

Extremwerttheorie für Finanzzeitreihen – ein unverzichtbares Werkzeug im Risikomanagement

Milan Borkovec

Claudia Klüppelberg

Zusammenfassung

Mathematik, vor allem die moderne Wahrscheinlichkeitstheorie und Statistik, haben das tägliche Geschäft der Finanzindustrie, insbesondere von Banken und Versicherungen, grundlegend verändert. Dies hat zu neuen Methoden geführt, Risiken abzusichern oder zu reduzieren. Die Kehrseite der Medaille zeigt allerdings auch die Versuchung, sehr grosse Wetten mit einem Potential für extreme Verluste einzugehen. Die Extremwerttheorie bietet hier Werkzeuge an, die zu einer Verbesserung des Risikomanagement führen.

1. Einleitung

Die Krise von LTCM im Spätsommer und Herbst 1998 zeigte wieder einmal, wie inadäquat traditionelles Risikomanagement mit extremen Ereignissen umgeht. Die kritischen Kommentare in anerkannten Zeitungen und Zeitschriften sind nicht unverdient. Einige Beispiele aus "Business Week" sind (vgl. Coy (1998) und Coy/Wolley (1998)):

- "Can you devise surefire ways to beat the markets? The rocket scientists thought they could. Boy were they wrong." (Headline).
- "The quants may have placed too much faith in their exquisitely tuned computer models... Sources say, Long-Term-Capital's worst-case scenario was only about 60% as bad as the one that actually happened."
- "In a certain sense, maybe the problem wasn't too much rocket science, but too little. Extreme, synchronized rises and falls in financial markets occur infrequently but they do occur. The problem with the models is that they did not assign a high enough chance of occurrences to the scenario in which many things go wrong at the same time – the perfect storm scenario."

Josef F. Wertschulte (Vorstand der Hypovereinsbank München) wird in einem Artikel des Handelsblatts vom 28.10.98 folgendermassen zitiert:

- "Fälle wie die folgenreichen Fehlspekulationen bei der Metallgesellschaft oder des amerikanischen Investment-Hauses Long Term Capital Management haben bewiesen, dass professionelles Risiko-Management bei immer stärker schwankenden Märkten zur Überlebensfrage für Unternehmen werden kann."

Ein wichtiges mathematisches Werkzeug, um Risiken zu modellieren und zu messen, liefert die Extremwerttheorie, die schon bei der Bestimmung der Deichhöhen in den Niederlanden erfolgreich eingesetzt wurde (de Haan (1990)). Auch im Versicherungsgeschäft spielt die Extremwerttheorie eine bedeutende Rolle, obwohl es manchen Aktuaren vielleicht nicht bewusst ist. Vor allem Rückversicherungen arbeiten seit Jahrzehnten mit Paretomodellen, um Grossschäden angemessen zu modellieren; Exzess-Loss (XL)-Verträge machen die Modellierung von Exzedenten notwendig. Im Vergleich dazu arbeiten Banken vorwiegend mit Gauss'schen Modellen, die insbesondere hohe Risiken viel zu optimistisch bewerten.

Ein Grund dafür ist sicherlich die komplexe Datenlage, hohe Dimensionalität und die Notwendigkeit schneller Algorithmen im Bankgeschäft. Ein weiterer Grund mag auch sein, dass die Normalverteilung und der zentrale Grenzwertsatz bei Praktikern und Ökonometrieprofessoren wohlbekannt sind (er gehört zur Standardausbildung für alle Ökonometriestudenten), während sogar ein Mathematikstudent sein Diplom erhalten kann, ohne die Extremwerttheorie kennengelernt zu haben. Das ist umso bedauerlicher, als beide Theorien sich ergänzen:

Der *zentrale Grenzwertsatz* beschreibt das Verhalten von Summen und Mittelwerten,

während die *Extremwerttheorie* das extreme Verhalten einer Zeitreihe beschreibt.

Summen und Mittelwerte sind für das Risikomanagement von untergeordneter Bedeutung, die Extremwerte einer Finanzzeitreihe sind es hauptsächlich, die das Risikoverhalten einfangen. Dies erkannte auch der Basler Ausschuss für die Bankenaufsicht, der Mitte der 90er Jahre neue Standards für das Risikomanagement formulierte, die insbesondere den Value-at-Risk als Risikomass vorschreiben, der als das p -Quantil der Gewinn/Verlustverteilung definiert ist. Geschätzt wird dieses Quantil allerdings meist entweder durch das empirische Quantil oder das Normalverteilungsquantil. In Emmer et al. (1998) wird gezeigt, wie man ein Quantil mit Hilfe von Extremwertmethoden schätzen kann, was zu einer höheren Risikoeinschätzung führt als die Schätzung mittels empirischer oder normaler Methode.

Eine realistische Einschätzung verschiedener Risiken mit Hilfe der Extremwerttheorie ist möglich für

- Marktrisiken
- Kreditrisiken
- Betriebsrisiken
- Versicherungsrisiken
- Rückversicherungsrisiken.

Zur Extremwerttheorie und (Rück)-Versicherungsrisiken verweisen wir auf Embrechts et al. (1997). Ein interessantes Beispiel eines Katastrophenbonds ist in Schmock (1999) behandelt. Ein einführendes Beispiel in VaR und Marktrisiken wird in Emmer et al. (1998) behandelt, weitere Arbeiten zu diesem Thema findet man auf unserer Webseite www.ma.tum.de/stat/; siehe z.B. Borkovec (1998), Borkovec/Klüppelberg (1998), Klüppelberg/Korn (1998) und Rootzén/Klüppelberg (1999). Embrechts, McNeil und Koautoren haben ebenfalls sehr lezenswerte Arbeiten zu diesen und verwandten Themen geschrieben, wir verweisen z.B. auf Embrechts et al. (1998a,1998b) McNeil (1997,1998) und McNeil/Frey (1998), sowie auf die Webseite www.math.ethz.ch/finance. Schliesslich wurde die Extremwerttheorie in Phoa (1999) und Cruz et al. (1998) verwendet, um Kredit- respektive Betriebsrisiken zu schätzen.

Zur Datenanalyse verwenden wir die Software EVIS von Alexander McNeil. Sie besteht aus S-Plus Routinen, d.h. sie benötigt als Basis das Statistik-Programmpaket S-Plus. EVIS steht auf der homepage www.math.ethz.ch/mcneil/software.html kostenlos zur Verfügung.

Während in Emmer et al. (1998) die klassische Extremwerttheorie für unabhängige und identisch verteilte (iid) Daten benutzt wird, um sehr grosse oder sehr kleine Quantile zu schätzen, wollen wir in dieser Arbeit erklären, welchen Einfluss zeitliche Abhängigkeit auf Quantilschätzer haben kann.

Es ist allgemein bekannt, dass Finanzzeitreihen nicht unabhängig und identisch verteilt sind. Ganz im Gegenteil, sie zeigen häufig folgende Eigenschaften (sogenannte “stylized

features”):

- Die Daten sind langschwänzig.
- Die Volatilität verändert sich in der Zeit.
- Die Daten neigen zu Clusterbildung in den Extrema.
- Die Daten sind (nahezu) unkorreliert, aber nicht unabhängig.

Volatilitätsmodelle versuchen, diese “stylized features” adäquat abzubilden, und das mit moderatem Erfolg. Als probabilistische Aufgabe scheint es schwierig zu sein, alle diese Eigenschaften in einem parametrischen Modell zu erfassen. Als statistische Aufgabe erfordert die Anpassung eines einigermaßen realistischen Modells oftmals die Entwicklung neuer computerintensiver Methoden (Parameter werden mit Hilfe von aufwendigen Simulationsverfahren geschätzt, z.B. mittels Markov Chain Monte Carlo (MCMC) Methoden); vgl. zum Beispiel Elerian et al. (1998).

Der Vorteil von Extremwertmethoden liegt darin, dass man für das Risikomanagement nicht die gesamte Zeitreihe modellieren muss. Es genügt, den Teil der Zeitreihe adäquat zu beschreiben, der über einer hohen Schwelle liegt.

Wir erläutern die Methoden anhand der Wechselkurse \$US/DEM vom 1.1.1996 bis zum 31.12.1996. Die Kurse wurden halbstündig gemessen. Die Daten haben die “stylized features” wie oben beschrieben.

