir wenden die Idee an, indem wir die GEV-Verteilung $H_{\xi, \mu, \sigma}$ zu den Daten des n -Block Maximums anpassen. Offensichtlich benötigt man wiederholende Beobachtungen eines n -Block Maximums und wir nehmen an, dass die Daten in m Blöcke mit der Größe n geteilt werden können. Dies macht nur dann Sinn, wenn sich die Daten auf eine natürliche Art in Blöcke aufteilen lassen. Diese Methodik stammt ursprünglich aus dem Bereich der Hydrologie, wo beispielsweise täglich der Wasserstand gemessen wird, welche in jährliche Blöcke aufgeteilt werden können.

Analog dazu untersuchen wir Finanz-Anwendungen wo tägliche Return-Daten (es werden dabei natürlich nur Handelstage gezählt) in jährliche, halbjährliche oder quartalsweise Blöcke aufgeteilt werden.

Wir bezeichnen das Block Maximum des j ten Blocks mit $M_{n,j}$, sodass unsere Daten $M_{n,1}, \dots, M_{n,m}$ sind. Die GEV-Verteilung kann mit verschiedenen Methoden geschätzt werden. Die am naheliegendsten Methode ist die Maximum-Likelihood Methode.

In der Ausführung des Maximum-Likelihoods ist unbedacht, ob die Daten abhängig oder unabhängig sind. Für ausreichend große Blöcke n ist anzunehmen, dass die Block Maxima Beobachtungen $M_{n,1}, \dots, M_{n,m}$ unabhängig werden.

Sei $h_{\xi, \mu, \sigma}$ die Dichte der GEV.

Für $\xi \neq 0$, $\sigma > 0$ und $1 + \xi \frac{M_{n,i} - \mu}{\sigma} > 0 \quad \forall i$ lässt sich die Log-Likelihood Funktion der GEV-Verteilung wie folgt ermitteln:

$$\begin{aligned}
 l(\xi, \mu, \sigma; M_{n,1}, \dots, M_{n,m}) &= \sum_{i=1}^m \ln h_{\xi, \mu, \sigma}(M_{n,i}) \\
 &= -m \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^m \ln \left(1 + \xi \frac{M_{n,i} - \mu}{\sigma}\right) \\
 &\quad - \sum_{i=1}^m \ln \left(1 + \xi \frac{M_{n,i} - \mu}{\sigma}\right)^{-\frac{1}{\xi}}
 \end{aligned} \tag{3.13}$$

Durch Maximieren von (3.13) erhält man den Maximum-Likelihood Schätzer der GEV-Verteilung. Aufgrund der Abhängigkeit der unbekannt Parameter zu den Daten, ist die Konsistenz des Schätzers (dies bedeutet, dass für große Stichproben n gegen einen tatsächlichen Wert konvergiert) und die asymptotische Effizienz für den Fall $\xi > \frac{1}{2}$ gewährleistet.

Bei der Bestimmung der Anzahl und Größe der Blöcke muss ein Kompromiss notwendigerweise stattfinden.

Fazit: Ein großer Wert von n führt zu einer genaueren Approximation der Block Maxima Verteilung durch eine GEV und bewirkt eine geringere Verzerrung bei der Parameterschätzung.

Ein großer Wert von m ergibt mehr Block Maxima Daten für die Maximum-Likelihood-Schätzung und führt zu einer geringen Abweichung in der Parameterschätzung.

Es ist ebenfalls zu beachten, dass im Falle von abhängigen Daten, die Auswahl von etwas größeren Blöcken sinnvoll sein kann - im Gegensatz zum iid-Fall. Die Abhängigkeit der Daten im Allgemeinen hat die Auswirkung, dass die Konvergenz der GEV langsamer ist, da die effektive Stichprobengröße $n\theta$ ist und dies ist kleiner als n (siehe [2, Beispiel 7.12]).

3.1.5 Return-Level und Return-Periode

In diesem Abschnitt gehen wir davon aus, dass ein Block Maxima Modell zu einem finanziellen Verlust von Daten unter Anwendung einer natürlichen Definition der Blocklänge n angepasst wurde.

Um Stress-Ereignisse mit derartigen Modellen zu analysieren gibt es eine Vielzahl von Möglichkeiten. Wir konzentrieren uns auf zwei:

- (i) Wir definieren die Häufigkeit des Auftretens des Stress-Ereignisses und schätzen seinen Umfang - bekannt als das *Return-Level*-Schätzproblem.
- (ii) Wir definieren die Größe des Stress-Ereignisses und schätzen die Häufigkeit ihres Auftretens - bekannt als das *Return-Periode*-Problem.

Angenommen die n -Block Maxima sind iid mit der echten Verteilung H . Wir stellen nun zwei Fragen auf:

- Welcher Level hat einmal alle k n -Blöcke im Durchschnitt überschritten?
- Wie oft wird der Level u durch ein n -Block Maximum überschritten?

Die erste Frage lässt sich durch die Berechnung des k n -Block Return-Levels interpretieren.

Der k -Jahr Return-Level ist eine Schwelle, welche durchschnittlich nur in einem n -Block im Zeitraum von k Jahren überschritten wird. Innerhalb dieses Blocks kann es aber auch zu mehreren Überschreitungen kommen. Dies widerspricht nicht der Definition des Return-Levels und in der Regel tritt dieser Fall ein. Die Log-Renditen unterliegen nämlich oft einer großen Cluster-Bildung. Der Block, indem der Return-Level überschritten wird, wird auch als Stress-Periode (*stress period*) bezeichnet und die Renditen nennt man dementsprechend Stress-Verluste (*stress losses*).

Da der Return-Level nur ein Punktschätzer ist, der nicht angibt wie groß die Log-Rendite ist und wenn dieser überschritten wird. Somit ist es sinnvoll Konfidenzintervalle zu verwenden. Interessant hierbei ist vor allem der rechte Randpunkt eines Konfidenzintervalls, denn dieser zeigt beispielsweise wie tief ein Aktiensturz sein könnte und modelliert sozusagen ein Worstcase-Szenario.

Definition 3.4. (Return-Level)

Sei H die Verteilungsfunktion der echten Verteilung des n -Block Maximums. Der k n -Block Return-Level $r_{n,k}$ ist das $(1 - \frac{1}{k})$ -Quantil von H mit $r_{n,k} = H^{\leftarrow} \left(1 - \frac{1}{k}\right) = q_{1-1/k}(H)$.

Wir könnten beispielsweise in Anbetracht ziehen, dass der 10 Handels-Jahre Return-Level $r_{260,10}$ jener Level ist, der durchschnittlich alle 10 Jahre überschritten wird - unter der Annahme, dass jedes Jahr im Schnitt 260 Handel-Tage hat.

Durch unser Modell können wir den Return-Level wie folgt schätzen:

$$\begin{aligned} \hat{r}_{n,k} &= H_{\hat{\xi}, \hat{\mu}, \hat{\sigma}}^{-1} \left(1 - \frac{1}{k}\right) \\ &= \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(-\ln \left(1 - \frac{1}{k}\right) \right)^{-\hat{\xi}} - 1 \right) \end{aligned} \tag{3.14}$$

Die zweite Fragestellung lässt sich durch die Schätzung einer Return-Periode beantworten. Hierbei stellt man sich ein extremes Ereignis vor - beispielsweise eine große negative Rendite. Man untersucht dabei, wie selten ein solches Szenario eintritt. Sprich, in welchem Zeitraum eine Schwelle u überschritten wird. Es lässt sich auch direkt untersuchen, wie wahrscheinlich es ist, dass bereits im nächsten Jahr ein neuer negativer Rekord eintritt - dass somit die maximale Rendite alle vorherigen übertritt.

Definition 3.5. (Return-Periode)

Sei H die Verteilungsfunktion der echten Verteilung des n -Block Maximums. Die Return-Periode des Ereignisses $\{M_n > u\}$ ist definiert durch $k_{n,u} = (\overline{H}(u))^{-1}$.

Die Return-Periode $k_{n,u}$ ist solchermaßen definiert, dass der $k_{n,u} - n$ -Block Return-Level u ist. In $k_{n,u}$ n -Blöcke würden wir wie erwartet einen einzelnen Block beobachten, in welchem der Level u überschritten wird. Bei einer starken Tendenz der Cluster-Extremwerte, erwarten wir mehrere Überschreitungen des Levels innerhalb dieses Blocks.

Unter Anwendung unseres Modells können wir die Return-Periode wie folgt schätzen:

$$\hat{k}_{n,u} = \left(\overline{H}_{\hat{\xi}, \hat{\mu}, \hat{\sigma}}(u) \right)^{-1} \tag{3.15}$$

Man beachte, dass $\hat{r}_{n,k}$ und $\hat{k}_{n,u}$ beide einfache Funktionale der geschätzten GEV-Parameter sind. Bei der Kalkulation des Punktschätzers für diese Größe ist, wie bereits erwähnt, die Verwendung von Konfidenzintervallen sinnvoll, sodass der Fehler im GEV-Parameterschätzer widergespiegelt werden kann. Eine gute Methode hierbei ist die Etablierung solcher Konfidenzintervalle in den Likelihood-Quotienten-Test.

Wir haben den Return-Level als ein Quantil einer Verteilung von einem einzelnen Block-Maximum definiert. Offensichtlich entspricht auch jeder gegebene Return-Level einem Quantil der zugrundeliegenden Verteilung F . Für iid Zufallsvariablen ist diese Beziehung offensichtlich.

Sei $\alpha \in (0, 1)$ und das α -Quantil x_α von einer stetigen Verteilung mit der Verteilungsfunktion F definiert durch:

$$x_\alpha = F^{-1}(\alpha) \tag{3.16}$$

wobei F^{-1} die Inverse der Verteilungsfunktion ist.

$\sum_{j=1}^k \mathbf{1}_{M_{n_j} > r_{n,k}}$ ist nun die Anzahl von Überschreitungen in k Jahren, wobei $r_{n,k}$ die gesuchte Schwelle ist, die innerhalb von k Jahren überschritten wird.

Der Return-Level ist nun so definiert, dass dieser nur in einem n -Block im Zeitraum von k Jahren überschritten wird.

Die erwartete Anzahl von Überschreitungen ist somit 1, sodass gilt:

$$\begin{aligned} \mathbb{E} \left[\sum_{j=1}^k \mathbf{1}_{M_{n_j} > r_{n,k}} \right] = 1 &\Leftrightarrow kP(M_n > r_{n,k}) = 1 \\ &\Leftrightarrow P(M_n > r_{n,k}) = \frac{1}{k} \end{aligned}$$

Somit ist

$$P(M_n \leq r_{n,k}) = 1 - \frac{1}{k} \tag{3.17}$$

Dies kennen wir bereits aus der Definition des Return-Levels.

Durch die iid Zufallsvariablen X_n folgt mit (3.16) und (3.17):

$$F(r_{n,k}) = P^n(M_n \leq r_{n,k}) = \left(1 - \frac{1}{k}\right)^{\frac{1}{n}} \tag{3.18}$$

Das heißt, der Return-Level ist das $\left(1 - \frac{1}{k}\right)^{\frac{1}{n}}$ -Quantil der zugrundeliegenden Verteilungsfunktion F .

Im Falle von Zufallsvariablen mit Extremalindex θ lässt sich die Beziehung zwischen den Return-Levels und Quantilen der zugrundeliegenden Verteilung nicht mehr exakt ableiten.

Durch (3.7) ist $r_{n,k}$ annähernd ein $\left(1 - \frac{1}{k}\right)^{\frac{1}{n\theta}}$ -Quantil. Dies wird jedoch nicht als Mittel zur Schätzung von Quantilen empfohlen, vor allem weil hier der Parameter θ benötigt wird, um eine zuverlässige Schätzung zu ermitteln.

3.2 Schranken-Überschreitungen

Die Anwendung der Block-Maxima-Methode erweist sich jedoch als relativ problematisch, da diese sehr „verschwenderisch“ im Umgang mit Datensätzen ist. Um unsere Analysen durchzuführen, behalten wir nur die maximalen Verluste in großen Blöcken. Aus diesem Grund wurde sie in der Praxis weitgehend von Methoden abgelöst, die auf Schranken-Überschreitungen basieren. Hierbei werden alle Daten benutzt, welche „extrem“ im weiteren Sinne ein einzelnes festgelegtes hohes Level überschreiten. Damit wir geeignete Modelle für Schranken-Überschreitungs-Daten aufbauen können, müssen wir vorher einige weitere Ergebnisse der Extremwerttheorie betrachten, welche das erwartete Verhalten solcher Daten zeigen können.

3.2.1 Verallgemeinerte Paretoverteilung - GPD

Das haupt-verbreitete Verteilungsmodell für Schranken-Überschreitungen ist die Verallgemeinerte Paretoverteilung (GPD).

Definition 3.6. (Verallgemeinerte Paretoverteilung) Die Verteilungsfunktion der verallgemeinerten Paretoverteilung ist definiert durch:

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-1/\xi} & \text{für } \xi \neq 0, \\ 1 - \exp\left(\frac{-x}{\beta}\right) & \text{für } \xi = 0, \end{cases}$$

für

$$x \in \mathcal{D}(\xi, \beta) := \begin{cases} [0; \infty) & \text{für } \xi \geq 0, \\ \left[0; -\frac{\beta}{\xi}\right] & \text{für } \xi < 0, \end{cases}$$

wobei $\beta > 0$ und $\xi \in \mathbb{R}$ sind.

Somit ist $G_{\xi,\beta}$ auch außerhalb von $\mathcal{D}(\xi, \beta)$ definiert und nimmt dort die Werte 0 bzw. 1 an.

Die Parameter ξ und β werden als den Form- und Skalenparameter bezeichnet.

Wie in der Definition der GEV, klassifiziert die allgemein parametrisierte Form der GPD abhängig vom Wert von ξ eine Reihe von Spezialfällen:

Für $\xi > 0$ entspricht $G_{\xi,\beta}$ einer Pareto-Verteilung mit $\alpha = 1/\xi$ und $\kappa = \beta/\xi$

Für $\xi = 0$ entspricht $G_{\xi,\beta}$ einer Exponential-Verteilung

Für $\xi < 0$ entspricht $G_{\xi,\beta}$ einer Pareto Typ II-Verteilung.

Ebenso wie für die GEV, ist die parametrisierte Form für fixe x stetig in ξ , sodass $\lim_{\xi \rightarrow 0} G_{\xi,\beta}(x) = G_{0,\beta}(x)$ ist.

Im Bezug auf den Anziehungsbereich ist $G_{\xi,\beta}(x) \in MDA(H_\xi)$ für alle $\xi \in \mathbb{R}$.

Für $\xi > 0$ und $\xi < 0$ lässt sich diese Aussage leicht aus dem Satz 3.2 und 3.3 ableiten.

Für $\xi = 0$ ist die Exponential-Verteilungsfunktion im Gumbel- MDA .

Im heavy-tailed-Fall (dh: die Verteilung hat keine vollständige Menge von endlichen Momenten), haben wir für $\xi > 0$ und $k \geq 1/\xi$ den $\mathbb{E}(X^k) = \infty$.

Der Erwartungswert der GPD ist definiert unter der Voraussetzung, dass $\xi < 1$ ist durch:

$$\mathbb{E}(X) = \frac{\beta}{1 - \xi} \quad (3.19)$$

Die Rolle der GPD in der Extremwerttheorie ist als Modell für *Exzess-Verteilungen* über eine hohe Schranke definiert. Wir bezeichnen dieses Konzept analog dazu mit der *mittleren Exzess-Funktion*.

Definition 3.7. (Exzess-Verteilung über die Schranke u) Sei X eine Zufallsvariable mit der Verteilungsfunktion F . Die Exzess-Verteilung über die Schranke u ist die bedingte Verteilung von $X - u$ bedingt auf $X > u$ und hat folgende Verteilungsfunktion:

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(x + u) - F(u)}{1 - F(u)} \quad (3.20)$$

für $0 \leq x < x_F - u$ wobei $x_F \leq \infty$ der rechte Randpunkt von F ist.

In Worte gefasst, entspricht dies der Wahrscheinlichkeit für die Überschreitung eines Verlustes über den Level u , jedoch höchstens begrenzt mit x - unter der Bedingung, dass u bereits überschritten wurde.

Definition 3.8. (Mittlere Exzess-Funktion) Die Mittlere Exzess-Funktion einer Zufallsvariable X mit endlichem Erwartungswert ist der Erwartungswert von F_u . Die Exzess-Verteilungsfunktion über die Schranke u wird ausgedrückt durch:

$$e(u) = \mathbb{E}(X - u | X > 0). \quad (3.21)$$

In Worten gefasst, gibt sie den erwarteten Überschreitungs-Verlust an - bedingt darauf, dass u bereits überschritten wurde.

