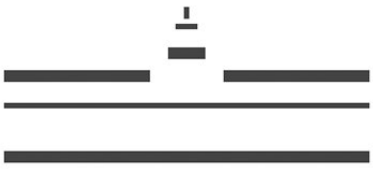

Westfälische
Wilhelms-Universität
Münster

Eine Anwendung der Block Maxima Methode im Risikomanagement

Diplomarbeit

vorgelegt von

Birgit Woeste

Betreuer: Privatdozent Dr. Volkert Paulsen
Mathematisches Institut für Statistik
Fachbereich Mathematik und Informatik
Westfälische Wilhelms-Universität Münster

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1 Einleitung

Im Alltag treten immer wieder Extremereignisse oder neue Rekorde auf. So spricht man häufig von einer Jahrhundertflut, wenn die Pegelstände einen noch nie da gewesenen Höchststand aufweisen. Auch Sturmfluten, Erdbeben oder Stürme, wie etwa 2007 der Sturm Kyrill, können Ausmaße annehmen, die zwar nur selten auftreten, dafür aber verheerende Auswirkungen haben. Solche Naturkatastrophen kann man freilich nicht gänzlich verhindern, aber es ist möglich, Maßnahmen zu ergreifen, die die Zerstörungskraft abschwächen. So ist ein guter Hochwasserschutz sinnvoll, um sich vor Hochwasser und Sturmfluten zu schützen. Doch wie hoch soll ein Deich gebaut werden, damit er nicht nur einer durchschnittlichen Flut sondern auch einer Jahrhundertflut standhält?
Wenn der Schutz nicht ausreichend war oder gar nicht möglich, wie bei starkem Sturm, bezahlen oft Versicherungen den finanziellen Schaden. Also sind auch sie daran interessiert, wie oft oder wie selten solche Extremereignisse auftreten. Denn sie müssen auch bei Großschäden ihren Verpflichtungen gegenüber den Versicherungsnehmern nachkommen. Deshalb müssen die Versicherungen und die Rückversicherungen genügend Risikokapital haben, um gegen Großschäden abgesichert zu sein.
Das Risikokapital spielt nicht nur bei Versicherungen eine große Rolle, sondern auch bei Banken und eigentlich in der ganzen Finanzwelt. Denn hier können Kurseinbrüche, Börsencrashs und Kaptialmarktkrisen zu erheblichen Verlusten führen. Dies hat man in der Finanzmarktkrise 2008/2009 gesehen, wo plötzlich ganze Banken am Rande des Ruins standen.
Doch wie kann man sich gegen Extremereignisse schützen? Wie oft und in welchem Ausmaß treten sie auf? Welche Gegenmaßnahmen können ergriffen werden, um sich gegen seltene, aber verheerende Ereignisse abzusichern?

Ein wichtiges Werkzeug, um obige Fragen zu beantworten, ist die Extremwerttheorie. Im Gegensatz zur Wahrscheinlichkeitstheorie, wo vor allem Mittelwerte von Zufallsvariablen eine Rolle spielen, fokussiert sich die Extremwerttheorie auf extreme Ereignisse, also auf die Maxima von Zufallsvariablen. Ein extremes Ereignis liegt vor, wenn die Flanken (*Tails*) der Verteilung angenommen werden. Da mit der Normalverteilung die Extrema nicht so gut modelliert werden können, denn sie hat zu dünne Tails, werden im Folgenden sogenannte Extremwertverteilungen eingeführt. Diese Verteilungen haben breitere Flanken.

Wie oben erwähnt finden sich viele Anwendungsgebiete für die Extremwerttheorie. In dieser Arbeit wird besonders auf die Bedeutung der Extremwerttheorie in der

Finanzwelt eingegangen. Dazu werden Renditen von Aktienkursen betrachtet. Es ist wichtig sich gegen große Verluste durch ein umsichtiges Risikomangagement zu schützen. Ziel des Risikomanagement ist es, Risikomodelle zu implementieren, die für seltene, aber verheerende Ereignisse geeignet sind. Diese Modelle sollen Risiken bewerten und messen können, um nützliche Maßnahmen zu treffen. Dafür sind folgende Fragen von Interesse:

- Wie tief kann ein Kurs abstürzen? Was ist der „Worst-Case“? Wie wahrscheinlich ist es, dass schon im nächsten Jahr ein neuer negativer Rekord eintritt?
- Oder wenn es sich nicht um einen neuen negativen Rekord handelt: Wie hoch ist die Schwelle, die von den negativen Renditen im Durchschnitt einmal alle k Jahre überschritten wird?
- Und andersherum gefragt: Wie häufig wird eine vorgegebene Schwelle von den negativen Renditen überschritten? In welchen Zeitabständen treten Extremereignisse auf?

Der erste Punkt behandelt ein „Worst-Case“-Szenario und die Fragen, die unter dem zweiten und dritten Punkt aufgelistet sind, werden durch die Berechnungen von Risikomaßen, des sogenannten *Return-Levels* und der sogenannten *Return-Periode*, beantwortet.

In der Extremwerttheorie werden vor allem Maxima betrachtet. Deswegen müssen aus einem Datensatz die größten Werte herausgefiltert werden. Eine Möglichkeit die Maxima zu erhalten, ist die *Block Maxima Methode*, die in dieser Arbeit angewendet wird. Diese Methode geht auf Gumbel [9] zurück und wurde zuerst in der Hydrologie, also im Hochwasserschutz, angewendet. Doch mittlerweile findet die Block Maxima Methode auch Beachtung in der Finanz- und Versicherungsbranche. Nach der Block Maxima Methode wird ein Datensatz in m Blöcke unterteilt, wobei jeder Block aus n Daten besteht. Seien also $X_1, X_2 \ldots, X_{n*m}$ die Zufallsvariablen eines Datensatzes. Dann sind $X_{1j}, \ldots, X_{nj}$ für alle $1 \leq j \leq m$ die Zufallsvariablen für den j-ten Block. Aus den einzelnen Blöcken wählt man dann das Maximum (vgl. [15]):

$$M_{nj} := \max\{X_{1j}, \ldots, X_{nj}\}.$$

Die Unterteilung der Daten in Blöcke richtet sich nach dem Datensatz. Häufig verwendet man jährliche Blöcke. Deswegen heißt diese Methode auch manchmal „Annual Blocks Method“. Dabei wird angenommen, dass alle Blöcke gleich lang sind, auch wenn die Jahre eine unterschiedliche Anzahl an Tagen haben.
Allerdings werden bei diesem Vorgehen sehr viele Daten gebraucht, damit genügend Maxima-Werte zur Verfügung stehen. Deswegen kann es auch sinnvoll sein, die Blöcke nach Halbjahren, Quartalen oder sogar Monaten auszuwählen. Dabei ist aber zu bedenken, dass bei zu kleinen Blöcken Beobachtungen als Maximum bestimmt werden, die eigentlich nicht extrem sind.

Der Aufbau der vorliegenden Arbeit bestimmt sich wie folgt: In Kapitel 2 wird die Block Maxima Methode behandelt unter der Annahme, dass die zu betrachtenden Zufallsvariablen unabhängig und identisch verteilt sind. Doch dies ist nicht immer gegeben, da Renditen oft voneinander abhängig sind. Denn gibt es heute einen Kursabsturz, werden sich die Kurse am nächsten Tag noch nicht wieder erholt haben. Deswegen werden in Kapitel 3 Folgen von stationären Zufallsvariablen betrachtet. Außerdem werden Zeitreihenmodelle eingeführt, die die Besonderheiten von Finanzzeitreihen widerspiegeln.
Jedes Kapitel besteht auch aus einem Anwendungsteil. In diesem werden die Ergebnisse an einem Datensatz veranschaulicht. Der Datensatz besteht aus den täglichen Renditen des Dow Jones Indexes. Die dazu notwendigen Berechnungen werden alle mit dem Statistikprogramm R durchgeführt.

An dieser Stelle möchte ich Dr. Volkert Paulsen für die Überlassung des interessanten Themas und die gute Unterstützung während der Bearbeitungszeit danken. Außerdem bedanke ich mich bei allen, die zum Gelingen dieser Arbeit beigetragen haben.

2 Block Maxima Methode mit iid Zufallsvariablen

2.1 Einführung

Wie in der Einleitung erwähnt, ist es das Ziel dieser Arbeit Methoden zu finden, mit denen man sich gegen Kursabstürze schützen kann. Die Ergebnisse sollen an einem Datensatz veranschaulicht werden. Dazu werden Werte des Dow Jones Indexes betrachtet. Die Renditen des Dow Jones Indexes spielen dabei eine wichtige Rolle, im Gegensatz zu den täglichen Eröffnungs- oder Abschlusswerten. Denn durch die Renditen werden die täglichen Veränderungen ausgedrückt. Darum werden die täglichen Log-Renditen der Abschlusswerte betrachtet, da durch sie starke Kursverluste besser modelliert werden. Eine Ein-Tages Rendite wird folgendermaßen gebildet ([15, S. 3]):

$$\tilde{X}_t = \frac{S_t - S_{t-1}}{S_{t-1}},$$

wobei S_t der Abschlusswert des Tages t ist und für die Log-Rendite gilt:

$$X_t = \log\left(\frac{S_t}{S_{t-1}}\right) = \log S_t - \log S_{t-1}$$

Es gilt zudem

$$\log\left(\frac{S_t}{S_{t-1}}\right) = \log\left(1 + \frac{S_t - S_{t-1}}{S_{t-1}}\right) \approx \frac{S_t - S_{t-1}}{S_{t-1}}$$

und daraus folgt $X_t \approx \tilde{X}_t$, da $\log(1+x) \approx x$ ist nach der Taylor Formel.
Also macht es kaum einen Unterschied, ob man die Renditen oder die Log-Renditen verwendet.
In dieser Arbeit werden aber immer die Log-Renditen betrachtet und so werden nur die Veränderungen oder Tendenzen des Aktienindexes dargestellt. Dadurch sind die Daten leichter miteinander zu vergleichen (s. [7, S. 403]).
Um sich gegen große Kursverluste abzusichern, ist es wichtiger sich mit den negativen Log-Renditen zu befassen als mit den Positiven. Deswegen werden die Renditen negiert, so dass die Verluste die Maxima sind.

In diesem Kapitel wird angenommen, dass die Zufallsvariablen unabhängig und identisch verteilt sind. Dass diese Annahme nicht immer gerechtfertigt ist, da Renditen häufig in sogenannten Clustern auftreten, ist Gegenstand des nächsten Kapitels.
Als erstes werden in diesem Kapitel die wichtigsten Grundlagen der Extremwerttheorie eingeführt. Anschließend werden dann in Kapitel 2.3 die Resultate veranschaulicht, indem sie auf den Dow Jones Datensatz angewendet werden.

Eine andere Methode, als die Block Maxima Methode, um die Extrema zu bestimmen, ist die „Peak over Threshold“ Methode (kurz: POT Methode). Dabei bestimmt man zu einem gegebenen Datensatz einen Schwellenwert (Threshold) und betrachtet anschließend nur die Beobachtungen, die über diesem Schwellenwert liegen. In dieser Arbeit wird die POT Methode nicht verwendet, allerdings wird in Abschnitt 2.4 diese Methode mit der Block Maxima Methode verglichen.

2.2 Grundlagen der Extremwerttheorie

Zu den Grundlagen der Extremwerttheorie gehören die Extremwertverteilungen, die geeignet sind extreme Ereignisse zu modellieren. Das wichtigste Resultat ist das sogenannte Fisher-Tippett Theorem, das eine ähnlich große Bedeutung in der Extremwerttheorie hat wie der zentrale Grenzwertsatz in der Wahrscheinlichkeitstheorie. Denn das Theorem zeigt, dass die Maxima, die nach der Block Maxima Methode ausgewählt werden, gegen eine Extremwertverteilung konvergieren. Die Extremwertverteilungen können zu einer Verteilung, der sogenannten verallgemeinerten Extremwertverteilung, zusammengefasst werden (siehe Abschnitt 2.2.2).
Anschließend werden mit der Maximum-Likelihood Schätzung die Parameter der Extremwertverteilungen geschätzt. Zum Schluss werden in Abschnitt 2.2.3 Return-Level und Return-Periode definiert.

2.2.1 Fisher-Tippett Theorem

Wenn es nicht anders gekennzeichnet ist, dann orientiert sich dieser Abschnitt im Wesentlichen an [13, Kap. 1].

Sei $(X_1, X_2, \ldots)$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit einer unbekannten zugrunde liegenden Verteilungsfunktion F.
Wie in der Einleitung erwähnt, sind vor allem die Extrema von Interesse, also die Maxima. Deshalb wird das Maximum der ersten $n \in \mathbb{N}$ Zufallsvariablen geschrieben als:

$$M_n := \max(X_1, \ldots, X_n).$$

Die besondere Aufmerksamkeit der Extremwerttheorie liegt auf der Verteilung von M_n und seinen Eigenschaften, wenn $n \to \infty$.
Analog zum Maximum wird auch das Minimum definiert:

$$m_n := \min(X_1, \ldots, X_n).$$

Im Folgenden wird immer nur das Maximum betrachtet, allerdings gelten alle Ergebnisse auch für Minima. Das ist nämlich durch

$$m_n = \min(X_1, \ldots, X_n) = -\max(-X_1, \ldots, -X_n)$$

gerechtfertigt.

Für die Verteilungsfunktion des Maximums F^{M_n} gilt:

$$\begin{aligned} F^{M_n}(x) &= P(M_n \leq x) \\ &= P(\max_{1 \leq i \leq n} X_i \leq x) \\ &= P(\bigcap_{i=1}^{n} \{X_i \leq x\}) \\ &= \prod_{i=1}^{n} P(X_i \leq x) = F^n(x), \end{aligned} \tag{2.1}$$

wobei $x \in \mathbb{R}$ und $n \in \mathbb{N}$. In den Gleichungen wurde verwendet, dass die X_i unabhängig und identisch verteilt sind.

Da Maxima in dem rechten Tail der Verteilung auftreten, muss man das asymptotische Verhalten von M_n an der oberen Grenze betrachten. Sei

$$x_F = sup\{x \in \mathbb{R} : F(x) < 1\}$$

der *rechte Randpunkt* von F. Für alle $x < x_F$ folgt nun:

$$\lim_{n \to \infty} P(M_n \leq x) = \lim_{n \to \infty} F^n(x) = 0.$$

Wenn $x_F < \infty$, dann gilt für $x \geq x_F$, dass

$$\lim_{n \to \infty} P(M_n \leq x) = \lim_{n \to \infty} F^n(x) = 1.$$

Da die Folge (M_n) nicht-fallend ist, folgt $M_n \to x_f$ *P-f.s.* (vgl. hierzu [7, S. 114ff]). Aber dieses Resultat ist nicht zufriedenstellend, weil es eigentlich keine weiteren Informationen liefert, deswegen muss M_n skaliert werden.
Mit geeigneten Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$, soll M_n so skaliert werden, dass das skalierte M_n gegen eine Grenzverteilung G konvergiert, also:

$$a_n(M_n - b_n) \xrightarrow{d} G, \quad n \to \infty.$$

Dieses kann man auch mit Hilfe von (2.1) umschreiben als:

$$F^n\Big(\frac{x}{a_n} + b_n\Big) \to G(x), \quad n \to \infty. \tag{2.2}$$

Daraus ergibt sich die erste Definition:

Definition 2.1.
Seien F und G Verteilungsfunktionen.
F liegt im Max-Anziehungsbereich *(Maximum domain of attraction (MDA)) von G, falls es Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$ gibt, so dass gilt*

$$F^n\Big(\frac{x}{a_n} + b_n\Big) \underset{n\to\infty}{\longrightarrow} G(x)$$

für alle x, in denen G stetig ist. Als Abkürzung wird die Schreibweise $F \in MDA(G)$ benutzt.

Das Ziel ist es Limesverteilungen für das skalierte M_n zu finden. Beim zentralen Grenzwertsatz ist die Normalverteilung die Limesverteilung. Doch welche Verteilungen kommen hier in Betracht?
Um eine Antwort zu finden, ist der Begriff *max-stabil* nützlich, denn eine max-stabile Verteilung hat in etwa die gleiche Eigenschaft, wie die Normalverteilung, die invariant unter Faltung ist.

Definition 2.2.
Sei G eine nicht-entartete Verteilungsfunktion, d.h. das dazugehörige Maß ist kein Dirac-Maß. G heißt genau dann max-stabil, *wenn es für jedes $n \in \mathbb{N}$ Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$ gibt, so dass*

$$G^n\Big(\frac{x}{a_n} + b_n\Big) = G(x) \quad \textit{für alle } x \in \mathbb{R} \tag{2.3}$$

gilt.
Eine nicht-entartete Verteilung heißt max-stabil, *wenn die dazugehörige Verteilungsfunktion max-stabil ist.*

Diese Definition ist hilfreich, denn die möglichen nicht-entarteten Verteilungsfunktionen, die als Grenzverteilung von (2.2) in Frage kommen, gehören zur Klasse der max-stabilen Verteilungen. Diese werden als Extremwertverteilungen bezeichnet. Dazu gehören folgende drei Verteilungen:

Definition 2.3.

1. *Die Verteilung mit der Verteilungsfunktion*

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

heißt Gumbel-Verteilung.

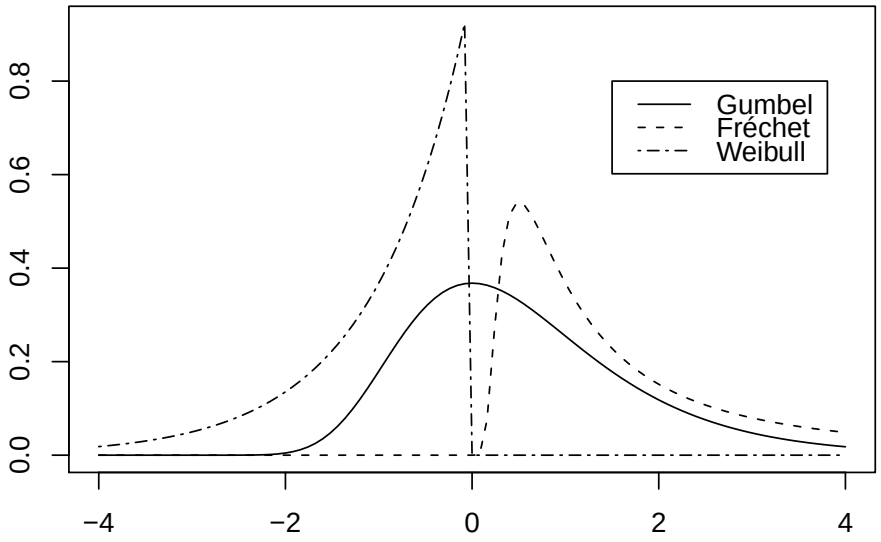

Abbildung 2.1: Die Dichten der Extremwertverteilungen mit $\alpha = 1$ für die Fréchet- und Weibull-Verteilung.

2. *Eine Verteilung mit Verteilungsfunktion*

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

heißt Fréchet-Verteilung.

3. *Die* Weibull-Verteilung *hat die Verteilungsfunktion:*

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

Diese Verteilungen werden manchmal auch als Extremwertverteilungen vom Typ I, II oder III bezeichnet (vgl. [12, S. 4]).

In Abbildung 2.1 sind die Dichten der drei Extremwertverteilungen dargestellt, dabei wurde $\alpha = 1$ für die Fréchet- und Weibull-Verteilung gewählt. Für die Weibull-Verteilung ist der rechte Randpunkt endlich, im Gegensatz zur Gumbel- und Fréchet-Verteilung, wo $x_F = \infty$ gilt. Alle drei Verteilungen sind nicht, wie die Normalverteilung, symmetrisch, haben dafür aber wesentlich breitere Tails.

Im nachfolgenden Beispiel wird gezeigt, dass die Gumbel-, Fréchet- und Weibull-Verteilung max-stabil sind:

Beispiel 2.1. *(vgl. [13, Bsp. 1.9])*
Sei F Fréchet-verteilt, für $a_n = n^{-1/\alpha}$ und $b_n = 0$ ergibt sich dann:

$$\begin{aligned} F^{n^{-1/\alpha}M_n}(x) &= F^{M_n}(xn^{1/\alpha}) \\ &\overset{(2.1)}{=} F^n(xn^{1/\alpha}) \\ &= \Phi_\alpha^n(xn^{1/\alpha}) \\ &= \left(e^{-(xn^{1/\alpha})^{-\alpha}}\right)^n \mathbf{1}_{(0,\infty)}(xn^{1/\alpha}) \\ &= \left(e^{\frac{-x^{-\alpha}}{n}}\right)^n \mathbf{1}_{(0,\infty)}(x) \\ &= e^{-x^{-\alpha}} \mathbf{1}_{(0,\infty)}(x) \\ &= \Phi_\alpha(x) \quad \textit{für alle } x \in \mathbb{R} \textit{ und } n \in \mathbb{N}. \end{aligned}$$

Die Fréchet-Verteilung ist also max-stabil.
Wenn man $a_n = 1$ und $b_n = \log n$ setzt, kann man auf gleicher Weise zeigen, dass die Gumbel-Verteilung max-stabil ist.
Wird $a_n = n^{1/\alpha}$ und $b_n = 0$ gesetzt, folgt, dass auch die Weibull-Verteilung max-stabil ist.

Eines der wichtigsten Theoreme in der Extremwerttheorie ist das Fisher-Tippett Theorem. Hier haben die max-stabilen Verteilungen eine sehr große Bedeutung, denn in dem Theorem wird bewiesen, dass als Grenzverteilung für normierte Maxima nur die Gumbel-, Fréchet- oder Weibull-Verteilung in Frage kommen. D.h. diese drei Verteilungen sind auch die einzigen Extremwertverteilungen, die es gibt.