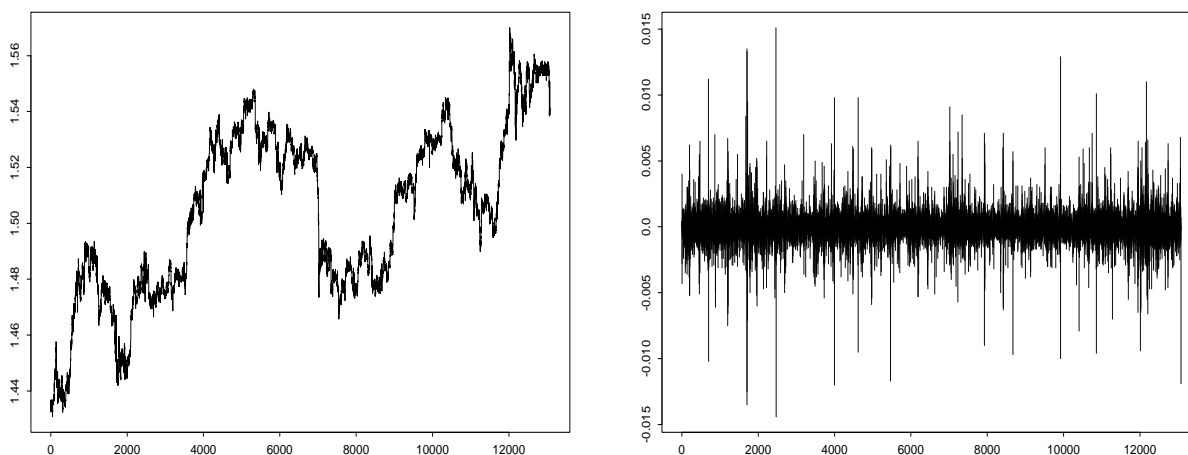


Abb. 1. Wechselkurse \$US/DEM vom 1.1.1996 bis 31.12.1996, jeweils halbstündig beobachtet und bereinigt (keine Wochenenden, Feiertage etc.). Daneben die zugehörigen Verluste (negativen Differenzen) der Wechselkurse.

2. Extremwerttheorie

Klassische Extremwerttheorie beschreibt das Verhalten der grössten Werte einer Stichprobe oder der Exzedenten über einer hohen Schwelle. Für eine Stichprobe X_1, \dots, X_n (iid Zufallsvariable), betrachtet man in einem ersten Schritt die partiellen Maxima

$$M_1 = X_1, \quad M_n = \max(X_1, \dots, X_n), \quad n \in \mathbb{N}.$$

Der folgende Satz ist *die* Grundlage klassischer Extremwerttheorie

Satz 1. (Satz von Fisher und Tippett)

$(X_n)_{n \geq 1}$ sei eine Folge von iid Zufallsvariablen mit Verteilungsfunktion F . Falls es normierende Konstanten $a_n > 0$, $b_n \in \mathbb{R}$ und eine nicht-degenerierte Verteilungsfunktion H gibt, so dass

$$(1) \quad \lim_{n \rightarrow \infty} P(M_n \leq a_n x + b_n) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x), \quad x \in \mathbb{R},$$

dann ist H eine der folgenden Verteilungsfunktionen:

$$\text{Fréchet:} \quad \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \quad \alpha > 0.$$

$$\text{Weibull:} \quad \Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0.$$

$$\text{Gumbel:} \quad \Lambda(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

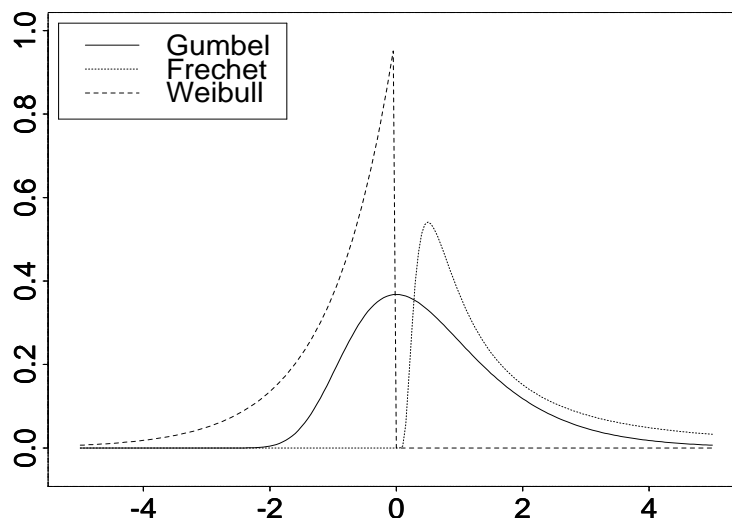


Abb. 2. Dichten der Standard-Extremwertverteilungen. Für die Fréchet und die Weibull Verteilung haben wir $\alpha = 1$ gewählt.

Die Grenverteilung H in (1) hängt ausschliesslich nur vom Tail der Verteilungsfunktion F ab. Dies folgt sofort aus der Tatsache, dass die untenstehenden Aussagen (i) – (iii) für iid Daten äquivalent sind. Dabei sei $u_n = a_n x + b_n$ und X eine Zufallsvariable mit der gleichen Verteilung wie X_1 .

(i) $\lim_{n \rightarrow \infty} P(M_n \leq u_n) = H(x), \quad x \in \mathbb{R},$

(ii) $\lim_{n \rightarrow \infty} nP(X > u_n) = -\ln H(x), \quad x \in \mathbb{R},$

(iii) es gibt eine messbare, positive Funktion $\beta(\cdot)$, für die gilt

$$(2) \quad \lim_{u \rightarrow \infty} P\left(\frac{X - u}{\beta(u)} > x \mid X > u\right) = \frac{\overline{F}(u + \beta(u)x)}{\overline{F}(u)} = (1 + \xi x)^{-1/\xi},$$

für $1 + \xi x > 0$ und $\xi = 1/\alpha$.

Die Äquivalenz dieser Aussagen werden wir später erläutern und auch verwenden.

3. Die Standard-POT-Methode

Die POT-Methode (POT steht für “Peaks Over Threshold”) ist eine Schätzmethode für einen Tail oder ein Quantil, basierend auf extrem grossen Beobachtungen.

Die Methode setzt sich aus drei Komponenten zusammen. Jede basiert auf einem probabilistischen Prinzip, das wir jeweils im folgenden erklären wollen. Zur Verdeutlichung soll Abbildung 3. dienen.

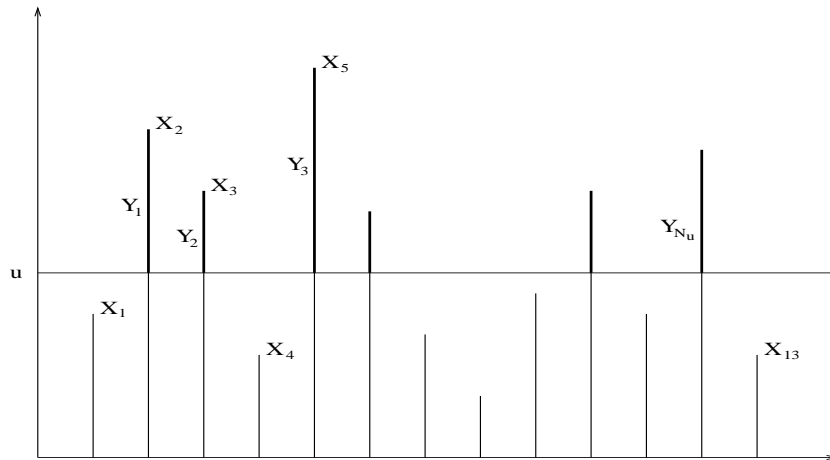


Abb. 3. Daten X_1, \dots, X_{13} mit zugehörigen Exzedenzen Y_1, \dots, Y_{N_u} .