Nach Definition von $e(u)$ und mit partieller Integration folgt:

$$e(u) = \frac{1}{\bar{F}(u)} \int_u^{x_F} (x - u) dF(x) = \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(x)$$

Weiters gilt

$$\bar{F}(x) = \frac{e(0)}{e(x)} e^{-\int_0^x \frac{1}{e(u)} du}$$

für die Berechnung aller Stetigkeitspunkte von F . Somit ist ein stetiges F durch seine Exzess-Funktion eindeutig bestimmt.

Besonders von Interesse ist diese Verteilung vor allem in der Finanz- und Versicherungswirtschaft. Anders als der Value-at-Risk, bestimmt sie nicht die Wahrscheinlichkeit eines Verlustes einer bestimmten Höhe, sondern auch den Erwartungswert, falls diese Grenze überschritten wird.

Beispiel 3.1. (Exzess-Verteilung von Exponential- und GPD)

Sei F die Verteilungsfunktion einer Exponential-Zufallsvariable. Dann gilt $F_u(x) = F(x)$ für alle x , welche die sogenannte Gedächtnislosigkeit der Exponential-Verteilung ist.

Wenn X die Verteilungsfunktion $F(x) = G_{\xi, \beta_0}(x)$ hat, dann ist für GPD die Exzess-Verteilungsfunktion unter Verwendung von (3.20) wie folgt:

$$\begin{aligned}
 F_u(x) &= \frac{1 - \left(1 + \frac{\xi(x+u)}{\beta_0}\right)^{-1/\xi} - \left(1 - \left(1 + \frac{\xi u}{\beta_0}\right)^{-1/\xi}\right)}{1 - \left(1 - \left(1 + \frac{\xi u}{\beta_0}\right)^{-1/\xi}\right)} \\
 &= \frac{-\left(\frac{\beta_0 + \xi(x+u)}{\beta_0}\right)^{-1/\xi} + \left(\frac{\beta_0 + \xi u}{\beta_0}\right)^{-1/\xi}}{\left(\frac{\beta_0 + \xi u}{\beta_0}\right)^{-1/\xi}} \\
 &= 1 - \left(\frac{\beta_0 + \xi(x+u)}{\beta_0 + \xi u}\right)^{-1/\xi} \\
 &= G_{\xi, \beta(u)}(x),
 \end{aligned} \tag{3.22}$$

mit $\beta(u) = \beta_0 + \xi u$,
wobei

$$\begin{cases} 0 \leq x < \infty & \text{wenn } \xi \geq 0, \\ 0 \leq x \leq -\frac{\beta_0}{\xi} - u & \text{wenn } \xi < 0. \end{cases}$$

Die Exzess-Verteilung hinterlässt der GPD denselben Formparameter ξ und einen Skalenparameter, welcher linear mit der Schranke u steigt.

Die Mittlere Exzess-Funktion der GPD ist dann unter Verwendung von (3.19) und (3.22)

$$e(u) = \frac{\beta(u)}{1 - \xi} = \frac{\beta_0 + \xi u}{1 - \xi} \tag{3.23}$$

wobei

$$\begin{cases} 0 \leq u < \infty & \text{wenn } 0 \leq \xi < 1, \\ 0 \leq u \leq -\frac{\beta_0}{\xi} & \text{wenn } \xi < 0. \end{cases}$$

Somit kann man behaupten, dass die Mittlere Exzess-Funktion *linear in der Schranke* u ist.

Bemerkung 3.1. Für spätere Berechnungen wird es sehr nützlich sein zu beobachten was passiert, wenn wir für einen bestimmten Schwellenwert u annehmen, dass die überschüssigen Verluste einer GPD mit dem Skalenparameter $\beta = \beta(u)$ unterliegen. Man betrachte dann eine höhere Schranke $v \geq u$. Wir wissen, dass $F_u(x) = G_{\xi, \beta(u)}(x)$ und

$$e(v) = \frac{\beta(v)}{1 - \xi}, \quad \text{mit } \beta(v) = \beta + \xi v,$$

wobei

$$\begin{cases} u \leq v < \infty & \text{wenn } 0 \leq \xi < 1, \\ u \leq v \leq u - \frac{\beta}{\xi} & \text{wenn } \xi < 0. \end{cases}$$

Das Beispiel 3.1 zeigt uns, dass die GPD eine Art von Stabilitäts-Eigenschaft hat. Wir geben nun ein mathematisches Ergebnis an, welches zeigt, dass die GPD eine natürliche Grenzverteilung für viele zugrundeliegende Verlust-Verteilungen ist. Dies kann ebenso als das Charakterisierungs-Theorem des Anziehungsbereichs für die GEV betrachtet werden. Das folgende Resultat liefert eine globale Charakterisierung von $MDA(H_\xi)$ für alle ξ im Grenzverhalten von Exzess-Verteilungen über Schranken, die ins Unendliche konvergieren.

Satz 3.4. (ohne Beweis) (**Pickands-Balkema-de Haan**)

Für eine geeignete positive messbare Funktion $\beta(u)$ gilt:

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

genau dann, wenn für alle $\xi \in \mathbb{R}$ die in x_F stetige Verteilungsfunktion $F \in MDA(H_\xi)$ ist.

Das Pickands-Balkema-de Haan-Theorem besagt, dass die Verteilung von Exzessen über eine ausreichend hohe Schranke im Fall von Wahrscheinlichkeitsverteilungen, die im Anziehungsbereich der GEV liegen, durch eine GPD approximiert werden kann.

Folglich stellen die Verteilungen, für welche die normierten Maxima gegen eine GEV konvergieren, ein Set von Verteilungen dar, für welches die Exzess-Verteilung gegen eine GPD konvergiert. Der Formparameter der Grenz-GPD für Exzesse ist dieselbe wie die der Grenz-GEV-Verteilung für die Maxima.

Im Kapitel 3.1.2 haben wir bereits kennen gelernt, dass im Wesentlichen die meist verwendeten stetigen Verteilungen der Statistik für gewisse ξ im $MDA(H_\xi)$ liegen. Satz 3.4 zeigt nun, dass die GPD die *kanonische Verteilung* für die Modellierung von Exzess-Verlusten über eine hohe Schranke ist.

Auf das Kapitel der Modellierung von Exzess-Verlusten wird hier nicht näher eingegangen. Mehr Informationen entnehmen Sie bitte aus [2, Kapitel 7.2.2].

3.2.2 Modellierung von Tails und Messgrößen von Tail-Risiken

Wir verwenden nun die GPD-Approximation zur Exzess-Verteilung über u , um einen Tail-Schätzer für die zugrundeliegende Verteilung $F(x)$ zu ermitteln. Für unsere Zwecke nehmen wir an, dass der Tail dieser Verteilung bei u beginnt. Wir betrachten, dass wir für $x \geq u$ folgendes erhalten:

$$\begin{aligned}\bar{F}(x) &= P(X > x \mid X > u) \cdot P(X > u) \\ &= P(X - u > x - u \mid X > u) \cdot \bar{F}(u) \\ &= \bar{F}_u(x - u) \cdot \bar{F}(u)\end{aligned}\tag{3.24}$$

Die Grundidee ist, unter dem GPD-Modell $G_{\xi, \hat{\beta}}$ die Verteilung $\bar{F}_u(x - u)$ zu schätzen, sodass die Parameter-Schätzer enthalten sind. Außerdem schätzen wir hier $\bar{F}(u)$ empirisch durch N_u/n . Dabei nehmen wir implizit an, dass ein ausreichender Anteil von Stichproben den Wert der Schranke u übersteigt, sodass $\bar{F}(u)$ zuverlässig geschätzt werden kann. Nichtsdestotrotz hoffen wir, unter Verwendung der geschätzten GPD-Form für die Exzess-Verteilung, die empirische Methode zu erlangen, um mehrere extreme Tail-Wahrscheinlichkeiten schätzen zu können. Wenn man nun die zwei Parameter-Schätzer mit dem Resultat aus (3.24) vergleicht, erhält man:

$$\hat{\bar{F}}(x) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x - u}{\hat{\beta}} \right)^{-1/\hat{\xi}}\tag{3.25}$$

nur für $x \geq u$.

Die Tail-Schätzer-Formel kann nun invertiert werden, sodass wir einen Schätzer eines hohen Quantils der zugrundeliegenden Verteilung erzielen können.

Für $\alpha \geq 1 - N_u/n$ erhalten wir:

$$\hat{q}_\alpha(x) = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1 - \alpha) \right)^{-\hat{\xi}} - 1 \right).\tag{3.26}$$

Folglich ist nun ein Schätzer einer Quantils-Risiko-Messgröße ein Resultat der Tail-Schätzer-Methode. Der Schätzer des dazugehörigen *Expected-Shortfalls* ist demnach ebenso einfach zu ermitteln.

Für eine Verlust-Variable X mit der Verteilungsfunktion F erhalten wir den Expected-Shortfall mit:

$$ES_\alpha = \mathbb{E}(X \mid X > q_\alpha(F)) = q_\alpha(F) + e(q_\alpha(F)) \quad (3.27)$$

wo wir nun die Risiko-Messgröße in Termen des Quantils und die Mittlere Exzess-Funktion darstellen.

Für $\alpha \geq F(u)$ ist das Quantil $q_\alpha(F)$ als eine (unbekannte) höhere Schranke als u zu betrachten. Nun schätzen wir die Mittlere Exzess-Funktion bei $v = q_\alpha(F)$ unter Verwendung von

$$\widehat{e}(v) = \frac{\hat{\xi}v}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1 - \hat{\xi}} \quad (3.28)$$

vgl. dazu [2, Kapitel 7.2.2 Formel (7.16)].

Wir kombinieren nun obige Formel mit dem Schätzer von $q_\alpha(F)$ in (3.26), um für $\alpha \geq 1 - N_u/n$ den Expected-Shortfall-Schätzer zu erhalten:

$$\widehat{ES}_\alpha = \widehat{q}_\alpha(F) + \widehat{e}(\widehat{q}_\alpha(F)) = \frac{\widehat{q}_\alpha(F)}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1 - \hat{\xi}}. \quad (3.29)$$

Zusammengefasst erhalten wir aus (3.25), (3.26) und (3.28) jeweils den Schätzer für Tail-Wahrscheinlichkeiten, Quantile und Expected-Shortfalls.

All diese Risiko-Messgrößen sind von der Form $g(\hat{\xi}, \hat{\beta}, N_u/n)$, wo diese eine Funktion $g(\xi, \beta, \overline{F}(u))$ schätzt, welche die genauen Werte der drei Quantilen unter der Annahme angibt, dass die Exzess-Verteilung über u exakt GPD ist.

Hätten wir den Likelihood-Ansatz gewählt, um ξ und β zu schätzen, ist es dann einfach, Konfidenzintervalle für $g(\hat{\xi}, \hat{\beta}, N_u/n)$ zu ermitteln. Diese Methode würde auch den Fehler in $\hat{\xi}$ und $\hat{\beta}$ erklären, aber vernachlässigt den Fehler in N_u/n als einen Schätzer von $\overline{F}(u)$. Wir verwenden nun den Ansatz, welcher im Kapitel 3.1.5 bezogen auf Return-Levels beschrieben ist, wobei der GPD-Modell in Termen von $\phi = g(\xi, \beta, N_u/n)$ re-parametrisiert ist und ein Konfidenzintervall für $\widehat{\phi}$ basierend auf den Likelihood-Ratio-Test konstruiert ist (vgl. [2, Appendix A.3.5]).

Es ist ebenso interessant zu beobachten, wie sich das Verhältnis zwischen den zwei Risiko-Messgrößen Expected-Shortfall und das Quantil für hohe Werte der Quantil-Wahrscheinlichkeit α verhalten wird. Aus (3.27) erhalten wir

$$\frac{ES_\alpha}{q_\alpha(F)} = 1 + \frac{e(q_\alpha(F))}{q_\alpha(F)},$$

sodass das asymptotische Verhältnis durch das Verhalten von $e(v)/v$ bestimmt ist, wobei v zu x_F , dem Endpunkt von F , tendiert. Wenn das GPD-Modell $G_{\xi,\beta}$ sich für Exzess-Verluste über eine Schranke u hält, dann impliziert Bemerkung 3.1 und (3.28) folgendes:

$$\frac{e(v)}{v} = \frac{\xi}{1-\xi} + \frac{\beta - \xi u}{(a - \xi)v},$$

wobei

$$\begin{cases} v \geq u & \text{wenn } 0 \leq \xi < 1, \\ u \leq v \leq u - \frac{\beta}{\xi} & \text{wenn } \xi < 0. \end{cases}$$

Es lässt nun einfach folgendes schließen:

$$\begin{cases} \frac{e(v)}{v} \rightarrow \frac{\xi}{1-\xi} & \text{wenn } 0 < \xi < 1, \\ \frac{e(v)}{v} \rightarrow 0 & \text{wenn } \xi \leq 0. \end{cases}$$

Folglich haben wir im Fall von heavy-tailed das asymptotische Verhältnis von

$$\begin{cases} \lim_{\alpha \rightarrow 0} \frac{ES_\alpha}{q_\alpha(F)} = (1 - \xi)^{-1} & \text{wenn } 0 < \xi < 1, \\ \lim_{\alpha \rightarrow 0} \frac{ES_\alpha}{q_\alpha(F)} = 1 & \text{sonst.} \end{cases}$$

In praktischen Anwendungen erhalten wir, wenn $\hat{\xi} > 0$ ist, aus dem Schätzer von einem hohen Quantil (3.26) den Schätzer des entsprechenden Expected-Shortfalls (3.29) durch Multiplikation des Ersteren mit einem Schätzer des asymptotischen Verhältnisses $(1 - \hat{\xi})^{-1}$ und Hinzufügen eines Korrelations-Terms.

Bevor wir das Kapitel GPD-Tail-Modellierung verlassen, ist es noch wichtig zu beobachten, wie empfindlich unsere Schätzer von Risiko-Messgrößen im Bezug auf die Wahl von Schranken sind. Bisher betrachteten wir einzelne Schranken u und beobachteten eine Reihe von stufenweisen Kalkulationen, die sich immer auf demselben GPD-Modell für Exzesse über dieser Schranke aufbauen.

3.2.3 Die Hill-Methode

Die GPD-Methode ist nicht die einzige Möglichkeit, um Tails von Verteilungen zu schätzen. Als eine Alternative beschreiben wir in diesem Kapitel die Hill-Methode, um Tails von heavy-tailed Verteilungen zu modellieren.

Schätzung des Tail-Indexes: Für diese Methode müssen wir den Glauben belassen, dass die zugrundeliegende Verlust-Verteilung einen Tail von regulärer Variation der Fréchet-Klasse hat. Aus dem Satz 3.2 wissen wir, dass derartige Verteilungen von der Form

$$\overline{F}(x) = L(x)x^{-\alpha} \quad (3.30)$$

sind - mit einer Funktion von langsamer Variation $L(x)$ und einem positiven Parameter α . Die Hill-Methode konzentriert sich traditionellerweise auf den Tail-Index α anstatt dessen Kehrwert ξ wie in (3.3) beschrieben. Das Ziel ist es, einen Schätzer von α basierend auf identisch verteilten Daten X_1, \dots, X_n .

Der Hill-Schätzer kann durch verschiedene Wege abgeleitet werden (vgl. [5, Kapitel 4.2.1 S 101 ff.]):

- Quantilsbasierte Ableitung
- Wahrscheinlichkeitsbasierte Ableitung
- Ableitung über Rényi's Exponentialdarstellung
- Ableitung über die Mittlere Exzess-Funktion.

Die möglicherweise eleganteste Variante ist es die Mittlere Exzess-Funktion der logarithmischen Verlustes $\ln X$ zu betrachten, wobei X eine Zufallsvariable ist mit der Verteilungsfunktion aus (3.30).

Sei e^* die Mittlere Exzess-Funktion von $\ln X$. Mit partieller Integration erhalten wir

$$\begin{aligned} e^*(\ln u) &= \mathbb{E}(\ln X - \ln u \mid \ln X > \ln u) \\ &= \frac{1}{\overline{F}(u)} \int_u^\infty (\ln x - \ln u) dF(x) \\ &= \frac{1}{\overline{F}(u)} \int_u^\infty \frac{\overline{F}(x)}{x} dx \\ &= \frac{1}{\overline{F}(u)} \int_u^\infty L(x)x^{-(\alpha+1)} dx . \end{aligned}$$

Für hinreichend große u kann die Funktion von langsamer Variation $L(x)$ für $x \geq u$ grundsätzlich als eine Konstante behandelt werden und somit auch vor das Integral gestellt werden.