Theorem 2.1. *Fisher-Tippett Theorem (vgl. [12, S. 11], [13, Satz 1.17])*
Sei $M_n = \max(X_1, \ldots, X_n)$ mit (X_n) iid und $n \to \infty$. Dann ist äquivalent:

(i) Eine nicht-entartete Verteilungsfunktion G ist eine Extremwertverteilung, d.h. es existieren Normierungskonstanten $a_n > 0$ und $b_n \in \mathbb{R}$, so dass gilt

$$a_n(M_n - b_n) \xrightarrow{d} G, \quad n \to \infty,$$

wobei G eine nicht-entartete Verteilungsfunktion ist.

(ii) G gehört zum Typ einer Gumbel-, Fréchet- oder Weibull-Verteilung.

In dem Buch von Leadbetter [12] und in dem Skript von Löwe [13] findet sich der ausführliche Beweis.
Im Wesentlichen werden dabei zwei Schritte berücksichtigt:
Zunächst wird gezeigt, dass G genau dann eine Extremwertverteilung ist, wenn G

max-stabil ist. Im zweiten Schritt wird dann bewiesen, dass eine Verteilung genau dann max-stabil ist, wenn sie vom Typ der Gumbel-, Fréchet- oder Weibull-Verteilung ist.

In Definition 2.1 wurde der Begriff Max-Anziehungsbereich eingeführt. Es stellen sich nun, im Hinblick darauf, dass die Extremwertverteilungen die Grenzverteilungen von normierten Maxima sind, folgende Fragen: Welche Verteilungen liegen im Anziehungsbereich der Extremwertverteilungen und welche Bedingungen müssen dafür erfüllt sein? Kann man sogar den Typ der Extremwertverteilung bestimmen?

Um diese Fragen zu beantworten, ist der rechte Randpunkt x_F nützlich, denn es ist wichtig zu wissen, welches Verhalten F in der Nähe von x_F hat, wenn das normierte M_n konvergiert.

Definition 2.4.
Sei $f : (0, \infty) \to (0, \infty)$ eine messbare Funktion.

1. *Die Funktion f heißt* langsam variierend in ∞, *wenn*

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = 1$$

   für alle $x > 0$ gilt.

2. *Die Funktion f heißt* regulär variierend in ∞ *mit Index $\alpha \in \mathbb{R}$, falls*

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = x^\alpha$$

   für alle $x > 0$ gilt. Mit $\mathcal{R}_\alpha^\infty$ wird die Klasse der regulär variierenden Funktionen zum Index $\alpha \in \mathbb{R}$ bezeichnet.

Eine regulär variierende Funktion verhält sich in ∞ asymptotisch wie eine Potenzfunktion.
Um das Verhalten der Verteilungsfunktion F in der Nähe des rechten Randpunktes besser zu beschreiben, betrachtet man die Tailfunktion und die Quantilfunktion von F.

Definition 2.5. *(vgl. [7, Def. 3.3.5], [13, Def. 1.23])*

1. *Sei F eine Verteilungsfunktion. Die* Tailfunktion von F *ist definiert als:*

$$\bar{F}(x) := 1 - F(x) \quad \textit{für alle } x \in \mathbb{R}.$$

2. *$\bar{F}$ heißt* regulär variierend *mit Index $\alpha \in \mathbb{R}$ in ∞, wenn gilt:*

$$\lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^\alpha \quad \textit{für alle } x > 0.$$

3. *Sei F eine Verteilungsfunktion. Die Pseudo-Inverse von F*

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad 0 < t < 1,$$

wird Quantil-Funktion *von F genannt. Das t-Quantil von F wird als $x_t := F^{-1}(t)$ definiert.*

Der nachfolgende Satz beantwortet jetzt die Frage, wann eine Verteilungsfunktion im Anziehungsbereich einer Extremwertverteilung liegt und klärt auch, um welche der drei Extremwertverteilungen es sich handelt.

Satz 2.1. *(vgl. [13, Satz 1.24])*
Sei F eine Verteilungsfunktion und γ_n sei für $n \geq 2$ definiert durch

$$\gamma_n := F^{-1}\Big(1 - \frac{1}{n}\Big).$$

Dann gilt:

1. *F liegt genau dann im Anziehungsbereich der Fréchet-Verteilung Φ_α (kurz: $F \in MDA(\Phi_\alpha)$), wenn*

$$x_F = +\infty \quad \textit{und} \quad \lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} \;=\; x^{-\alpha} \quad \textit{für alle } x > 0.$$

D.h. $\bar{F}$ ist in ∞ regulär variierend mit Index $-\alpha$. Man kann hier die Folgen (a_n) und (b_n) als $a_n = \gamma_n^{-1}$ und $b_n = 0$ wählen für die schwache Konvergenz von $a_n(M_n - b_n)$.

2. *Sei $\bar{F}^*(x) = 1 - F^*(x)$ die zu*

$$F^*(x) := F\Big(x_F - \frac{1}{x}\Big), \quad x > 0$$

gehörende Tailfuntkion. Dann gilt:
F liegt genau dann im Anziehungsbereich der Weibull-Verteilung Ψ_α ($F \in MDA(\Psi_\alpha)$), wenn gilt:

$$x_F < +\infty \quad \textit{und} \quad \lim_{t\to\infty} \frac{\bar{F}^*(tx)}{\bar{F}^*(t)} \;=\; x^{-\alpha} \quad \textit{für alle } x \in \mathbb{R}^+.$$

$\bar{F}^*$ ist also in ∞ regulär variierend mit Index $-\alpha$. Man kann nun

$$a_n = \frac{1}{x_F - \gamma_n} \quad \textit{und} \quad b_n = x_F$$

wählen für die Konvergenz von $a_n(M_n - b_n)$.

3. *F liegt genau dann im Anziehungsbereich der Gumbel-Verteilung* Λ *(*$F \in MDA(\Lambda)$*), wenn eine positive, messbare Funktion g existiert, so dass gilt:*

$$\lim_{t \uparrow x_F} \frac{\bar{F}(t + xg(t))}{\bar{F}(t)} = e^{-x} \quad \textit{für alle } x \in \mathbb{R}.$$

Die Folgen a_n *und* b_n *werden als* $a_n = \frac{1}{g(\gamma_n)}$ *und* $b_n = \gamma_n$ *für die schwache Konvergenz von* $a_n(M_n - b_n)$ *gewählt.*

In [13] findet sich der Beweis zu diesem Satz.
Die nachfolgende Aufzählung gibt einen Überblick, welche Verteilungen im Anziehungsbereich der Extremwertverteilungen liegen (siehe [7, S. 153ff] und [17, S. 265ff]):

- Im Anziehungsbereich der Fréchet-Verteilung liegen u.a. die Cauchy-, Pareto-, Burr- und Loggamma-Verteilung.

Diese Verteilungen und die Fréchet-Verteilung sind sogenannte *heavy-tailed* Verteilungen bei denen nicht alle Momente existieren. Es gilt $E[X^j] = \infty$ für $j \geq \alpha$. Denn:

Sei X eine Zufallsvariable und $E[X^j]$ das j-te Moment einer Fréchet-Verteilung $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$, wie in Definition 2.3. Dann ist $\Phi'_\alpha(x) = \alpha x^{-\alpha-1} e^{-x^{-\alpha}}$ die Dichte der Fréchet-Verteilung. Es folgt für das j-te Moment

$$E[X^j] = \int_0^\infty x^j \ \Phi'(x) dx = \int_0^\infty x^j \alpha x^{-\alpha-1} e^{-x^{-\alpha}} dx$$

Durch Substitution von $y = x^{-\alpha}$ erhält man

$$\int_0^\infty y^{-j/\alpha} e^{-y} dy = \Gamma\Big(1 - \frac{j}{\alpha}\Big) \quad \text{für } j < \alpha,$$

wobei Γ die Gamma-Funktion

$$\Gamma(\lambda) = \int_0^\infty x^{\lambda-1} e^{-x} dx, \quad \lambda > 0$$

ist. Daraus folgt, dass die Momente nur existieren, wenn $\alpha > j$ ist ([19, S. 17]).

- Im Anziehungsbereich der Weibull-Verteilung liegen u.a. die Gleich- und Beta-Verteilung.

Die Verteilungen im Anziehungsbereich der Weibull-Verteilung werden als *short-tailed* Verteilungen bezeichnet. Hier existieren alle Momente. Da auch eine Weibull-verteilte Zufallsvariable als Gamma-Funktion geschrieben werden kann, ergibt sich für das j-te Moment:

$$E[X^j] = (-1)^j \Gamma\Big(1 + \frac{j}{\alpha}\Big).$$

Daraus folgt:

$$E[X^j] < \infty \quad \text{für alle } j \geq 1.$$

Die Herleitung ähnelt dem Fréchet-Fall.

- Und im Anziehungsbereich der Gumbel-Verteilung liegen u.a. die Gamma-, Exponential-, Normal- und Lognormalverteilung.

Solche Verteilungen sind sogenannte *thin-tailed* Verteilungen und für gewöhnlich existieren hier alle Momente. (Der Beweis dazu findet sich in [7, S. 148].)

2.2.2 Verallgemeinerte Extremwertverteilung

Wenn man eine Extremwertverteilung an einen Datensatz anpassen möchte, muss man sich vorher entscheiden, welche der drei Extremwertverteilungen am besten dafür geeignet sein könnte. So eine Entscheidung kann aber schon zu einer Fehleinschätzung führen. Um das zu umgehen, wird stattdessen die *verallgemeinerte Extremwertverteilung* benutzt, die alle drei Verteilungen zu einer zusammenfasst:

Definition 2.6. *(vgl. [13, Def. 2.1])*
Die durch

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

für alle ξ und x mit $1 + \xi x > 0$ definierte Funktion H_ξ heißt verallgemeinerte Extremwertverteilung *(Generalized Extreme Value (GEV) Distribution). Aus der GEV erhält man durch Substitution von x die dazugehörige skalierte Familie:*

$$H_{\xi,\mu,\sigma}(x) := H_\xi\Big(\frac{x-\mu}{\sigma}\Big), \quad \mu \in \mathbb{R},\ \sigma > 0$$

für alle x mit $1 + \xi\frac{x-\mu}{\sigma} > 0$.

Es ist zu bemerken, dass H_ξ stetig ist in ξ, denn für festes x gilt $\lim\limits_{\xi\to 0} H_\xi(x) = H_0(x)$ wegen $(1+\xi x)^{1/\xi} \longrightarrow \exp(x)$, für $\xi \to 0$.

Die Parameter μ und σ aus der Definition werden auch als Lageparameter (Location-Parameter) bzw. Skalierungsparameter (Scale-Parameter) bezeichnet. ξ ist der sogenannte Form-Parameter (Shape-Parameter) und mit $1/\xi$ wird der Tail-Index bezeichnet (siehe [24, S. 144]). (Manchmal wird der Tail-Index auch als ξ definiert, vgl. [15]).

Eine Veränderung des Lageparamters μ bewirkt eine Verschiebung der Dichte um denselben Wert. Denn μ beschreibt, wo die Extrema im Durchschnitt liegen. Wenn der Wert des Skalierungsparameters ansteigt, wird die Dichtefunktion gestaucht. Der Parameter σ ähnelt der Varianz und zeigt an, wie weit die Extrema streuen (vgl. [11, S. 269]).

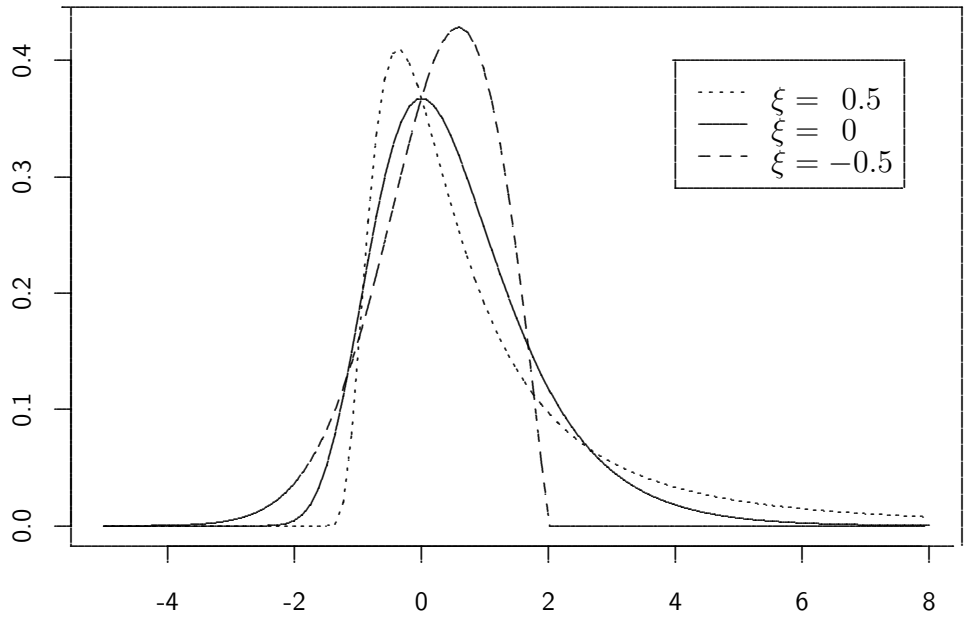

Abbildung 2.2: Die Dichten der GEV-Verteilung mit $\mu = 0$ und $\sigma = 1$.

Der Shape-Parameter ξ bestimmt, zu welchem Typ einer Extremwertverteilung die GEV-Verteilung gehört.

Die verallgemeinerte Extremwertverteilung ist für:

- $\xi > 0$ vom Typ der Fréchet-Verteilung $\Phi_{\xi^{-1}}$ mit

$$H_\xi(x) = \Phi_{\xi^{-1}}(1 + \xi x) \quad \text{für alle } 1 + \xi x > 0,$$

- $\xi = 0$ vom Typ der Gumbel-Verteilung Λ mit

$$H_\xi(x) = \Lambda(x),$$

- $\xi < 0$ vom Typ der Weibull-Verteilung $\Psi_{-\xi^{-1}}$ mit

$$H_\xi(x) = \Psi_{-\xi^{-1}}(-(1 + \xi x)) \quad \text{für alle } 1 + \xi x > 0.$$

In Abbildung 2.2 sind die Dichten der GEV-Verteilung für verschiedene Werte von ξ gezeichnet, wobei $\mu = 0$ und $\sigma = 1$ ist. Für $\xi < 0$ gehört die GEV-Verteilung zum Typ einer Weibull-Verteilung. Sie hat einen endlichen rechten Endpunkt und wird deshalb, wie am Ende von Abschnitt 2.2.1 erwähnt, als short-tailed Verteilung bezeichnet. Wegen des kurzen rechten Tails, ist sie für den weiteren Verlauf dieser

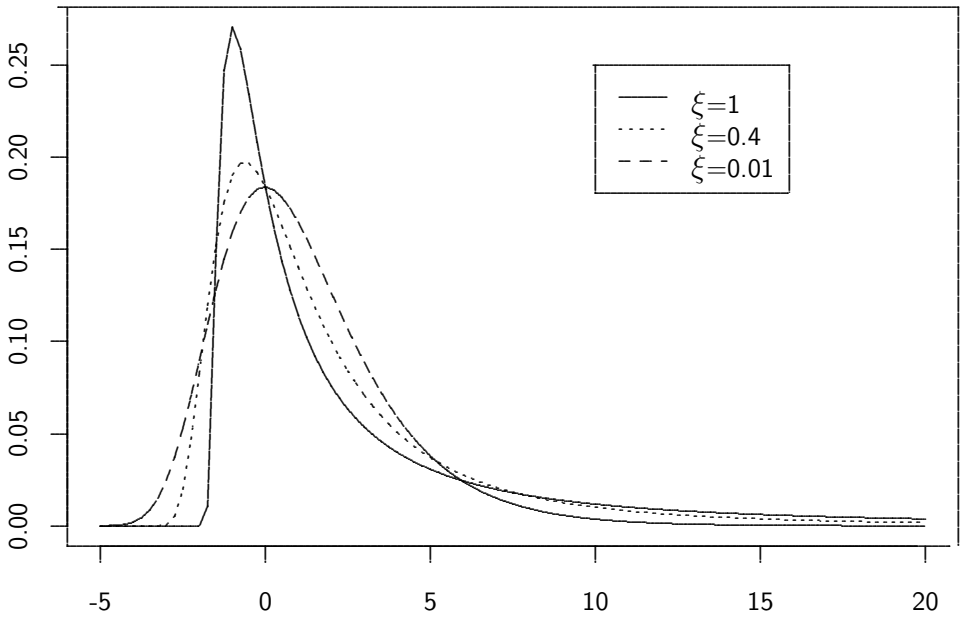

Abbildung 2.3: Die Dichten der GEV-Verteilung mit verschiedenen Shape-Parametern, $\mu = 0$ und $\sigma = 2$.

Arbeit nicht relevant.
Wichtiger sind stattdessen die GEV-Verteilung mit $\xi = 0$, die zu den thin-tailed Verteilungen gehört, und besonders die heavy-tailed GEV-Verteilung mit $\xi > 0$. Denn dann hat die Verteilung sehr breite Flanken, wie es in Abbildung 2.2 zu sehen ist.
Der Shape-Parameter ξ gibt an, ob die Verteilung heavy-tailed ist, und je größer ξ wird, desto mehr nimmt die Wahrscheinlichkeitsmasse im rechten Tail zu. Dies ist auch in Abbildung 2.3 zu erkennen, wo unterschiedliche, aber nur positive Werte für ξ in die GEV-Verteilung eingesetzt worden sind. Für alle drei Verteilungen gilt $\mu = 0$ und $\sigma = 2$.
Aufgrund dieses heavy-tailed Verhalten ist die GEV-Verteilung vom Typ einer Fréchet-Verteilung gut geeignet, um extreme Ereignisse, wie z.B. seltene, aber große Kursverluste an der Börse zu modellieren (s. [15], [17, S. 265ff]).

Das Fisher-Tippett Theorem aus dem vorherigen Abschnitt 2.2.1 kann man auch im Hinblick auf die GEV-Verteilung $H_{\xi,\mu,\sigma}$ umschreiben. Dabei haben μ und σ in etwa die gleiche Bedeutung wie die Normalisierungskonstanten a_n und b_n (vgl. [15]):

Theorem 2.2. *(vgl. [3, S. 48])*
Wenn es Konstantenfolgen $a_n > 0$ und b_n gibt, so dass

$$a_n(M_n - b_n) \longrightarrow H(x),\ n \to \infty$$

gilt für eine nicht-entartete Verteilungsfunktion H. Dann gehört H zur GEV-Familie

$$H_{\xi,\mu,\sigma}(x) = \exp\Big\{ -\Big(1+\xi\Big(\frac{x-\mu}{\sigma}\Big)\Big)^{-1/\xi}\Big\},$$

mit $1+\xi\frac{x-\mu}{\sigma} > 0$, wobei $\mu \in \mathbb{R}$, $\sigma > 0$ und $\xi \in \mathbb{R}$.

$X_1, X_2, \ldots$ sind die unabhängigen und identisch verteilten Zufallsvariablen aus denen man, gemäß der Block Maxima Methode, die Maxima auswählt. Für diesen „neuen“ Datensatz, also $\{M_{n1}, \ldots, M_{nm}\}$, will man eine GEV-Verteilung anpassen. Der nächste Schritt ist nun, die Parameter μ, σ und ξ zu schätzen. Dies erfolgt hier mit der Maximum-Likelihood Methode (ML Methode), die auch angewendet werden darf, da auch die Maxima iid sind. Aber es gibt natürlich verschiedene Methoden, unter anderem die Momentenmethode, um eine Parameterschätzung durchzuführen. Diese werden z.B. in [2, S. 133ff] und [11, S. 274ff] beschrieben, aber hier wird nicht näher darauf eingegangen.

Sei $\vartheta = (\xi, \mu, \sigma) \in \Theta = \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ der Vektor der unbekannten Parameter und schreibe $H_\vartheta = H_{\xi,\mu,\sigma}$. Dann ist h_ϑ die Dichte von H_ϑ mit

$$h_\vartheta(M_n) = \begin{cases} \frac{1}{\sigma}\Big(1+\xi\frac{M_n-\mu}{\sigma}\Big)^{-\frac{1}{\xi}-1}\exp\Big(-\Big(1+\xi\frac{M_n-\mu}{\sigma}\Big)^{-\frac{1}{\xi}}\Big), & \xi \neq 0, \\ \frac{1}{\sigma}\exp\Big(-\frac{M_n-\mu}{\sigma} - e^{-\frac{M_n-\mu}{\sigma}}\Big), & \xi = 0, \\ 0 & \text{sonst.} \end{cases}$$

Für $\xi \neq 0$ und $1+\xi\frac{M_{nj}-\mu}{\sigma} > 0$ ist die Log-Likelihood Funktion der GEV-Verteilung:

$$\begin{aligned} \ell(\vartheta; M_{nj}, j = 1, \ldots, m) &= \sum_{j=1}^{m} \log h_\vartheta(M_{nj}) \qquad (2.4) \\ &= -m\log\sigma - \Big(1+\frac{1}{\xi}\Big)\sum_{j=1}^{m}\log\Big(1+\xi\frac{M_{nj}-\mu}{\sigma}\Big) - \sum_{j=1}^{m}\Big(1+\xi\frac{M_{nj}-\mu}{\sigma}\Big)^{-1/\xi}. \end{aligned}$$

Die Log-Likelihood Funktion für den Fall $\xi = 0$ ist:

$$\begin{aligned} \ell(\vartheta; M_{nj}, j = 1, \ldots, m) &= \sum_{j=1}^{m} \log h_\vartheta(M_{nj}) \qquad (2.5) \\ &= -m\log\sigma - \sum_{j=1}^{m}\Big(\frac{M_{nj}-\mu}{\sigma}\Big) - \sum_{j=1}^{m}\exp\Big(-\Big(\frac{M_{nj}-\mu}{\sigma}\Big)\Big). \end{aligned}$$

Den Maximum-Likelihood Schätzer $\hat{\vartheta} = (\hat{\xi}, \hat{\mu}, \hat{\sigma})$ für die GEV-Verteilung bekommt man nun indem man die Log-Likelihood Funktionen (2.4) und (2.5) maximiert. Da das bei diesen Gleichungen einen hohen Rechenaufwand bedeutet, wird das in Abschnitt 2.3 mit Hilfe des Statistikprogramms R durchgeführt.