(1) Der Punktprozess der Exzedenten

Wir wollen einen Grenzprozess für den Punktprozess der Überschreitungen von hohen

Schwellen herleiten. Für eine hohe Schwelle u_n markieren wir die Beobachtungen der Stichprobe X_1, \dots, X_n , die grösser als u_n sind. (In Abbildung 3. sind das die Beobachtungen Nummer 2, 3, 5, 6, 10, 12.) Für eine Grenzwertaussage soll der Umfang n der Stichprobe gegen unendlich streben und gleichzeitig die Schwelle u_n immer höher werden, und das im richtigen Verhältnis. Für iid Daten hat jeder einzelne Datenpunkt die gleiche Chance, die Schwelle u_n zu überschreiten, nämlich $P(X_i > u_n)$ für $i = 1, \dots, n$. (Für die endliche Menge A ist $\#A$ die Anzahl der Elemente in A .) Dann ist die Anzahl der Beobachtungen, die diese Schwelle überschreiten,

$$\#\{i : X_i > u_n, i = 1, \dots, n\}$$

binomialverteilt mit Parametern n und $P(X > u_n)$. Wenn nun

$$(3) \quad \lim_{n \rightarrow \infty} nP(X > u_n) = \tau \in (0, \infty),$$

dann konvergiert nach dem klassischen Satz von Poisson die Verteilung von $\#\{i : X_i > u_n, i = 1, \dots, n\}$ gegen eine Poissonverteilung mit Parameter τ . Falls X eine absolut stetige Verteilung hat, also eine Dichte besitzt, ist (3) eine sehr schwache Bedingung: Für alle bekannten absolut stetigen Verteilungen und jedes $\tau > 0$ kann man eine geeignete Folge (u_n) finden (siehe z.B. Embrechts et al. (1997), Chapter 3). Markiert man in dem Intervall $[0, n]$ die Punkte $\{i : X_i > u_n, i = 1, \dots, n\}$, so wird das Intervall immer grösser, die markierten Punkte werden immer seltener (da die Schwelle u_n mit n steigt). Eine bessere Darstellung erhält man, wenn man die Punkte nicht im Intervall $[0, n]$ aufträgt, sondern reskaliert im Intervall $[0, 1]$. Ein Exzedent X_i von u_n wird dann also nicht bei i aufgetragen, sondern bei i/n . Bezeichnet man für $n \in \mathbb{N}$

$$(4) \quad N_n((a, b]) = \#\{i/n \in (a, b] : X_i > u_n, i = 1, \dots, n\},$$

für alle Intervalle $(a, b] \subset (0, 1]$, so definiert N_n einen Punktprozess auf dem Intervall $(0, 1]$. Er heisst der *zeitnormierte Punktprozess der Exzedenten*. Wenn u_n so gewählt wird, dass (3) gilt, dann konvergiert die Folge N_n von Punktprozessen für $n \rightarrow \infty$ in Verteilung gegen einen Poissonprozess mit Parameter τ .

(2) Die Verallgemeinerte Paretoverteilung

Bei den Exzedenten einer hohen Schwelle interessiert uns nicht nur, wann sie auftreten oder wie oft das passiert, sondern auch, wie gross jeweils der Exzess $X - u | X > u$ ist (gegeben $X > u$, wird der Exzess $X - u$ betrachtet). (In Abbildung 3. sind die Exzesse mit Y_1, \dots, Y_{N_u} bezeichnet, dort ist $N_u = 6$ die Anzahl der Exzedenten.) Unter der Bedingung (3) mit $\tau = -\ln H(x)$ und $u_n = a_n x + b_n$ wie im Theorem 1. kann man zeigen, dass es eine messbare, positive Funktion β gibt, so dass für $1 + \xi y > 0$ gilt

$$(5) \quad \lim_{u \rightarrow \infty} P\left(\frac{X - u}{\beta(u)} > y \mid X > u\right) = (1 + \xi y)^{-1/\xi},$$

falls die linke Seite überhaupt konvergiert. Falls $\xi = 0$ ist, wird die rechte Seite als e^{-y} interpretiert. Für alle $\xi \in \mathbb{R}$ ist die rechte Seite der Tail einer Verteilungsfunktion, der sogenannten *verallgemeinerten Paretoverteilung*. Für $\xi \geq 0$ ist $[0, \infty)$ der Träger der Verteilung, für $\xi < 0$ ist der Träger das kompakte Intervall $[0, -1/\xi]$.

(3) Die Unabhängigkeit

Zuletzt lässt sich zeigen, dass der Punktprozess der Überschreitungen und die Exzesse, also die Höhe der Überschreitungen, im Limes unabhängig sind.

Wie kann man diese Grenzwertsätze nutzen, um Tails und Quantile zu schätzen?

Der folgende Abschnitt beschreibt die POT-Methode für eine Stichprobe X_1, \dots, X_n . Für eine hohe Schwelle u sei

$$N_u = \#\{i : X_i > u, i = 1, \dots, n\}.$$

Mit Y_1, \dots, Y_{N_u} bezeichnen wir die Exzesse von X_1, \dots, X_n , wie auch in Abbildung 3. dargestellt. Sei X eine Zufallsvariable mit der gleichen Verteilungsfunktion $F(x) = P(X \leq x)$, $x \in \mathbb{R}$, wie jedes X_k , $k = 1, \dots, n$. Mit $\bar{F} = 1 - F$ bezeichnen wir den Tail von F . Weiter sei $\bar{F}_u(y) = P(Y_1 > y \mid X > u)$, $y \geq 0$, dann gilt

$$\bar{F}_u(y) = P(X - u > y \mid X > u) = \frac{\bar{F}(u + y)}{\bar{F}(u)}, \quad y \geq 0.$$

Folglich gilt

$$(6) \quad \bar{F}(u + y) = \bar{F}(u)\bar{F}_u(y), \quad y \geq 0.$$

Einen Schätzer für den Tail (für Werte grösser als u) erhält man, indem man beide Tails auf der rechten Seite der obigen Gleichung schätzt.

Dazu verwendet man (1)–(3).

Man schätzt $\bar{F}(u)$ durch das empirische Gegenstück, die relative Häufigkeit von Überschreitungen von u durch X_1, \dots, X_n :

$$(7) \quad \widehat{\bar{F}(u)} = \frac{N_u}{n}.$$

Beachte, dass dieses Vorgehen für $\bar{F}(u + y)$ selber nicht sinnvoll ist, da im allgemeinen dazu zu wenige Beobachtungen zur Verfügung stehen. Der zweite Faktor in (6) $\bar{F}_u(y)$ wird durch die verallgemeinerte Paretoverteilung approximiert, wobei man die Skalenfunktion $\beta(u)$ berücksichtigen muss. Sie wird als Parameter β in die Grenzverteilung integriert.

$$(8) \quad \bar{F}_u(y) \approx \left(1 + \xi \frac{y}{\beta}\right)^{-1/\xi},$$

wobei ξ und β (durch $\widehat{\xi}$ und $\widehat{\beta}$) geschätzt werden müssen.

Man erhält dann einen Tailschätzer der Form

$$(9) \quad \widehat{F}(u+y) = \frac{N_u}{n} \left(1 + \widehat{\xi} \frac{y}{\widehat{\beta}} \right)^{-1/\widehat{\xi}}, \quad y \geq 0.$$

Für gegebenes $q \in (0, 1)$ erhält man durch Inversion einen Schätzer für das q -Quantil von der Form

$$(10) \quad \widehat{x}_q = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left(\left(\frac{n}{N_u} (1-q) \right)^{-\widehat{\xi}} - 1 \right).$$

Schätzung des Value-at-Risk

Eine gängige Methode, die Parameter ξ und β zu schätzen, ist die Maximum Likelihood (ML) Schätzung. Diese beruht auf der numerischen Maximierung der Likelihood-Funktion zu den betrachteten Datenpunkten, hier also den Y_1, \dots, Y_{N_u} . Allerdings ist dabei zu beachten, dass die Schätzungen oft nur auf wenigen Daten basieren, da lediglich die Exzesse in das Schätzverfahren einfließen. Aus diesem Grund kann man sich nicht auf die asymptotischen Optimalitätseigenschaften der ML-Schätzer verlassen.