Um die weitere Schlussfolgerung darzustellen, führen wir zunächst nachstehenden Satz ein:

Satz 3.5. (ohne Beweis) (**Satz von Karamata**)

Sei L eine Funktion von langsamer Variation und lokal beschränkt auf $[x_0, +\infty)$ für alle $x_0 \in \mathbb{R}$, dann gilt:

(i) Für $\kappa > -1$:

$$\int_{x_0}^x t^\kappa L(t) dt \sim \frac{1}{\kappa + 1} x^{\kappa+1} L(x) \quad \text{für } x \rightarrow \infty$$

(ii) Für $\kappa < -1$:

$$\int_x^{+\infty} t^\kappa L(t) dt \sim -\frac{1}{\kappa + 1} x^{\kappa+1} L(x) \quad \text{für } x \rightarrow \infty$$

Aus dem Satz von Karamata folgt nun:

$$\begin{aligned} e^*(\ln u) &= \frac{1}{\overline{F}(u)} \int_u^\infty L(x) x^{-(\alpha+1)} dx \\ &\sim \frac{L(u) u^\alpha \alpha^{-1}}{\overline{F}(u)} = \alpha^{-1}, \quad \text{für } u \rightarrow \infty, \end{aligned}$$

sodass die Mittlere Exzess-Funktion des logarithmischen Verlustes über die logarithmische Schranke $\ln u$ zu eine Konstante konvergiert für große u .

Für eine Stichprobe von Mittleren Exzess-Funktionen e_n^* würden wir dasselbe Verhalten erwarten, mit

$$e_n(v) = \frac{\sum_{i=1}^n (X_i - v) \mathbf{1}_{\{X_i > v\}}}{\sum_{i=1}^n \mathbf{1}_{\{X_i > v\}}} \quad (3.31)$$

für positive Verlust-Daten X_1, \dots, X_n (vgl. [2, Seite 274]).

Sei nun $X_{n,n} \leq \dots \leq X_{1,n}$ die Ordnungsstatistik, dann würden wir insbesondere folgendes für hinreichend kleine k erwarten:

$$e_n^*(\ln X_{k,n}) \approx \alpha^{-1}$$

Wenn man nun die Stichprobe von Mittleren Exzess-Funktionen mit der gegebenen Ordnungsstatistik auswertet, erhalten wir den Schätzer

$$\hat{\alpha}^{-1} = \left((k-1)^{-1} \sum_{j=1}^{k-1} \ln X_{j,n} - \ln X_{k,n} \right).$$

Die standardisierte Form des Hill-Schätzers wird durch eine geringe Modifikation wie folgt dargestellt:

$$\hat{\alpha}_{k,n}^{(H)} = \left(\frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1}, \text{ mit } 2 \leq k \leq n. \quad (3.32)$$

Der Hill-Schätzer ist einer der best-studierten Schätzer in der Extremwerttheorie. Folgender Satz beschreibt die asymptotischen Eigenschaften dieses Schätzers:

Satz 3.6. (ohne Beweis)

Es sei X_1, \dots, X_n eine iid Folge mit der Randverteilung F , welche vom Typ Pareto sei. $\hat{\alpha}_{k,n}^{(H)}$ bezeichne den Hill-Schätzer.

(i) **Konsistenz:**

Für $\frac{k}{n} \rightarrow 0$, wenn $k, n \rightarrow \infty$ gilt:

$$\hat{\alpha}_{k,n}^{(H)} \xrightarrow{P} \alpha, \quad n \rightarrow \infty.$$

(ii) **Asymptotische Normalverteiltheit:**

Unter gewissen weiteren Bedingungen an k und F gilt:

$$\sqrt{k} \left(\hat{\alpha}_{k,n}^{(H)} - \alpha \right) \xrightarrow{d} \mathcal{N}(0, \alpha^2), \quad n \rightarrow \infty.$$

Im Idealfall von iid Daten einer Verteilung mit regulär-variierendem Tail (insbesondere wenn die langsam-variierende Funktion sehr nahe an der Konstante des Tails ist), ist der Hill-Schätzer oft eine sehr gute Schätzung von α oder dessen Kehrwert ξ . In der Praxis ist die allgemeine Strategie Hill-Schätzer für verschiedene Werte von k zu plotten. Dies zeigt den Hill-Plot $\left\{ \left(k, \widehat{\alpha}_{k,n}^{(H)} \right) : k = 2, \dots, n \right\}$.

Im Idealfall gehen wir von einem stabilen Bereich aus, wo Schätzungen von einer kleinen Anzahl von Ordnungsstatistiken aufgebaut werden, die weitgehend vergleichbare Schätzungen abgeben.

In der Praxis können verschiedene vom Idealfall abweichende Situationen auftreten. In [2, Seite 284 ff.] werden einige „Hill-horror-plots“ beschrieben, die auf simulierte Daten basieren und die jene Situationen darstellen, welche für den Schätzer problematisch sein können:

1. Die echte zugrundeliegende Verteilung hat keinen regulär-variierenden Tail. In diesem Fall ist die Hill-Methode nicht zweckmäßig und auf derartige Daten angewendete Hill-Plots sind eher irreführend.
2. Die Verteilung hat einen regulär-variierenden Tail, aber die Funktion von langsamer Variation bewirkt einen nicht vernachlässigbaren Einfluss auf den Bereich, von dem wir unsere Daten haben.
3. Die Daten sind von einem streng stationären Zeitreihen-Prozess mit einer stationären Verteilung, welche einen regulär-variierenden Tail hat. Dies könnte die Hill-Plots undeutlich machen.

Punkt (2) und (3) beeinträchtigen ebenso die Annäherung über den GPD. In diesen Fällen ist der theoretische Tail-Index eher schwierig zu schätzen.

Für Anwendungen im Risikomanagement ist jedoch der Tail-Index von heavy-tailed Daten allein ein störender Parameter, welcher vernachlässigt werden darf. Hierbei konzentriert man sich eher auf die Tails und Schätzer von Risikomaßen. Nun betrachten wir die Standard-Strategie, welche auf den Hill-Schätzer basiert, um Tails von Verteilungen zu modellieren.

Hill-basierender Tail-Schätzer: Gesetzt den Fall, dass wir ein besonders realistisches k gewählt haben, ist es nun möglich, einen Schritt weiter zu gehen. Wir konstruieren einen Tail-Schätzer unter Verwendung des Hill-Schätzers des Tail-Indexes und geben somit das gewöhnlich heuristische Argument für dessen Konstruktion.

Die Wahl von k und die Anzahl der Ordnungsstatistik definieren die Schranke $X_{k,n}$, wobei

die Funktion von langsamer Variation in (3.30) als Konstante behandelt werden kann. Im Wesentlichen nehmen wir ein Tail-Modell dieser Form an:

$$\bar{F}(x) = Cx^{-\alpha}, \quad x \geq u > 0.$$

Die Schranke u kann durch $X_{k,n}$ ersetzt werden. Der Tail-Index α kann durch den Hill-Schätzer ermittelt werden. Nun bleibt nur noch die Konstante C zu schätzen.

Da C als $C = u^\alpha \bar{F}(u)$ dargestellt werden kann, müssen wir nur noch $\bar{F}(u)$ schätzen und der eindeutige empirische Schätzer ist $\frac{k}{n}$. Wenn man nun diese Annahmen zusammenfasst, ergibt dies den Hill-Tail-Schätzer in seiner standardisierten Form:

$$\widehat{\bar{F}}(x) = \frac{k}{n} \left(\frac{x}{X_{k,n}} \right)^{\widehat{\alpha}_{k,n}^{(H)}}. \quad (3.33)$$

Für alle Paare k und n werden sowohl der Hill-Schätzer als auch der dazugehörige Tail-Schätzer als Funktionen der k hohen Ordnungsstatistiken von der Stichprobe der Größe n betrachtet. Offensichtlich ist es möglich die Inverse des Schätzer zu ermitteln, um den Quantils-Schätzer zu erhalten. Als nächstes kann man dann den Schätzer des Expected-Shortfalls entwickeln unter der Verwendung der Argumente über die regulär-variierenden Tails.

Nun stellt sich die Frage wie man den Schätzer aus (3.33) mit dem Tail-Schätzer aus (3.25) basierend auf die GPD-Approximation zur Exzess-Verteilung vergleichen kann? Im letzteren Schätzer haben wir stets die Überlegung gehabt, dass die Stichprobe der Größe n gegeben und die Schranke u fix angenommen ist. Der Tail-Schätzer wird immer als eine Funktion der Zufallszahl N_u der höheren Ordnungsstatistiken dargestellt. Dennoch lassen sich die beiden Schätzer leicht in einer ähnlichen Darstellung umformen. Angenommen, wir umschreiben die Darstellung aus (3.33) in die aus (3.25) indem wir $\widehat{\xi}^{(H)}$, u und N_u jeweils für $1/\widehat{\alpha}_{k,n}^{(H)}$, $X_{k,n}$ und k substituieren. Dann erhalten wir:

$$\widehat{\bar{F}}(x) = \frac{N_u}{n} \left(1 + \widehat{\xi}^{(H)} \frac{x - u}{\widehat{\xi}^{(H)} u} \right)^{-1/\widehat{\xi}^{(H)}}.$$

Bei diesem Schätzer fehlt der zusätzliche Skalierungsparameter wie aus (3.25) und ist außerdem nicht so gut durchführbar. Im [2, Kapitel 7.2.5] werden einige simulierte Beispiele dargestellt, welche aufzeigen sollen, wie performierbar die beschriebenen Tail-Schätzungsmethoden in der Extremwerttheorie sind, auf die hier nicht mehr näher eingegangen wird (vgl. dazu [2, Seite 286 ff.]).

Wir verlassen nun das Kapitel der Schranken-Überschreitungen und zeigen im nachstehenden Abschnitt einige Tails von speziellen Modellen.

3.3 Tails aus speziellen Modellen

In diesem Abschnitt werden die Tails von einigen wichtigen Verteilungen und Modellen aufgezeigt, welche teilweise bereits in vorherigen Abschnitten erwähnt wurden.

3.3.1 Fréchet-Anziehungsbereich

Wie wir bereits im Abschnitt 3.1.2 gezeigt haben, setzt sich der Anziehungsbereich der Fréchet-Verteilung aus Verteilungen mit regulär-variierenden Tails der Form $\bar{F}(x) = x^{-1/\xi}L(x)$ für $\xi > 0$ zusammen. Diese sind heavy-tailed Modelle, wo die höheren Momente nicht mehr oft existieren. Die Theorie besagt außerdem, dass normierte Maxima von genug großen Blöcken sich gleich verhalten sollen wie Daten einer Fréchet-Verteilung mit dem Formparameter ξ und dass Exzesse über hinreichend hohen Schranken sich gleich verhalten sollen wie die Beobachtungen eines GPDs mit positivem Formparameter ξ .

Wir zeigen nun, dass die Student-t-Verteilung und die inverse Gamma-Verteilung zu dieser Klasse gehören. Die erstere analysieren wir aufgrund ihrer allgemeinen Wichtigkeit in Finanz-Modellen und die zweitere weil sie als Mischverteilung auftritt, welche die Student-t-Verteilung in der Klasse der Normal-Varianz-Misch-Modelle liefert (vgl. [2, Beispiel 3.6]). Im Abschnitt 3.3.3 werden wir noch sehen, dass die Mischverteilung in einem Normal-Varianz-Misch-Modell im Grunde den Tail des Modells festlegt.

Beide Verteilungen werden in Termen von deren Dichten dargestellt und die Untersuchung deren Tails zeigt eine simple Anwendung aus dem Satz von Karamata (vgl. [2, A.1.3]). Aus diesem Resultat kann man somit die langsam variierende Funktion aus dem Integral herausziehen.

Beispiel 3.2. (Student-t-Verteilung)

Eine Standard univariate t-Verteilung mit $v \geq 1$ Freiheitsgraden gehört zur allgemeinen Familie wie in [2, Beispiel 3.6] definiert. Mit der Anwendung dieser Definition erhalten wir den Fall $t_1(v, 0, 1)$ und die Dichte kann leicht so dargestellt werden:

$$f_v(x) = x^{-(v+1)}L(x).$$

Dem Satz von Karamata zufolge kann man nun

$$\bar{F}_v(x) = \int_x^\infty f_v(x)dx$$

berechnen, indem man die langsam variierende Funktion als eine Konstante behandelt. Somit erhalten wir:

$$\bar{F}_v(x) = \int_x^\infty x^{-(v+1)}L(x)dx \sim v^{-1}x^{-v}L(x), \quad x \rightarrow \infty$$

Daraus schließen wir, dass $F_v \in MDA(H_{1/v})$ ist.

Beispiel 3.3. (Inverse Gamma-Verteilung)

Die Inverse Gamma-Verteilung hat ihre Dichtefunktion der Form:

$$f_{\alpha,\beta}(x) = x^{-(\alpha+1)}L(x),$$

da $\exp(-\beta/x) \rightarrow 1$ für $x \rightarrow \infty$ und somit langsam variierend. Nach dem selben Berechnungs-Prinzip wie im vorherigen Beispiel folgt auch, dass $F_{\alpha,\beta} \in MDA(H_{1/\alpha})$ ist.

3.3.2 Gumbel-Anziehungsbereich

Eine mathematische Charakteristik der Gumbel-Klasse ist die, dass sie die sogenannte *Von-Mises-Verteilungsfunktion* enthält und auch andere Verteilungen, welche *tail-äquivalent* zu von-Mises Verteilungen sind. Die Definitionen dieser beiden Bezeichnungen folgen unten. Man beachte, dass die Verteilungen dieser Klasse beide endliche und unendliche rechte Randpunkte enthalten kann.

Man bezeichne wie gehabt $x_F = \sup \{x \in \mathbb{R} : F(x) < 1\} \leq \infty$ als den rechten Randpunkt von F .

Definition 3.9. (Von-Mises-Verteilungsfunktion)

Angenommen, es existieren gewisse $z < x_F$, sodass F folgende Darstellung hat:

$$\bar{F}(x) = c \cdot \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\}, \quad z < x < x_F,$$

wobei c eine positive Konstante, $a(\cdot)$ eine positive und absolut stetige Funktion (bezüglich eines Lebesgue-Maßes) und a' die Dichte ist mit $\lim_{x \rightarrow x_F} a'(x) = 0$.

Dann ist F die sogenannte Von-Mises-Verteilungsfunktion.

Definition 3.10. (Tail-Äquivalenz)

Zwei Verteilungsfunktionen F und G sind *tail-äquivalent*, wenn sie dieselben rechten Randpunkte (geschrieben $x_F = x_G$) und wenn

$$\lim_{x \rightarrow x_F} \frac{\overline{F}(x)}{\overline{G}(x)} = c$$

für gewisse Konstanten $0 < c < \infty$.

Für die Entscheidung, ob eine bestimmte Verteilungsfunktion F eine Von-Mises-Verteilungsfunktion ist, ist die folgende Bedingung von Bedeutung. Angenommen, es existieren gewisse $z < x_F$, sodass F auf (z, x_F) zweimal differenzierbar ist mit der Dichte $f = F'$ und $F'' < 0$ in (z, x_F) .

Dann ist F eine Von-Mises-Verteilungsfunktion genau dann, wenn gilt:

$$\lim_{x \rightarrow x_F} \frac{\overline{F}(x)F''(x)}{f^2(x)} = -1. \tag{3.34}$$

Wir verwenden diese Bedingung um zu zeigen, dass die Gamma-Verteilungsfunktion auch zu dieser Klasse gehört.

Beispiel 3.4. (Gamma-Verteilung)

Die Dichte $f = f_{\alpha, \beta}$ der Gamma-Verteilung ist so gegeben

$$f_{\alpha, \beta} = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp(-x/\beta).$$

Eine simple Berechnung liefert

$$F''(x) = f'(x) = -f(x) (\beta - (1 - \alpha)/x) < 0$$

für $x > \max((\alpha - 1)/\beta, 0)$. Offensichtlich ist nun

$$\lim_{x \rightarrow \infty} \frac{F''(x)}{f(x)} = -\beta.$$

Durch den l'Hospital erhalten wir

$$\lim_{x \rightarrow \infty} \frac{\overline{F}(x)}{f(x)} = \lim_{x \rightarrow \infty} -\frac{f(x)}{f'(x)} = \beta^{-1}.$$

Kombinieren wir nun die zwei Limiten erhalten wir (3.34).

Beispiel 3.5. (GIG-Verteilung)

Die Verallgemeinerte inverse Gaussverteilung (GIG) ist in [2, Appendix A.2.3] definiert. Sei $H_{\lambda,\chi,\psi}(x)$ die Verteilungsfunktion der $N^-(\lambda, \chi, \psi)$ -Verteilung und besitze den Fall wo $\psi > 0$ ist.

Für den Fall mit $\psi = 0$ ist die GIG eine inverse Gamma-Verteilung, welche wie im Beispiel 3.3 gezeigt wurde, zur Fréchet-Klasse gehört.