Ein Problem, das bei der Maximum-Likelihood Schätzung einer GEV-Verteilung auftreten kann, ist, dass die Regularitätsbedingungen nicht erfüllt sind. Denn die Endpunkte der zugrunde liegenden Verteilungsfunktion hängen von den unbekannten Parametern ab. So ist $\mu - \sigma/\xi$ ein rechter Endpunkt, wenn $\xi < 0$ ist, und ein linker Endpunkt, wenn $\xi > 0$ ist (s. [3, S. 54]).
Aber die Regularitätsbedingungen sind erfüllt, wenn $\xi > -1/2$, denn dann sind die Schätzer konsistent, das bedeutet, dass der Schätzwert für große Stichproben gegen den tatsächlichen Wert konvergiert. Außerdem sind die asymptotische Effizienz und die asymptotische Normalität gewährleistet, es gilt also:

$$\sqrt{m}(\hat{\vartheta} - \vartheta) \to \mathcal{N}(0, I_1(\vartheta)^{-1}) \quad \text{für } \xi > -1/2,$$

wobei $I_1(\vartheta)^{-1}$ das Inverse der Fisher-Informations Matrix ist (vgl. [2, S. 133]). Der Beweis dafür, dass die asymptotischen Eigenschaften erfüllt sind, wird ausführlich in [22] behandelt.
Im Folgenden kann die Maximum-Likelihood Methode bedenkenlos angewendet werden, denn für Finanzrenditen gilt gewöhnlich, dass sie heavy-tailed sind, oder manchmal auch thin-tailed, und dann ist immer $\xi \geq 0$ (vgl. [11, S. 275]).

2.2.3 Return-Level und Return-Periode

Nachdem im letzten Abschnitt die GEV-Verteilung eingeführt wurde, werden nun die Risikomaße, wie das Return-Level und die Return-Periode, definiert.
Im weiteren Verlauf dieses Abschnittes wird davon ausgegangen, dass die Blöcke, die anhand der Block Maxima Methode ausgewählt werden, jährliche Blöcke sind.

Ein k-Jahr *Return-Level* ist eine Schwelle, welche, im Durchschnitt, nur in einem n-Block im Zeitraum von k Jahren überschritten wird. Innerhalb dieses Blockes kann es aber durchaus zu mehreren Überschreitungen kommen. Dies widerspricht nicht der Definition des Return-Levels und in der Regel tritt dieser Fall ein. Denn die Log-Renditen unterliegen oft einer großen Clusterbildung.
Der Block, indem das Return-Level überschritten wird, wird auch als *Stress-Periode* (stress period) bezeichnet und die Renditen nennt man demzufolge *Stress-Verluste* (stress losses) (s. [16], [24, S. 154]).
Da das Return-Level nur ein Punktschätzer ist, welcher nicht angibt, wie groß die Log-Rendite ist, wenn das Return-Level überschritten wird, ist es sinnvoll Konfidenzintervalle zu betrachten. Dabei ist vor allem der rechte Randpunkt eines Konfidenzintervalles interessant, denn er zeigt, wie tief ein Aktiensturz sein könnte und

modelliert so eine Art *Worst-Case*-Szenario. Dies ist wichtig, da es auch bei einem Prozess, der normale tägliche Renditen erzeugt, nicht widersprüchlich ist, wenn es zu gelegentlichen Abstürzen kommt ([14], [15]).

Eine andere Herangehensweise ist die Methode der *Return-Periode* (vgl. [17, S. 273ff]). Hier stellt man sich ein extremes Ereignis vor, also eine große negative Rendite, und untersucht dann, wie selten so ein Szenario eintritt. Das heißt, in welchem Zeitraum wird eine Schwelle u überschritten?
Es lässt sich auch direkt untersuchen, wie wahrscheinlich es ist, dass bereits im nächsten Jahr ein neuer negativer Rekord eintritt, d.h. dass die maximale Rendite alle vorherigen übertrifft.
Auch bei der Return-Periode ist es sinnvoll Konfidenzintervalle zu bestimmen, allerdings ist hier der linke Randpunkt wichtiger, denn er zeigt an, wie kurz der Abstand zwischen zwei Stress-Perioden sein kann ([17], [24]).

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

$$x_\alpha = F^{-1}(\alpha), \tag{2.6}$$

wobei F^{-1} das Inverse der Verteilungsfunktion ist (s. [15]).
Sei $R_{n,k}$ die gesuchte Schwelle, die innerhalb von k Jahren übertreten wird, dann ist

$$\sum_{j=1}^{k} \mathbf{1}_{\{M_{nj} > R_{n,k}\}}$$

die Anzahl von Überschreitungen in k Jahren. Das Return-Level ist so definiert, dass es nur einmal in einem n-Block innerhalb von k Jahren überschritten wird. Die erwartete Anzahl an Überschreitungen ist also eins. Es gilt:

$$\begin{aligned} E\sum_{j=1}^{k} \mathbf{1}_{\{M_{nj} > R_{n,k}\}} = 1 \Leftrightarrow\ & k\ P(M_{nj} > R_{n,k}) = 1 \\ \Leftrightarrow\quad & P(M_{nj} > R_{n,k}) = \frac{1}{k} \end{aligned}$$

Daraus folgt:

$$P(M_{nj} \leq R_{n,k}) = 1 - \frac{1}{k}. \tag{2.7}$$

Damit ergibt sich dann die Definition für das Return-Level:

Definition 2.7. *(vgl. [15] und [17, S. 273])*
Seien $M_{n1}, M_{n2}, \ldots$ die n-Block Maxima mit zugrunde liegender Verteilungsfunktion F^n (siehe (2.1)). Dann ist das k n-Block Return-Level $R_{n,k}$ *das (1 - 1/k)-Quantil der Verteilung.*

Da die Log-Renditen (X_n) unabhängig und identisch verteilt sind mit Verteilungsfunktion F, kann man mit Hilfe von (2.6) und (2.7) leicht folgern, dass

$$\Big(1-\frac{1}{k}\Big)^{1/n} = P^{1/n}(M_n \leq R_{n,k}) = F(R_{n,k}) \tag{2.8}$$

gilt. Also ist $R_{n,k} = x_{(1-1/k)^{1/n}}$, d.h. das Return-Level ist das $(1-1/k)^{1/n}$-Quantil der zugrunde liegenden Verteilungsfunktion F ([15]).

Wenn die n-Block Maxima GEV-verteilt sind, dann ist das Return-Level

$$R_{n,k} = H^{-1}_{\xi,\mu,\sigma}\Big(1-\frac{1}{k}\Big) = \begin{cases} \mu - \frac{\sigma}{\xi}\Big(1-\Big(-\log\Big(1-\frac{1}{k}\Big)\Big)^{-\xi}\Big) & \text{für } \xi \neq 0, \\ \mu - \sigma\log\Big(-\log\Big(1-\frac{1}{k}\Big)\Big) & \text{für } \xi = 0, \end{cases} \tag{2.9}$$

wobei H die GEV-Verteilungsfunktion ist.
Durch Gebrauch der Maximum-Likelihood Parameter $\hat{\xi}$, $\hat{\mu}$ und $\hat{\sigma}$ kann man $R_{n,k}$ schätzen:

$$\hat{R}_{n,k} = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\Big(1-\Big(-\log\Big(1-\frac{1}{k}\Big)\Big)^{-\hat{\xi}}\Big) & \text{für } \hat{\xi} \neq 0, \\ \hat{\mu} - \hat{\sigma}\log\Big(-\log\Big(1-\frac{1}{k}\Big)\Big) & \text{für } \hat{\xi} = 0. \end{cases} \tag{2.10}$$

Im Gegensatz zum Return-Level, stellt man sich bei der Return-Periode die Frage, in welchem Jahr überschreiten die Log-Renditen eine bestimmte Schwelle? Oder andersherum, wie wahrscheinlich ist es, dass schon im nächsten Jahr eine Log-Rendite auftritt, die alle anderen übertrifft?

Sei u als vorgegebene Schwelle eine extrem große Log-Rendite und $M_{n1}, M_{n2}, \ldots$ eine Folge von Maxima mit zugrunde liegender Verteilungsfunktion F^n. Die erste Überschreitungszeit bei u ist

$$\tau_u = \min\{j : M_{nj} > u\} \quad j \geq 1,$$

wobei die Schwelle u kleiner ist als der rechte Endpunkt der Verteilungsfunktion.

Die erste Zeit der Überschreitung der Schwelle u ist eine unabhängige, geometrisch verteilte Zufallsvariable. Die Wahrscheinlichkeit, dass im l-ten Block die Schwelle u überschritten wird, ist

$$\begin{aligned} P(\tau_u = l) &= P(M_{n,1} \leq u, \ldots, M_{n,l-1} \leq u, M_{n,l} > u) \\ &= p(1-p)^{l-1}, \quad l = 1, 2, \ldots \end{aligned} \tag{2.11}$$

wobei $p = 1 - F^n(u)$, und in der letzten Gleichung wurde ausgenutzt, dass die M_{nj} unabhängig und identisch verteilt sind (vgl. [19, S. 11]).
Der Erwartungswert von τ_u ist dann die Return-Periode:

Definition 2.8. *(vgl. [17, S. 273])*
Seien $M_{n1}, M_{n2}, \ldots$ die n-Block Maxima mit zugrunde liegender Verteilungsfunktion F^n. Die Return-Periode *des Ereignisses $\{M_{nj} > u\}$ ist durch*

$$k_{n,u} := E(\tau_u) = \frac{1}{1 - F^n(u)}$$

gegeben, wobei u eine vorgegebene Schwelle ist und $k_{n,u}$ geht gegen ∞, wenn $u \to \infty$.

Die Return-Periode wird so definiert, dass es innerhalb von $k_{n,u}$ n-Blöcken einen Block gibt, wo u übertreten wird. D.h. das $k_{n,u}$ n-Block Return-Level ist u.

Aber der Wert der errechneten Return-Periode kann zu hoch sein. Denn da

$$P(\tau_u \leq l) = p \sum_{i=1}^{l} (1-p)^{i-1} = 1 - (1-p)^l, \quad l \in \mathbb{N}$$

gilt, ist die Wahrscheinlichkeit, dass die Schwelle u schon vor der Return-Periode überschritten wird,

$$P(\tau_u \leq k_{n,u}) = P(\tau_u \leq [1/p]) = 1 - (1-p)^{[1/p]}, \tag{2.12}$$

wobei [ ] die Gaußklammer ist und $p = 1 - F^n(u)$. Je größer die Schwelle u gewählt wird, desto kleiner wird p. Deswegen kann man den Grenzwert von (2.12) bestimmen:

$$\lim_{u \to \infty} P(\tau_u \leq k_{n,u}) = \lim_{p \to 0} (1 - (1-p)^{[1/p]}) = 1 - e^{-1} = 0.6321206.$$

Das heißt für eine hohe Schwelle u wird diese mit einer Wahrscheinlichkeit von über 60% schon vor der errechneten Return-Periode übertreten (vgl. [7, S. 306]).

Wenn die n-Block Maxima GEV-verteilt sind, dann ist

$$k_{n,u} = E(\tau_u) = \frac{1}{1 - H_{\xi,\mu,\sigma}(u)}$$

die Return-Periode mit GEV-Verteilungsfunktion H.
Die Return-Periode kann, genauso wie das Return-Level, auch mit den Maximum-Likelihood Parametern der GEV-Verteilung geschätzt werden:

$$\hat{k}_{n,u} = \frac{1}{1 - H_{\hat{\xi},\hat{\mu},\hat{\sigma}}(u)}.$$

Bleibt noch die Frage zu klären, wie man beim Return-Level k bestimmt, d.h. wie viele Jahre will man betrachten. Soll es ein 10-, 50- oder sogar 100-Jahr Return-Level sein?
Das hängt zum einen davon ab, wie groß der zu untersuchende Datensatz ist. Wie viele Daten hat man zur Verfügung und wie viele Jahre umfasst der Datensatz?

Zum anderen bleibt die Bestimmung von k eine Risikoentscheidung, die das Risikomanagement treffen muss. Dabei ist zu berücksichtigen, wie oft eine Stress-Periode tolerierbar ist, wenn sie eintritt.
Ähnliches Vorgehen ergibt sich bei der Return-Periode, auch hier lässt sich die Wahl von u nach den eben genannten Gesichtspunkten treffen ([15]).

Zu beachten ist aber immer, dass das Return-Level und die Return-Periode mit geschätzten Parametern berechnet werden, d.h. sie sind mit einer gewissen Unsicherheit behaftet. Deswegen ist es sinnvoll, nicht nur Punktschätzungen zu betrachten, sondern auch Konfidenzintervalle. Die Delta-Methode und die Profile Likelihood Methode sind zwei Möglichkeiten, um die 95% Konfidenzintervalle für das Return-Level und die Return-Periode zu berechnen.

Mit der Delta-Methode wird gezeigt, dass der Maximum-Likelihood Schätzer des Return-Levels $\hat{R}_{n,k}$ gegen die Normalverteilung konvergiert. Es gilt (vgl. [2, S. 137]):

$$\sqrt{m}(\hat{R}_{n,k} - R_{n,k}) \xrightarrow{d} \mathcal{N}(0, Var(\hat{R}_{n,k})), \quad m \to \infty$$

wobei

$$Var(\hat{R}_{n,k}) \approx \nabla R_{n,k}^T V \nabla R_{n,k}$$

ist. V ist die Varianz-Kovarianz Matrix von $(\hat{\xi}, \hat{\mu}, \hat{\sigma})$ und

$$\begin{aligned}
\nabla R_{n,k}^T &= \Big[\frac{\partial R_{n,k}}{\partial \xi}, \frac{\partial R_{n,k}}{\partial \mu}, \frac{\partial R_{n,k}}{\partial \sigma}\Big] \\
&= \Big[\frac{\sigma}{\xi^2}\Big(1 - \Big(-\log\Big(1-\frac{1}{k}\Big)\Big)^{-\xi}\Big) \\
&\qquad - \frac{\sigma}{\xi}\Big(-\log\Big(1-\frac{1}{k}\Big)\Big)^{-\xi}\log\Big(-\log\Big(1-\frac{1}{k}\Big)\Big), \\
&\qquad 1, -\frac{1}{\xi}\Big(1 - \Big(-\log\Big(1-\frac{1}{k}\Big)\Big)^{-\xi}\Big)\Big],
\end{aligned}$$

ausgewertet bei $(\hat{\xi}, \hat{\mu}, \hat{\sigma})$.
Näheres zur Delta-Methode findet sich in [3, S. 33ff und S. 56].
Mit dieser Methode wird ein $100(1-\alpha)\%$ Konfidenzintervall durch

$$\hat{R}_{n,k} \pm \Phi^{-1}\Big(1 - \frac{\alpha}{2}\Big)\sqrt{Var(\hat{R}_{n,k})}$$

gegeben, wobei hier mit Φ die Normalverteilung bezeichnet wird.

Allerdings ist die Delta-Methode nicht so gut geeignet, denn die Normalapproximation kann erheblich von der Verteilung des Maximum-Likelihood Schätzers abweichen. Außerdem erhält man nur symmetrische Konfidenzintervalle. Aber ein

Konfidenzintervall, das asymmetrisch über dem Schätzer $\hat{R}_{n,k}$ liegt, wäre geeigneter, da der rechte Randpunkt des Intervalls auch ein Worst-Case-Szenario darstellt. Denn dieser Punkt zeigt an, wie tief die Log-Renditen abstürzen könnten.
Deswegen wird besser die Profile Likelihood Methode verwendet, denn sie erzeugt asymmetrische Konfidenzintervalle (vgl. [15]):

Es wird angenommen, die Blöcke der Größe n seien GEV-verteilt, dann kann man (2.9) umschreiben als:

$$\mu = \begin{cases} R_{n,k} + \frac{\sigma}{\xi}\Big(1 - \Big(-\log\Big(1-\frac{1}{k}\Big)\Big)^{-\xi}\Big) & \xi \neq 0 \\ R_{n,k} - \sigma \log\Big(-\log\Big(1-\frac{1}{k}\Big)\Big) & \xi = 0. \end{cases} \tag{2.13}$$

So lässt sich μ aus den GEV-Likelihood-Funktionen (2.4) und (2.5) eliminieren, indem (2.13) dort für μ eingesetzt wird. Damit erhält man dann auch das Log-Likelihood $\ell(\xi, \sigma, R_{n,k})$ von m beobachteten Block Maxima.
Nun betrachtet man einen Likelihood Ratio Test mit der Nullhypothese $R_{n,k} = r$. Sei

$$\ell(\hat{\xi}_r, \hat{\sigma}_r, r) = \max_{\xi,\sigma} \ell(\xi, \sigma, r)$$

das Maximum Profile Log-Likelihood unter der Nullhypothese und $\ell(\hat{\xi}, \hat{\sigma}, \hat{R}_{n,k})$ sei das Maximum Log-Likelihood mit den Schätzern, die man aus (2.4) bzw. (2.5) und (2.10) erhalten hat.
Dann gilt unter der Nullhypothese, wenn $m \to \infty$

$$-2(\ell(\hat{\xi}_r, \hat{\sigma}_r, r) - \ell(\hat{\xi}, \hat{\sigma}, \hat{R}_{n,k})) \sim \chi^2_1.$$

Die Hypothese wird also mit Hilfe der χ^2-Verteilung getestet. Die χ^2-Verteilung hat hier einen Freiheitsgrad. (Nähere Beschreibung eines Likelihood Ratio Tests findet sich in [3, S. 34].)
Besonders wichtig sind die Konfidenzintervalle. Ein $\alpha\%$ Konfidenzintervall für $R_{n,k}$ ist die Menge von r für welche die Nullhypothese bei einem $\alpha\%$ Niveau nicht abgelehnt wird. Diese Menge wird durch

$$\{r : \ell(\hat{\xi}_r, \hat{\sigma}_r, r) \geq \ell(\hat{\xi}, \hat{\sigma}, \hat{R}_{n,k}) - 0.5\chi^2_1(\alpha)\}$$

gegeben. Mit $(r, \ell(\hat{\xi}_r, \hat{\sigma}_r, r))$ wird die Profile Log-Likelihood Kurve gezeichnet und sie ist für gewöhnlich nicht symmetrisch über ihr Maximum $\hat{R}_{n,k}$. Deswegen ergeben sich so die asymmetrischen Konfidenzintervalle (vgl. [15, S. 15]).

Die Konfidenzintervalle wurden hier jetzt nur für das Return-Level bestimmt, aber die Profile Likelihood Methode ist auch geeignet die Intervalle für die Return-Periode auszurechnen. Die Delta-Methode ist eigentlich nicht geeignet, um die Konfidenzintervalle zu bestimmen. Das liegt nicht nur an der Normalapproximation, die

von der Verteilung des Maximum-Likelihood Schätzers abweichen kann, sondern vor allem an den symmetrischen Konfidenzintervallen. Denn für eine große Schwelle u ergibt sich eine hohe Return-Periode mit einem sehr großen Konfidenzintervall. Dieses Intervall kann so groß sein, dass die untere Grenze des Intervalls schon negativ ist und das ist nicht sinnvoll.

2.3 Datenanalyse der Log-Renditen

Nachdem in Abschnitt 2.2 die theoretischen Grundlagen hergeleitet wurden, sollen nun die Ergebnisse mit einem Beispiel aus der Praxis veranschaulicht werden. Dazu betrachte ich die täglichen Abschlusswerte des Dow Jones Indexes aus den Jahren 1960 bis 2007[1]. Wie schon zu Beginn dieses Kapitels erwähnt, werde ich nicht die absoluten Werte betrachten, sondern die Log-Renditen, denn dadurch werden die täglichen Veränderungen des Indexes besser dargestellt.

Da es besonders wichtig ist, sich gegen Kurseinbrüche abzusichern und nicht gegen Kursgewinne, werden die Log-Renditen negiert. Denn dadurch können die Resultate der Extremwerttheorie direkt angewendet werden, da so die größten negativen Renditen die Maxima sind.
Alle Berechnungen, Abbildungen und Diagramme in diesem Abschnitt werden mit dem Statistikprogramm R durchgeführt.

In der oberen Grafik von Abbildung 2.4 sind die täglichen Abschlusswerte und in der unteren Grafik die negativen Log-Renditen des Dow Jones Indexes dargestellt. Insgesamt sind es 12081 Werte. Hier kann man sehr gut die prozentualen Veränderungen erkennen. Ein besonders extremes Ereignis fand am 19. Oktober 1987 statt. An diesem Tag stürzte der Dow Jones Index um 500 Punkte ab. Dies entsprach einer Log-Rendite von etwa 25%.