Wie im letzten Absatz zu sehen war, liegen der ML-Schätzung die Daten Y_1, \dots, Y_{N_u} zugrunde, d.h. es muss ein Schwellenparameter u gewählt werden. Eine exzellente Möglichkeit für die Wahl von u bietet wegen (8) die *Exzessfunktion*

$$(11) \quad e(u) = E(X - u \mid X > u), \quad u \geq 0.$$

Man rechnet leicht nach, dass die Exzessfunktion der Exponentialverteilung eine Konstante ist, nämlich gerade ihr Parameter. Die Exzessfunktion von Verteilungen, deren Tail leichter ist als der Tail der Exponentialverteilung, fällt gegen 0; bei Verteilungen mit fetterem Tail strebt die Exzessfunktion gegen unendlich.

Es lässt sich zeigen, dass die Exzessfunktion (11) im Falle einer verallgemeinerten Paretoverteilung mit $\xi < 1$ und beliebigem $\beta > 0$ linear steigt. Damit ist infolge (8) klar, wie man theoretisch eine vernünftige Schwelle u wählt: Wähle u so, dass die empirische mittlere Exzessfunktion in einer grösstmöglichen rechtsseitigen Umgebung von u linear wächst. Abb. 4. zeigt eine steigende empirische mittlere Exzessfunktion, wobei es schwierig ist, eine eindeutige Schwelle u zu finden. Wir legen deshalb an dieser Stelle die Schwelle noch nicht fest, sondern schätzen sämtliche Parameter in Abhängigkeit von u .

Mit $z^+ = \max(z, 0)$ bezeichnen wir wie üblich den Positivteil von z . Die empirische Funktion

$$(12) \quad e_n(u) = \frac{1}{\#\{i : X_i > u, i = 1, \dots, n\}} \sum_{i=1}^n (X_i - u)^+, \quad u \geq 0,$$

schätzt die *Exzessfunktion*

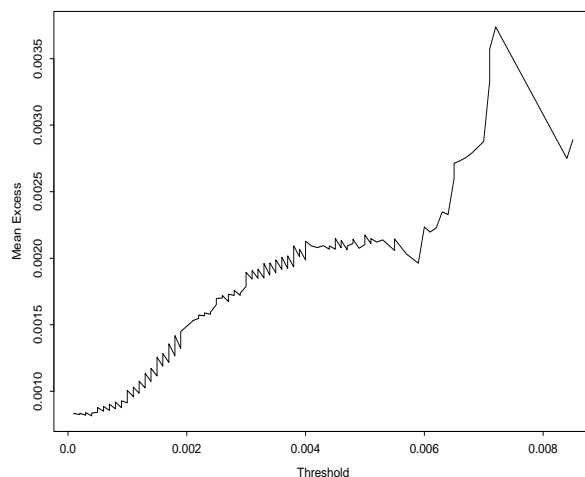


Abb. 4. Die mittlere empirische Exzessfunktion der Verluste.

Wie in Emmer et al. (1998) gezeigt wurde, führt eine Quantilschätzung mit Hilfe der Extremwerttheorie für iid Daten normalerweise zu einer sehr viel höheren (realistischeren) Risikoeinschätzung als die empirische und die normale Methode.

4. Die POT-Methode für stationäre Zeitreihen

Wie wir in Abschnitt 3. gesehen haben, muss man zum Risikomanagement nicht die gesamte Zeitreihe modellieren. Es genügt, den Teil der Zeitreihe adäquat zu beschreiben, der über einer hohen Schwelle liegt. Die POT-Methode kann unter schwachen Bedingungen an eine stationäre Zeitreihe modifiziert werden, so dass auch Cluster in den Extrema erfasst werden.

Stationäre Zeitreihen können sowohl Langzeitabhängigkeit als auch lokale Abhängigkeiten aufweisen. Der Einfluss von Langzeitabhängigkeit auf das extreme Verhalten von Zeitreihen wird in dieser Arbeit nicht untersucht. Die hier vorgeschlagene Methode funktioniert unter Mischungsbedingungen, die Langzeitabhängigkeit ausschliessen, aber lokale Abhängigkeiten zulassen; vgl. Embrechts et al. (1997), Abschnitt 4.4.

Intuitiv dürfte klar sein, was Cluster sind; Volatilitätscluster ist ein häufig gebrauchter Terminus. Cluster in den Extrema bedeutet, dass Volatilitätscluster so grosse Ausschläge zeigen, dass sie über einer hohen Schwelle sichtbar werden. Das hat offensichtlich Konsequenzen für das Risikomanagement. Wesentlich für die mathematische Behandlung ist es, Cluster geeignet zu beschreiben. Dies kann auf verschiedene Art und Weise geschehen. Der vielleicht einfachste Weg ist die “Blockdefinition”. Hierbei werden die Beobachtungen X_1, \dots, X_n einer stationären Zeitreihe mit Randverteilungsfunktion $F(x) = P(X \leq x)$, $x \in \mathbb{R}$, für $i = 1, \dots, n$, in k Blöcke der Grösse r aufgeteilt. Wir nehmen der Einfachheit halber an, dass $n = kr$ (sonst nehmen wir $r = \lceil n/k \rceil$, d.h. die grösste

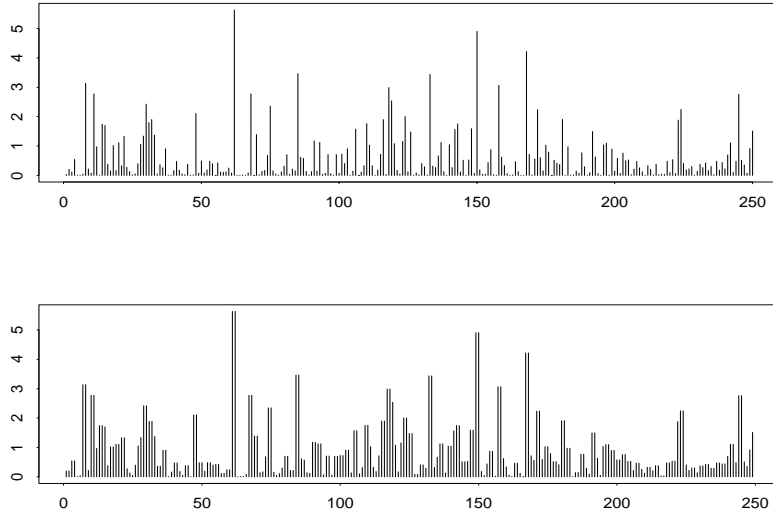


Abb. 5. $X_k = \max(Y_k, Y_{k+1})$ für iid Y_k . Dann ist $\theta = 0.5$, alle extremen Cluster haben deterministische Grösse 2. Für mehr Details über dieses Beispiel verweisen wir auf Embrechts et al. (1997), Example 4.4.4.

natürliche Zahl $\leq n/k$). Beachte, dass k und r von n abhängen. Unsere Daten haben nun die Form

$$X_1, \dots, X_r; \dots; X_{(k-1)r+1}, \dots, X_{kr}.$$

Für jeden Block wird nun dessen Maximum

$$(13) \quad M^{(i)} = \max(X_{(i-1)r+1}, \dots, X_{ir}), \quad i = 1, \dots, k.$$

bestimmt. Schliesslich wählen wir eine Folge von Schwellenwerten (u_n) so, dass (3) gilt, d.h.

$$(14) \quad \lim_{n \rightarrow \infty} nP(X_i > u_n) = \tau \in (0, \infty).$$

Die Definition der Cluster ist jetzt denkbar simpel: Sobald es Blöcke gibt, in denen mindestens eine Beobachtung über der Schwelle u_n liegt (d.h. wo $M^{(i)} > u_n$ gilt), werden alle Exzedenten in diesem Block zu einem Cluster zusammenfasst. Die Clusterposition kann unterschiedlich definiert werden. Üblicherweise nimmt man immer den Beginn oder das Ende des zugehörigen Blockes. Wenn man r respektive k geeignet wählt, kann man mit Hilfe der oben genannten Mischungseigenschaften die (asymptotische) Unabhängigkeit der Cluster zeigen. In der Praxis ist die Wahl von r nicht ganz einfach. Zum einen kennt man die genauen Mischungseigenschaften der zu untersuchenden Zeitreihe nicht. Zum anderen steht man vor dem folgenden Dilemma: Für die asymptotische Unabhängigkeit muss man r hinreichend gross wählen, andererseits darf man r nicht zu gross wählen, da nur noch

die Clustermaxima $M^{(i)}$, $i = 1, \dots, k$, zur Schätzung herangezogen werden, wie wir später sehen werden (vgl. (23) und (26)).