Wenn nun $\psi > 0$ ist, dann ist $\lambda \geq 0$ und unter Anwendung einer ähnlichen Technik wie im vorherigen Beispiel kann bewiesen werden, dass die GIG eine Von-Mises-Verteilungsfunktion ist. Für den Fall wenn $\lambda > 0$ ist, ist es leichter, die Tail-Äquivalenz mit einer Gamma-Verteilung zu demonstrieren, die den speziellen Fall wenn $\chi = 0$ darstellt. Man beachte, dass

$$\lim_{x \rightarrow \infty} \frac{\overline{H}_{\lambda,\chi,\psi}(x)}{\overline{H}_{\lambda,0,\psi}(x)} = \lim_{x \rightarrow \infty} \frac{h_{\lambda,\chi,\psi}(x)}{h_{\lambda,0,\psi}(x)} = c_{\lambda,0,\psi}$$

ist, für gewisse Konstante $c_{\lambda,\chi,\psi}$. Somit folgt, dass $H_{\lambda,\chi,\psi} \in MDA(H_0)$ ist.

3.3.3 Misch-Modelle

Eine 1-dimensionale Normal-Varianz-Mischung (oder die Martingal-Verteilung einer d-dimensionalen Normal-Varianz-Mischung) hat denselben Typ wie die Verteilung einer Zufallsvariable X , welche folgendes erfüllt:

$$X \stackrel{d}{=} \sqrt{W}Z, \tag{3.35}$$

wobei $Z \sim N(0,1)$ und W eine unabhängige, positive Skalar-Zufallsvariable ist. Nun würden wir etwas über die Tails von Verteilungen erfahren, welche (3.35) erfüllen.

Genauer genommen - um die Tails der Martingal-Verteilungen von elliptischen Verteilungen zu verstehen, reicht es *Sphärische Verteilungen* zu betrachten.

Satz 3.7. (ohne Beweis)

Die Zufallsvariable X hat genau dann eine sphärische Verteilung, wenn sie die stochastische Repräsentation

$$X \stackrel{d}{=} RS, \tag{3.36}$$

hat. Dabei ist S eine Komponente aus einem Zufallsvektor $\mathbf{S} = (S_1, \dots, S_d)^T$ und gleichverteilt auf der Einheitskugel $\mathcal{S}^{d-1} = \{\mathbf{s} \in \mathbb{R}^d : \mathbf{s}^T \mathbf{s} = 1\}$. R ist eine von S unabhängige und positive radiale Zufallsvariable.

Es stellt sich nun die Frage, welche Art von Tails X haben kann bzw. unter welchen Umständen wir Tails von regulärer Variation erhalten.

In [2, Abschnitt 4.3] werden streng stationäre stochastische Prozesse $(X_t)_{t \in \mathbb{Z}}$ betrachtet - ähnlich den GARCH-Prozessen, welche dieses erfüllen:

$$X_t = \sigma_t Z_t, \quad (3.37)$$

wobei Z_t ein striktes weißes Rauschen ist - typischerweise mit einer Gauss'schen oder (realistischerweise) einer skalierten Student-t-Verteilung, und σ_t eine \mathcal{F}_{t-1} messbare Zufallsvariable, welche die Volatilität darstellt.

Um nun die Tails von Misch-Modellen zu analysieren, ist der folgende Satz von Bedeutung. In weiterer Folge werden wir diesen verwenden, um die Beispiele aus (3.35), (3.36) und (3.37) zu diskutieren.

Satz 3.8. (ohne Beweis) **(Tails aus Misch-Modellen)**

Sei X gegeben durch $X = YZ$ für unabhängige, nicht-negative Zufallsvariablen Y und Z , sodass gilt

1. Y hat einen regulär-variiierenden Tail mit dem Tail-Index α ,
2. $\mathbb{E}(Z^{\alpha+\epsilon}) < \infty$ für $\epsilon > 0$.

Dann hat X einen Tail von regulärer Variation mit dem Tail-Index α und es gilt:

$$P(X > x) \sim \mathbb{E}(Z^\alpha) P(Y > x), \quad x \rightarrow \infty.$$

Beispiel 3.6. **(Tails aus der Normal-Varianz-Mischung)**

Man betrachte eine univariate Normal-Varianz-Mischverteilung, welche (3.35) erfüllt für $Z \sim N(0, 1)$ und eine unabhängige positive Zufallsvariable W . Der Tail der symmetrischen Verteilung von X kann nun durch Betrachtung von $\sqrt{W} |Z|$ analysiert werden. Angenommen, W ist regulär-variiierend mit dem Tail-Index α , ähnlich der inversen Gamma-Zufallsvariable aus dem Beispiel 3.3.

Es ist nun leicht prüfbar, dass $P(\sqrt{W} > x) = L(x^2)x^{2\alpha}$ gilt, sodass \sqrt{W} einen regulär-variiierenden Tail mit dem Tail-Index 2α hat. Aus dem Satz 3.8 folgt nun, dass X einen regulär-variiierenden Tail hat mit demselben Tail-Index. Somit stimmen die Beispiele 3.2 und 3.3 überein, da eine Zufallsvariable $W \sim Ig(v/2, v/2)$ als eine Misch-Variable verwendet werden kann, um eine $t_1(v, 0, 1)$ -Verteilung zu kreieren.

Es ist nun eine hinreichende Bedingung für X , um einen regulär-variierenden Tail zu besitzen, wenn W einen regulär-variierenden Tail hat. Dies bedeutet, dass man den Tail-Index des ersteren durch Verdoppeln des zweiten erhält. Genauer genommen, ist dies sogar eine notwendige Bedingung.

Wenn die Verteilungsfunktion F_W von W dies erfüllt:

$$F_W \in MDA(H_0), \text{ dann ist } \mathbb{E}(W^k) < \infty \text{ für alle } k > 0,$$

folgt daraus

$$\mathbb{E}|X|^k < \infty \text{ für alle } k > 0.$$

Somit kann dann X keinen regulär-variierenden Tail haben. Dies würde bedeuten, dass beispielsweise die univariate hyperbolische und Normal-Inverse Gauss'sche Verteilung keine Power-Tails haben, da die GIG im Maximalen Anziehungsbereich von Gumbel liegt für alle Fälle ausgenommen dem Spezial-Fall wenn $\psi = 0$. Dies entspricht einer inversen Gamma-Verteilung.

Beispiel 3.7. (Tails von Sphärischen Verteilungen)

Man betrachte die univariate Verteilung einer Einzel-Komponente X eines sphärischen Vektors, welcher in (3.36) gegeben ist. Da $|S|$ eine Zufallsvariable mit endlichem Träger $[0, 1]$ ist (wie im Beispiel 3.6 bewiesen), können wir nur einen Tail von regulärer Variation erhalten, wenn die Zufallsvariable R einen regulär variierenden Tail besitzt. Satz 3.8 beschreibt außerdem, dass R und X denselben Tail-Index haben müssen.

Analysen von Tails von stationären Verteilungen eines GARCH-Typ-Modells bringen mehr Herausforderung. In Anbetracht des Satzes 3.8 und den obigen Beispielen, ist es klar, dass wenn Z_t „Gaussisch“ ist, dann hat X_t (wie in (3.37) beschrieben) einen regulär variierenden Tail, wenn die Volatilität σ_t einen regulär variierenden Tail besitzt. Welche Art von Tails würde σ_t haben, wenn wir beispielsweise ein $GARCH(1,1)$ -Modell mit $\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ betrachten würden? Eine allgemeine Lösungsvariante für derartige Prozesse unter beliebigen Annahmen von Verteilungs-Innovationen ist im folgenden Satz gegeben:

Satz 3.9. (ohne Beweis) (Tails von GARCH(1,1))

Sei $(X_t)_{t \in \mathbb{Z}}$ ein streng stationärer $GARCH(1,1)$ -Prozess. Angenommen die Parameter α_1 und β_1 und die Verteilungs-Innovation von Z_0 erfüllen die folgenden Bedingungen:

- (i) Z_0 hat eine positive Dichte,
- (ii) $\mathbb{E}(\ln A_t) = \mathbb{E}(\ln(\alpha_1 Z_t^2 + \beta_1)) < 0$, welches die Bedingung für eine einheitliche stationäre Lösung zu $GARCH(1,1)$ -Gleichheiten ist, und
- (iii) Es existiert ein $s_0 < \infty$, sodass $E|Z_t|^s < \infty$ für alle $s < s_0$ und $E|Z_t|^{s_0} = \infty$ oder $E|Z_t|^{s_0} < \infty$ für alle $s \geq 0$ und $\lim_{s \rightarrow \infty} |Z_t|^s = \infty$.

Dann folgt:

(A) Die Gleichheit

$$EA_t^{s/2} = 1 \quad (3.38)$$

hat eine einheitliche positive Lösung κ .

(B) Die stationären Verteilungen von X_t und σ_t haben Tails, welche folgendes erfüllen:

$$P(|X| > x) \sim \mathbb{E}(|Z_t|^\kappa) P(\sigma_t > x) \sim c_0 x^{-\kappa}, \quad x \rightarrow \infty, \quad (3.39)$$

für gewisse $c_0 > 0$.

Unter den Bedingungen des Satzes, hat folglich ein streng stationärer $GARCH$ -Prozess eine heavy-tailed stationäre Verteilung und der Tail-Index ist als Lösung aus (3.38) gegeben.

Kapitel 4

Prämienkalkulation für heavy-tailed Risiken

Das nachstehende Kapitel basiert auf dem Paper „Premium Calculation for Fat-tailed Risk“ von Roger Gay.

Versicherungsschäden, die einer heavy-tailed Verteilung zugrundeliegen, verursachen erhebliche Fehler über den Wert des Tail-Indexes. Roger Gay stellt in seinem Paper „Premium Calculation for Fat-tailed Risk“ unter Verwendung der Theorie von Risikoaversionen ein neues Prämienprinzip vor, das Potenzprinzip - analog zum Exponentialprinzip für thin-tailed Schäden. Für Schäden, die einer Verallgemeinerten Pareto-Verteilung zugrundeliegen, ist die daraus resultierende Prämie als das Verhältnis der zwei größten erwarteten Schäden definiert, welches für das Verhältnis des aktuellen Schadens ein erwartungstreuer und konsistenter Schätzer ist. Während thin-tailed Schadens-Prämien weitgehend durch die ersten zwei Momente der Schadens-Verteilung bestimmt sind, werden heavy-tailed Schadens-Prämien durch die ersten zwei Extrema bestimmt.

Das Konzept der Risikoaversion ist ein natürliches Modell, das Tail-Index-Unsicherheiten in den Prämien verursacht, jedoch die Basis-Verhältnis-Struktur unverändert lässt.

Das Potenzprinzip kann verwendet werden, um ein zusammenhängendes System zu bestimmen, in welchem die Prämien festgelegt werden, wenn der Fehler im Tail-Index geeignet modelliert wurde. Die „Fehler-Prämien“ bzw. Risikoaversions-Prämien, die auf die Tail-Index-Unsicherheit zurückzuführen sind, können entweder als gewichtetes Mittel von Risikoaversions-Prämien, die vom Potenzprinzip abgeleitet werden, oder als den „Bayes’schen posterior Mittelwert“ betrachtet werden. Die neue Prämienkalkulationsmethode - das Potenzprinzip für heavy-tailed Risiken - wird als analoge Kalkulationsmethode zum Exponentialprinzip für thin-tailed Risiken beschrieben.

4.1 Thin-tailed Risiken und das Exponentialprinzip

Wie bereits im Abschnitt 2.1.7 beschrieben, wird die Prämie \mathcal{P} nach dem Exponentialprinzip für einen Schaden X wie folgt berechnet:

$$\mathcal{P} = \frac{1}{a} \ln M_X(a), \quad a > 0 \quad (4.1)$$

und

$$\mathcal{P} \approx \mu + \frac{1}{2}a\sigma^2 + o(a) \quad (4.2)$$

wobei $M_X(\cdot)$, μ , σ^2 jeweils die Momenterzeugende Funktion, der Mittelwert und die Varianz von X sind. Der Parameter a bezeichne die konstante absolute Risikoaversion des Versicherers. Und für $o(a)$ gilt:

$$\frac{\mathcal{P} - (\mu + \frac{1}{2}a\sigma^2)}{a} \rightarrow 0 \quad \text{für } a \rightarrow 0. \quad (4.3)$$

Das Exponentialprinzip wird für den jährlichen Gesamtschaden S_N angewendet, wobei

$$S_N = X_1 + X_2 + \dots + X_N \quad (4.4)$$

mit $\{X_i\}$ iid verteilt und N die von $\{X_i\}$ unabhängige Anzahl der Schäden. Häufig wird eine poisson- oder negativ binomial-verteilte Zufallsvariable angenommen, da

$$M_{S_N}(a) = M_N[\ln \{M_X(a)\}] \quad (4.5)$$

mit X als ein beliebiger repräsentative Schaden, sodass

$$\mathcal{P} = \mathbb{E}[X] \times \mathbb{E}[N] + \frac{1}{2}a \{ \text{Var}(X)\mathbb{E}[N] + \mathbb{E}[X]^2\text{Var}(N) \} + o(a). \quad (4.6)$$

In diesem Sinne erhalten wir eine einigermaßen vollständige Theorie für thin-tailed Risiken zur Berechnung der Gesamtprämie, welche von einzelnen Verteilungs-Annahmen von X und N abgeleitet sind und ebenso die Ergebnisse für stop-loss-Versicherungen enthält.

Das Exponentialprinzip verwendet praktischerweise die *Pricing-Funktion*

$$m_\theta(x) = \exp(\theta x) \quad (4.7)$$

mit dem Risikoparameter $\theta = a$ und der *Pricing-Regel*:

$$m_\theta(\mathcal{P}) = \mathbb{E}[m_\theta(X)]. \quad (4.8)$$

Satz 4.1. (ohne Beweis)

Verwendet man nun die *Pricing-Regel* (4.8) in Verbindung mit einer *Pricing-Funktion* $m_\theta(x)$, so kann man folgende Vereinheitlichungen von Prämien-Prinzipien festlegen:

- (i) Mit $m(x) = x$ führt dies zu einer risikoneutralen Prämie $\mathcal{P} = \mathbb{E}[X]$.
- (ii) Mit $m_a(x) = \exp(ax)$ führt dies zum Exponentialprinzip und ebenso die daraus abgeleiteten Variationen.
- (iii) Für Schäden $X > 0$ mit absolut stetiger Verteilungsfunktion $F(x)$ und

$$m_\theta(x) = \{F(x)\}^\theta, \quad \theta > 0 \quad (4.9)$$

in Verbindung mit (4.8) führt dies zum Quantilsprinzip. Dabei ist die Prämie \mathcal{P} als ein geeignetes Quantil der Schadensverteilung gewählt mit

$$F(\mathcal{P}) = q(\theta) = (\theta + 1)^{-1/\theta}. \quad (4.10)$$

Beachte, dass $q(\theta)$ streng monoton steigend ist mit θ und $e^{-1} < q(\theta) < 1$.

- (iv) Das „Potenzprinzip“ für heavy-tailed Schäden, welche im folgendem Abschnitt beschrieben wird, hat somit die *Pricing-Funktion*

$$m_\alpha(x) = x^{\alpha+1}, \quad \alpha > 0. \quad (4.11)$$

4.2 Heavy-tailed Risiken und das Potenzprinzip

Individuelle Schäden X_i aus (4.4) liegen einer heavy-tailed Verteilung $F(x)$ zugrunde wenn gilt:

$$1 - F(x) = x^{-\delta} L(x), \quad \text{für große } x > 0, \delta > 0. \quad (4.12)$$

$L(x)$ sei langsam variierend bei ∞ . Somit gilt für $a > 0$

$$\lim_{x \rightarrow \infty} \frac{L(ax)}{L(x)} = 1.$$

Daher hat die Schadens-Zufallsvariable X die Momente $\mathbb{E}[X^k]$ nur für $k < \delta$.

4.2.1 Pareto- und Verallgemeinerte Pareto-Verteilung

Das Potenzprinzip behandelt Pareto-Schäden von Verteilungen der Form:

$$F(v) = 1 - \left(1 + \frac{v}{\lambda}\right)^{-\delta}, \quad v > 0, \delta > 0. \quad (4.13)$$

Die Formel (4.13) beschreibt virtuell die einfachste Form einer reinen Tail-Verteilung. Wie bereits im Abschnitt 3.2 beschrieben, hat diese in der Extremwerttheorie eine viel wichtigere Bedeutung. Für ausreichend große Schranken u nähert sich die Funktion aus (4.13) für eine umfassende Klasse von Zufallsvariablen einer Schranken-Verteilung über u . In diesem Fall ist δ unabhängig von u und λ nicht.

Wenn der Schaden V der Verteilung (4.13) genügt, so ergibt sich für $X = \frac{V}{\lambda}$ die vereinfachte Verteilung von X mit:

$$F(x) = 1 - (1 + x)^{-\delta}, \quad v > 0, \delta > 0. \quad (4.14)$$

Heavy-tailed Schäden mit limitierten Momenten haben keine Momenterzeugende Funktionen und daher keine konstante risiko-averse Prämien unter (4.1).