Doch so ein Absturz ist natürlich eher selten, trotzdem kann so ein Ereignis erhebliche Auswirkungen auf die Finanzwelt haben. Deswegen wird im Folgenden auch untersucht, wie wahrscheinlich es ist, dass so ein Absturz noch einmal eintritt. Ist so ein Sturz ein Jahrhundertereignis, oder muss man dafür nicht erst hundert Jahre warten?
Es wird zwar in diesem Kapitel angenommen, die Log-Renditen seien unabhängig und identisch verteilt, aber es ist in Abbildung 2.4 zu erkennen, dass dies nicht immer gewährleistet ist. Denn es kommt manchmal zu einer Clusterbildung.
Die größte Rendite ist in diesem Datensatz 25,63%, die kleinste -9,66% (dies ist eigentlich ein Gewinn, da das in Wirklichkeit eine positive Rendite ist).
Der Erwartungswert der Log-Renditen ist ungefähr Null (-0.000246), d.h. über einen längeren Zeitraum überwiegen weder die positiven noch die negativen Renditen. Außerdem gibt es nur eine geringe Streuung der Daten, denn die Varianz liegt bei

[1] Die Daten stammen von der Website http://de.finance.yahoo.com/

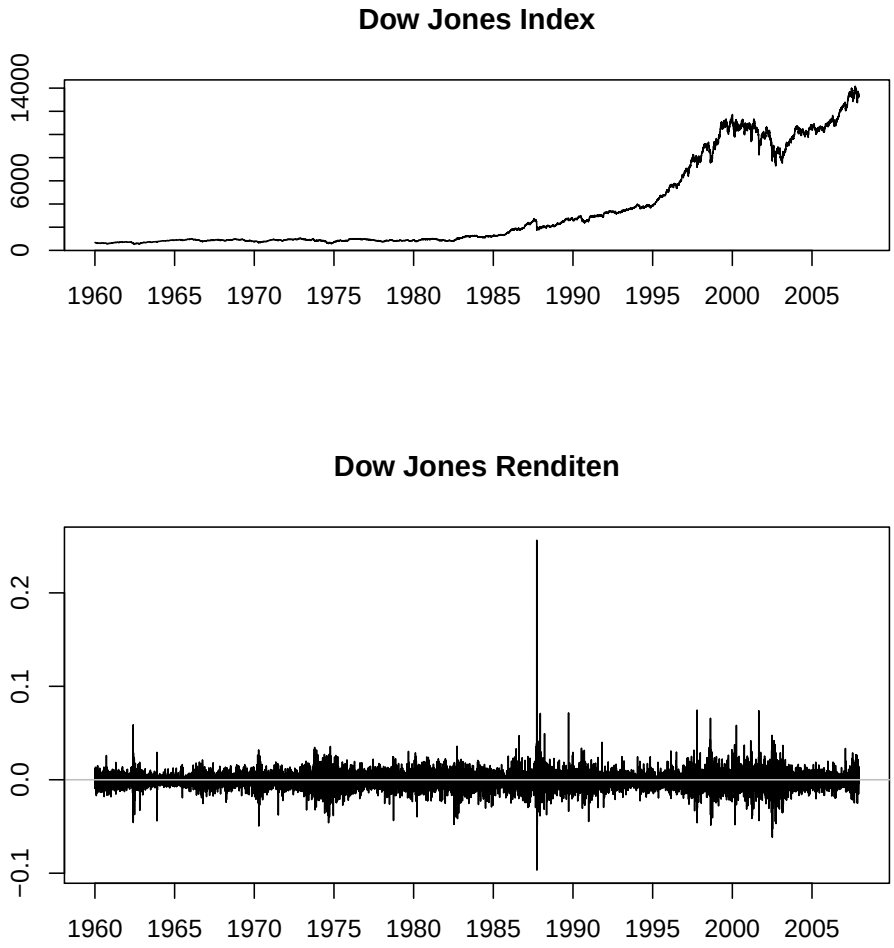

Abbildung 2.4: Die täglichen Dow Jones Werte von 1960-2007. In der oberen Grafik sind die Abschlusswerte und in der unteren Grafik sind die negativen Log-Renditen eingezeichnet.

0.00009067. Allerdings zeigt der positive Wert der Schiefe (1.6622), dass die Daten leicht rechtsschief (oder linkssteil) sind, also ist die rechte Seite der Verteilung flacher als die linke.

In dem QQ-Plot (Quantil-Quantil-Plot) in Abbildung 2.5 wird eine Standard-Normalverteilung an die Dow Jones Daten angepasst und es ist hier deutlich zu erkennen, dass die Renditen wesentlich breitere Tails haben als die Normalverteilung. Dieser QQ-Plot besteht aus den Punkten (vgl. [3, S. 37]):

$$\left\{\left(F^{-1}\left(\frac{i}{s+1}\right), X_{(i)}\right) : i = 1, \ldots, s\right\}.$$

Dabei wird hier eine Standard-Normalverteilung zugrunde gelegt und $X_{(1)} \leq \ldots \leq$

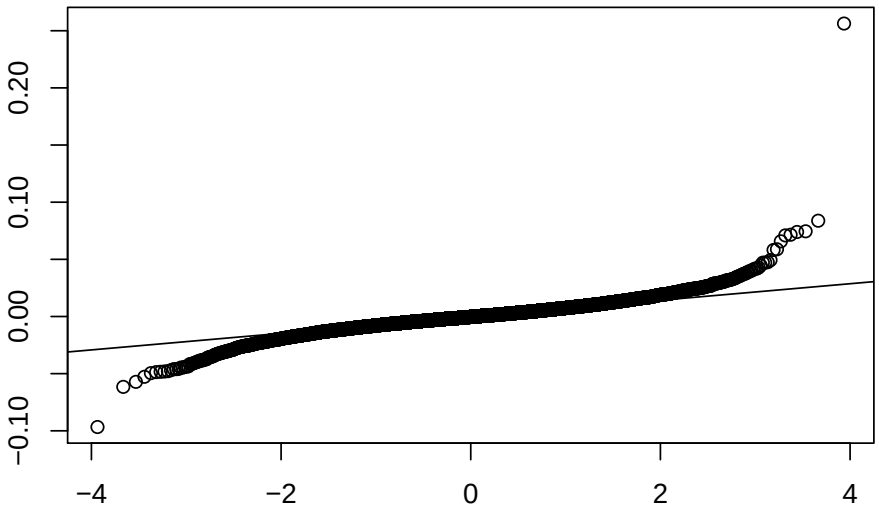

Abbildung 2.5: Normal QQ-Plot für die täglichen Log-Renditen des Dow Jones Index von 1960-2007

$X_{(s)}$ sind die geordneten Log-Renditen mit $s = m * n$.
Die Kurtosis einer Normalverteilung beträgt ungefähr 3, dagegen haben diese Renditen eine Kurtosis von 51.307 (Erklärungen zu Schiefe und Kurtosis finden sich in [11, S. 10ff]). Dass die Kurtosis einen so hohen Wert hat, liegt vor allem an dem Extremereignis vom Oktober 1987. Dies zeigt deutlich die Schwächen der Normalverteilung, um sie an die Daten anzupassen, denn die Normalverteilung hat zu dünne Tails. Somit würden gerade die Extremereignisse nicht gut genug modelliert werden.
Deswegen wird besser die in Abschnitt 2.2.2 vorgestellte GEV-Verteilung benutzt.

2.3.1 GEV-Schätzung

In diesem Abschnitt wird nun eine GEV-Verteilung an die Daten angepasst. Da diese Verteilung nur an die größten Werte angeglichen wird, werden nun die Maxima gemäß der Block Maxima Methode aus dem Datensatz herausgefiltert. D.h. ich unterteile die Daten in gleich große Blöcke und wähle aus jedem Block den maximalen Wert. Seien $X_1, \ldots, X_{n*m}$ die unabhängig und identisch verteilten negativen Log-Renditen des Dow Jones Indexes, mit $m * n = 12081$. Für den j-ten Block gilt

dann

$$M_{nj} = \max\{X_{1j}, \ldots, X_{nj}\} \quad 1 \leq j \leq m.$$

$M_{n1}, \ldots, M_{nm}$ ist somit der Datensatz, der nur aus den Maxima besteht.
Als Blockgröße nehme ich nicht nur die Zeitspanne von einem Jahr, sondern ich werde auch Blöcke betrachten, die nur ein Halbjahr, ein Quartal oder einen Monat umfassen.
Allerdings bringt die Wahl einer geeigneten Blockgröße auch einige Schwierigkeiten mit sich. Denn wenn n klein ist, hat man zwar viele Maxima, so dass die Maximum-Likelihood Schätzer nur eine geringe Unsicherheit haben. Aber die Approximation der Maxima durch die GEV-Verteilung kann darunter leiden, da eine Annahme des Fisher-Tippett Theorems war, dass $n \to \infty$ gilt. Bei größeren Blöcken ist das schon eher gegeben. Doch dann stehen auch weniger Maxima-Werte zur Verfügung, demzufolge kann es zu einer großen Varianz in der Parameter-Schätzung kommen. Also bleibt die Wahl der Blockgröße immer ein Kompromiss (vgl. [17, S. 272]).
Alle Blöcke sollen natürlich gleich groß sein. Allerdings ist dies nicht immer der Fall, da jedes Jahr eine unterschiedliche Anzahl von Handelstagen hat. Doch im weiteren wird angenommen, dass jedes Jahr $n = 252$ Handelstage hat. (Ein Halbjahr hat somit 126, ein Quartal 63 und ein Monat 21 Handelstage.)

Nachdem die Maxima aus den Daten zu einem neuen Datensatz zusammengefasst worden sind, wird nun eine Maximum-Likelihood Schätzung, wie in Abschnitt 2.2.2 beschrieben, durchgeführt, um Schätzwerte für die Parameter ξ, μ und σ zu bekommen.

	$\approx n$	m	ξ	μ	σ
Monat	21	576	**0.1877** (0.1313-0.244)	**0.0117** (0.0112-0.0122)	**0.0055** (0.0053-0.0058)
Quartal	63	192	**0.3106** (0.196-0.4251)	**0.0158** (0.0148-0.0167)	**0.0062** (0.0057-0.0069)
Halbjahr	126	96	**0.4731** (0.2633-0.683)	**0.0185** (0.0171-0.0199)	**0.0066** (0.0057-0.0076)
Jahr	252	48	**0.5024** (0.2099-0.7948)	**0.0227** (0.0199-0.0256)	**0.0091** (0.0067-0.0114)

Tabelle 2.1: Parameter-Schätzer der GEV-Verteilung, angepasst an die Log-Renditen des Dow Jones Indexes (95% Konfidenzintervalle in Klammern)

Die Ergebnisse für die Parameter-Schätzer sind in Tabelle 2.1 aufgelistet.
Das wichtigste Resultat ist, dass der Shape-Parameter ξ für alle Datensätze größer als Null ist und selbst die 95% Konfidenzintervalle liegen im positiven Bereich. Also brauche ich die ML-Schätzung für den Gumbel-Fall nicht berechnen. Die GEV-Verteilungen sind heavy-tailed und es existieren nicht alle Momente. Bei den Jahresmaxima ist der Shape-Parameter so hoch, dass sogar schon die Varianz unendlich ist.
Auffällig ist außerdem, dass ξ mit der Blockgröße zunimmt, also je mehr Daten

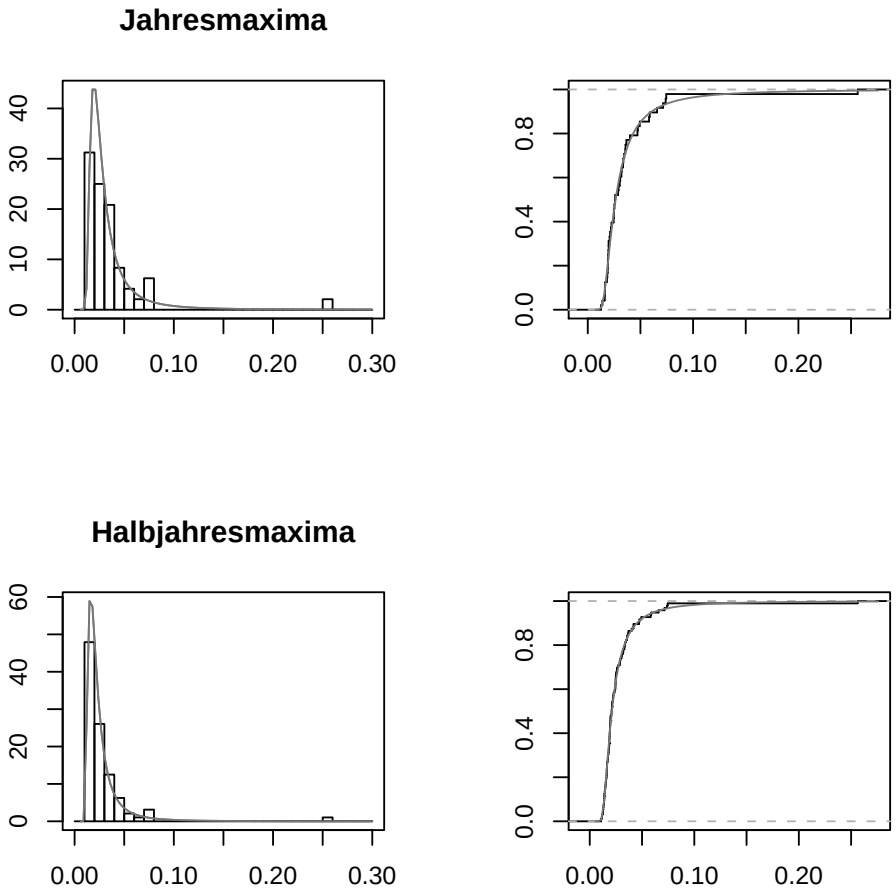

Abbildung 2.6: Histogramm und empirische Verteilungsfunktion der Jahres- und Halbjahresmaxima. Grau eingezeichnet ist die angepasste GEV-Verteilung.

ein einzelner Block umfasst, desto größer ist der Shape-Parameter. Die Ursache ist, dass man bei den Monatsmaxima 576 Blöcke hat, die aber nur aus 21 Werten bestehen. Dies hat dann zur Folge, dass Log-Renditen als Maximum bestimmt werden, obwohl sie eigentlich keine Extremwerte sind. Genau umgekehrt verhält es sich dann bei den Jahresmaxima, hier werden eventuell tatsächliche Extremereignisse nicht beachtet, weil in einer Periode mit großen Renditen nur die Größte ausgewählt wird. Außerdem kommt es bei der Parameter-Schätzung zu einer großen Unsicherheit, wenn man nur so wenige Werte zur Verfügung hat. Dadurch erklären sich auch die großen Konfidenzintervalle der Shape-Parameter bei den Jahres- und Halbjahresmaxima.
Im Gegensatz zum Shape-Parameter haben die unterschiedlichen Blöcke auf die Lage- und Skalierungsparameter kaum einen Einfluss. μ und σ werden zwar auch größer, je größer die Blöcke sind, aber sie erhöhen sich nur leicht.

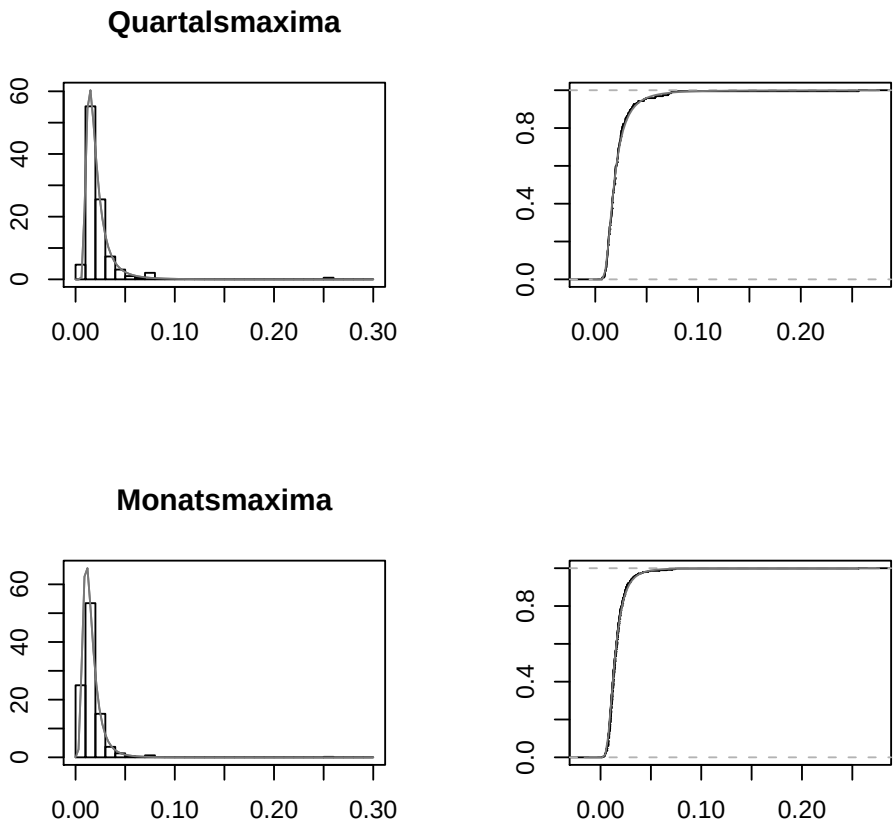

Abbildung 2.7: Histogramm und empirische Verteilungsfunktion der Quartals- und Monatsmaxima. Grau eingezeichnet ist die angepasste GEV-Verteilung.

Es bleibt nun die Frage zu klären, ob die Annahme von GEV-verteilten Maxima überhaupt gerechtfertigt ist und, wenn ja, wie gut die GEV-Verteilung die Daten beschreibt. Dazu habe ich jeweils vier Histogramme und empirische Verteilungsfunktionen aus den vier verschiedenen Datensätzen (Jahresmaxima, Halbjahresmaxima, usw.) gezeichnet und darin die GEV-Dichten bzw. GEV-Verteilungsfunktionen mit den geschätzten Parametern eingezeichnet. Die Überprüfung der Anpassung mittels empirischer Verteilungsfunktion ist durch den Satz von Glivenko/Cantelli gerechtfertigt. Denn dort wird gezeigt, dass die unbekannte Verteilungsfunktion durch die empirische Verteilungsfunktion approximiert werden kann (vgl. [1, S. 197ff]).
Das Ergebnis ist in den Abbildungen 2.6 und 2.7 zu sehen und es zeigt sich, dass die Daten, bis auf ein paar kleine Abweichungen, sehr gut durch die GEV-Verteilung beschrieben werden. Zu den Abweichungen kommt es vor allem bei den Histogrammen, denn hier ist die „Spitze“ der GEV-Verteilung immer größer als der größte Balken des jeweiligen Histogramms.
Es gibt aber, neben Histogramm und empirischer Verteilungsfunktion, noch andere

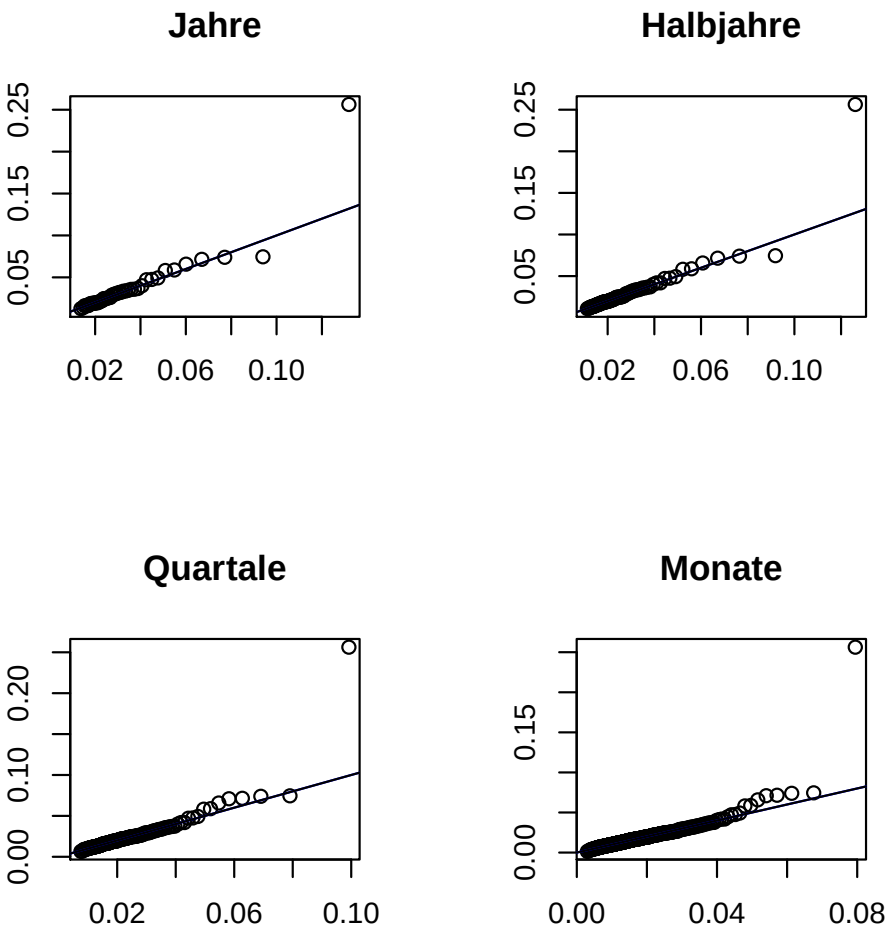

Abbildung 2.8: Quantil-Plot für die GEV-Verteilung

Möglichkeiten die Güte der Anpassung zu überprüfen, z.B. durch einen Quantil-Plot. Statt einer Normalverteilung, wie in Abbildung 2.5, wird jetzt die GEV-Verteilung an den Maxima-Werten angepasst.
Der Quantil-Plot besteht aus den Punkten

$$\left\{\left(\hat{H}^{-1}\left(\frac{j}{m+1}\right), M_{n(j)}\right),\ j = 1, \ldots, m\right\},$$

wobei $\hat{H}^{-1}$ das Inverse der GEV-Verteilung ist, in der die geschätzten Parameter $\hat{\xi}, \hat{\mu}$ und $\hat{\sigma}$ eingesetzt werden, und $M_{n(1)} \leq M_{n(2)} \leq \ldots \leq M_{n(m)}$ bezeichnen die geordneten Block Maxima (vgl. [3, S. 58]).
In Abbildung 2.8 sind die Quantil-Plots für alle vier Blockgrößen abgebildet und auch hier zeigt sich, dass die Anpassung gut gelungen ist, da die geplotteten Punkte fast linear sind. Allerdings ist in allen vier Grafiken ein „Ausreißer“ zu sehen, das

ist das Extremereignis vom Oktober 1987.
Bei den Halbjahresmaxima werden alle Punkte, bis auf zwei, von der Geraden getroffen. Dagegen kommt es bei den Quartals- und Monatsmaxima bei mehreren Punkten zu leichten Abweichungen. Doch sind diese nicht so gravierend, dass man nicht mit den geschätzten Parametern weiterrechnen könnte.