In Analogie zu (4) definieren wir den zeitnormierten Punktprozess der Clusterexzedenten durch

$$(15) \quad N_n^C((a, b]) = \#\{i r/n \in (a, b] : M^{(i)} > u_n, i = 1, \dots, k\}$$

für alle Intervalle $(a, b] \subset (0, 1]$. Wenn u_n so gewählt wird, dass (14) gilt, sowie einige weitere schwache technische Bedingungen für die Zeitreihe gelten, dann konvergiert die Folge N_n^C von Punktprozessen für $n \rightarrow \infty$ in Verteilung gegen einen Poissonprozess mit Intensität $\theta\tau$ für ein $\theta \in (0, 1]$, was wiederum zeigt, dass die Cluster (asymptotisch) unabhängig sind.

Wichtig ist dabei, dass die Zeitreihe stationär ist. Der stationäre Parameterraum und der Tail der stationären Verteilung ist explizit bekannt für lineare Modelle, ARCH(1), GARCH(1,1), AR(1)+ARCH(1) (vgl. Embrechts et al. (1997), de Haan et al. (1989), Mikosch/Starica (1998) und Borkovec/Klüppelberg (1998)). Für GARCH(p,q), $p, q > 1$, sind nur hinreichende Bedingungen für Stationarität bekannt, die leider für viele Finanzzeitreihen nicht erfüllt sind.

Der Parameter θ heisst der *extreme Index* der Zeitreihe. Im nächsten Abschnitt wollen wir die Bedeutung des extremen Index θ erläutern. Er spielt eine wesentliche Rolle bei der Modellierung und Interpretation für die Tail- und Quantilschätzung.

Der extreme Index

Wir beginnen mit einem einfachen Beispiel, um die Bedeutung des extremen Index zu erläutern.

Beispiel 1. Wir wollen die Eigenkapitalreserven berechnen, um ein Portfolio für 10 Tage abzusichern, so dass es mit 95% Wahrscheinlichkeit als Reserve ausreicht. Angenommen, dass die 99% und 99.5% Quantile der täglichen Verlusthöhe 10 Mio bzw. 11 Mio betragen. Falls die Tagesdifferenzen iid mit Erwartungswert 0 sind, dann sind Reserven in Höhe von 11 Mio nötig ($0.995^{10} \approx 0.95$). Falls aber die Tagesdifferenzen abhängig sind mit extremem Index $\theta = 0.5$, dann sind Reserven in Höhe von 10 Mio ausreichend ($0.99^5 \approx 0.95$).

Dieses Beispiel zeigt, dass die Schätzung des extremen Indexes ein zentrales Problem der Extremwertstatistik für abhängige Daten ist. Eine solche Schätzung hat, wie auch die Schätzverfahren für Quantile, mit dem Problem von kleinen Stichproben zu kämpfen. Wir nehmen wieder probabilistische Eigenschaften des extremen Indexes zuhelfe, um zuverlässige Schätzverfahren herzuleiten. Diese Eigenschaften führen auch zu Interpretationen des extremen Indexes.

Zuerst einmal beschreibt der extreme Index den Unterschied zwischen der Verteilung des Maximums der Zeitreihe und dem Maximum einer unabhängigen Stichprobe mit

gleicher Randverteilung F . Wir beginnen mit der technischen Definition: Falls man für $\tau \in (0, \infty)$ eine Folge (u_n) findet, so dass

$$(16) \quad \lim_{n \rightarrow \infty} nP(X > u_n) = \tau,$$

$$(17) \quad \lim_{n \rightarrow \infty} P(M_n \leq u_n) = \exp(-\theta\tau),$$

gelten, dann heisst θ der extreme Index der Zeitreihe (X_n) . Man kann dann zeigen, dass $\theta \in [0, 1]$ liegt. Für iid Daten ist $\theta = 1$, der Fall $\theta = 0$ ist ein pathologischer Fall, der für uns keine Rolle spielt.

Um einen Vergleich der Zeitreihe und den entsprechenden iid Beobachtungen zu ziehen, ignorieren wir die Abhängigkeit der Zeitreihe, d.h. wir schauen die sogenannte unabhängige assoziierte Folge (\tilde{X}_n) an ((\tilde{X}_n) ist die iid Folge mit Verteilungsfunktion F).

Falls nun (X_n) den extremen Index $\theta \in (0, 1]$ hat, dann sind für Folgen (u_n) und Werte $\tau \in (0, \infty)$ äquivalent:

$$\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau,$$

$$(18) \quad \lim_{n \rightarrow \infty} P(\tilde{M}_n \leq u_n) = \exp(-\tau),$$

$$(19) \quad \lim_{n \rightarrow \infty} P(M_n \leq u_n) = \exp(-\theta\tau),$$

Der extreme Index hat eine schöne, anschauliche Interpretation; er stellt gerade den reziproken Wert der mittleren Clustergrösse dar. Betrachte dazu wieder den Punktprozess der Exzedenten N_n in (4), d.h.

$$N_n((a, b]) = \#\{i/n \in (a, b] : X_i > u_n, i = 1, \dots, n\},$$

für alle Intervalle $(a, b] \subset (0, 1]$, wobei die X_i jetzt nicht mehr unabhängig sein müssen. Mit Hilfe des Konvergenzresultates von N_n^C kann man zeigen, dass N_n für eine stationäre Zeitreihe mit gleichen Mischungseigenschaften wie bisher in Verteilung gegen einen zusammengesetzten Poissonprozess konvergiert, der sich wie folgt beschreiben lässt:

$$N((a, b]) = \sum_{n=1}^{N^C(b) - N^C(a)} Y_n, \quad (a, b] \subset (0, 1],$$

wobei $\sum_{n=1}^0 Y_n = 0$ gesetzt wird. N^C ist der Grenzprozess der Punktprozessfolge N_n^C , also ein Poissonprozess mit Intensität $\theta\tau$, und die Y_n sind iid Zufallsvariable, die die Clustergrösse beschreiben, d.h. die Y_n nehmen Werte in \mathbb{N} an und haben eine diskrete Wahrscheinlichkeitsverteilung $(\pi_n)_{n \in \mathbb{N}}$.

Die folgende Rechnung erklärt die Interpretation von θ :

$$(20) \quad \tau = \lim_{n \rightarrow \infty} nP(X > u_n) = \lim_{n \rightarrow \infty} EN_n(0, 1] = EN(0, 1] = \theta\tau EY_1.$$

Damit ist $EY_1 = 1/\theta$, d.h. die mittlere Clustergrösse ist gerade $1/\theta$.

Um θ zu schätzen, betrachten wir wieder die Blockmaxima $M^{(i)}$, $i = 1, \dots, k$, aus (13). Dabei nehmen wir an, dass r so gross gewählt ist, dass die Maxima der einzelnen Blöcke als unabhängig angesehen werden können (sie haben auch die gleiche Verteilung, da die Zeitreihe stationär ist). Wegen dieser Unabhängigkeit kann man nun das Gesamtmaximum der Daten bestimmen durch ($M^{(i)} \stackrel{d}{=} M$ for $i = 1, \dots, k$)

$$(21) \quad P(M_n \leq u_n) = P(\max_{1 \leq i \leq k} M^{(i)} \leq u_n) = (P(M_r \leq u_n))^k \approx \left(1 - \frac{K}{k}\right)^k,$$

wobei $K = N_n^C((0, 1])$, d.h. die Anzahl Cluster respektive die Anzahl Blöcke mit mindestens einem Exzedenten darstellt. Für die Approximation in (21) wurde die theoretische Verteilungsfunktion durch die empirischen Verteilungsfunktion der iid $M^{(1)}, \dots, M^{(k)}$ ersetzt. Die Folge (u_n) wurde wieder so gewählt, dass (14) gilt.

Die Approximation (19) bedeutet, dass

$$P(M_n \leq u_n) \approx F^{n\theta}(u_n) \quad \left(= P(\widetilde{M}_n \leq u_n)^\theta \right).$$

Das lässt sich übersetzen in

$$\lim_{n \rightarrow \infty} \frac{\ln P(M_n \leq u_n)}{n \ln F(u_n)} = \theta.$$

Setzt man nun obige Approximation (21) in diese Beziehung ein, erhält man den Schätzer

$$(22) \quad \widehat{\theta}_n = \frac{k \ln(1 - K/k)}{n \ln(1 - N/n)}.$$

Dabei ist $N = N_n((0, 1])$ die Anzahl der Exzedenten von u_n in X_1, \dots, X_n und k ist wie zuvor die Anzahl der Blöcke.