4.2.2 Typ I Extremwert - Fréchet-Verteilung

Eine andere Verteilung, die zur Modellierung von heavy-tailed Schäden verwendet werden kann und vor allem eine umfassende Klasse von heavy-tailed Verteilungen $F(\cdot)$ vom Typ aus (4.12) enthält, ist die Fréchet-Verteilung.

Die Dichte des k ten Extremas $X_{(k)}^*$, der Grenzwert der normierten Ordnungsstatistik $X_{(n-k+1)}/v_n$ der Verteilung $F(\cdot)$ für große n ist definiert durch:

$$f_k(x) = \frac{\delta x^{-k\delta-1} \exp\{-x^{-\delta}\}}{\Gamma(k)}, \quad (k \geq 1, x > 0, \delta > 0) \quad (4.15)$$

Die Abfolge für normierende Konstanten $\{v_n\}$ - der Tail-Quantil-Funktion - wird von $n\{1 - F(v_n)\} = 1$ für alle k abgeleitet (siehe [6, Seite 167]).

Die Tail-Quantil-Funktion spielt eine große Rolle in der Extremwerttheorie - unter anderem die Bestimmung der Zugehörigkeit von $F(\cdot)$ innerhalb des Anziehungsbereichs von (4.15). Somit ist $v_n = v_n(\rho)$, $\rho = 1/\delta$ verteilungs-spezifisch. Ein Beispiel: $v_n = \lambda n^\rho$ für Pareto aus (4.13).

4.2.3 Relation zwischen Verallgemeinerte Pareto- und Fréchet-Prämien

Der Fokus in diesem Paper bezieht sich auf große Schäden von heavy-tailed Verteilungen. Die Pareto-Verteilung aus (4.13) ermöglicht die Bestimmung der Prämie sofern die Schäden eine hinreichend große Schranke überschreiten. Die Fréchet-Dichte beschreibt das Verhalten des Grenzwertes von gereihten Schäden.

Wie bereits im Abschnitt 3.3.1 erwähnt, ist die Klasse der heavy-tailed Verteilungen von (4.12), die einer verallgemeinerten Pareto-Verteilung zugrundeliegt wenn die Exzess-Überschreitungen ein hinreichend hohes Level erreicht, genau die Klasse innerhalb des Fréchet-Anziehungsbereichs.

4.2.4 Das Potenzprinzip

Für heavy-tailed Risiken kann die Prämie \mathcal{P} mithilfe des Potenzprinzips ermittelt werden. Abgeleitet von der Theorie der Risikoaversionen ist die Versicherungs-Prämie impliziert mit der konstanten Risikoaversion und gemessen durch den Parameter α über die Pricing-Funktion

$$m(x) = x^{\alpha+1} \quad (x > 0, \alpha \geq 0) \quad (4.16)$$

und über die Pricing-Regel aus (4.8).

Für die Schadens-Zufallsvariable X führt das Potenzprinzip zu einer Prämie \mathcal{P} , die durch den Risikoaversions-Koeffizienten α bestimmt ist, mit

$$\mathcal{P}(\alpha) = \left\{ \mu'_{\alpha+1} \right\}^{\frac{1}{\alpha+1}} \quad (4.17)$$

wobei μ'_k das k te Moment von X ist mit $\mu'_1 = \mu = \mathbb{E}[X]$.

Satz 4.2. (Eigenschaften der Potenzprinzip-Prämien) (nach [6, Seite 168])
Prämien, welche durch (4.17) festgelegt werden, haben gewisse wünschenswerte Eigenschaften für eine „gute Prämie“ (siehe [7, Seite 79]).

- (i) $\mathcal{P}(0) = \mu$; $\mathcal{P}(\alpha)$ ist wachsend in $\alpha + 1 \geq 0$, wobei die Risikoaversion wachsend in α ist.
- (ii) Für $\alpha \geq 0$ ist \mathcal{P} subadditiv: $\mathcal{P}_{X+Y}(\alpha) \leq \mathcal{P}_X(\alpha) + \mathcal{P}_Y(\alpha)$. Geteilte Risiken reduzieren aber nicht die Prämien.
- (iii) *Stochastische Dominanz:*
 Für $\alpha \geq 0$, $X > Y \Rightarrow \mathcal{P}_X(\alpha) > \mathcal{P}_Y(\alpha)$
 (dies folgt aus der Tatsache, dass für $X > 0$ gilt, $\mathbb{E}[X^{\alpha+1}] = \int_0^\infty (1 - F_X(x^{\frac{1}{\alpha+1}})) dx$).
- (iv) Für $\alpha \geq 0$: $\mathcal{P}_a(\alpha) = a$ für alle Konstanten $a > 0$ (keine „ungerechtfertigte Prämien-Gewichtung“).
- (v) *Proportionalität:* Für $\alpha \geq 0$: $\mathcal{P}_{aX}(\alpha) = a\mathcal{P}_X(\alpha)$

4.3 Risikoaverse Prämien

4.3.1 Prämien für große heavy-tailed Schäden

Wenn der Schaden X die Verteilung $F(\cdot)$ aus (4.12) besitzt, dann ist seine bedingte Verteilung über ein hinreichend großes Level u die Pareto aus (4.13) mit dem Parameter λ . Um die weiteren Ziele zu erreichen, ist es hilfreich den Parameter λ mit in den Schaden einzubinden. Somit erhalten wir unter Verwendung von (4.17) in der Pareto-Verteilung aus (4.13) für den Schaden $(1 + \frac{X}{\lambda})$ die Prämie \mathcal{P} durch:

$$\begin{aligned}\mathcal{P}^{\alpha+1} &= \lambda^{-1} \int_0^\infty \left(1 + \frac{x}{\lambda}\right)^{\alpha+1} \delta \left(1 + \frac{x}{\lambda}\right)^{-\delta-1} dx \\ &= \delta \{\delta - (\alpha + 1)\}^{-1}\end{aligned}$$

woraus wir folgendes erhalten:

$$\mathcal{P}(\alpha) = \left\{1 - \frac{\alpha + 1}{\delta}\right\}^{-1/(\alpha+1)} \quad (4.18)$$

Nun re-parametrisieren wir (4.18) als

$$\mathcal{P}_\beta = \left(1 - \frac{\rho}{\beta}\right)^{-\beta} \quad (4.19)$$

wobei $\beta = (\alpha + 1)^{-1}$ die Gebühr der Versicherer ausmacht, welche für tragende Schadensrisiken verrechnet wird und einer Verteilung mit Tail-Index $\rho = 1/\delta$ zugrundeliegt.

Beispiel 4.1. Für $\rho = 0,75$ und $\beta = 0,95$ ist $\mathcal{P}_{0,95} = 4,39$. Vergleicht man sie mit $\mathcal{P}_{1,0} = \mu = 4$ impliziert die Gebühr von $\beta = 0,95$ eine etwa 10%ige Überschreitung der risikoneutralen Prämie wenn $\rho = 0,75$ ist.

Bemerkung 4.1.

(i) *Maximale Tail-Heaviness:*

Die Formel (4.19) zeigt, dass der Level der konstanten Risikoaversion α des Versicherers den *maximalen Tail-Index* $\beta = (\alpha + 1)^{-1} = \rho_{\max} = 1/\delta_{\min}$ bestimmt, den der Versicherer für eine bestimmte Klasse von heavy-tailed Versicherungsschäden einsetzen kann.

(ii) *Wenn kein erwarteter Schaden $\mathbb{E}[X]$ existiert:*

Die Formel (4.19) setzt voraus, dass $\delta > 1$ ($\rho < 1$) ist und $E[X]$ existiert. Wenn $\delta \leq 1$ ($\rho \geq 1$) ist, kann die Versicherungsdeckung für einen Teil von einem Pareto-Schaden $(1 + X/\lambda)^\varphi$, ($\varphi\rho < 1$) angeboten werden.

Mit der Verwendung von (4.17) ergibt die Prämie

$$\mathcal{P}_\beta = \left(1 - \frac{\rho\varphi}{\beta}\right)^{-\beta} \quad (4.20)$$

für die Teil-Pareto-Schäden.

(iii) *Risiko-neutrale Prämien:*

Für Pareto aus (4.13) zeigt die direkte Integration, dass $\mathbb{E}[1 + X/\lambda] = (1 - \rho)^{-1}$ ist und konform mit der Prämienformel aus (4.19) mit $\beta = 1$ ist.

Für $X_{(k)}$ als die k te Pareto-Ordnungsstatistik gilt:

$$(1 - \rho)^{-1} = \frac{\mathbb{E}[1 + X_{(n)}/\lambda]}{\mathbb{E}[1 + X_{(n-1)}/\lambda]} \quad (4.21)$$

wobei die Relation unabhängig von der Stichproben-Größe n ist.

Der Beweis dazu folgt nachstehend.

Beweis von Bemerkung 4.1. (iii):

Für Pareto aus (4.14) und jene Werte v mit $1 \leq k \leq v < k + 1 \leq n$ ist die v te fraktionelle Ordnungsstatistik $X_{(k,n)}^{(v)}$, interpoliert von $v = k$ nach $v < k + 1$, definiert durch ihre Dichte auf $(0, \infty)$, welche gegeben ist durch

$$f_v(x) = \frac{\delta [1 - (x + 1)^{-\delta}]^{v-1} (x + 1)^{-(n-v)\delta} (x + 1)^{-\delta-1}}{B(v, n + 1 - v)} \quad (4.22)$$

und analog zur k ten Ordnungsstatistik ist, wobei $B(v, n + 1 - v) = \Gamma(v)\Gamma(n + 1 - v)/\Gamma(n + 1)$.

Somit ist die v te Ordnungsstatistik gleich der k ten Ordnungsstatistik, wenn $v = k$ ein Integer ist, jedoch interpoliert von der k ten ausgenommen der $(k + 1)$ ten Ordnungsstatistik ($1 \leq k \leq n - 1$) und impliziert durch ihre Dichte aus (4.22).

Verwendet man nun (4.22), ist es einfach folgendes zu zeigen:

$$\begin{aligned} \mathbb{E} \left[1 + X_{(k,n)}^{(v)} \right] &= B(v, n + 1 - v - \rho) / B(v, n + 1 - v) \\ &= \{ \Gamma(v) \Gamma(n + 1 - v - \rho) / \Gamma(n + 1 - \rho) \} \times \Gamma(n + 1) / \Gamma(v) / \Gamma(n + 1 - v) \\ &= \Gamma(n + 1 - v - \rho) \Gamma(n + 1) / \{ \Gamma(n + 1 - \rho) \Gamma(n + 1 - v) \} \end{aligned}$$

Ein Beispiel, wenn $v = n - k + 1$ ist:

$$\mathbb{E} [1 + X_{(n-k+1)}] = \Gamma(k - \rho) \Gamma(n + 1) / \{ \Gamma(n + 1 - \rho) \Gamma(k) \} \quad (4.23)$$

$$\sim n^\rho \Gamma(k - \rho) / \Gamma(k) \quad (4.24)$$

$$= \mathbb{E} [X_{(1)}^*]$$

und

$$\mathbb{E} [1 + X_{(n-1)}] = (1 - \rho) \mathbb{E} [1 + X_{(n)}] \quad (4.25)$$

Wir erhalten ebenso für Pareto:

$$\begin{aligned} \frac{1 + \mathbb{E} [X_{(n)}]}{1 + \mathbb{E} [X_{(n-k,n)}^{(v)}]} &= \frac{\Gamma(1 - \rho) \Gamma(n + 1) / \Gamma(n + 1 - \rho)}{\Gamma(v + 1 - \rho) \Gamma(n + 1) / \{ \Gamma(n + 1 - \rho) \Gamma(v + 1) \}} \\ &= \{ \Gamma(1 - \rho) \Gamma(v + 1) \} / \Gamma(v + 1 - \rho) \\ &= v B(1 - \rho, v) \end{aligned}$$

,

unabhängig von n .

Insbesondere ist für $v = 1$, $k = n - 1$:

$$\frac{\mathbb{E} [1 + X_{(n)}]}{\mathbb{E} [1 + X_{(n-1)}]} = (1 - \rho)^{-1} \quad (4.26)$$

Es gibt einen echten Grenzwert für das Verhältnis der zwei größten Extremas für alle heavy-tailed Verteilungen im Fréchet-Anziehungsbereich. Für $n \rightarrow \infty$ gilt:

$$\frac{\mathbb{E} [1 + X_{(n)}/\lambda]}{\mathbb{E} [1 + X_{(n-1)}/\lambda]} \rightarrow \frac{\mathbb{E} [v_n X_{(1)}^*]}{\mathbb{E} [v_n X_{(2)}^*]} = \frac{\mathbb{E} [X_{(1)}^*]}{\mathbb{E} [X_{(2)}^*]} = (1 - \rho)^{-1}$$

□

Für die Verhältnisse aus anderen erwarteten Ordnungsstatistiken und Extremas würden wir ähnliche Resultate erhalten. Ferner gilt allgemein für $n \rightarrow \infty$:

$$\frac{\mathbb{E} [1 + X_{(n)}/\lambda]}{\mathbb{E} [1 + X_{(n-k)}/\lambda]} \rightarrow \frac{\mathbb{E} [X_{(1)}^*]}{\mathbb{E} [X_{(1+k)}^*]} = kB(k, 1 - \rho) \quad (4.27)$$

wobei $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$, $a > 0$, $b > 0$.

Da $\rho \uparrow 1$ wird das Verhältnis aus (4.21) groß. Der erwartete Wert des zweitgrößten Schadens ist, verglichen mit dem erwarteten Wert des größten Schadens, klein.

4.3.2 Total erwartete Schäden und größt-erwartete Schäden

Aus (4.21) erhalten wir die Beziehung

$$\frac{\mathbb{E} [1 + X_{(n-1)}/\lambda]}{\mathbb{E} [1 + X_{(n)}/\lambda]} = (1 - \rho)$$

$$\frac{\mathbb{E} [1 + X_{(n-2)}/\lambda]}{\mathbb{E} [1 + X_{(n)}/\lambda]} = (1 - \rho)(1 - \rho/2)$$

$$\frac{\mathbb{E} [1 + X_{(n-3)}/\lambda]}{\mathbb{E} [1 + X_{(n)}/\lambda]} = (1 - \rho)(1 - \rho/2)(1 - \rho/3)$$

etc.

In diesem Sinne können die jährlichen Gesamtschäden S_N aus (4.4) für Versicherungen mit $\rho \in [\frac{1}{2}, 1)$ als vorwiegend bedingt durch den größten oder die größten Schäden erwartet werden. Aus diesen Beziehungen folgen diese numerischen Werte:

- (i) für $\rho = \frac{1}{2}$ folgen die Verhältnisse $\frac{1}{2}, \frac{3}{8}, \frac{5}{16}, \dots$ im Bezug auf in „absteigender Ordnung“ erwartete große Schäden zum größten erwarteten Schaden,
- (ii) für $\rho = \frac{3}{4}$ folgen die Verhältnisse $\frac{1}{4}, \frac{5}{32}, \frac{15}{128}, \dots$
- (iii) *Existenz eines erwartungstreuen Schätzers:*
Besonders bemerkenswert für Pareto-Schäden ist, dass die risikoneutrale Prämie für den Schaden $(1 + X/\lambda)$ einen erwartungstreuen Schätzer zulässt. Sei der Einfachheit halber $\lambda = 1$. Dann gilt:

$$\begin{aligned} \mathbb{E} \left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}} \mid X_{(n-k)}, X_{(n-k-1)}, \dots \right] &= \mathbb{E} \left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}} \mid X_{(n-k)} \right] \\ &= \mathbb{E} \left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}} \right] \\ &= \frac{\mathbb{E} [1 + X_{(n)}]}{\mathbb{E} [1 + X_{(n-k)}]} \\ &\stackrel{*}{=} kB(k, 1 - \rho) \end{aligned}$$

* siehe [6, Seite 184 ff.]

und

$$\mathbb{E} \left[\frac{1 + X_{(n)}}{1 + X_{(n-k)}} \right] = \mathbb{E} \left[\frac{X_{(1)}^*}{X_{(1+k)}^*} \right] = \frac{\mathbb{E} [X_{(1)}^*]}{\mathbb{E} [X_{(1+k)}^*]} = kB(k, 1 - \rho) \quad (4.28)$$

(Beweis dazu siehe [6, Seite 184 ff.])

Ein konsistenter und erwartungstreuer Schätzer \hat{P} für Pareto-Schäden in der Form $(1 + X/\lambda)$ ist

$$\hat{P} = \frac{X_{(1)}^*}{X_{(2)}^*} \quad (4.29)$$

Aus (4.29) resultiert der folgende Satz:

Satz 4.3.

Prämien für thin-tailed Schäden lassen sich weitgehend aus den ersten zwei Momenten der Verteilung bestimmen. Für heavy-tailed Schäden hingegen, werden die Prämien überwiegend aus den ersten zwei Extremas ermittelt.