Die Güte einer Anpassung kann man aber nicht nur graphisch darstellen, sondern auch mittels eines Anpassungstests.
Hierfür habe ich den Kolmogorov-Smirnov-Test benutzt. Dieser Test überprüft mit Hilfe der empirischen Verteilungsfunktion, ob die vermutete Verteilung vorliegt. (Näheres zum KS-Test siehe [18, S. 154ff].) Die p-Werte, die sich aus dem Test ergeben, liegen, wie in Tabelle 2.2 aufgelistet, weit über 0.9.
Der p-Wert zeigt an, dass die Hypothese nur bei einem p-Wert, der kleiner ist als das vorgegebenen Signifikanzniveau, abgelehnt werden würde. Da die p-Werte hier so hoch sind, kann die Hypothese der Verteilungsgleicheit zu fast allen Niveaus nicht abgelehnt werden. Somit zeigt auch der KS-Test, dass die Anpassung gut gelungen ist. Den höchsten p-Wert erreichen die Halbjahresmaxima, also ist nach dem KS-Test eine Halbjahres-Blockgröße die beste Aufteilung der Daten.

	Monat	Quartal	Halbjahr	Jahr
p-Wert	0.961	0.973	0.9985	0.9591

Tabelle 2.2: p-Werte des KS-Tests

Die Annahme von GEV-verteilten Maxima ist gerechtfertigt, denn die Anpassung durch die GEV-Verteilung ist für alle vier Blockgrößen gut gelungen. Das wird durch die Abbildungen und dem Anpassungstest deutlich. Auch wenn es zu leichten Unterschieden zwischen den Blockgrößen kommt, können die Parameter der Maximum-Likelihood Schätzung weiter verwendet werden, um im nächsten Abschnitt die Return-Level und Return-Perioden zu berechnen.

2.3.2 Return-Level und Return-Periode

Nachdem die GEV-Verteilung an die Daten angepasst wurde, folgt nun, im nächsten Schritt, die Berechnung des Return-Levels. Man bestimmt also die Schwelle, die nur einmal alle k Jahre überschritten wird. Wichtig dabei ist auch die Berechnung der Konfidenzintervalle, denn sie zeigen auf, wie tief ein Aktiensturz eigentlich sein kann, wenn das Return-Level überschritten wird. Sie modellieren ein Worst-Case-Szenario. Aber es ist nicht nur bedeutsam zu wissen, welche großen Werte die Log-Renditen annehmen können, sondern auch, in welchen Zeitabständen solche Extremereignisse auftreten.

Das Return-Level wird mit (2.9) aus Definition 2.7 bestimmt. Ich berechne das Return-Level für alle vier verschiedenen Blockgrößen. Dabei betrachte ich immer

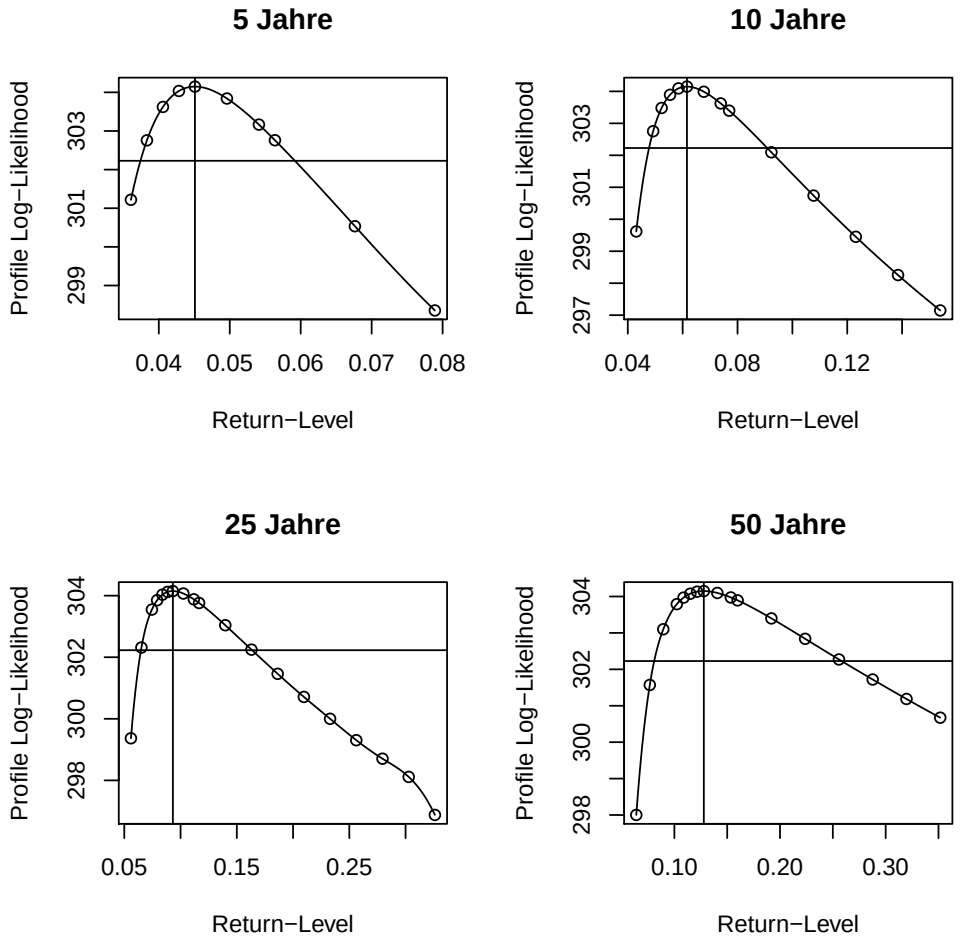

Abbildung 2.9: In allen vier Grafiken wird das 95% Konfidenzintevall für die Halbjahresmaxima berechnet. Die vertikale Linie markiert das Return-Level.

denselben Zeitraum, d.h. es gilt $k \in \{5, 10, 25, 50\}$, wobei $k = 5$ einen Zeitraum von fünf Jahren entspricht. Dies gilt auch für die anderen Blockgrößen, also wenn bei den Jahresmaxima fünf Jahre betrachtet werden, sind es bei den Halbjahresmaxima zehn Halbjahre usw.
Die 95% Konfidenzintervalle werden mit der Profile Log-Likelihood Methode bestimmt.

In Abbildung 2.9 sind die vier Profile Log-Likelihood Kurven eingezeichnet, die mit den Halbjahresmaxima berechnet wurden. Anhand dieser Kurven lassen sich die asymmetrischen 95% Konfidenzintervalle ablesen. Das sind die Schnittpunkte mit der eingezeichneten horizontalen Geraden. Die vertikale Linie markiert das Return-Level. Es ist zu erkennen, dass der Wert des Return-Levels und die entsprechenden Intervalle immer größer werden, je größer k wird. So liegt das 5-Jahr Return-Level der Halbjahresmaxima bei 0.0451, das 50-Jahr Return-Level dagegen ist 0.1279.

Hier umfasst das Konfidenzintervall sogar einen Bereich von 0.0813 bis 0.2577.
Alle Ergebnisse, auch die für die anderen Blockgrößen, sind in der Tabelle 2.3 aufgelistet. In den Klammern stehen jeweils die Konfidenzintervalle.

	5 Jahre	10 Jahre	25 Jahre	50 Jahre
Monat	**0.0457** (0.0413-0.0515)	**0.0546** (0.0484-0.0632)	**0.0682** (0.0587-0.0819)	**0.0802** (0.0675-0.0991)
Quartal	**0.0464** (0.0403-0.0563)	**0.0589** (0.0491-0.0755)	**0.08** (0.0628-0.1114)	**0.1004** (0.0757-0.1492)
Halb-jahr	**0.0451** (0.0376-0.0591)	**0.0616** (0.0481-0.0904)	**0.0932** (0.0654-0.1632)	**0.1279** (0.0813-0.2577)
Jahr	**0.0431** (0.0352-0.0575)	**0.0607** (0.0464-0.0954)	**0.0949** (0.0644-0.1909)	**0.1332** (0.0807-0.3296)

Tabelle 2.3: Return-Level für die verschiedenen Blockgrößen, 95% Konfidenzintervalle in Klammern.

Bei dem 5-Jahr Return-Level gibt es zwischen den einzelnen Blockgrößen kaum Unterschiede. Das Return-Level liegt ungefähr bei 0.045. Auch die Konfidenzintervalle weisen keine großen Unterschiede auf, zudem decken sie auch nur einen kleinen Bereich ab. Anders ist dies beim 25-Jahr und besonders beim 50-Jahr Return-Level. Hier weichen die Ergebnisse der Monats- und Quartalsmaxima von den Halbjahres- und Jahresmaxima ab. Besonders deutlich wird es bei den Konfidenzintervallen, denn sie sind bei den Halbjahren und Jahren wesentlich größer, als bei den anderen Blockgrößen.
Die Gründe hierfür sind die Gleichen, die schon in Abschnitt 2.3.1 im Zusammenhang mit der Maximum-Likelihood Schätzung aufgeführt wurden, denn auch dort gab es große Unterschiede zwischen den Shape-Parametern der einzelnen Blockgrößen. Es ist deshalb zu beachten, dass das Return-Level mit den geschätzten Parametern der GEV-Verteilung gebildet wird und demzufolge spiegeln die Konfidenzintervalle die Fehler der Parameterschätzung wider (s. [17, S. 274]).

Wenn man nun nur die Monatsmaxima betrachten würde, dann wäre eine Log-Rendite von 0.0991 der größte mögliche Absturz innerhalb von 50 Jahren. Dies ist jedoch weit entfernt von dem Absturz, der sich am 19. Oktober 1987 ereignet hat. Aber bei den Halbjahres- und Jahresmaxima liegt die Log-Rendite von 0.2563 innerhalb der entsprechenden Konfidenzintervalle. Bei den Jahresmaxima ist damit noch gar nicht der Worst-Case eingetreten. Wenn man besonders vorsichtig und konservativ kalkuliert, ist sogar mit einem Absturz von 0.3296 zu rechnen. So ein Ereignis tritt dann im Durchschnitt aber nur einmal innerhalb von 50 Jahren auf. Natürlich lassen sich die Bedingungen noch weiter verschärfen, so ist es auch möglich 99% Konfidenzintervalle zu betrachten. Die Intervalle werden dann noch größer und man erhält wirklich eine sehr vorsichtige Kalkulation.

In Abbildung 2.10 sind nur die Halbjahresmaxima der Dow Jones Log-Renditen von 1960 bis 2007 zu sehen. Mit der schwarzen horizontalen Geraden ist bei 0.0616

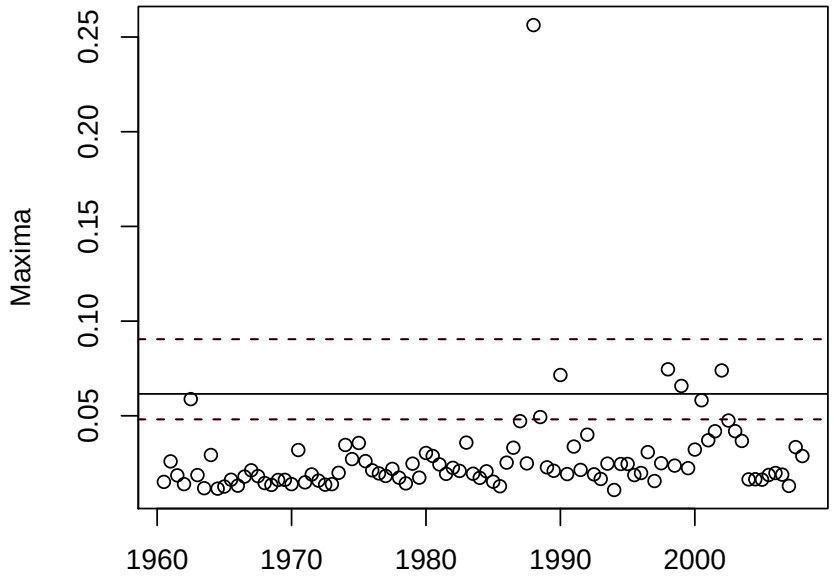

Abbildung 2.10: Halbjahresmaxima der Dow Jones Log-Renditen von 1960-2007. Die schwarze Linie ist das 20-Halbjahr Return-Level mit den entsprechenden 95% Konfidenzintervallen (schwarz-gestrichelte Linien)

das 20-Halbjahr Return-Level eingezeichnet. Die schwarz-gestrichelten Linien markieren den Bereich des 95% Konfidenzintervalls von 0.0481 bis 0.0904.
Insgesamt wird fünf Mal das Return-Level übertreten, es gibt also fünf Stress-Verluste. Es kann natürlich sein, dass innerhalb eines Halbjahres mehrere Log-Renditen größer waren als das Return-Level, diese Werte sind in Abbildung 2.10 dann aber nicht eingezeichnet, weil ja nur das Maximum gesucht war.
In den ersten zwanzig Jahren treten überhaupt keine Stress-Perioden auf, dafür kommt es es in den achtziger und auch in den neunziger Jahren jeweils zu zwei Übertretungen des Return-Levels. Nach der Definition des Return-Levels soll es aber nur zu einer Stress-Periode kommen innerhalb von 10 Jahren. Deswegen ist hier auch die Bedeutung der Konfidenzintervalle festzustellen, da alle Überschreitungen des Return-Levels, bis auf einen Wert, innerhalb des Konfidenzintervalls liegen. Der einzige „Ausreißer“, der nicht im Konfidenzintervall liegt, ist die Log-Rendite, die sich am 19. Oktober 1987 ereignet hat. Um ein solches Extremereignis auch durch die Konfidenzintervalle abdecken zu können, reicht es nicht, nur ein 10-Jahr Return-Level zu betrachten. Aber wenn man ein 50-Jahr oder 100-Halbjahr Return-Level betrachtet, wird, wie schon in Tabelle 2.3 zu sehen, auch ein Sturz von 0.2563 durch das 95% Konfidenzintervall eingeschlossen.

Wird die Rendite von 1987 aber außer Acht gelassen, dann gibt das 20-Halbjahr Return-Level eine gute Schätzung für auftretende Extremereignisse.
Wenn man nun alle Log-Renditen des Dow Indexes von 1960-2007 betrachtet, dann ist das 20-Halbjahr Return-Level das 0.9996 Quantil. Dieses ergibt sich mit Hilfe von (2.8) aus Abschnitt 2.2.3 für $k = 20$ und $n = 126$:

$$P^{1/n}(M_n \leq R_{n,k}) = \left(1 - \frac{1}{k}\right)^{1/n} = 0.999593$$

In Abschnitt 2.2.3 wurde zur Berechnung der Konfidenzintervalle nicht nur die Profile Likelihood Methode vorgestellt, sondern auch die Delta-Methode, mit der man allerdings nur symmetrische Konfidenzintervalle erhält. Diese sind aber nicht so gut geeignet, denn bei dem 50-Jahr Return-Level der Jahresmaxima, das einen Wert von 0.1332 hat, ergibt sich als Konfidenzintervall (0.0418-0.2246). D.h. der Worst-Case ist viel kleiner als bei der Profile Likelihood Methode, dies hat zur Folge, dass die größte Log-Rendite von 0.2563 gar nicht mehr vom Konfidenzintervall abgedeckt wird. Zu klein ist auch der linke Randpunkt des Konfidenzintervalls, denn eine Log-Rendite von 0.041 tritt häufiger auf als nur einmal innerhalb von 50 Jahren. Deswegen ist es sinnvoller die Konfidenzintervalle nur mit der Profile Likelihood Methode zu bestimmen.

Nachdem ich bis jetzt nur das Return-Level berechnet habe, welches eine Schwelle angibt, die im Durchschnitt nur einmal alle k Jahre überschritten wird, wird nun die Frage andersherum gestellt. D.h. wie oft, bzw. wie selten, tritt ein Extremereignis, also eine große negative Rendite, auf? Dazu wählt man eine Rendite, die als extrem angesehen wird, und berechnet damit die Return-Periode wie in Abschnitt 2.2.3.

Als Schwelle u, die eine große Log-Rendite markieren soll, wähle ich vier verschiedene mögliche Extrem-Renditen aus. Als erstes liegt es natürlich nahe die Log-Rendite vom 19. Oktober 1987 zu nehmen, um die Frage zu klären, in wie vielen Jahren ist es zu erwarten, dass es noch einmal einen so großen Absturz gibt. Eine Schwelle liegt also bei 0.2563. Als nächstes habe ich die zweithöchste Rendite des Datensatzes genommen. Diese Log-Rendite liegt bei 0.0838. Da sich diese Rendite auch im Oktober 1987 ereignet hat, wird sie allerdings nicht berücksichtigt, wenn die Maxima gemäß der Block Maxima Methode ausgesucht werden. Die dritthöchste Log-Rendite hat einen Wert von 0.0745. Da es einen so großen Abstand zwischen der höchsten und der zweithöchsten Rendite gibt, nehme ich als vierten Wert eine Log-Rendite von 0.12. Als Blockgrößen ziehe ich sowohl die Halbjahres- als auch die Jahresmaxima heran. Die Ergebnisse sind in Tabelle 2.4 aufgelistet, wobei die Angaben immer in Jahre gezählt werden. Zudem werden die 95% Konfidenzintervalle wieder mit Hilfe der Profile Log-Likelihood Methode berechnet.

Zuerst fällt auf, dass es bei den Return-Perioden zwischen den verschiedenen Blockgrößen bei den zwei kleineren Log-Renditen keine wesentlichen Unterschiede gibt. Aber die Return-Perioden für die beiden größten Renditen unterscheiden sich und

	0.0745	0.0838	0.12	0.2563
Halbjahr	**15.3** (7.4-38.3)	**19.8** (8.9-55.3)	**43.5** (16.0-172.8)	**225.0** (52.5-2067.4)
Jahr	**15.2** (7.2-40.1)	**19.4** (8.4-58.0)	**40.4** (13.8-188.4)	**189.0** (45.0-2746.3)

Tabelle 2.4: In der ersten Zeile stehen die vorgegebenen Log-Renditen. Dazu werden die entsprechenden Return-Perioden berechnet. Die Angabe entspricht Jahren. In Klammern stehen die 95% Konfidenzintervalle.

auch bei den in Klammern stehenden Konfidenzintevallen kommt es zu Abweichungen.
Insgesamt sind die Return-Perioden der Halbjahresmaxima immer leicht höher als die der Jahresmaxima. Diese Entwicklung wäre auch bei den Quartals- und Monatsmaixma zu sehen. Dort wären die Return-Perioden und auch die Konfidenzintervalle wesentlich größer als die der Jahresmaxima. Deswegen habe ich sie hier nicht aufgelistet, da sie eigentlich keine weiteren Informationen liefern.
Es ist aber auch hier zu beachten, dass die Return-Perioden der Jahres- und Halbjahresmaxima, genau wie bei dem Return-Level, mit geschätzten Parametern berechnet werden und damit anfällig für Fehler sind (vgl. [17, S. 274]).
Bei einer Log-Rendite von 0.0745 ergibt sich ein Wert von etwa 15 Jahren, d.h. innerhalb von 15 Jahren, bzw. 30 Halbjahren, wird einmal in einem n-Block die Schwelle von 0.0745 überschritten. Eine Log-Rendite von 0.0745 ist somit ein 15-Jahr Ereignis. Allerdings verringert es sich sogar zu einem 7-Jahr Ereignis, wenn man das Konfidenzintervall betrachtet. Dieses umfasst einen Bereich von 7 bis etwa 40 Jahre. Bei der Return-Periode ist also, im Gegensatz zum Return-Level, die untere Grenze des Intervalls wichtiger. Denn dadurch wird deutlich, dass ein Extremereignis häufiger auftreten kann, als es sich durch die Return-Periode vermuten lässt.
Je größer die betrachtete Log-Rendite, desto größer ist natürlich auch die Return-Periode. So ist eine Rendite vom 19. Oktober 1987 ein 189- bzw. 225-Jahr Ereignis. Dieses ist eine sehr große Zeitspanne, wo man geneigt ist, sie nicht weiter zu beachten, da wohl niemand seine Finanzplanung über einen Zeitraum von 200 Jahren berechnet. Aber werden hier die entsprechenden Konfidenzintervalle betrachtet, dann reduziert sich das 200-Jahr Ereignis zu einem 45- bzw. 52-Jahr Ereignis. Allerdings ist der Konfidenzbereich sehr groß, denn die obere Grenze liegt bei ungefähr 2000 Jahren. Wird also der Eintritt einer extremen Rendite geschätzt, so ist das Ergebnis mit einer sehr großen Unsicherheit behaftet.