Nun kann man zeigen, dass unter Bedingung (3), wenn also (4) für den Parameter ξ und Skalenfunktion $\beta(u)$ gilt, auch

$$(23) \quad \lim_{u \rightarrow \infty} P\left(\frac{M^{(i)} - u}{\beta(u)} > y \mid M^{(i)} > u\right) = (1 + \xi y)^{-1/\xi},$$

für $1 + \xi x > 0$, gilt und zwar für den gleichen Parameter ξ und gleiche Skalenfunktion $\beta(u)$.

Wie verändert sich der Quantilschätzer (9) im Fall einer stationären Zeitreihe mit extremem Index θ ? Die POT-Methode basiert nach wie vor auf der Zerlegung (5)

$$(24) \quad \overline{F}(u + y) = \overline{F}(u)\overline{F}_u(y), \quad y \geq 0.$$

Es gibt verschiedene Ansätze, um die beiden Faktoren $\overline{F}(u)$ und $\overline{F}(u + y)$ zu schätzen (vgl. Coles/Tawn (1998)):

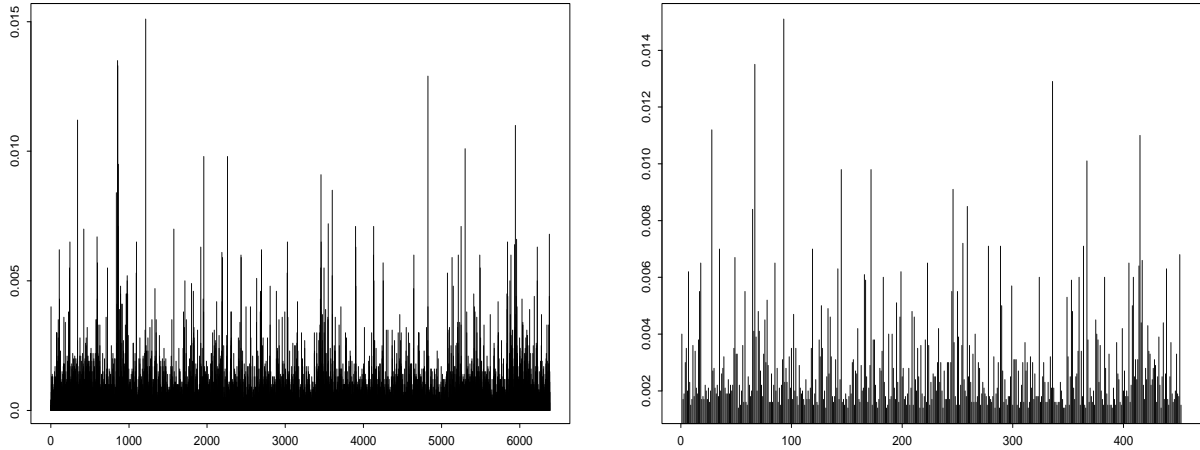


Abb. 6. *Ursprüngliche Daten (nur Verluste) und die Blockmaxima. Als Blockgrösse wurde $r = 20$ gewählt.*

1. Ignoriere die Abhängigkeit und benutze die Standard-POT-Methode, wie sie in (7)-(10) beschrieben wurde. Diese Schätzmethode ergibt unter den entsprechenden Mischungsbedingungen konsistente Schätzer, allerdings sind die Konvergenzraten erheblich langsamer (die Varianz der Schätzer ist grösser). Präzise Konvergenzraten findet man in Drees (1999).
2. Identifiziere die Cluster der extremen Beobachtungen und nehme als Daten nur die Clusterpositionen und die Clustermaxima. Modelliere die Clusterpositionen als homogenen Poissonprozess mit Intensität $\tau\theta$ und approximiere die Verteilungsfunktion der Exzedenten der Clustermaxima durch eine verallgemeinerte Paretoverteilung, unabhängig von dem Poissonprozess der Clusterpositionen. Der Vorteil dieser Methode ist es, dass sie einfach anzuwenden ist, der Nachteil, dass sie Datenmaterial verschwendet. Ausserdem ist die Identifikation der Cluster meist willkürlich.
3. Modelliere die Zeitreihe als lineares Modell, (G)ARCH oder stochastisches Volatilitätsmodell, für das man den extremen Index kennt. Der Vorteil ist, dass man den extremen Index theoretisch berechnen kann, also nicht aus den Daten schätzen muss. Allerdings wird die ganze Zeitreihe gefittet, nicht nur ihr extremes Verhalten. Simulationen und Studien zeigen, dass diese Methode nicht zu empfehlen ist (vgl. z.B. Mikosch/Starica (1998)).
4. Eine weitere Alternative, die wir in dieser Arbeit nicht berücksichtigen, besteht darin, die Abhängigkeit in den Exzedenten explizit zu modellieren, z.B. durch ei-

ne Markovkette (Smith et al. (1994)). Ein Vorteil dieser Methode besteht darin, dass die Likelihoodfunktion explizit berechnet werden kann, und die Abhängigkeit berücksichtigt wird. Ausserdem kann der extreme Index explizit berechnet werden, muss also nicht aus den Daten geschätzt werden. Ein Nachteil der Methode ist die Rechenkomplexität sowie ein Mangel an Robustheit durch die Modellspezifikation.

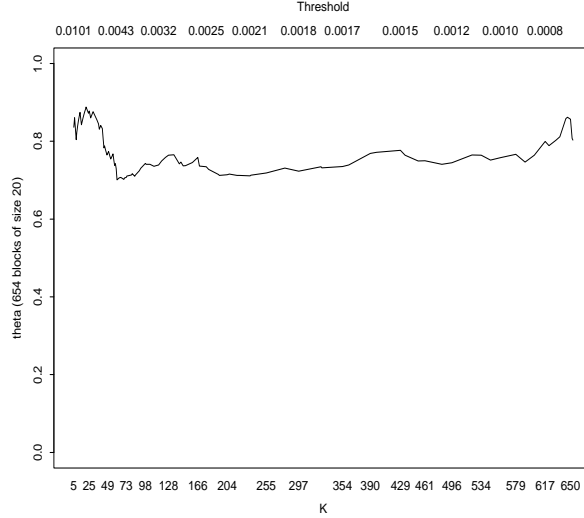


Abb. 7. $\hat{\theta}(u)$ wurde mittels der Blockmethode für $r = 20$ geschätzt.

Die erste Methode haben wir in (7)-(10) beschrieben. Nun zeigen wir, wie man die 2. Methode statistisch umsetzt. Für eine hohe Schwelle u sei nun

$$N_u^C = \#\{\text{Cluster von Exzedenten von } u\}.$$

Wir schätzen den extremen Index θ mittels der Blockmethode. Dazu teilen wir die Beobachtungen in Cluster mit jeweils r Beobachtungen auf und schätzen θ mittels $\hat{\theta}_u$ als Funktion von verschiedenen Schwellenwerten u .

Dann schätzt man $\overline{F}(u)$ durch

$$(25) \quad \widehat{\overline{F}}(u) = \frac{N_u^C}{n\hat{\theta}_u},$$

d.h. wir ersetzen die Anzahl Exzedenten N_u in (7) durch die Anzahl Clusterexzedenten, multipliziert mit der mittleren Clustergrösse.

Als nächstes schätzen wir $\overline{F}_u(y)$ basierend auf den Exzessen der Clustermaxima, die als iid modelliert werden. Für iid Beobachtungen haben wir $\overline{F}_u(y)$ zuerst durch eine GPD approximiert und dann die Parameter mittels ML-Schätzung geschätzt. Die Likelihoodfunktion ist in diesem Fall das Produkt der Randdichten.