4.4 Risikoaverse Prämien beim ungewissen Tail-Index

4.4.1 Heavy-tailed Schadensmodelle

Eine prinzipielle Anwendung von heavy-tailed Verteilungen ist die Modellierung der Klassen von Versicherungen, die einige potentiell sehr große individuelle Risiken aufweisen. Dabei ist an Katastrophen, Naturgewalten, Betriebshaftpflicht, Schadenersatzleistungen bzw. Gewerbliche Feuerversicherung zu denken. Um für (4.4) ein sinnvolles Modell für sehr große Risiken zu bestimmen, muss die Varianz dieses individuellen Schadens unendlich sein. Andernfalls unterliegen die Gesamtschäden dem Zentralen Grenzwertsatz. Das heißt, dass der Tail-Index δ aus (4.12), (4.13) und (4.14) den Wert 2 nicht übersteigt (oder $\rho \geq \frac{1}{2}$). Die größte Anordnung der Ordnungsstatistik hat keinen zweiten Moment. Alle anderen Ordnungsstatistiken hingegen schon (vgl. [6, Seite 171]).

Tail-Index-Schätzungen sind umfangreich erforscht - für gewöhnlich unter der Annahme, dass große Stichproben-Größen verfügbar sind. (vgl. [6, Seite 171 ff]).

Viele vertraute Prinzipien für die Bestimmung von thin-tailed Risiken (das Exponentialprinzip bzw. das Varianzprinzip) können nicht für Schadenmodellierungen herangezogen werden, wenn sie keinen zweiten Moment besitzen. Außerdem ist für Schäden X von der Form Pareto (4.13) der Mittelwert $\mu = \mathbb{E}[1 + X/\lambda] = (1 - \rho)^{-1}$ ein seltenes Ereignis wie $\rho \uparrow 1$ im Sinne von $P(X > \mu) \rightarrow 0$.

Der Mittelwert ist größer als alle Quantile der Verteilung, da $\rho \rightarrow 1$. Da wir ρ (also auch μ) nicht exakt kennen, ist es eher gewagt, das Quantilprinzip aus (4.10) zu verwenden, um die Prämie mittels Schätzen von ρ zu bestimmen, wenn ρ im Bereich von $[\frac{1}{2}, 1)$ liegt. Die Prämie \mathcal{P} muss dann als ein Quantil $q(\alpha)$ mit $\alpha > \alpha_\mu$ gewählt werden, wobei $F(\mu) = q(\alpha_\mu)$ ist und μ unbekannt. Bei geeigneten Termen kann dieses Defizit durch Modellierung von „Tail-Index-Ungewissheit“ aufgehoben werden. Somit ist es anzunehmen, dass der Schadens-Mittelwert $\mathbb{E}[X]$ existiert, sodass $\rho < 1$ ist.

4.4.2 Ein Modell bei Ungewissheit von Tail-Heaviness

Das Funktional der Prämie $\mathcal{P}_\beta = (1 - \rho/\beta)^{-\beta}$ aus (4.19), welches durch $\beta = \rho_{max}$ begrenzt ist, deutet auf die Familie von transformierten Beta-Variablen hin:

$$f_\rho(x) = \frac{vx^{v-1}}{\beta^v} \quad \text{mit } 0 \leq x < \beta; v \geq 1 \quad (4.30)$$

als ein natürliches Modell für die Ungewissheit in ρ .

Die Berücksichtigung von Ungewissheit in ρ kann bei der Berechnung von „unsichere Prämien“ angewendet werden wie die Mittelwerte von risikoaversen Prämien \mathcal{P}_β , gewichtet durch (4.30).

Als Alternative kann die Verteilung (4.30) als Bayes'sche Prior für ρ mit Prämien $\mathcal{P}(v, \beta) = \mathbb{E}[\mathcal{P}_\beta]$ und das resultierende Posterior-Mittel für festgelegtes v betrachtet werden.

4.4.3 Eigenschaft von unsicheren Prämien

Die Prämie $\mathcal{P}_{v,\beta}$ für den Schaden $[1 + X/\lambda]$ kann als den Mittelwert oder Posterior-Mittel von einer Pareto-Prämie erlangt werden (i.e. der Erwartungswert der risikoaversen Prämie \mathcal{P}_β unter Verwendung der Dichte aus (4.30).

Somit folgt:

$$\mathcal{P}(v, \beta) = \mathbb{E}[\mathcal{P}_\beta] \tag{4.31}$$

$$\mathcal{P}(v, \beta) = \int_0^\beta v\beta^{-v} x^{v-1} (1 - x/\beta)^{-\beta} dx = vB(\theta, 1 - \beta) \tag{4.32}$$

Beweis:

Durch Substitution von $u = x/\beta$ erhalten wir:

$$\begin{aligned}
 \mathcal{P}(v, \beta) &= v \int_0^1 u^{v-1} (1-u)^{-\beta} du \\
 &= v \left(\underbrace{-\frac{(1-u)^{-\beta+1} u^{v-1}}{-\beta+1} \Big|_0^1}_{=0} + \frac{v-1}{-\beta+1} \int_0^1 u^{v-2} (1-u)^{-\beta+1} du \right) \\
 &= v \cdot \frac{v-1}{-\beta+1} \left(\underbrace{-\frac{(1-u)^{-\beta+2} u^{v-2}}{-\beta+2} \Big|_0^1}_{=0} + \frac{v-2}{-\beta+2} \int_0^1 u^{v-3} (1-u)^{-\beta+2} du \right) \\
 &= \dots \\
 &= v \cdot \frac{v-1}{-\beta+1} \cdot \frac{v-2}{-\beta+2} \cdot \dots \cdot \frac{1}{-\beta+v-1} \int_0^1 (1-u)^{-\beta+v-1} du \\
 &= v \cdot \frac{v-1}{-\beta+1} \cdot \frac{v-2}{-\beta+2} \cdot \dots \cdot \frac{1}{-\beta+v} \\
 &= v \cdot \frac{(v-1)!}{(-\beta+1) \cdot (-\beta+2) \cdot \dots \cdot (-\beta+v)} \\
 &= v \cdot B(\theta, 1-\beta)
 \end{aligned}$$

□

Ersetzt man den Exponenten β durch 2β aus dem Integral im (4.32) erhalten wir

$$\mathbb{E}[\mathcal{P}_\beta^2] = vB(\theta, 1-2\beta) \quad \text{für } x \in [0, \beta)$$

Es gibt somit keinen zweiten posterior Moment (und keine posterior Varianz) der Prämie $\mathcal{P}_\beta = (1 - \rho/\beta)^{-\beta}$, für alle $\beta > \frac{1}{2}$. Die Schäden haben ebenso keine Varianz dieser Werte. Man bezeichnet $\mathcal{P}(v, \beta)$ als *die unsichere Prämie* (vgl. [6, Seite 173]).

4.4.4 Relations-Struktur der unsicheren Prämie $\mathcal{P}(v, \beta)$

Für Pareto-Schäden, wenn k ein Integer und $\mathcal{P}(k, \beta) = kB(k, 1 - \beta)$ ist, wobei

$$kB(k, 1 - \beta) = \frac{\mathbb{E}[1 + X_{(n)}]}{\mathbb{E}[1 + X_{(n-k)}]} = \mathbb{E}\left[\frac{X_{(1)}^*}{X_{(1+k)}^*}\right] \quad (4.33)$$

unabhängig von n ist (siehe [6, Appendix 2]) und die Ordnungsstatistik basierend auf den Tail-Index $\beta = \rho_{max}$ ist.

Die Relations-Struktur von Prämien ist in den unsicheren Prämien enthalten. (vgl. [6, Seite 174])

Allerdings werden für die Prämien $\mathcal{P}(v, \beta)$ nur jene Integer-Werte von v eingesetzt, die einen großen „Gap“ haben.

Wenn nun v kein Integer ist, kann die fiktive Struktur von $\mathcal{P}(v, \beta)$, als das Verhältnis von Extremas, durch die Definition der fraktionellen Ordnungsstatistik und fraktionellen Extremas gegenüber deren Dichten dargestellt werden.

Beispielsweise definiert man von der Verteilung $F(\cdot)$ (mit der Dichte $f(\cdot)$) die v te fraktionelle Ordnungsstatistik $X_{(k,n)}^{(v)}$ durch Interpolation von $X_{(k)}$ und $X_{(k+1)}$, wobei $k = [v]$ ist (der Integer-Teil von v) für alle v , sodass $1 \leq k \leq v < k + 1 \leq n$ gegenüber deren Dichte $f_v(x)$ mit

$$f_v(x) = \frac{\{F(x)\}^{v-1} \{1 - F(x)\}^{n-v} f(x)}{B(v, n + 1 - v)}, \quad (4.34)$$

wobei $B(v, n + 1 - v) = \Gamma(v)\Gamma(n + 1 - v)/\Gamma(n + 1)$.

Fraktionelle Extremas können ähnlich definiert werden. Die Konvergenz des Erwartungswertes von normierten Ordnungsstatistiken zu dem dazugehörigen Erwartungswert von Extremas wird in [16, Kapitel 2] gezeigt.

Die Relations-Struktur wurde nun bestimmt und liefert einen stetigen Wechsel von unsicheren Prämien der Form

$$\begin{aligned}
 \mathcal{P}(v, \beta) &= vB(v, 1 - \beta) \\
 &= \frac{\mathbb{E} [1 + X_{(n)}/\lambda]}{\mathbb{E} [1 + X_{(n-k,n)}^{(v)}/\lambda]} \\
 &= \mathbb{E} \left[\frac{X_{(1)}^*}{X_{(k)}^{*(v)}} \right]
 \end{aligned} \tag{4.35}$$

Nun untersuchen wir, wie die Beziehungen zwischen (4.28) und (4.35) aufgestellt werden können, sodass sie in der Tail-Index-Schätzung und Prämien-Bestimmung unterstützend angewendet werden können, wenn ein konsistentes Schema unter einigen großen Extremas existiert. Üblicherweise benötigen Schätz-Methoden im Allgemeinen eine große Stichprobengröße für die Konsistenz und Fehler-Analyse (vgl. [17, Kapitel 6.4]).

Die gewöhnlichen Annahmen benötigen folgende Kriterien (vgl. [6, Seite 175]):

- (i) die Nummer von Extremas $k \rightarrow \infty$,
- (ii) verglichen mit der Stichprobengröße n , gilt $k/n \rightarrow 0$.

Unter diesen Bedingungen folgt

$$\frac{X_{(k,n)}}{X_{(k+1,n)}} \xrightarrow{P} 1$$

verglichen mit (4.33).

4.5 Berechnung der Prämie

Als Beispiel der risikoaversen Methode, untersuchen wir mögliche Prämien anhand von US-Hurrikan-Daten laut unten stehender Tabelle 4.1.

Die Daten zeigen ein klassisches Muster von Extremas. Die Statistik zeigt die 15 teuersten Hurrikan-Versicherungsschäden weltweit im Zeitraum von 1970 bis 2013 nach der Höhe des versicherten Schadens (in US-Dollar). Abgebildet wurden die versicherten Sach- und Betriebsunterbrechungsschäden (ausgenommen der Haftpflicht- und Lebensversicherungsschäden).

Unser Ziel ist es auf Basis dieser Daten die Prämien für die nächsten 15 größten Schäden der nächsten Jahre zu bestimmen.

Eine Validierung der Pareto-Annahmen für Naturkatastrophen-Daten dieser Art wird in ([18]) beschrieben.

KAPITEL 4. PRÄMIENKALKULATION FÜR HEAVY-TAILED RISIKEN

Ranking	Datum	Land	Event	Versicherter Schaden
1	28. August 2005	Orleans, Mississippi, Alabama, Florida, Ostküste USA	Hurrikan Katrina	89.373
2	29. Oktober 2012	Jamaika, Kuba, Bahamas, New Jersey, Ostküste USA	Hurrikan Sandy	36.890
3	24. August 1992	Bahamas, Süd-Florida, Louisiana, Süden USA	Hurrikan Andrew	27.594
4	01. September 2008	Turks- und Caicosinseln, Bahamas, Haiti, Dom. Republik, Kuba, Florida, Mississippi, Louisiana, Texas, Ohio River, Ontario, Québec	Hurrikan Ike	22.751
5	02. September 2004	Windward Islands, Venezuela, Jamaika, Grand Cayman, Kuba, Alabama, Florida, Ostküste USA	Hurrikan Ivan	17.218
6	15. Oktober 2005	Jamaika, Haiti, Cayman Islands, Kuba, Honduras, Nicaragua, Belize, Yucatán, Florida, Bahamas, Atlantische Provinzen Kanadas	Hurrikan Wilma	15.570
7	18. September 2005	Bahamas, Florida, Kuba, Yucatán, Louisiana, Texas, Mississippi, Arkansas	Hurrikan Rita	11.869
8	09. August 2004	Jamaika, Cayman Islands, Kuba, Florida, South Carolina, North Carolina	Hurrikan Charly	10.313
9	09. September 1989	Guadeloupe, Montserrat, Dominica, British Virgin Islands, U.S. Virgin Islands, Puerto Rico, South Carolina, North Carolina, Ostküste USA	Hurrikan Hugo	8.924
10	24. August 2004	Turks- und Caicosinseln, Bahamas, Florida	Hurrikan Frances	6.593
11	20. August 2011	Puerto Rico, Dom. Republik, Bahamas, Ostküste USA	Hurrikan Irene	6.274
12	15. September 1998	Leeward Islands, Puerto Rico, Dom. Republik, Haiti, Kuba, Florida Keys, Mississippi, Alabama, Louisiana, Florida Panhandle	Hurrikan George	5.240
13	13. September 2004	Virgin Islands, Puerto Rico, Dom. Republik, Haiti, Bahamas, Florida, Osten USA;	Hurrikan Jeanne	4.872
14	08. September 1999	Bahamas, Norfolk, Virginia, New Jersey, Ostküste USA	Hurrikan Floyd	4.100
15	27. September 1995	Guatemala, Yucatán Peninsula, Alabama, Florida Panhandle, Georgia, Nordosten USA	Hurrikan Opal	3.979

Tabelle 4.1: Hurrikanschäden von 1970 bis 2013 nach versichertem Schaden
(Quelle: <http://de.statista.com>, Okt. 2014)

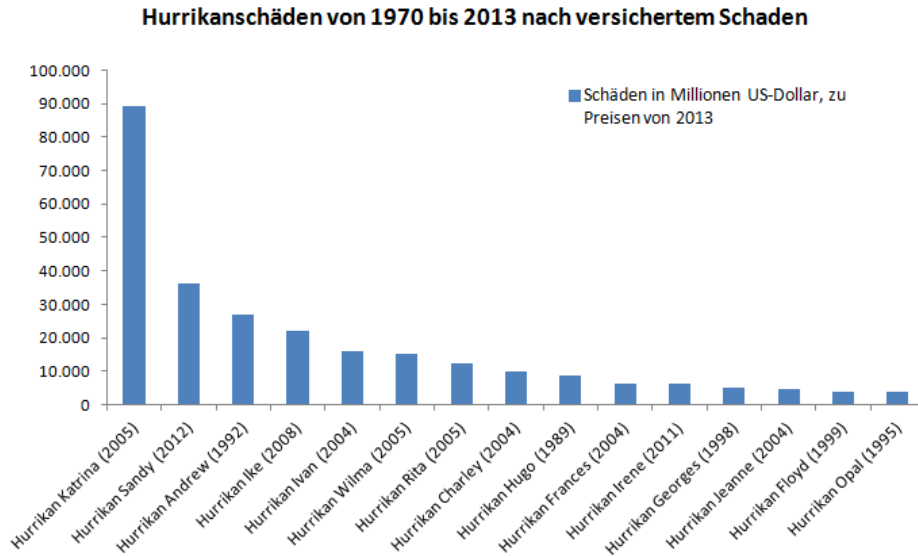


Abbildung 4.1: Die 15 teuersten versicherten Schäden in Millionen US-Dollar zu Preisen von 2013

Das konsistente Muster von Extremas ist hilfreich für Versicherer bei der Wahl des Wertes $\rho_{max} = \beta$ für Prämien-Bestimmungen.

4.5.1 Schätzungen des Tail-Index von Extrem-Relationen

Die Beziehung zwischen den Erwartungswerten der Relationen der größten Extremas, aus der Äquivalenz (4.28), kann bei der Bestimmung von einem maximalen β für tail-heaviness ρ hilfreich sein, welche wiederum für die Berechnung der Prämie notwendig ist.