Die Return-Periode hat zum Ziel einen Zeitraum abzuschätzen, indem eine bestimmte Schwelle überschritten wird. Dabei kann es, wie eben gesehen, zu großen Schwankungen kommen. Deswegen kann man aber auch ganz konkret fragen, wie wahrscheinlich es ist, dass es schon im nächsten Jahr einen neuen negativen Rekord gibt. D.h. es tritt eine Rendite auf, die alle vorherigen noch übertrifft. Dazu wird (2.11) aus Abschnitt 2.2.3 benutzt.

Die größte Rendite ist 0.2563 und wenn man jetzt die Jahresmaxima betrachtet, dann ist die Wahrscheinlichkeit, dass es im nächsten Jahr einen neuen Rekord gibt,

$$P(M_{n,m+1} > \max\{M_{n,1}, \ldots, M_{n,m}\}) = 1 - H_{\hat{\xi},\hat{\mu},\hat{\sigma}}(0.2563) = 0.005291805,$$

wobei $n = 252$ und $m = 48$ ist. Die Wahrscheinlichkeit ist mit 0.529% sehr niedrig, dass es im nächsten Jahr (mit dem gegebenen Datensatz des Dow Jones Indexes wäre es das Jahr 2008) eine Rendite gibt, die sogar größer ist als 25.63%.
Werden stattdessen die Halbjahresmaxima benutzt, dann besteht mit 0.222% die Möglichkeit, dass es im ersten Halbjahr von 2008 eine Log-Rendite von mehr als 25,63% gibt. Für das zweite Halbjahr liegt die Wahrscheinlichkeit wieder bei 0.222%.

Aber eine Log-Rendite von 25.63% ist ein sehr extremes Ereignis. Deshalb stellt sich die Frage, wie wahrscheinlich ist es, dass im nächsten Jahr eine Log-Rendite von 8.38% eintritt?
Diese Wahrscheinlichkeit liegt bei 0.05164165, also ungefähr 5.15%.

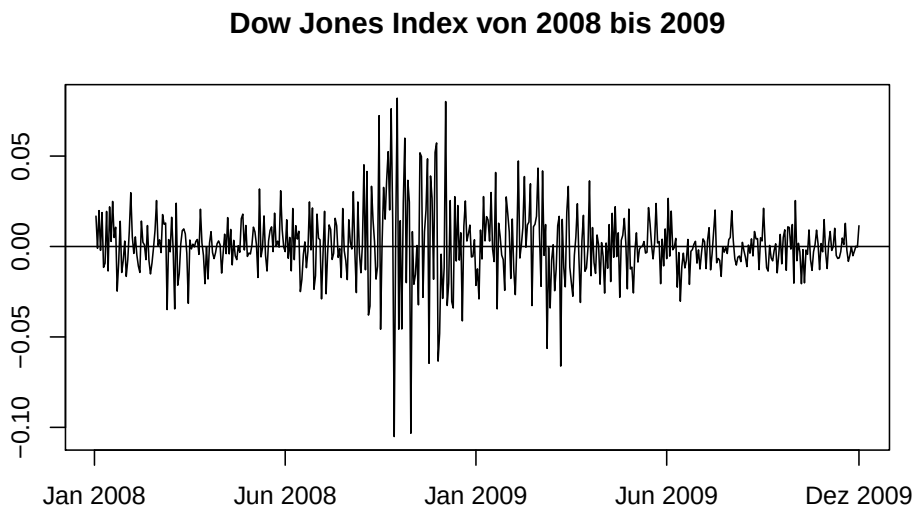

Abbildung 2.11: Die täglichen negativen Log-Renditen des Dow Jones Indexes von Januar 2008 bis Dezember 2009.

Betrachtet man jetzt die täglichen Dow Jones Daten des Jahres 2008, dann ist zu erkennen, dass dieser Fall fast eingetreten wäre. Denn die Finanzmarktkrise hat auch beim Dow Jones Index Spuren hinterlassen. So liegt die höchste Log-Rendite des Jahres 2008 bei 0.082 und damit gar nicht einmal so weit entfernt von der zweithöchsten jemals eingetretenen Log-Rendite im Zeitraum von 1960 bis 2007.

In Abbildung 2.11 sind die täglichen Log-Renditen des Dow Jones Indexes zu sehen.

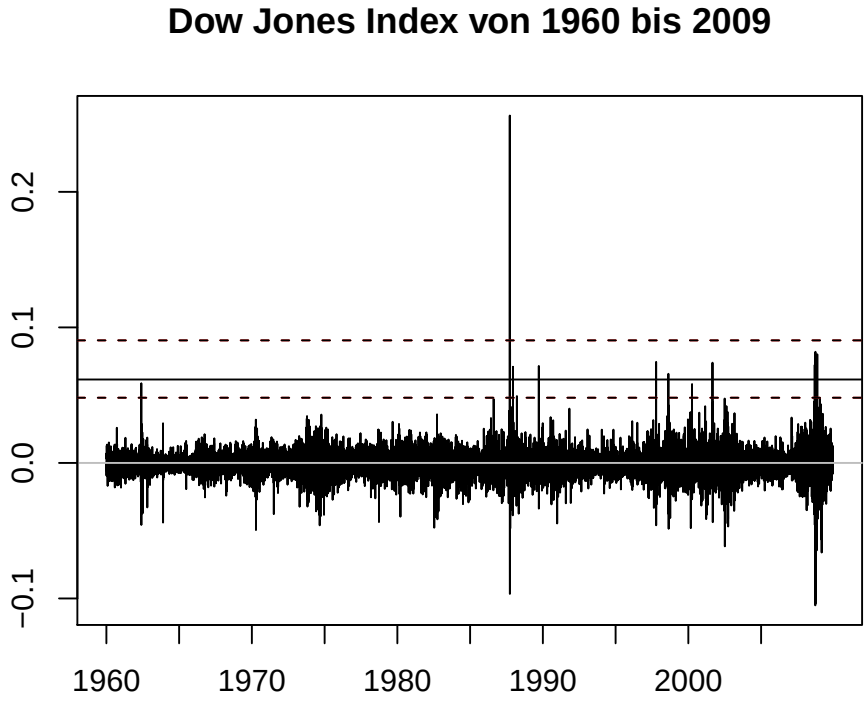

Abbildung 2.12: Die täglichen negativen Log-Renditen des Dow Jones Indexes. Eingezeichnet ist auch das 20-Halbjahr Return-Level (schwarze Linie) mit den entsprechenden 95% Konfidenzintervallen (schwarz-gestrichelte Linien).

Der abgebildete Zeitraum umfasst die Jahre 2008 bis 2009, in denen die Auswirkungen der Finanzkrise deutlich zu erkennen sind. Denn die Log-Renditen schwanken in dieser Zeit sehr stark.

Alle Dow Jones Log-Renditen von 1960 bis 2009 sind in Abbildung 2.12 zu sehen. Zusätzlich dazu ist das 20-Halbjahr Return-Level von 0.0616 mit den entsprechenden 95% Konfidenzintervallen (0.0481-0.0904) eingezeichnet (siehe Tabelle 2.3). Auch im Jahr 2008 tritt eine Stress-Periode auf, denn innerhalb eines Halbjahres wird das Return-Level mehrmals überschritten. Aber die Log-Renditen liegen alle im Bereich des Konfidenzintervalls.
In dieser Abbildung ist auch zu erkennen, dass seit den achtziger Jahren ungefähr alle zehn Jahre Stress-Perioden eintreten, denn das Return-Level wird immer öfter überschritten. Auch die Abstände zwischen den einzelnen Stress-Perioden am Aktienmarkt werden immer kleiner. Es stellt sich die Frage, ob es in Zukunft immer öfter und in kleineren Zeiträumen zu solchen „Krisensituationen“ in der Finanzwelt kommen kann, so dass die Aktienkurse deutlich größeren Schwankungen ausgesetzt sind?

In Tabelle 2.4 sind die Return-Perioden aufgelistet und es wurde berechnet, dass eine Schwelle von 0.0745 einmal innerhalb von 15 Jahren übertreten wird. Betrachtet man jetzt allerdings die Dow Jones Daten bis 2009, dann sind nach dem Eintritt von der letzten großen Rendite im Jahr 2001 bis hin zur Kapitalmarktkrise 2008/2009 nur 7 bis 8 Jahre vergangen, bis die Schwelle von 0.0745 wieder übertreten wurde. Der errechnete Wert der Return-Periode ist also zu hoch, weil es auch schon 1998 zu einer Überschreitung kam.
Doch dies ist nicht sehr verwunderlich, denn benutzt man (2.12) aus Abschnitt 2.2.3, liegt die Wahrscheinlichkeit bei 0.6317, dass die Schwelle vor der Return-Periode übertreten wird. Die Wahrscheinlichkeit ist somit sehr hoch. Trotzdem ist die Schätzung der Return-Periode nicht wertlos, denn ein Zeitabstand von sieben Jahren wird von dem Konfidenzintervall mit eingeschlossen.

Zusammenfassend kann gesagt werden, dass das Return-Level und die Return-Periode zwei Methoden sind, um das Risiko eines großen Kurseinbruchs einzugrenzen. Natürlich kann die Zukunft nicht vorhergesagt werden, aber es ist möglich, die extremen Werte, die die Log-Renditen annehmen können, abzuschätzen. Dabei ist es bei der Berechnung des Return-Levels und der Return-Periode aber nicht nur wichtig verschiedene Zeiträume zu betrachten, sondern auch verschiedene Blockgrößen. Denn die Ergebnisse können sehr voneinander abweichen.

2.4 Unterschied zur POT Methode

In der Einleitung zu diesem Kapitel wurde schon erwähnt, dass es neben der traditionellen Block Maxima Methode noch eine andere Möglichkeit gibt, die Maxima aus einem Datensatz auszuwählen. Das ist die sogenannte „Peak over Threshold“ Methode (POT Methode), die noch jünger ist als die Block Maxima Methode. Hierbei wird ein Schwellenwert (Threshold) bestimmt und alle Werte, die über dieser Schwelle liegen, sind die zu untersuchenden Maxima. An diese Maxima wird dann aber keine verallgemeinerte Extremwertverteilung angepasst, sondern eine verallgemeinerte Pareto-Verteilung (s. [16]).
Eine Schwierigkeit bei dieser Methode ist die Wahl eines geeigneten Thresholds, denn er darf nicht zu hoch sein, denn dann sind möglicherweise nur noch wenige Daten vorhanden. Doch wenn er zu niedrig gewählt wird, werden Renditen zu Maxima, die keine Extremwerte sind. Eine Voraussetzung der POT Methode ist, dass die Renditen alle iid sind. In diesem Kapitel wurde das auch vorausgesetzt. Aber im nächsten Kapitel wird gezeigt, dass auch für nicht unabhängige Zufallsvariablen die Block Maxima Methode anwendbar ist. Die Annahme der Unabhängigkeit ist bei der POT Methode natürlich schwer zu halten, denn bei einer starken Clusterbildung kann man eigentlich nicht von Unabhängigkeit sprechen. Zur Clusterbildung kommt es vor allem in Zeiten von Kapitalmarktkrisen und dann wird der Threshold

oft überschritten. Also werden alle Überschreitungen als Maxima gewertet. Bei der Block Methode hingegen wird nur der höchste Wert des Monats oder Jahres ausgewählt, deswegen kann hier schon von unabhängigen Werten ausgegangen werden (vgl. [11, S. 266]). Wichtig ist es aber hier, einen langen Betrachtungszeitraum zu haben, damit auch genügend Werte zur Verfügung stehen ([16, S. 3]).
Mit der POT Methode können als Risikomaße der Value-at-Risk (VaR) und der Expected Shortfall (ES) berechnet werden. (Näheres zum VaR und ES siehe [16].) Sowohl der VaR, als auch der ES basieren auf einer Verlustverteilung und nicht auf einer Verteilung des Maximumverlustes wie bei der Block Methode (vgl. [3, S. 157]). Doch das in dieser Arbeit verwendete Return-Level ist ein vorsichtigeres Maß, um das Risiko zu schätzen als der VaR und der ES. Außerdem kann mit Hilfe der Konfidenzintervalle auch ein Worst-Case-Szenario bestimmt werden. Dagegen ist bei der POT Methode der ES zwar ein konservativeres Risikomaß als der VaR, aber dieses Risikomaß gibt nur einen Mittelwert der Werte an, die über dem VaR liegen. So wird der größte mögliche Absturz nicht bestimmt ([15, S. 2ff]).

3 Block Maxima Methode mit stationären Zufallsvariablen

3.1 Einführung

Im vorherigen Kapitel wurde immer davon ausgegangen, dass die Zufallsvariablen unabhängig und identisch verteilt sind. Allerdings war im Anwendungsteil in Kapitel 2.3 zu erkennen, dass diese Annahme nicht immer gerechtfertigt ist. Denn die Log-Renditen des Dow Jones Indexes bilden oft Cluster und so kann man nicht mehr von unabhängigen Zufallsvariablen sprechen.

Fast alle Renditen von Finanzdaten haben besondere statistische Eigenschaften gemeinsam. Diese Eigenschaften werden als *stilisierte Fakten* (engl. stylized facts) bezeichnet. Einige davon spielten schon im letzten Kapitel eine Rolle: So hat eine Rendite-Verteilung für gewöhnlich breitere Flanken als die Normalverteilung (heavy-tail Verhalten). Außerdem ist die Verteilung der Renditen leptokurtisch, d.h. sie ist spitzer als die Normalverteilung, denn die Kurtosis ist größer als drei (s. Abschnitt 2.3).
Zu den stilisierten Fakten gehört aber auch, dass besonders große oder besonders kleine Renditen meistens in Clustern auftreten. Betrachtet man einen längeren Zeitraum, ist zu erkennen, dass sich Phasen hoher und Phasen niedriger Volatilität immer abwechseln. Das ist die sogenannte *bedingte Heteroskedastiziät*, denn die bedingte Varianz ändert sich im Zeitablauf (vgl. [21, S. 12ff]).

Es werden nun nicht mehr unabhängige Zufallsvariablen betrachtet, sondern es wird berücksichtigt, dass die Renditen Cluster bilden. Die Zeitreihenanalyse ist hierfür ein wichtiges Werkzeug. Dazu werden im Folgenden sogenannte ARCH-Prozesse eingeführt, zu deren Eigenschaften auch die bedingte Heteroskedastiziät zählt. Deswegen können ARCH-Prozesse gut Finanzzeitreihen modellieren. Es wird gezeigt, dass die Log-Renditen durch ein Zeitreihenmodell simuliert werden können.

Auch in diesem Kapitel spielen die extremen Renditen eine wichtige Rolle, denn es steht immer noch im Vordergrund sich gegen starke Kurseinbrüche und Finanzmarktkrisen abzusichern. Wenn man nun keine unabhängigen Zufallsvariablen betrachtet, ist es in der Extremwerttheorie hilfreich den *extremalen Index* (engl. extremal index) zu bestimmen. Denn dieser Index ist eine reziproke mittlere Größe der Cluster und ist nützlich, um die Abhängigkeit in den Daten zu modellieren.

Zu Beginn dieses Kapitels werden erst die wichtigsten Begriffe und Definitionen der Zeitreihenanalyse eingeführt. Danach liegt der Fokus wieder auf der Extremwerttheorie, die dann mit Hilfe der Zeitreihen weiter ausgeführt wird. Zum Ende werden die Ergebnisse wieder auf die Dow Jones Daten angewendet. Abschließend wird ein GARCH-Prozess an die Log-Renditen angepasst.
In diesem ganzen Kapitel werden die zu betrachtenden Maxima wieder gemäß der Block Maxima Methode ausgewählt.

3.2 Grundlagen der Zeitreihenanalyse

Dieser Abschnitt soll einen kurzen Überblick über die Grundlagen der Zeitreihenanalyse geben, dazu dient insbesondere [17, S. 125ff] als Literaturquelle. Zuerst werden elementare Begriffe, wie Stationarität und Autokorrelationsfunktion, definiert. Zum Schluss wird als stochastischer Prozess der ARCH-Prozess vorgestellt, der für den weiteren Verlauf dieses Kapitels bedeutsam ist.

Grundbegriffe

Sei $(X_t)_{t\in\mathbb{Z}}$ eine Folge von Zufallsvariablen über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Dann ist $(X_t)_{t\in\mathbb{Z}}$ ein *stochastischer Prozess*.

Definition 3.1.
Sei $(X_t)_{t\in\mathbb{Z}}$ ein stochastischer Prozess und die Momente der Zeitreihe existieren. Dann gilt:

- $\mu(t)$ *ist die* Erwartungswertfunktion *von* $(X_t)_{t\in\mathbb{Z}}$ *mit*

$$\mu(t) = E(X_t), \quad t \in \mathbb{Z}.$$

- $\gamma(t,s)$ *ist die* Autokovarianzfunktion *von* $(X_t)_{t\in\mathbb{Z}}$ *mit*

$$\gamma(t,s) = Cov(X_t, X_s) = E((X_t - \mu(t))(X_s - \mu(s))), \quad t,s \in \mathbb{Z}.$$

  Es gilt zudem $\gamma(t,t) = Var(X_t)$.

- $\rho(t,s)$ *ist die* Autokorrelationsfunktion *(ACF) von* $(X_t)_{t\in\mathbb{Z}}$ *mit*

$$\rho(t,s) = \frac{\gamma(t,s)}{\sqrt{Var(X_t)}\sqrt{Var(X_s)}}, \quad t,s \in \mathbb{Z}.$$

Wenn die Annahme der Unabhängigkeit gelockert wird, geht man davon aus, dass die zu betrachtenden Prozesse stationär sind. Dabei gibt es zwei Arten von Stationarität:

Definition 3.2.
Die Zeitreihe $(X_t)_{t\in\mathbb{Z}}$ *ist* strikt stationär*, wenn gilt*

$$(X_{t_1}, \ldots, X_{t_n}) \stackrel{d}{=} (X_{t_1+k}, \ldots, X_{t_n+k}),$$

für alle $t_1, \ldots, t_n,\ k \in \mathbb{Z}$ *und für alle* $n \in \mathbb{N}$.

Strikte Stationarität bedeutet, dass die gemeinsamen Verteilungen von $X_{t_1}, \ldots, X_{t_n}$ und $X_{t_1+k}, \ldots, X_{t_n+k}$ gleich sind.
Eine etwas schwächere Annahme ist, wenn die Zufallsvariablen schwach stationär sind:

Definition 3.3.
Die Zeitreihe $(X_t)_{t\in\mathbb{Z}}$ *ist* schwach stationär *(oder auch: kovarianz-stationär), wenn die ersten beiden Momente existieren und den folgenden Bedingungen genügen:*

$$\begin{aligned} \mu(t) &= \mu, && t \in \mathbb{Z}, \\ \gamma(t,s) &= \gamma(t+k, s+k), && t,\ s,\ k \in \mathbb{Z}. \end{aligned}$$

Eine strikt stationäre Zeitreihe ist auch schwach stationär, wenn die Varianz endlich ist. Dies ist aber nicht immer der Fall, so gibt es viele Prozesse, die eine unendliche Varianz haben.

Aus der schwachen Stationarität folgt:

$$\gamma(t-s, 0) = \gamma(t, s) = \gamma(s, t) = \gamma(s-t, 0), \quad \text{für alle } t,\ s \in \mathbb{Z}.$$

Das bedeutet, dass die Kovarianz zwischen X_t und X_s nur von dem Zeitabstand $|s-t|$ abhängt. Das wird auch als Lag (engl. Zeitverschiebung) bezeichnet. Damit kann die Autokovarianzfunktion eines schwach-stationären Prozesses mit nur einer Variablen h geschrieben werden als:

$$\gamma(h) = \gamma(h, 0), \quad \text{für alle } h \in \mathbb{Z}.$$

Damit ergibt sich, dass $\gamma(0) = Var(X_t) = \sigma^2$ für alle $t \in \mathbb{Z}$ gilt. Also ist die Varianz konstant unter der Annahme schwacher Stationarität.
Auch die Autokorrelationsfunktion eines schwach stationären Prozesses kann umgeschrieben werden:

$$\rho(h) = \rho(X_h, X_0) = \frac{\gamma(h)}{\gamma(0)}, \quad \text{für alle } h \in \mathbb{Z}.$$

Seien $(X_t)_{1\le t\le T}$ gegebene Zufallsvariablen, dann ist die empirische Autokovarianz-

bzw. Autokorrelationsfunktion definiert als (vgl. [21, S. 118]):

$$\hat{\gamma}(s) := \frac{1}{T}\sum_{t=1}^{T-s}(X_t - \bar{X})(X_{t+s} - \bar{X}), \quad 0 \leq s < T-1,$$

$$\hat{\rho}(s) := \frac{\sum_{t=1}^{T-s}(X_t - \bar{X})(X_{t+s} - \bar{X})}{\sum_{t=1}^{T}(X_t - \bar{X})^2} = \frac{\hat{\gamma}(s)}{\hat{\gamma}(0)}, \quad 0 \leq s < T-1,$$

mit $\bar{X} = (\sum_{t=1}^{T} X_t)/T$.