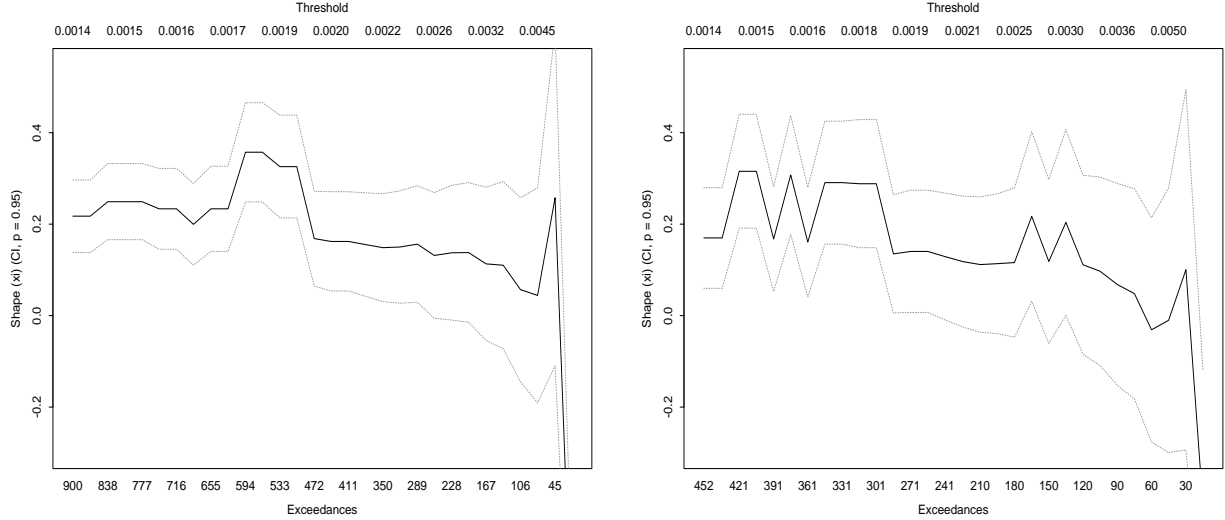


Abb. 8. *Punktweise Maximum Likelihood Schätzer $\hat{\xi}_u$ für Werte $u \geq 0.0014$ mit punktweisen asymptotischen 95%-Konfidenzbändern. Die linke Schätzung basiert auf allen Verlusten, die rechte nur auf den Clustermaxima. Cluster wurden mittels der Blockmethode für $r = 20$ definiert.*

Im abhängigen Fall haben wir die Cluster so gewählt, dass die Clustermaximas iid angesehen werden können. Diese werden nun zur Schätzung von $\bar{F}_u(y)$ herangezogen. Für $i = 1, \dots, N_u^C$ sei $M^{(i)}$ das Clustermaximum.

Nun schätzt man ξ und β durch ML-Schätzung, wobei als Daten die pseudo-iid Clustermaxima $M^{(n_i)}$, $i = 1, \dots, N_u^C$, herangezogen werden. Konkret sind die Schätzer gegeben durch

$$(26) \quad (\hat{\xi}, \hat{\beta}) = \operatorname{argmax}_{(\xi, \beta)} \left(-N_u \ln \beta - (1 + 1/\xi) \sum_{i=1}^{N_u^C} \ln \left(1 + \frac{\xi(M^{(n_i)} - u)}{\beta} \right) \right).$$

Mit dem gleichen Vorgehen wie im iid Fall erhalten wir einen Tailsschätzer von der Form

$$(27) \quad \widehat{\bar{F}}(u + y) = \frac{N_u^C}{n \hat{\theta}_u} \left(1 + \hat{\xi} \frac{y}{\hat{\beta}} \right)^{-1/\hat{\xi}}, \quad y \geq 0.$$

und durch Inversion schliesslich einen Schätzer für das q -Quantil ($q \in (0, 1)$ beliebig) von der Form

$$(28) \quad \hat{x}_q = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{n \hat{\theta}_u}{N_u^C} (1 - q) \right)^{-\hat{\xi}} - 1 \right).$$

Schranke u	$\hat{\xi}$	$\hat{\beta}$	$\hat{\xi}$	$\hat{\beta}$
0.0010	0.2630	0.00065	0.1470	0.00127
0.0015	0.2030	0.00098	0.1710	0.00126
0.0020	0.1700	0.00121	0.1360	0.00149
0.0025	0.1380	0.00140	0.1150	0.00166
0.0030	0.1010	0.00167	0.1070	0.00177
0.0035	0.0935	0.00175	0.0673	0.00198
0.0040	0.0573	0.00196	-0.0374	0.00250
0.0045	0.0741	0.00192	0.0098	0.00227
0.0050	0.0815	0.00193	0.0188	0.00227
0.0055	0.2609	0.00148	0.2009	0.00164
0.0060	0.6580	0.00089	0.5608	0.00101

Tab. 1. $\hat{\xi}$ und $\hat{\beta}$ in Abhängigkeit von u : Vergleich der Standard-POT-Methode (links) mit der POT-Methode mit abhängiger Modellierung (rechts).

Beachte, dass unser Schätzer für das q -Quantil noch entscheidender von der Schranke u abhängt als im iid Fall, da jetzt noch zusätzlich der geschätzte extreme Index $\hat{\theta}_u$ mit hinzukommt.

5. Ergebnisse der Datenanalyse

In diesem Abschnitt wollen wir die Standard-POT-Methode, welche auf der (falschen) Annahme beruht, dass die Daten unabhängig sind, mit der POT-Methode für stationäre Zeitreihen vergleichen. Als zugrunde liegenden Datensatz haben wir die Wechselkurse \$US/DEM vom 1.1.1996 bis 31.12.1996, jeweils halbstündig beobachtet und bereinigt (keine Wochenenden, Feiertage etc.), genommen und haben die zugehörigen 95%, 97.5%, 99% und 99.9% Quantile der Verluste (negative Differenzen) berechnet. Die negativen Verluste zeigen dabei die typischen ‘stylized features’ wie Clusterbildung in den Maxima und grosse Fluktuationen (vgl. Abbildung 1.).

Um die POT-Methode für stationäre Zeitreihen anwenden zu können, muss man die Blocklänge (d.h. die Länge eines Clusters) von Anfang an festlegen. Es hat sich gezeigt, dass für unseren Datensatz $r = 20$ eine vernünftige Wahl darstellt. Der Schätzer $\hat{\theta}_u$ für den extremen Index in Abhängigkeit von der Schwelle u ist in Abbildung 7. dargestellt. In Tabelle 1. sind die geschätzten Werte für ξ und β für beide Methoden wiederum in Abhängigkeit von u festgehalten. Abbildung 8. gibt einen weiteren Eindruck von der Abhängigkeit des Schätzers $\hat{\xi}$ von der Schwelle u bei beiden Verfahren. es zeigt sich, dass ξ und β sehr stark von der Wahl der Schwelle u abhängen.

Mit Hilfe der Formel (10) respektive (28) können wir schliesslich die Quantile in

Abhängigkeit von der Schwelle u berechnen. Abbildung 9. und Tabelle 2. geben die Resultate wider. Zwei Dinge sind dabei bemerkenswert: Zum einen zeigt es sich, dass die Wahl von u keinen so grossen Einfluss, wie vielleicht nach der Berechnung von ξ und β erwartet, auf die geschätzten Quantile hat. Dies ist sowohl bei der Standard-POT-Methode wie auch bei der POT-Methode für stationäre Zeitreihen der Fall. Der Quantilschätzer scheint also bei beiden Verfahren relativ robust gegenüber Schwellenwechseln zu sein. Die zweite Beobachtung ist sehr interessant. Aus Abbildung 9. ist deutlich sichtbar, dass die Standard-POT-Methode für “kleinere” Quantile (95%, 97.5%) viel risikoaverser ist als die POT-Methode basierend auf den Clustermaxima. Falls man aber Quantile schätzen will, die sehr nahe an 1 liegen, dann verändert sich das Bild drastisch. Dies zeigt, dass man bei der Quantilbestimmung sehr vorsichtig sein sollte und immer mehrere Verfahren anwenden sollte. Will man sehr grosse Quantile schätzen, dann liefert die POT-Methode für stationäre Zeitreihen die konservativere Schätzung.