Wir erhalten:

$$\begin{aligned}
 \mathbb{E} [\nu_n X_{(1)}^* / \nu_n X_{(2)}^*] &= (1 - \rho)^{-1} \\
 \mathbb{E} [\nu_n X_{(1)}^* / \nu_n X_{(3)}^*] &= (1 - \rho/2)^{-1} (1 - \rho)^{-1} \\
 \mathbb{E} [\nu_n X_{(1)}^* / \nu_n X_{(4)}^*] &= (1 - \rho/3)^{-1} (1 - \rho/2)^{-1} (1 - \rho)^{-1} \\
 &\vdots \\
 \mathbb{E} [\nu_n X_{(1)}^* / \nu_n X_{(1+9)}^*] &= 14B(14, 1 - \rho)
 \end{aligned}$$

Die 14 Beziehungen bieten eine Schätzung für $\hat{\rho}$ für ρ . Mit unseren Daten erhalten wir folgende Relations-Werte für ρ :

$$\begin{aligned}
 &0,5872; 0,5686; 0,5636; 0,5970; 0,5905; 0,6286; 0,6384; \\
 &0,6507; 0,6931; 0,6892; 0,7089; 0,7108; 0,7292; 0,7256;
 \end{aligned}$$

Der Mittelwert dieser Schätzungen ist $\hat{\rho} = 0,6487$ mit einer mittleren absoluten Abweichung von 0,0524.

4.5.2 Maximum Likelihood auf Basis der k -größten Extremas

Folgende Annahmen sind notwendig:

- (i) Die k -größten Schäden sind Pareto-verteilt mit dem fixen Tail-Index δ wobei $\rho = 1/\delta$.
- (ii) Die Gesamtzahl n von Schäden ist groß, ohne der Annahme n zu kennen - sodass die Verteilung des k ten größten Schadens annähernd Fréchet ist (4.15).
- (iii) Andere Schäden in dem Daten-Set können von anderen Verteilungen stammen.

Im Folgenden wird die Bedingte Maximum Likelihood Schätzung, basierend auf Überschreitungen von einem hinreichend hohen Level von Verallgemeinerte Pareto Verteilungen, beschrieben (vgl. [19] und [20, Seite 17]):

Die Bedingte Maximum Likelihood Funktion ist wie folgt definiert:

Definition 4.1. (Bedingte Maximum Likelihood Funktion)

$$L_1 [\mathbf{X} | X_{(n-k)} = x_{(n-k)}] = \prod_{j=0}^{k-1} f(y_{(n-j)}) \quad (4.36)$$

Wobei $\mathbf{X} = [X_{(n)}, X_{(n-1)}, \dots, X_{(n-k+1)}]$, $y_{(n-j)} = x_{(n-j)} - x_{(n-k)}$ und $f(\cdot)$ die Dichte der Pareto-Verteilungsfunktion aus (4.13) ist.

Eine alternative Form der bedingten Maximum Likelihood-Funktion, welche auf den Markov-Charakter von Ordnungsstatistiken basiert, sieht wie folgt aus:

$$L_2 [X_{(n)}, X_{(n-1)}, \dots, X_{(n-k+1)} | X_{(n-k)} = x_{(n-k)}] = \prod_{j=0}^{k-1} \frac{f(x_{(n-j)})}{[1 - F(x_{(n-k)})]^k} \quad (4.37)$$

Wobei $F(\cdot)$ die Pareto-Verteilungsfunktion (4.13) und $f(\cdot)$ seine Dichte ist.

Für die Anwendung unseres Daten-Sets ist $L_1(\cdot)$ generell auf beliebige Verteilungsfunktionen aus (4.12) anwendbar. Während die zweite Likelihood-Funktion $L_2(\cdot)$ Pareto als Verlust-Verteilung voraussetzt.

Dennoch liefern beide Annahmen die gleichen Fréchet-Extremas.

4.5.3 Ergebnisse des Bedingten Maximum Likelihood-Schätzers (CMLE)

Unter Anwendung des CLMEs auf unser Daten-Set erhalten wir folgende Ergebnisse:

- (i) Mit $L_1(\cdot)$ ist $\hat{\rho}_1 = 0,5010$ und $\hat{\lambda}_1 = 16.717,61$
Überschreitungen über $x_{(15)} = 3.979$ werden als Pareto (4.13) angenommen.
- (ii) Mit $L_2(\cdot)$ ist $\hat{\rho}_2 = 0,5010$ und $\hat{\lambda}_2 = 12.738,61$

Die Ergebnisse sind schlüssig, sodass $\hat{\lambda}_1 = \hat{\lambda}_2 + x_{(15)}$ unter bestimmten Voraussetzungen erwartet wird. Vergleicht man beispielsweise Pareto $f(x)$ mit $f(x - k|X > k)$, so ist das Maximum von $f(x)$ an der Stelle $x = 0$ und beträgt $\frac{\delta}{\lambda}$ und das Maximum von $f(x - k|X > k) = \frac{\delta}{\lambda} \frac{(1+\frac{x}{\lambda})^{-\delta-1}}{(1+\frac{k}{\lambda})^{-\delta}}$ an der Stelle $x = k$ und beträgt $\frac{\delta}{(\lambda+k)}$.

4.5.4 Vergleich zwischen Relationen und MLE-Ergebnisse

Die folgende Grafik stellt die drei Relationen-Sets für $k = 1$ bis 14 dar:

- (i) 1. Balken: Erwartungswerte der Relationen wenn $\rho = 0,5010$ für $kB(k; 1 - 0,5010)$,
- (ii) 2. Balken: Erwartungswerte der Relationen wenn $\rho = 0,6487$ für $kB(k; 1 - 0,6487)$,
- (iii) 3. Balken: Aktuelle Relationen $\frac{X_{(1)}^*}{X_{(1+k)}^*}$,

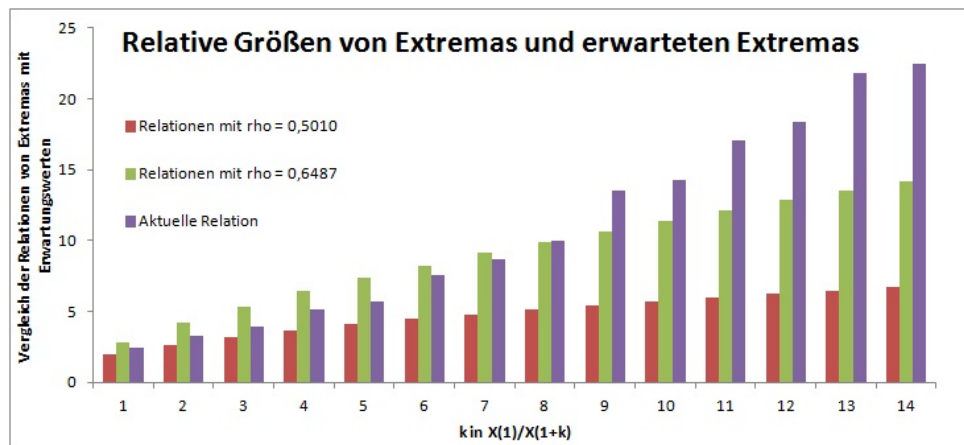


Abbildung 4.2: Relationen von erwarteten Extrem-Schäden mit $\rho = 0,5010$ und Moment-Schätzer $\rho = 0,6487$, gemeinsam mit Relationen der aktuellen Extremas

Die unzureichende Anpassung durch den CLME ($\rho = 0,5010$) ist durch die obige Grafik gut erkennbar. Insbesondere ist es schwierig die Prämien anhand von Maximum Likelihood-Schätzer zu bestimmen, vor allem wenn $\rho \in [\frac{1}{2}; 1)$ ist (vgl. [6, Seite 178]).

4.5.5 Prämien für Extrem-Schäden

Unter Berücksichtigung, dass die höchsten Schäden am Gefährlichsten sind, ist die Wahl von $\beta = \rho_{max}$ umso wichtiger für die Bestimmung der Prämien. β sollte insbesondere für die Prämienberechnung des größten Schadens herangezogen werden. Erhöht man β so verringern sich relativ gesehen die Prämien für weniger hohe Extrem-Schäden. Somit setzt man die Prämie für den größten Schaden fest und die restlichen Prämien folgen aus den Relationen von (4.33).

Die Prämien aus der folgenden Grafik wurden mit $\beta = 0,61; 0,62; 0,63; 0,64; 0,65$ berechnet. Die Prämie für den zweiten Extrem-Schaden aus dem Daten-Set wurde mit seiner aktuellen Schadenshöhe bewertet.

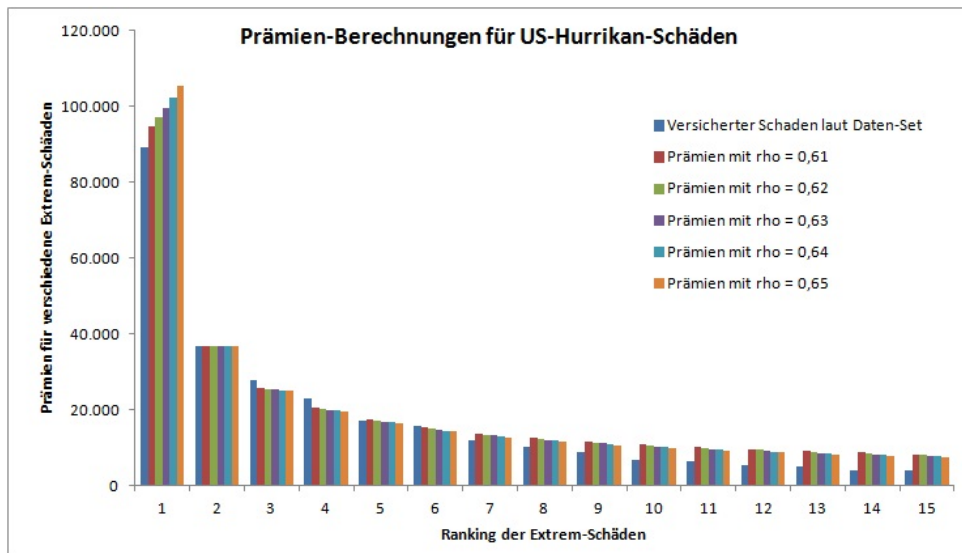


Abbildung 4.3: Prämien für US-Hurrikane unter Verwendung der Tail-Index-Werten $\beta = \rho_{max} = 0,61; 0,62; 0,63; 0,64; 0,65$ und der aktuellen Schadenshöhe von $X_{(2)}^* = 36.890$.

KAPITEL 4. PRÄMIENKALKULATION FÜR HEAVY-TAILED RISIKEN

Prämien-Ergebnisse:

	Aktueller Schaden	$\beta = 0,61$	$\beta = 0,62$	$\beta = 0,63$	$\beta = 0,64$	$\beta = 0,65$
	89.373	94.590	97.079	99.703	102.472	105.400
	36.890	36.890	36.890	36.890	36.890	36.890
	27.594	25.639	25.454	25.270	25.085	24.901
	22.751	20.425	20.194	19.963	19.734	19.506
	17.218	17.311	17.064	16.819	16.576	16.336
	15.570	15.199	14.948	14.700	14.455	14.212
	11.869	13.653	13.403	13.156	12.913	12.673
	10.313	12.464	12.216	11.972	11.732	11.496
	8.924	11.513	11.269	11.029	10.794	10.562
	6.593	10.733	10.493	10.257	10.026	9.799
	6.274	10.078	9.842	9.611	9.384	9.162
	5.240	9.519	9.288	9.061	8.838	8.621
	4.872	9.035	8.808	8.585	8.367	8.154
	4.100	8.611	8.388	8.169	7.955	7.746
	3.979	8.236	8.016	7.801	7.591	7.386
Total	271.560	303.897	303.351	302.986	302.812	302.843

Tabelle 4.2: Prämienaufstellungen anhängig von β für unser Daten-Set

4.6 Vergleich mit klassischen Prämienkalkulationsprinzipien

Verwendet man nun dasselbe Daten-Set und berechnet man die daraus resultierenden Prämien anhand der „klassischen“ Prämienkalkulationsprinzipien aus Kapitel 2, erkennt man schnell, dass diese Kalkulationsprinzipien in vielerlei Hinsicht unzureichende Prämien liefern.

Zunächst betrachten wir das empirische Mittel unseres Daten-Sets mit $\mathbb{E}[S] = 18.104$ und die empirische Varianz mit $\text{Var}[S] = 449.569.353,73$ und zuletzt die Standardabweichung mit $\sqrt{\text{Var}[S]} = 21.203,05$.

Wendet man nun das Erwartungswertprinzip an mit einem zu $\mathbb{E}[S]$ proportionalen Schwankungszuschlag $\theta > 0$ auf das Daten-Set, erhält man folgende Prämienaufstellung:

Erwartungswertprinzip:

θ	\mathcal{P} Erwartungswertprinzip
0,200	21.724,80
0,400	25.345,60
0,600	28.966,40
0,800	32.587,20
1,000	36.208,00
1,200	39.828,80
1,400	43.449,60
1,600	47.070,40
1,800	50.691,20
2,000	54.312,00
2,200	57.932,80
2,400	61.553,60
2,600	65.174,40
2,800	68.795,20
3,000	72.416,00
3,200	76.036,80
3,400	79.657,60
3,600	83.278,40
3,800	86.899,20

Tabelle 4.3: Prämienaufstellung anhängig von θ nach dem Erwartungswertprinzip

KAPITEL 4. PRÄMIENKALKULATION FÜR HEAVY-TAILED RISIKEN

$\mathbb{E}[S]$ stellt kein Risikomaß dar, daher ist die alleinige Abhängigkeit davon aus langfristiger Sicht zu „gefährlich“. Die Variabilität des Risikos wird auch mithilfe des Schwankungszuschlags θ nicht berücksichtigt.

Betrachte man nun die Ergebnisse aus dem Varianzprinzip bzw. Standardabweichungsprinzip erhält man folgende Prämienaufstellungen:

Varianzprinzip:

α	\mathcal{P} Varianzprinzip
1,00E-07	18.148,96
4,00E-07	18.283,83
7,00E-07	18.418,70
1,00E-06	18.553,57
4,00E-06	19.902,28
7,00E-06	21.250,99
1,00E-05	22.599,69
2,00E-05	27.095,39
3,00E-05	31.591,08
4,00E-05	36.086,77
5,00E-05	40.582,47
6,00E-05	45.078,16
7,00E-05	49.573,85
8,00E-05	54.069,55
9,00E-05	58.565,24
1,00E-04	63.060,94
1,10E-04	67.556,63
1,30E-04	76.548,02
1,50E-04	85.539,40

Tabelle 4.4: Prämienaufstellung anhängig von α nach dem Varianzprinzip

Obwohl das Varianzprinzip Schwankungen berücksichtigt, ist dieses Prämienkalkulationsprinzip nicht „fair“ gegenüber dem Versicherungsnehmer. Denn Abweichungen des Risikos S unterhalb des Erwartungswertes $\mathbb{E}[S]$ entsprechen theoretisch kleinen Schäden - gehen jedoch im selben Umfang in das Varianzprinzip ein, wie Abweichungen über dem Erwartungswert.

Um nicht Größen verschiedener Dimensionen zu addieren, wird statt der Varianz eher die Standardabweichung verwendet, um die Schwankungen des Gesamtschadens in die Prämienkalkulation mit einzubeziehen.

Standardabweichungsprinzip:

β	\mathcal{P} Standardabweichungsprinzip
1,00E-04	18.106,12
6,00E-04	18.116,72
1,00E-03	18.125,20
6,00E-03	18.231,22
1,00E-02	18.316,03
6,00E-02	19.376,18
1,00E-01	20.224,31
6,00E-01	30.825,83
1,00E+00	39.307,05
1,20E+00	43.547,66
1,40E+00	47.788,27
1,60E+00	52.028,88
1,80E+00	56.269,49
2,00E+00	60.510,10
2,20E+00	64.750,71
2,40E+00	68.991,32
2,60E+00	73.231,93
2,80E+00	77.472,54
3,00E+00	81.713,15

Tabelle 4.5: Prämienaufstellung anhängig von β nach dem Standardabweichungsprinzip

Quantilsprinzip:

Die Prämien-Bestimmung mittels dem Quantilsprinzip wird aus (2.15) abgeleitet. Für eine obere Schranke $\epsilon > 0$ gilt:

$$\mathcal{P} = F^{-1}(\epsilon) = \inf \{x \in \mathbb{R} | F(x) \leq \epsilon\}$$

wobei $F(\cdot)$ die Pareto-Verteilungsfunktion aus (4.13) ist, mit:

$$F(x) = 1 - \left(1 + \frac{x}{\lambda}\right)^{-\delta}.$$

Wir erhalten folgende Quantilsfunktion:

$$\mathcal{P} = F^{-1}(\epsilon) = \lambda \left((1 - \epsilon)^{-\frac{1}{\delta}} - 1 \right)$$

Mit unserem Daten-Set erhalten wir folgende Prämien mittels dem Quantilsprinzip:

KAPITEL 4. PRÄMIENKALKULATION FÜR HEAVY-TAILED RISIKEN

(1 - ϵ) - Quantil	\mathcal{P} Quantilsprinzip
0,75-Quantil	19.984,50
0,80-Quantil	23.719,60
0,85-Quantil	27.109,70
0,90-Quantil	33.171,60
0,99-Quantil	82.025,38

Tabelle 4.6: Prämienaufstellung nach dem Quantilsprinzip

Auch das Quantilsprinzip eignet sich nicht für die Prämienberechnung von heavy-tailed Verteilungen.