Die grundlegenden Bausteine einer Zeitreihe sind die sogenannten White Noise Prozesse (oder: Weißes Rauschen). Diese sind unkorreliert und modellieren den als zufällig angesehenen Teil der Zeitreihe.

Definition 3.4.
$(X_t)_{t\in\mathbb{Z}}$ *ist ein* White Noise Prozess*, wenn er schwach stationär ist mit Autokorrelationsfunktion*

$$\rho(h) = \begin{cases} 1, & h = 0, \\ 0, & h \neq 0, \end{cases} \quad \textit{für alle } h \in \mathbb{Z}.$$

Dieser Prozess hat Erwartungswert Null und Varianz $\sigma^2 = Var(X_t)$*. Er wird mit* $WN(0, \sigma^2)$ *bezeichnet.*

Definition 3.5.
$(X_t)_{t\in\mathbb{Z}}$ *ist ein* strikter White Noise Prozess*, wenn die Zufallsvariablen unabhängig und identisch verteilt sind mit Erwartungswert Null und Varianz* σ^2*. Dieser Prozess wird mit* $SWN(0, \sigma^2)$ *bezeichnet.*

3.2.1 ARCH-Prozesse

Finanzzeitreihen unterliegen der Bildung von Clustern, denn es wechseln sich immer wieder Phasen von hoher und Phasen von niedriger Volatilität ab. Das wird auch als bedingte Heteroskedastizität bezeichnet und somit unterscheiden sich die bedingte und die unbedingte Varianz:

$$Var(X_t|\mathcal{F}_{t-1}) \neq Var(X_t).$$

Deswegen wird in diesem Abschnitt der ARCH-Prozess eingeführt, denn eine Eigenschaft des ARCH-Prozesses ist, dass sich die bedingte Varianz im Zeitablauf ändert.
Im Folgenden wird ein White Noise Prozess mit $(\epsilon_t)_{t\in\mathbb{Z}}$ und ein strikter White Noise Prozess mit $(Z_t)_{t\in\mathbb{Z}}$ bezeichnet.

ARCH(p)

Definition 3.6.
Sei $(Z_t)_{t\in\mathbb{Z}}$ ein SWN(0,1). Der Prozess $(X_t)_{t\in\mathbb{Z}}$ ist ein ARCH(p)-Prozess *(Autoregressive Conditional Heteroskedasticity) der Ordnung p, wenn er strikt stationär ist und er den folgenden Gleichungen genügt*

$$\begin{aligned} X_t &= \sigma_t Z_t && \text{für alle } t \in \mathbb{Z}, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 && \text{für alle } t \in \mathbb{Z}, \end{aligned} \tag{3.1}$$

wobei $\alpha_0 > 0$, $\alpha_i \geq 0$ für $i = 1, \ldots, p$ und der Startwert X_0 ist unabhängig von (Z_t).

Durch die Konstruktion des ARCH-Prozesses sind Z_t von $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$ unabhängige Zufallsvariablen für jedes $t \in \mathbb{N}$.
Sei $\mathcal{F}_t = \sigma(\{X_s : s \leq t\})$ die Sigma-Algebra, die alle Informationen bis zur Zeit t repräsentiert. Dann ist σ_t wie in (3.1) messbar bezüglich $\mathcal{F}_{t-1}$ und der bedingte Erwartungswert des ARCH-Prozesses ist Null, wenn $E(|X_t|) < \infty$:

$$E(X_t|\mathcal{F}_{t-1}) = E(\sigma_t Z_t|\mathcal{F}_{t-1}) = \sigma_t E(Z_t|\mathcal{F}_{t-1}) = \sigma_t E(Z_t) = 0.$$

Im Folgenden sei vorausgesetzt, dass (X_t) schwach stationär ist. Dann ist $E(X_t^2) < \infty$ und es gilt:

$$Var(X_t|\mathcal{F}_{t-1}) = E(\sigma_t^2 Z_t^2|\mathcal{F}_{t-1}) = \sigma_t^2 Var(Z_t) = \sigma_t^2.$$

Die bedingte Varianz $Var(X_t|\mathcal{F}_{t-1}) = \sigma_t^2$ des ARCH-Prozesses, oder auch Volatilität genannt, ist eine Zufallsvariable, die von der Zufallsvariable X_{t-i} abhängt. Wenn ein oder mehrere Werte von $|X_{t-1}|, \ldots, |X_{t-p}|$ besonders groß sind, dann hat X_t eine große Varianz und tendiert dazu, selbst einen großen Wert anzunehmen. Dadurch entstehen die Volatilitäts-Cluster, die häufig bei Finanzrenditen zu sehen sind (s. [17, S. 139ff]).

Der unbedingte Erwartungswert ist bei einem ARCH-Prozess gleich Null:

$$E(X_t) = E(E(X_t|\mathcal{F}_{t-1})) = E(E(\sigma_t Z_t|\mathcal{F}_{t-1})) = E(\sigma_t(E(Z_t)) = 0.$$

Die unbedingte Varianz ist bei ARCH-Prozessen ungleich der bedingten Varianz, denn für die unbedingte Varianz gilt, wenn $\sum_{i=1}^{p} \alpha_i < 1$,

$$\begin{aligned} Var(X_t) &= E(X_t^2) \\ &= \frac{\alpha_0}{1 - \sum_{i=1}^{p} \alpha_i} < \infty. \end{aligned}$$

Für den ARCH(1)-Prozess wird diese Aussage später in Satz 3.2 in Abschnitt 3.3.2 bewiesen.
Für die Kovarianz des ARCH-Prozesses gilt:

$$E(X_t X_{t-h}) = E(E(X_t X_{t-h}|\mathcal{F}_{t-h})) = E(X_{t-h} E(X_t|\mathcal{F}_{t-h})) = 0.$$

Die bedingte Varianz ist für den ARCH-Prozess nicht konstant. Im Gegensatz dazu existiert die unbedingte Varianz nur, wenn $\sum_{i=1}^{p} \alpha_1 < 1$ ist. In Abschnitt 3.3.2 wird diese Eigenschaft noch einmal aufgegriffen, allerdings steht dann im Vordergrund, unter welchen Bedingungen der ARCH-Prozess strikt stationär ist.

In der Regel sind die Parameter des ARCH-Prozesses unbekannt und müssen deswegen geschätzt werden.
Die am häufigsten benutzte Methode ist die Maximum-Likelihood Schätzung, die in [17, S. 150ff] und [23, S. 107], vorgestellt wird. Aber auch die Methode der kleinsten Quadrate kann angewendet werden.
Außerdem ist die Ordnung des ARCH-Prozesses oft nicht bekannt. Eine Möglichkeit, um eine passende Ordnung zu finden, wird in [23, S. 106] beschrieben.

Für die Innovationen $(Z_t)_{t\in\mathbb{Z}}$ wurde hier nur vorausgesetzt, dass sie Erwartungswert Null und Varianz Eins haben. Wenn die Parameter mit einer Maximum-Likelihood Schätzung bestimmt werden, wird für die Innovationen oft eine Normalverteilung oder eine t-Verteilung zugrunde gelegt.

Nach dieser kurzen Einführung über die ARCH-Prozesse erklärt sich auch der Name ARCH. Denn der Prozess ist autoregressiv, da X_t immer von den vorherigen X_{t-i} abhängt. Außerdem ist er bedingt heteroskedastisch, weil sich die bedingte Varianz stetig ändert.

3.3 Extremwerttheorie mit stationären Zufallsvariablen

Nachdem im letzten Abschnitt der ARCH-Prozess eingeführt wurde, steht jetzt wieder die Extremwerttheorie im Vordergrund. Dabei werden einige Ergebnisse aus dem zweiten Kapitel aufgegriffen, dieses Mal aber für den Fall stationärer Zufallsvariablen. Um auch für diese Zufallsvariablen zu einem ähnlichen Resultat wie dem Fisher-Tippett Theorem zu gelangen, wird der extremale Index definiert. Anschließend werden auch die ARCH-Prozesse wieder gebraucht.

3.3.1 Extremaler Index

Es wird nun angenommen, $(X_n)_{n\in\mathbb{N}}$ sei eine strikt stationäre Folge von Zufallsvariablen. Diese ist auch schwach stationär, wenn die ersten beiden Momente endlich sind. Im Weiteren wird strikte Stationarität einfach nur als Stationarität bezeichnet. Dieser Abschnitt bezieht sich vor allem auf Embrechts [7, S. 209ff], Leadbetter [12, S. 51ff] und McNeil [15].

In Kapitel 2 wurden für die iid Zufallsvariablen als Grenzverteilung die verallgemeinerte Extremwertverteilung eingeführt. Die Grundlage hierfür ist das Fisher-Tippett Theorem: Wenn es Folgen $(a_n) > 0$ und $(b_n) \in \mathbb{R}$ gibt, dann haben die Maxima als Grenzverhalten

$$\lim_{n\to\infty} P(a_n(M_n - b_n) \leq x) = \lim_{n\to\infty} F^n(a_n^{-1}x + b_n) = H_\xi(x),$$

wobei H eine verallgemeinerte Extremwertverteilung ist. Wenn man auf beiden Seiten den Logarithmus nimmt, dann ergibt sich

$$\lim_{n\to\infty} n(1 - F(a_n^{-1}x + b_n)) = -\log H_\xi(x). \tag{3.2}$$

In der Gleichung wurde benutzt, dass $-\log(1-x) \sim x$ gilt für $x \to 0$. (Das ergibt sich, wenn der Logarithmus um 1 entwickelt wird.)
Diese Aussage lässt sich verallgemeinern:
Sei (u_n) eine Folge reeller Zahlen. Für $0 \leq \tau \leq \infty$ ist äquivalent:

(i)
$$\lim_{n\to\infty} n(1 - F(u_n)) = \tau,$$

(ii)
$$\lim_{n\to\infty} P(M_n \leq u_n) = \exp(-\tau). \tag{3.3}$$

Beweis:
1. Fall: $\tau < \infty$
"$(i) \Rightarrow (ii)$"

$$P(M_n \leq u_n) = F^n(u_n) = \left[1 - \frac{1}{n}(n(1 - F(u_n)))\right]^n \underset{n\to\infty}{\longrightarrow} \exp(-\tau).$$

"$(ii) \Rightarrow (i)$" Zuerst ist zu zeigen:

$$P(M_n \leq u_n) \to \exp(-\tau) \Rightarrow 1 - F(u_n) \to 0 \tag{3.4}$$

Dies wird mit Widerspruch gezeigt, denn angenommen die Folgerung gilt nicht, dann existiert eine Teilfolge (u_{n_k}) von (u_n) und $\beta \in (0,1)$ mit $1 - F(u_n) \geq \beta$ für alle k. Daraus ergibt sich $F^{n_k}(u_{n_k}) \leq (1-\beta)^{n_k} \to 0$ und damit $P(M_{n_k} \leq u_{n_k}) \to 0$. Dies ist ein Widerspruch zu $\tau < \infty$. Nun ist die eigentliche Behauptung zu zeigen: Es gilt $P(M_n \leq u_n) = F^n(u_n) \to \exp(-\tau)$ und wenn auf beiden Seiten der Logarithmus angewendet wird, ergibt sich, da $\tau < \infty$,

$$n \log F(u_n) = n \log(1 - (1 - F(u_n))) \to -\tau.$$

Da für $x \to 0$, $-\log(1-x) \sim x$ gilt, folgt mit (3.4) die Behauptung.
2. Fall: $\tau = \infty$
" $\Rightarrow$ " Auch dies wird mit Widerspruch gezeigt. Sei also angenommen, (i) gilt, aber

nicht (ii). Dann gibt es eine Teilfolge (n_k), so dass $P(M_{n_k} \leq u_{n_k}) \to \exp(-\tilde{\tau})$, wenn $k \to \infty$ und $\tilde{\tau} < \infty$. Aber dann folgt aus (ii), ähnlich wie im Fall $\tau < \infty$, (i) mit

$$\lim_{k\to\infty} n_k(1 - F(u_{n_k})) = \tilde{\tau} < \infty.$$

Dies ist aber ein Widerspruch zu (i) mit $\tau = \infty$.
" $\Leftarrow$ " Das folgt ähnlich wie die Hinrichtung. (Vgl. [7, S. 116].) □

Das Ziel dieses Abschnittes ist es, einen ähnlichen „Grenzwertsatz" für eine stationäre Folge (X_n) zu finden, wie es ihn für iid Zufallsvariablen gibt.

Sei (X_n) eine stationäre Folge und von einer Beobachtung X_i ist F die Marginal-Verteilung. Dann wird mit

$$M_n = \max(X_1, \ldots, X_n)$$

das Maximum der ersten n Zufallsvariablen bezeichnet. $(\tilde{X}_n)$ ist eine assoziierte Folge mit iid Zufallsvariablen und gleicher Marginal-Verteilung F und es gilt

$$\tilde{M}_n = \max(\tilde{X}_1, \ldots, \tilde{X}_n).$$

Unter folgenden zwei Bedingungen haben die Maxima von (X_n) dasselbe Grenzverhalten wie die Maxima von $(\tilde{X}_n)$:

- Die stationäre Folge darf nur eine schwache lang andauernde Abhängigkeit haben (engl: weak long-range dependence).
- Die Folge darf keine Tendenz aufweisen, dass große Werte Cluster bilden.

Die erste Bedingung ist eine sogenannte Mischungseigenschaft und bedeutet formal:

Definition 3.7. *(vgl. [3, S. 93])*
Eine stationäre Folge $X_1, X_2, \ldots$ erfüllt die Mischungsbedingung $\mathcal{D}(u_n)$ für eine Folge u_n, falls für alle $i_1 < \ldots < i_p < j_1 < \ldots < j_q$ mit $j_1 - i_p > l$ gilt:

$$\begin{aligned}&|P(X_{i_1} \leq u_n, \ldots, X_{i_p} \leq u_n, X_{j_1} \leq u_n, \ldots, X_{j_q} \leq u_n)\\ &-P(X_{i_1} \leq u_n, \ldots, X_{i_p} \leq u_n)P(X_{j_1} \leq u_n, \ldots, X_{j_q} \leq u_n)| \leq g(n, l),\end{aligned}$$

wobei $g(n, l_n) \to 0$ für eine Folge l_n, so dass $l_n/n \to 0$, wenn $n \to \infty$.

Wenn die Zufallsvariablen unabhängig sind, dann ist der obige Ausdruck für jede Folge u_n Null. Sind die Variablen nicht unabhängig, aber die Bedingung gilt und sie sind weit voneinander entfernt, so dass die Differenz der Wahrscheinlichkeiten gegen Null geht, dann hat das keinen Einfluss auf das Grenzverhalten der Maxima, wie es im nachfolgenden Theorem gezeigt wird.
Mit der Definition von $\mathcal{D}(u_n)$ gilt folgende Eigenschaft (s. [12, S. 56]):

$$P(M_n \leq u_n) - P^k(M_{[n/k]} \leq u_n) \to 0, \quad n \to \infty, \tag{3.5}$$

Theorem 3.1. *(vgl. [3, S. 93])*
Sei $(X_n)_n$ eine stationäre Folge von Zufallsvariablen. Wenn es Folgen $(a_n) > 0$ und $(b_n) \in \mathbb{R}$ gibt, so dass

$$P(a_n(M_n - b_n) \leq x) \to G(x) \tag{3.6}$$

gilt, wobei G eine nicht-entartete Verteilungsfunktion ist und die $\mathcal{D}(u_n)$ - Bedingung ist erfüllt mit $(u_n) = (a_n^{-1}x + b_n)$ für alle $x \in \mathbb{R}$. Dann ist G eine Extremwertverteilung.

Beweisskizze: (vgl. [12, S. 57])
Mit (3.5) ergibt sich, wenn dort nk anstelle von n gesetzt wird,

$$P(M_{nk} \leq a_{nk}^{-1}x + b_{nk}) - P^k(M_n \leq a_{nk}^{-1}x + b_{nk}) \to 0.$$

Daraus kann mit (3.6) gefolgert werden, dass

$$P(a_{nk}(M_n - b_{nk}) \leq x) \to G^{1/k}(x)$$

gilt und das bedeutet, dass G max-stabil ist. (Dies ergibt sich aus der Äquivalenz: G ist max-stabil genau dann, wenn es eine Folge F_n gibt und Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$, so dass gilt $F_n(a_{n_k}^{-1}x + b_{n_k}) \to G^{1/k}(x)$, vgl. [12, S. 8].) G ist aber auch max-stabil genau dann, wenn G eine Extremwertverteilung ist. Diese Folgerung ergibt sich aus dem Beweis des Fisher-Tippett Theorems aus Kapitel 2. □

Nach dem Theorem konvergieren die skalierten Maxima einer stationären Folge von Zufallsvariablen, bei erfüllter $\mathcal{D}(u_n)$-Bedingung, gegen eine Extremwertverteilung. Das Theorem besagt aber nicht, dass aus $P(a_n(M_n - b_n) \leq x) \to G(x)$ und $P(a_n(\tilde{M}_n - b_n) \leq x) \to H(x)$ folgt, dass $G = H$ gilt.
Damit man auch für stationäre Folgen von Zufallsvariablen ein ähnliches Resultat wie (3.3) erhält, muss noch eine zweite Bedingung erfüllt sein. Die sogenannte „Anti-Cluster"-Bedingung.

Definition 3.8. *(vgl. [7, S. 213])*
Sei (X_n) eine stationäre Folge von Zufallsvariablen und (u_n) eine Folge in $\mathbb{R}$. Die Bedingung $\mathcal{D}'(u_n)$ ist für (X_n) erfüllt, falls gilt

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} P(X_1 > u_n, X_j > u_n) = 0.$$

Satz 3.1.
Sei (X_n) eine stationäre Folge von Zufallsvariablen und (u_n) eine Folge. (X_n) erfüllt die Bedingungen $\mathcal{D}(u_n)$ und $\mathcal{D}'(u_n)$ und es gibt ein $0 \leq \tau < \infty$. Dann ist äquivalent:

(i)

$$\lim_{n\to\infty} n(1 - F(u_n)) = \tau$$

(ii)

$$\lim_{n\to\infty} P(M_n \leq u_n) = \exp(-\tau).$$

Der Beweis findet sich in [7, S. 213] und [13, S. 78]. Mit diesem Resultat ergibt sich jetzt als Grenzverteilung für stationäre Folgen:

Theorem 3.2. *(vgl. [7, S. 215])*
Sei (X_n) eine stationäre Folge mit Verteilungsfunktion F, wobei $F \in MDA(H_\xi)$ und H_ξ eine verallgemeinerte Extremwertverteilung ist. D.h. es existieren Folgen $(a_n) > 0$ und $(b_n) \in \mathbb{R}$, so dass gilt

$$\lim_{n\to\infty} n(1 - F(a_n^{-1}x + b_n)) = -\log H_\xi(x), \quad x \in \mathbb{R}. \tag{3.7}$$

Die Folge $(u_n) = (a_n^{-1}x + b_n)$ erfüllt die Bedingungen $\mathcal{D}(u_n)$ und $\mathcal{D}'(u_n)$ für alle $x \in \mathbb{R}$. Dann ist (3.7) äquivalent zu jeder der beiden folgenden Aussagen:

$$a_n(M_n - b_n) \to H_\xi, \tag{3.8}$$

$$a_n(\tilde{M}_n - b_n) \to H_\xi. \tag{3.9}$$

Beweis:
Die Äquivalenz von (3.7) und (3.8) folgt direkt aus Satz 3.1.
Die Äquivalenz von (3.7) und (3.9) folgt aus (3.3) und (3.2). □

Wenn die Bedingungen $\mathcal{D}$ und $\mathcal{D}'$ erfüllt sind, dann haben die Maxima einer stationären Folge dasselbe Grenzverhalten wie die Maxima einer assoziierten unabhängigen Folge.
Doch werden nun die Log-Renditen von Finanzzeitreihen betrachtet, dann ist es vertretbar, dass die Renditen nur eine schwache lange Abhängigkeit haben, aber die zweite Bedingung ist wenig realistisch. Denn es ist bekannt, dass Log-Renditen häufig Cluster bilden und somit ist die Bedingung $\mathcal{D}'$ nicht erfüllt und das obige Theorem ist nicht mehr gültig. Um aber auch dann noch eine Grenzverteilung für die Maxima zu erhalten, bietet der *extremale Index* einen Ausweg.