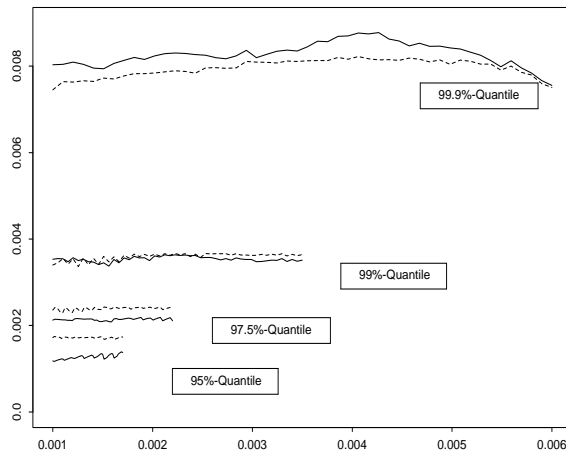


Abb. 9. Schätzwerte $\hat{x}_p(u)$, für grosse Werte von u . Die Schätzung basiert auf der POT-Methode, die gestrichelte Linie auf allen Daten (ohne Berücksichtigung der Abhängigkeit), die durchgezogene Linie auf den Clustermaxima. Da Quantile höherer Ordnung höher sind, gibt es für sehr grosse Quantile ein grösseres Intervall möglicher Schwellenwerte ($u > x_p$ zu wählen, ist nicht möglich).

u	95%	97.5%	99%	99.9%	95%	97.5%	99%	99.9%
0.0010	0.00170	0.00235	0.00339	0.00745	0.00118	0.00213	0.00353	0.00803
0.0015	0.00174	0.00240	0.00358	0.00773	0.00133	0.00210	0.00344	0.00794
0.0020		0.00239	0.00363	0.00785		0.00216	0.00360	0.00785
0.0025			0.00361	0.00796			0.00357	0.00826
0.0030			0.00362	0.00810			0.00348	0.00820
0.0035			0.00366	0.00813			0.00353	0.00838
0.0040				0.00823				0.00877
0.0045				0.00815				0.00862
0.0050				0.00817				0.00846
0.0055				0.00799				0.00801
0.0060				0.00779				0.00762

Tab. 2. *Schätzungen des 95%, 97.5%, 99% und 99.9% Quantils in Abhängigkeit von u : Vergleich der Standard-POT-Methode (links) mit der POT-Methode mit abhängiger Modellierung (rechts).*

Autorenverzeichnis

- Borkovec, M. (2000) Extremal behavior of the autoregressive process with ARCH(1) errors. *Stoch. Proc. Appl.* **85**, 189–207.
- Borkovec, M. and Klüppelberg, C. (1998) The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. Preprint. TU München.
<http://www.ma.tum.de/stat/>
- Borkovec, M. and Klüppelberg, C. (1998) Extremal behaviour of diffusion models in finance. *Extremes* **1**, 47-80. <http://www.ma.tum.de/stat/>
- Coy, P. 1998. How the game was played – and why Long-Term lost. *Business Week* (Oct 12, 1998), 40-41.
- Coy, P. and Wooley, S. (1998) Failed wizards of Wall Street. *Business Week* (Sept 21, 1998), 54-59.
- Coles, S. and Tawn, J. (1998) Statistical methods for extreme values. A course presented at the 1998 RSS conference, Strathclyde. <http://www.maths.lancs.ac.uk/coles>
- Cruz, M., Coleman, R. and Salkin, G. (1998) Modeling and measuring operational risk. *Journal of Risk*, **1**, 63-72.
- Drees, H. (1999) Weighted approximations of tail processes under mixing conditions. Preprint. University of Cologne. <http://www.euklid.mi.uni-koeln.de/hdrees/>
- Elerian, O., Chib, S. and Shephard, N. (1998) Likelihood inference for discretely observed non-linear diffusions. Preprint. Oxford University.
<http://www.nuff.ox.ac.uk/economics/papers/>
- Embrechts, P., Klüppelberg, C. und Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- Embrechts, P., Resnick, S.I. and Samorodnitsky, G. (1998) Extreme value theory as a risk management tool. Preprint. ETH Zurich. <http://www.math.ethz.ch/finance/>
- Embrechts, P., Resnick, S.I. and Samorodnitsky, G. (1998) Living on the edge. *RISK* **11**, 96-100.
- Emmer, S., Klüppelberg, C. and Trüstedt, M. (1998) VaR - ein Mass für das extreme Risiko. *Solutions* **2**, 53-63. <http://www.ma.tum.de/stat/>
- Mc Neil, A.J. and Frey, R. (1999) Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Preprint. ETH Zurich.
<http://www.math.ethz.ch/finance/>

- Haan, L. de (1990) Fighting the arch-enemy with mathematics. *Statist. Neerlandica* **44**, 45-68.
- Haan, L. de, Resnick, S.I., Rootzén, H. and Vries, C. de (1989) Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.* **32**, 213-224.
- Unsignierter Artikel (1998) Uni baut an Dämmen gegen Finanzdesaster. *Handelsblatt* (28.10.98)
- Klüppelberg, C. and Korn, R. (1998) Optimal portfolios with bounded value-at-risk. *Math. Finance*. To appear. <http://www.ma.tum.de/stat/>
- Leadbetter, M.R. (1991) On a basis for 'peaks over threshold' modeling. *Statist. Probab. Letters* **12**, 357-362.
- Mikosch, T. and Starica, C. (1998) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. Preprint. University of Groningen. <http://www.math.rug.nl/~mikosch>
- McNeil, A.J. (1997) History Repeating. *RISK* **11**, 99.
- McNeil, A.J. (1998) Calculating quantile risk measures for financial time series using extreme value theory. Preprint. ETH-Zürich. <http://www.math.ethz.ch/finance/>
- Phoa, W. (1999) Estimating credit spread risk using extreme value theory. *Journal of Portfolio Management* Spring 1999, 69-73.
- Rootzén, H. und Klüppelberg, C. (1999) A single number can't hedge against economic catastrophes. *Ambio* **28** No 6, 550-555. Royal Swedish Academy of Sciences. <http://www.ma.tum.de/stat/>
- Schmock, U. (1999) Estimating the value of the wincat coupons of the Winterthur insurance convertible bond: a study of the model risk. *ASTIN Bulletin* **29**, 101-163.
- Smith, R.L., Tawn, J.A. and Coles, S.G. (1997) Markov models for threshold exceedances. *Biometrika* **84**, 249-268.

Dr. Milan Borkovec studierte an der Eidgenössischen Technischen Hochschule Zürich Mathematik mit Schwerpunkt Stochastik. Für seine Diplomarbeit wurde er mit dem Walter-Saxer Versicherungs-Hochschulpreis ausgezeichnet. Von 1995-1997 hatte Milan Borkovec eine Assistentenstelle an der Johannes-Gutenberg Universität Mainz, seit SS 1997 ist er als Wissenschaftlicher Angestellter an der Technischen Universität München tätig. Im SS 99 promovierte Milan Borkovec in Mathematik mit der Arbeit "Large Fluctuations in Financial Models". Seine Forschungsinteressen liegen im Bereich der Extremwerttheorie, Finanzmathematik und Telekommunikation. Für das akademische Jahr 1999/2000 ist Dr. Borkovec beurlaubt und verbringt es an der Cornell University (USA).
Adresse: School of Operations Research and Engineering, Cornell University, Ithaca, NY 14853-3801, USA, email: borkovec@orie.cornell.edu

Prof. Dr. Claudia Klüppelberg ist Inhaberin des Lehrstuhls für Mathematische Statistik am Zentrum Mathematik der Technischen Universität München. Nach ihrem Studium zur Diplom-Mathematikerin und der Promotion zum Dr. sc. math. an der Universität Mannheim hatte sie fünf Jahre lang eine Oberassistentenstelle an der ETH-Zürich inne, wo sie sich habilitierte. Von 1995-1997 war sie Professorin für Angewandte Statistik am Fachbereich Mathematik der Universität Mainz. Zum SS 1997 folgte Prof. Claudia Klüppelberg einem Ruf an die TU München, wo sie an der Einführung des Studiengangs Finanz- und Wirtschaftsmathematik zum WS 1997/98 maßgeblich beteiligt war. Ihre Forschungsinteressen liegen innerhalb der Finanzmathematik in der Analyse von Finanzdaten, insbesondere im Hinblick auf das Risikomanagement.

Adresse: Zentrum Mathematik, Technische Universität München, 80290 München, email: cklu@ma.tum.de, url: <http://www.ma.tum.de/stat/>