Die am häufigsten verwendete Prämienberechnungsmethode bei thin-tailed-Schadensverteilungen ist das Exponentialprinzip, zumal auch weil diese Berechnungsmethode laut Kapitel 2 die meisten Gütekriterien erfüllt. Da aber für diese Prämienberechnungsmethode die Momenterzeugende Funktion benötigt wird, welche für heavy-tailed Verteilungen nicht existiert, ist das Exponentialprinzip für eine Prämienberechnung von heavy-tailed-Schadensverteilungen nicht geeignet.

Vergleicht man die berechneten Prämien aus Kapitel 4.5.5 beispielsweise mit den Ergebnissen nach dem Standardabweichungsprinzip mit geeignet ausgewählten β s, erkennt man anhand der nachfolgenden Tabelle bzw. Grafik ganz gut, dass die Prämien nach dem Standardabweichungsprinzip nur durch einen sehr hohen Schwankungszuschlag β für hohe Schäden angeglichen werden können.

Bei den weniger hohen Schäden kann die Prämie wiederum nur durch einen negativen Schwankungszuschlag β konkurrenzfähig werden.

KAPITEL 4. PRÄMIENKALKULATION FÜR HEAVY-TAILED RISIKEN

Schadenshöhe	Prämien mit Tail-Index $\rho = 0,61$	β	Prämien nach Standardabweichungsprinzip mit β
89.373	94.590	3,50E+00	92.314,68
36.890	36.890	9,00E-01	37.186,75
27.594	25.639	4,00E-01	26.585,22
22.751	20.425	2,00E-01	22.344,61
17.218	17.311	1,00E-01	20.224,31
15.570	15.199	1,00E-02	18.316,03
11.869	13.653	9,00E-03	18.294,83
10.313	12.464	8,00E-03	18.273,62
8.924	11.513	7,00E-03	18.252,42
6.593	10.733	6,00E-03	18.231,22
6.274	10.078	5,00E-03	18.210,02
5.240	9.519	4,00E-03	18.188,81
4.872	9.035	3,00E-03	18.167,61
4.100	8.611	2,00E-03	18.146,41
3.979	8.236	1,00E-03	18.125,20

Tabelle 4.7: Prämienvergleich mit dem Standardabweichungsprinzip für unser Daten-Set

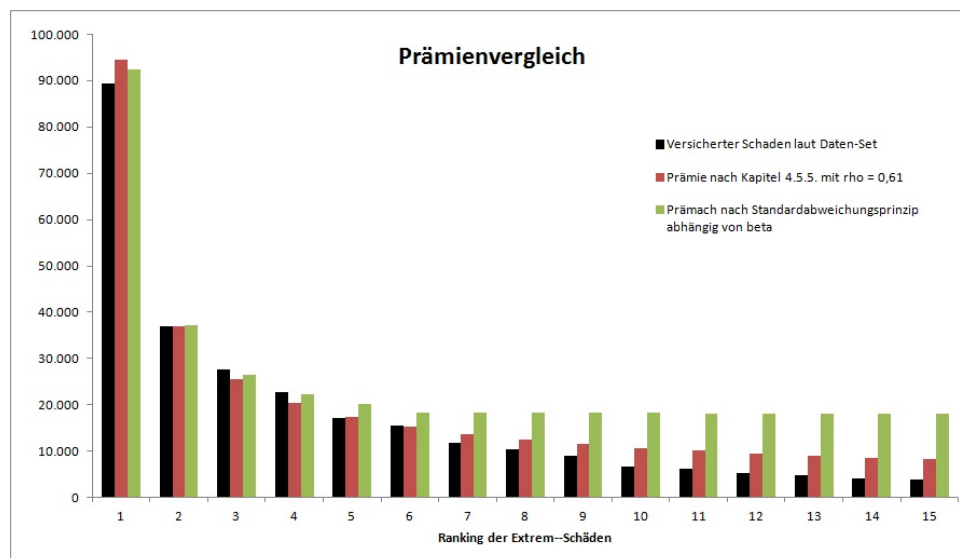


Abbildung 4.4: Prämienvergleich für US-Hurrikane

KAPITEL 4. PRÄMIENKALKULATION FÜR HEAVY-TAILED RISIKEN

Es ist allgemein schwierig die Prämie anhand eines theoretischen Prinzips zu bestimmen. Ist die Prämienzahlung zu hoch, so ist das Versicherungsunternehmen nicht konkurrenzfähig. Ist sie zu niedrig, so ist die Ruinwahrscheinlichkeit zu hoch. In der Praxis werden die Risiken geteilt und durch verschiedenste Rückversicherungsmodelle (Summenexzedenten, Quoten-Versicherung, etc.) abgedeckt.

Anhang

Berechnung von CMLE

Aus dem Abschnitt 4.5.2 ist die Bedingte Maximum Likelihood-Funktion wie folgt definiert:

$$L_1 [\mathbf{X} | X_{(n-k)} = x_{(n-k)}] = \prod_{j=0}^{k-1} f(x_{(n-j)} - x_{(n-k)})$$

Wobei $\mathbf{X} = [X_{(n)}, X_{(n-1)}, \dots, X_{(n-k+1)}]$ und $f(\cdot)$ sei die Dichte der Pareto-Verteilungsfunktion aus (4.13) mit:

$$F(x_{(n-j)} - x_{(n-k)}) = 1 - \left(1 + \frac{x_{(n-j)} - x_{(n-k)}}{\lambda}\right)^{-\delta} \quad \text{und}$$

$$f(x_{(n-j)} - x_{(n-k)}) = \frac{\delta}{\lambda} \cdot \left(1 + \frac{x_{(n-j)} - x_{(n-k)}}{\lambda}\right)^{-\delta-1}$$

Somit ist

$$L_1 = \prod_{j=0}^{k-1} \frac{\delta}{\lambda} \cdot \left(1 + \frac{x_{(n-j)} - x_{(n-k)}}{\lambda}\right)^{-\delta-1}$$

Daraus ergibt sich die Log-Likelihood-Funktion:

$$\begin{aligned}
 l_1 &= \sum_{j=0}^{k-1} \left[\ln \left(\frac{\delta}{\lambda} \right) + \ln \left(1 + \frac{x_{(n-j)} - x_{(n-k)}}{\lambda} \right)^{-\delta-1} \right] \\
 &= \sum_{j=0}^{k-1} \left[\ln(\delta) - \ln(\lambda) + (-\delta - 1) \cdot \ln \left(1 + \frac{x_{(n-j)} - x_{(n-k)}}{\lambda} \right) \right] \\
 &= \sum_{j=0}^{k-1} \left[\ln(\delta) - \ln(\lambda) - (\delta + 1) \cdot (\ln(\lambda + x_{(n-j)} - x_{(n-k)}) - \ln(\lambda)) \right] \\
 &= \sum_{j=0}^{k-1} \left[\ln(\delta) + \delta \ln(\lambda) - (\delta + 1) \cdot \ln(\lambda + x_{(n-j)} - x_{(n-k)}) \right] \\
 &= k \cdot \ln(\delta) + \delta k \cdot \ln(\lambda) - (\delta + 1) \cdot \sum_{j=0}^{k-1} \ln(\lambda + x_{(n-j)} - x_{(n-k)})
 \end{aligned}$$

Den CML-Schätzer erhält man somit mit:

$$\begin{aligned}
 \frac{\partial l_1}{\partial \delta} &= \frac{k}{\delta} + k \ln(\lambda) - \sum_{j=0}^{k-1} \ln(\lambda + x_{(n-j)} - x_{(n-k)}) := 0 \\
 \hat{\delta} &= \frac{k}{\sum_{j=0}^{k-1} \ln(\lambda + x_{(n-j)} - x_{(n-k)}) - k \ln(\lambda)}
 \end{aligned}$$

und

$$\begin{aligned} \frac{\partial l_1}{\partial \lambda} &= \frac{\delta k}{\lambda} - (\delta + 1) \cdot \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)} - x_{(n-k)}} := 0 \\ \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)} - x_{(n-k)}} &= \frac{\delta k}{\lambda} \cdot \frac{1}{\delta + 1} \\ \lambda \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)} - x_{(n-k)}} &= \frac{\delta k}{\delta + 1} \end{aligned}$$

Die Lösung für $\hat{\delta}_1$ und $\hat{\lambda}_1$ erhält man nur numerisch. Zu unseren Ergebnissen gelangen wir mittels Newton-Verfahren mit der Jakobi-Matrix

$$J_1 = \begin{pmatrix} \frac{\partial f_1}{\partial \lambda} & \frac{\partial f_1}{\partial \delta} \\ \frac{\partial f_2}{\partial \lambda} & \frac{\partial f_2}{\partial \delta} \end{pmatrix}$$

mit

$$\frac{\partial f_1}{\partial \lambda} = -\frac{\delta k}{\lambda^2} + (\delta + 1) \cdot \sum_{j=0}^{k-1} \frac{1}{(\lambda + x_{(n-j)} - x_{(n-k)})^2}$$

$$\frac{\partial f_1}{\partial \delta} = \frac{k}{\lambda} - \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)} - x_{(n-k)}}$$

$$\frac{\partial f_2}{\partial \lambda} = \frac{k}{\lambda} - \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)} - x_{(n-k)}}$$

$$\frac{\partial f_2}{\partial \delta} = \frac{-k}{\delta^2}$$

Analog zu L_1 erhalten wir für die alternative Form des CMLE mit:

$$L_2 [X_{(n)}, X_{(n-1)}, \dots, X_{(n-k+1)} | X_{(n-k)} = x_{(n-k)}] = \prod_{j=0}^{k-1} \frac{f(x_{(n-j)})}{[1 - F(x_{(n-k)})]^k}$$

mit

$$F(x_{(n-k)}) = 1 - \left(1 + \frac{x_{(n-k)}}{\lambda}\right)^{-\delta} \quad \text{und}$$

$$f(x_{(n-j)}) = \frac{\delta}{\lambda} \cdot \left(1 + \frac{x_{(n-j)}}{\lambda}\right)^{-\delta-1}$$

Somit ist

$$L_2 = \left(1 + \frac{x_{(n-k)}}{\lambda}\right)^{k\delta} \cdot \prod_{j=0}^{k-1} \frac{\delta}{\lambda} \cdot \left(1 + \frac{x_{(n-j)}}{\lambda}\right)^{-\delta-1}$$

Daraus ergibt sich die Log-Likelihood-Funktion:

$$\begin{aligned} l_2 &= k\delta \cdot \ln\left(1 + \frac{x_{(n-k)}}{\lambda}\right) + \sum_{j=0}^{k-1} \left[\ln(\delta) - \ln(\lambda) - (\delta + 1) \cdot \ln\left(1 + \frac{x_{(n-j)}}{\lambda}\right) \right] \\ &= k\delta \cdot [\ln(\lambda + x_{(n-k)}) - \ln(\lambda)] + \sum_{j=0}^{k-1} [\ln(\delta) - \ln(\lambda) - (\delta + 1) \cdot (\ln(\lambda + x_{(n-j)}) - \ln(\lambda))] \\ &= k\delta \cdot [\ln(\lambda + x_{(n-k)}) - \ln(\lambda)] + \sum_{j=0}^{k-1} [\ln(\delta) - (\delta + 1) \cdot \ln(\lambda + x_{(n-j)}) + \delta \ln(\lambda)] \\ &= k\delta \cdot \ln(\lambda + x_{(n-k)}) + k \cdot \ln(\delta) - (\delta + 1) \cdot \sum_{j=0}^{k-1} \ln(\lambda + x_{(n-j)}) \end{aligned}$$

Den CML-Schätzer erhält man somit mit:

$$\frac{\partial l_2}{\partial \delta} = k \cdot \ln(\lambda + x_{(n-k)}) + \frac{k}{\delta} - \sum_{j=0}^{k-1} (\lambda + x_{(n-j)}) := 0$$

$$\frac{k}{\sum_{j=0}^{k-1} \ln(\lambda + x_{(n-j)}) - k \ln(\lambda + x_{(n-k)})} = \hat{\delta}$$

und

$$\frac{\partial l_2}{\partial \lambda} = \frac{k\delta}{\lambda + x_{(n-k)}} - \sum_{j=0}^{k-1} \frac{\delta + 1}{\lambda + x_{(n-j)}} := 0$$

$$\frac{k\delta}{(\lambda + x_{(n-k)}) (\delta + 1)} = \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)}}$$

Die Lösung für $\hat{\delta}_2$ und $\hat{\lambda}_2$ erhält man ebenso nur numerisch. Dazu wenden wir wieder das Newton-Verfahren mit der folgenden Jakobi-Matrix an

$$J_2 = \begin{pmatrix} \frac{\partial f_1}{\partial \lambda} & \frac{\partial f_1}{\partial \delta} \\ \frac{\partial f_2}{\partial \lambda} & \frac{\partial f_2}{\partial \delta} \end{pmatrix}$$

mit

$$\frac{\partial f_1}{\partial \lambda} = -\frac{k\delta}{(\lambda + x_{(n-k)})^2} + (\delta + 1) \cdot \sum_{j=0}^{k-1} \frac{1}{(\lambda + x_{(n-j)})^2}$$

$$\frac{\partial f_1}{\partial \delta} = \frac{k}{\lambda + x_{(n-k)}} - \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)}}$$

$$\frac{\partial f_2}{\partial \lambda} = \frac{k}{\lambda + x_{(n-k)}} - \sum_{j=0}^{k-1} \frac{1}{\lambda + x_{(n-j)}}$$

$$\frac{\partial f_2}{\partial \delta} = \frac{-k}{\delta^2}$$

Literaturverzeichnis

- [1] P. GRANDITS, A. ZWIRCHMAYR: Sachversicherungsmathematik, Technische Universität Wien, 2008
- [2] A. J. MCNEIL, R. FREY, P. EMBRECHTS: Quantitative Risk Management: Concepts, Techniques and Tools, 2004
- [3] M. LÖWE: Extremwerttheorie, Universität Münster, WS 2008/09
verfügbar in <http://www.math.uni-muenster.de/statistik/loewe/extrem.pdf> (abgerufen am 17.09.2012)
- [4] J. FRANKE, W. HÄRDLE, C. HAFNER: Einführung in die Statistik der Finanzmärkte , 2.Auflage
verfügbar in http://sfb649.wiwi.hu-berlin.de/fedc_homepage/xplore/ebooks/html/sfm (abgerufen am 31.10.2013)
- [5] J. BEIRLANT, Y. GOEGEBEUR , J. SEGERS, J. TEUGELS: Statistics of Extremes: Theory and Applications, 2004
- [6] R. GAY: Premium Calculation for Fat-tailed Risk
- [7] T. ROLSKI, H. SCHMIDLI, V. SCHMIDT, J. TEUGELS: Stochastic processes for Insurance and Finance, 1999
- [8] K.D. SCHMIDT: Versicherungsmathematik, 3.Auflage 2009
- [9] M. RIEDLE: Risikotheorie Teil IV, Humboldt-Universität zu Berlin, 2005
verfügbar in <http://www.mathematik.hu-berlin.de/riedle/winter04/teil4.pdf> (abgerufen am 13.06.2012)
- [10] W. SCHOTT: Preise für versicherungstechnische Risiken: Zu Rationalität und Realitätsnähe bei Anwendung aktuarielle und ökonomischer Modelle, 1997
- [11] T. MACK: Schadenversicherungsmathematik, 1997
- [12] T. MIKOSCH: Non-Life Insurance Mathematics, 2003

- [13] H. U. GERBER: An Introduction to Mathematical Risk Theory, 1979
- [14] H. FÖLLNER, A. SCHIED: Stochastic Finance, 2002
- [15] W. SCHOTT: Preise für Versicherungstechnische Risiken, 1997
- [16] S.I. RESNICK: Extreme Values - Regular Variation and Point Processes, 1987
- [17] P. EMBRECHTS, C. KLUPPELBERG, T. MIKOSCH: Modelling Extremal Events for Insurance and Finance, 1997
- [18] P.-H. HSIEH: A data-analytic method for forecasting the next record catastrophic loss, Journal of risk and insurance, 2004
- [19] H. DREES, A. FERREIRA, L. DE HAAN: On Maximum Likelihood estimation of the Extreme Value Index Seiten 1179 - 2101, The Annals of Applied Probability, 2004
- [20] D.R. COX, D.V. HINKLEY: Theoretical Statistics, Chapman and Hall, 1974