Definition 3.9.
Sei (X_n) eine stationäre Folge und $0 \leq \theta \leq 1$ eine reelle Zahl. Falls für jedes $\tau > 0$ eine Folge $(u_n(\tau))$ existiert, so dass

$$\begin{aligned} \lim_{n\to\infty} n(1 - F(u_n(\tau))) &= \tau \qquad \textit{und} \\ \lim_{n\to\infty} P(M_n \leq u_n(\tau)) &= \exp(-\theta\tau) \end{aligned}$$

gilt. Dann wird θ als extremaler Index *der Folge (X_n) bezeichnet.*

Die Definition von θ ist unabhängig von der Wahl der Folge $(u_n(\tau))$ (siehe [12, S. 66]). Wenn (X_n) den extremalen Index $\theta > 0$ hat, dann sind für (u_n) und $\tau > 0$ folgende Aussagen äquivalent (vgl. [15, S. 7]):

(i)
$$\lim_{n\to\infty} n(1-F(u_n)) \;=\; \tau$$

(ii)
$$\lim_{n\to\infty} P(\tilde{M}_n \le u_n) = \; \exp(-\tau)$$

(iii)
$$\lim_{n\to\infty} P(M_n \le u_n) = \; \exp(-\theta\tau). \tag{3.10}$$

Diese äquivalenten Aussagen können mit denen aus (3.3) verglichen werden. So erhält man für große n

$$P(M_n \le u_n) \approx P^{\theta}(\tilde{M}_n \le u_n) = F^{n\theta}(u_n), \tag{3.11}$$

denn

$$\begin{aligned} P(M_n \le u_n) &\approx \exp(-\theta\tau) \approx \exp(-\theta n\bar{F}(u_n)) \\ &= (\exp(-\bar{F}(u_n)))^{n\theta} \approx (1-\bar{F}(u_n))^{n\theta} = F^{n\theta}(u_n). \end{aligned}$$

Eine stationäre Folge mit obiger Eigenschaft hat den extremalen Index θ und dieser ist ein Maß für die Abhängigkeitstruktur der Daten. Je kleiner der extremale Index ist, desto größer ist die Abhängigkeit der Zufallsvariablen, und umgekehrt. So haben iid Zufallsvariablen einen extremalen Index von eins. Allerdings folgt aus $\theta = 1$ nicht unbedingt, dass die Zufallsvariablen unabhängig sind. Sie haben dann aber nur eine schwache Abhängigkeit.
Mit der Aussage (3.11) lässt sich ebenfalls folgern, dass sich das Maximum von n Beobachtungen einer stationären Folge mit extremalen Index θ genauso verhält, wie das Maximum von $n\theta$ Beobachtungen der assoziierten iid Zufallsvariablen.
$n\theta$ ist die Anzahl von unabhängigen Clustern in n Beobachtungen, demzufolge ist θ der Kehrwert der durchschnittlichen Cluster-Größe. Das heißt, θ^{-1} ist die durchschnittliche Anzahl von extremen Werten in einem Cluster.
Das wichtigste Ergebnis für stationäre Folgen ist das folgende Theorem, das aus (3.10) und (3.11) folgt.

Theorem 3.3. *(vgl. [15, S. 7])*
Sei (X_n) eine stationäre Folge mit extremalen Index $\theta > 0$ und $(\tilde{X}_n)$ eine Folge von unabhängigen Zufallsvariablen. Dann gilt

$$\lim_{n\to\infty} P(a_n(\tilde{M}_n - b_n) \le x) = H(x),$$

für eine nicht-entartete Verteilungsfunktion H und Folgen $(a_n) > 0$ und $(b_n) \in \mathbb{R}$ genau dann, wenn

$$\lim_{n\to\infty} P(a_n(M_n - b_n) \le x) = H^{\theta}(x),$$

wobei auch H^{θ} eine nicht-entartete Verteilungsfunktion ist.

Wenn F im Max-Anziehungsbereich einer Extremwertverteilung liegt, dann konvergieren auch die skalierten Maxima einer stationären Folge gegen eine GEV-Verteilung. Es gilt dann für $\xi \neq 0$

$$\begin{aligned} H^{\theta}_{\xi,\mu,\sigma}(x) &= \exp\Big\{ -\theta\Big(1+\xi\Big(\frac{x-\mu}{\sigma}\Big)\Big)^{-1/\xi}\Big\} \\ &= \exp\Big\{ -\Big(\theta^{-\xi}\Big(1+\xi\Big(\frac{x-\mu}{\sigma}\Big)\Big)\Big)^{-1/\xi}\Big\} \\ &= \exp\Big\{ -\Big(1+\xi\Big(\frac{x-\mu^*}{\sigma^*}\Big)\Big)^{-1/\xi}\Big\}, \end{aligned}$$

wobei

$$\mu^* = \mu - \frac{\sigma}{\xi}(1-\theta^{\xi}) \quad \text{und} \quad \sigma^* = \sigma\theta^{\xi}.$$

Die angepasste Verteilung der stationären Folge hat denselben Shape-Parameter ξ wie im unabhängigen Fall. Das Tail-Verhalten bleibt gleich, nur der Loc- und der Scale-Parameter verändern sich, wenn die Folge (X_n) stationär ist.
Für den Fall $\xi = 0$, ändert sich nur der Loc-Parameter μ, denn dann ist

$$\begin{aligned} H^{\theta}_{\xi,\mu,\sigma}(x) &= \exp\Big\{ -\theta e^{-\left(\frac{x-\mu}{\sigma}\right)}\Big\} \\ &= \exp\Big\{ -e^{-\left(\frac{x-\mu^*}{\sigma^*}\right)}\Big\}, \end{aligned}$$

wobei

$$\mu^* = \mu + \sigma\log\theta \quad \text{und} \quad \sigma^* = \sigma.$$

Das Theorem 3.3 zeigt, dass die Abhängigkeit der (X_n) einen Einfluss auf die Konvergenz gegen die GEV-Verteilung hat. Die Konvergenzgeschwindigkeit ist langsamer als bei unabhängigen Zufallsvariablen, denn $n\theta$ ist kleiner als n. Deswegen sollte darauf geachtet werden, dass die Maximum-Blöcke groß genug sind, um sie an eine GEV-Verteilung anpassen zu können (vgl. [15, S. 7]).

Über den extremalen Index ist noch zu bemerken, dass nicht jede stationäre Folge einen extremalen Index besitzt. In [7, S. 418] findet sich dazu ein Gegenbeispiel. In dieser Arbeit wird aber angenommen, dass der extremale Index immer existiert und er auch echt größer ist als Null, denn nur der Fall $0 < \theta \leq 1$ ist von praktischem Interesse. Aber in [12, S. 68ff] wird auch der Fall $\theta = 0$ betrachtet.

Schätzmethoden für den extremalen Index

Der extremale Index spielt bei stationären Folgen eine wichtige Rolle, deshalb werden nun zwei Methoden vorgestellt, um den extremalen Index aus einem gegebenen Datensatz zu schätzen.
Als erstes wird die sogenannte *Block Methode* (siehe [7, S. 419ff]) beschrieben und anschließend wird kurz die *Run Methode* erklärt.

In (3.11) wurde schon erwähnt, dass

$$P(M_n \leq u_n) \approx P^{\theta}(\tilde{M}_n \leq u_n) = F^{\theta n}(u_n)$$

gilt, wobei $n(1 - F(u_n)) \to \tau$ vorausgesetzt wird. Daraus folgt dann

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

Da weder $P(M_n \leq u_n)$ noch $F(u_n)$ bekannt sind, müssen beide Größen geschätzt werden.
Seien $X_1, \ldots, X_k$ die gegebenen Log-Renditen aus denen gemäß der Block-Maxima Methode die Maxima ausgewählt werden, wobei $k = mn$. Die $X_1, \ldots, X_k$ werden in m Blöcke der Größe n unterteilt.

Ein möglicher Schätzer für den Tail $(1 - F(u_k)) = P(X_i > u_k)$ ergibt sich durch den Gebrauch der empirischen Verteilungsfunktion:

$$\frac{N_u}{k} = \frac{1}{k} \sum_{i=1}^{k} \mathbf{1}_{\{X_i > u_k\}}.$$

Um nun einen empirischen Schätzer für $P(M_k \leq u_k)$ zu finden, wird folgende Approximation berücksichtigt, die schon in (3.5) erwähnt wurde,

$$P(M_k \leq u_k) \approx P^m(M_n \leq u_k).$$

Damit lässt sich folgern:

$$\begin{aligned} P(M_k \leq u_k) &= P\Big(\max_{1\leq j\leq m} M_{n,j} \leq u_k\Big) \\ &\approx P^m(M_n \leq u_k) \\ &= \Big(\frac{1}{m} \sum_{j=1}^{m} \mathbf{1}_{\{M_{n,j} \leq u_k\}}\Big)^m \\ &= \Big(1 - \frac{K_u}{m}\Big)^m. \end{aligned}$$

Nun kann ein Schätzer für den extremalen Index konstruiert werden, wenn zuvor ein hoher Threshold u gewählt wurde:

$$\hat{\theta}_1 = \frac{m}{mn} \frac{\log(1 - K_u/m)}{\log(1 - N_u/(mn))} = \frac{1}{n} \frac{\log(1 - K_u/m)}{\log(1 - N_u/(mn))}, \tag{3.12}$$

wobei N_u die Anzahl von Überschreitungen des Thresholds aller X_i ist und K_u ist die Anzahl von Blöcken, in denen der Threshold überschritten wird.

Wenn K_u/m und $N_u/(mn)$ klein sind, dann lässt sich $\hat{\theta}_1$ zu einem einfacheren Schätzer reduzieren (s. [15, S. 12])

$$\hat{\theta}_2 = \frac{1}{n}\frac{K_u/m}{N_u/(mn)} = \frac{K_u}{N_u}.$$

Eine Schwierigkeit, um einen guten Schätzer für den extremalen Index zu finden, ist die Wahl eines geeigneten Thresholds. Eine Möglichkeit ist es, u in Abhängigkeit von N_u zu setzen. Das heißt, es werden mehrere Thresholds gewählt, so dass die Anzahl der Überschreitungen in einem bestimmten Intervall liegen. Die Größe des Intervalls hängt dann von der Länge des Datensatzes ab. So erhält man mehrere Werte für den extremalen Index und erkennt, wie sich θ ändert, bei wechselndem u. Der Durchschnittswert bildet dann den extremalen Index.

Die zweite Methode, um einen Schätzer für θ zu finden, ist die sogenannte *Run Methode*, die hier nur kurz vorgestellt werden soll. Eine ausführliche Darstellung findet sich in [7, S. 422ff].
Dort wird gezeigt, dass

$$\lim_{k\to\infty} P(M_{2,r} \leq u_k | X_1 > u_k) = \theta$$

gilt, wobei $M_{2,r} = \max(X_2, \ldots, X_r)$ ist und $r/k \to 0$, wenn $r \to 0$.

θ ist eine bedingte Wahrscheinlichkeit und ein Maß, das angibt, ob eine Tendenz besteht, dass die X_i Cluster bilden. $M_{2,r}$ kann nur dann kleiner sein als u_k, wenn X_1 das letzte Element eines Clusters ist, welches u_k überschreitet. In einem Cluster gibt es keine r aufeinander folgende Werte, die unter dem gegebenen Threshold liegen. Damit kann dann ein Schätzer für θ gebildet werden, wenn zuvor ein hoher Threshold u und eine Run-Länge r festgelegt wurden:

$$\hat{\theta}_3 = \frac{\sum_{i=1}^{mn-r} \mathbf{1}_{\{X_i>u, X_{i+1}\leq u, \ldots, X_{i+r}\leq u\}}}{\sum_{i=1}^{mn} \mathbf{1}_{\{X_i>u\}}} = \frac{\sum_{i=1}^{mn-r} \mathbf{1}_{\{X_i>u, X_{i+1}\leq u, \ldots, X_{i+r}\leq u\}}}{N_u},$$

wobei N_u die Anzahl der Überschreitungen über dem Threshold u ist.

Als erstes werden bei dieser Methode die Cluster der Überschreitungen identifiziert und zwar so, dass sie von mindestens r „Nicht-Überschreitungen“ getrennt werden und dann wird noch durch die gesamte Anzahl an Überschreitungen geteilt.

3.3.2 Eigenschaften von ARCH-Prozessen

In diesem und dem nächsten Teil des Kapitels werden die ARCH-Prozesse noch gründlicher betrachtet. Es dreht sich vor allem um die Frage, wie gut extreme Ereignisse mit ARCH-Prozessen modelliert werden können. Dafür wird nun untersucht, wann ein ARCH-Prozess heavy-tailed ist und wie sich die Maxima eines

ARCH-Prozesses verhalten.
Im weiteren Verlauf wird ein ARCH(1)-Prozess betrachtet, der den folgenden Gleichungen genügt (s. Definition 3.6)

$$\begin{aligned} X_t &= \sigma_t Z_t, \qquad t \geq 1, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2, \end{aligned}$$

wobei (Z_t) eine Folge von iid standard-normalverteilten Zufallsvariablen ist und α_0 und α_1 sind nicht-negative Konstanten. X_0 ist die Start-Variable, die unabhängig von (Z_t) ist.

Da der Begriff der Stationarität eine zentrale Rolle spielt, wird nun geklärt, welche Bedingungen erfüllt sein müssen, damit der ARCH(1)-Prozess strikt stationär ist (vgl. [7, S. 455ff]).

Es wird ein quadrierter ARCH(1)-Prozess $X_t^2 = \sigma_t^2 Z_t^2$ betrachtet. Daraus ergibt sich

$$X_t^2 = \alpha_0 Z_t^2 + \alpha_1 X_{t-1}^2 Z_t^2 \tag{3.13}$$

und im Folgenden wird untersucht, wann diese Gleichung eine stationäre Lösung hat. Die Gleichung (3.13) bezeichnet man auch als stochastische Rekurrenzgleichung (stochastic recurrence equation). Sie gehört zu einer Klasse von Rekurrenzgleichungen der Form

$$Y_t = A_t Y_{t-1} + B_t, \tag{3.14}$$

wobei $(A_t)_{t\in\mathbb{Z}}$ und $(B_t)_{t\in\mathbb{Z}}$ Folgen von iid Zufallsvariablen sind. Durch Iteration ergibt sich

$$Y_t = Y_0 \prod_{j=1}^{t} A_j + \sum_{m=1}^{t} B_m \prod_{j=m+1}^{t} A_j. \tag{3.15}$$

Wenn nun gilt

$$E(\max\{0, \log|B_t|\}) < \infty \text{ und } E(\log|A_t|) < 0, \tag{3.16}$$

dann gilt $Y_t \xrightarrow{d} Y$, für eine Zufallsvariable Y, und

$$Y \stackrel{d}{=} AY + B,$$

wobei $(A, B) = (A_1, B_1)$ und Y unabhängig sind. Diese Gleichung hat eine eindeutige Lösung, da der erste Term von (3.15) verschwindet, wenn die Bedingungen (3.16) erfüllt sind. Es gilt

$$Y \stackrel{d}{=} \sum_{m=1}^{\infty} B_m \prod_{j=1}^{m-1} A_j.$$

Sei nun $Y_0 \stackrel{d}{=} Y$, dann ist der Prozess (Y_t) strikt stationär, denn es gilt für jedes $t \in \mathbb{N}$:

$$(Y_t, Y_{t+1}) = (Y_t, B_{t+1} + A_{t+1}Y_t) \stackrel{d}{=} (Y_0, B_1 + A_1 Y_0) = (Y_0, Y_1).$$

Diese Aussage kann jetzt auf den ARCH(1)-Prozess übertragen werden. Denn der quadrierte ARCH-Prozess kann mit $A_t = \alpha_1 Z_t^2$ und $B_t = \alpha_0 Z_t^2$ auch in der Form (3.14) geschrieben werden. Seien nun

$$E(\max(0, \log|\alpha_0 Z_t^2|)) < \infty \text{ und } E(\log(\alpha_1 Z_t^2)) < 0 \tag{3.17}$$

erfüllt. Dann gilt (siehe [7, S. 465]): Der Prozess (X_t) ist strikt stationär, wenn

$$X_0^2 \stackrel{d}{=} \alpha_0 \sum_{m=1}^{\infty} Z_m^2 \prod_{j=1}^{m-1} (\alpha_1 Z_j^2) = \alpha_0 \sum_{m=0}^{\infty} \alpha_1^m \prod_{j=1}^{m+1} Z_j^2. \tag{3.18}$$

Ob die Bedingung (3.17) erfüllt und der ARCH-Prozess strikt stationär ist, lässt sich direkt aus dem Parameter α_1 ableiten, wenn die Innovationen $(Z_t)_{t\in\mathbb{Z}}$ standardnormalverteilt sind. Denn dann ergibt sich

$$\begin{aligned} E(\log(\alpha_1 Z_t^2)) &= \log(\alpha_1) + E(\log Z_t^2) \\ &= \log(\alpha_1) + \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \log(x^2) e^{-x^2/2} dx \\ &= \log(\alpha_1) + \frac{1}{\sqrt{\pi}} \Big(\log(2)\Gamma\Big(\frac{1}{2}\Big) + \Gamma'\Big(\frac{1}{2}\Big)\Big) \\ &= \log(\alpha_1) + \log(2) - 2\log(2) - \gamma \\ &= \log(\alpha_1) - (\gamma + \log(2)), \end{aligned} \tag{3.19}$$

wobei $\sqrt{\pi} = \Gamma(1/2)$ und $\gamma = \lim \sum_{n=1}^{N} \frac{1}{n} - \log(N) \approx 0.5772$ für $N \to \infty$ die Euler-Mascheroni Konstante ist. Da für die strikte Stationarität $E(\log(\alpha_1 Z_t^2)) < 0$ gelten muss, folgt dann:
Der ARCH(1)-Prozess ist genau dann strikt stationär, wenn $\alpha_1 < 2\exp(\gamma) \approx 3.562$ ist.

Im nachfolgenden Satz wird bewiesen, wann der ARCH(1)-Prozess schwach stationär ist. Mit diesem Resultat lässt sich dann auch die Varianz des ARCH-Prozesses berechnen, die schon in Abschnitt 3.2.1 erwähnt wurde.

Satz 3.2. *(s. [17, S. 142])*
Der ARCH(1)-Prozess ist schwach stationär genau dann, wenn $\alpha_1 < 1$ ist. Die Varianz des schwach stationären Prozesses ist dann durch $\alpha_0/(1-\alpha_1)$ gegeben.

Beweis:
Wenn vorausgesetzt wird, dass der ARCH(1)-Prozess schwach stationär ist, dann folgt aus (3.13) und $E(Z_t^2) = 1$, dass

$$Var(X_t) = E(X_t^2) = \alpha_0 + \alpha_1 E(X_{t-1}^2) = \alpha_0 + \alpha_1 Var(X_{t-1}),$$

da die Varianz eines schwach stationären Prozesses konstant ist, gilt $Var(X_t) = Var(X_{t-1})$. Es ergibt sich $Var(X_t) = \alpha_0/(1-\alpha_1)$, wobei $\alpha_1 < 1$ sein muss.
Sei nun $\alpha_1 < 1$ dann gilt mit der Jensenschen Ungleichung

$$E(\log(\alpha_1 Z_t^2)) \leq \log(E(\alpha_1 Z_t^2)) = \log(\alpha_1) < 0,$$

da $E(Z_t^2) = 1$, und mit (3.18) folgt

$$E(X_t^2) = \alpha_0 \sum_{m=0}^{\infty} \alpha_1^m = \frac{\alpha_0}{1-\alpha_1}.$$

Da das zweite Moment existiert und die Varianz für alle $t \in \mathbb{Z}$ konstant ist, folgt die Behauptung. □

Mit diesem Satz folgt, dass auch der strikt stationäre ARCH-Prozess die Varianz $\alpha_0/(1-\alpha_1)$ besitzt, wenn $\alpha_1 < 1$ ist. α_1 ist somit ein wichtiger Parameter für den ARCH-Prozess.
Zusammenfassend lässt sich sagen, der ARCH(1)-Prozess ist

- für $\alpha_1 = 0$ ein White Noise,
- für $\alpha_1 \in (0,1)$ strikt stationär mit endlicher Varianz,
- für $1 \leq \alpha_1 < 2\exp(\gamma) \approx 3.562$ strikt stationär mit unendlicher Varianz.

Außerdem beeinflusst α_1 das Verhalten der Tails des ARCH-Prozesses. Denn ein Grund dafür, dass der ARCH-Prozess besonders gut Finanzzeitreihen modelliert, ist, dass er heavy-tailed ist, selbst dann, wenn die Innovationen es nicht sind.
Eine einfache Möglichkeit zu zeigen, dass der Prozess heavy-tailed ist, ist die Kurtosis auszurechnen. So ergibt sich mit standard-normalverteilten Innovationen (Z_t), die eine Kurtosis von 3 haben, für das bedingte vierte Moment

$$E(X_t^4|\mathcal{F}_{t-1}) = E(\sigma_t^4 Z_t^4|\mathcal{F}_{t-1}) = 3E(\sigma_t^4|\mathcal{F}_{t-1}) = 3(\alpha_0 + \alpha_1 X_{t-1}^2)^2.$$

Damit lässt sich dann das unbedingte vierte Moment berechnen

$$E(X_t^4) = E(E(\sigma_t^4 Z_t^4|\mathcal{F}_{t-1})) = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}. \tag{3.20}$$

Das vierte Moment ist aber nur dann endlich, wenn $\alpha_1^2 < 1/3$. Für die Kurtosis ergibt sich

$$\frac{E(X_t^4)}{\sigma^4} = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \times \frac{(1-\alpha_1)^2}{\alpha_0^2} = \frac{3(1-\alpha_1^2)}{1-3\alpha_1^2}. \tag{3.21}$$

Die Kurtosis eines ARCH(1)-Prozesses ist also für $\alpha_1^2 < 1/3$ größer als 3. Damit hat der ARCH-Prozess breitere Flanken als die Normalverteilung und ist somit heavy-tailed.

3.3.3 Extremales Verhalten von ARCH-Prozessen

Nachdem nun hergeleitet wurde, wann ein ARCH(1)-Prozess strikt stationär ist und wann die Momente existieren, wird in diesem Abschnitt zuerst der Tail der Marginal-Verteilung eines ARCH(1)-Prozesses beschrieben. Denn gerade der Tail ist wichtig, um Extrema zu modellieren, wie aus dem vorherigen Kapitel bekannt ist. Danach wird untersucht, welches extremale Verhalten ein strikt stationärer ARCH-Prozess hat.
Dieser Abschnitt orientiert sich insbesondere an [5] und [7, S. 463ff], sofern keine andere Literatur angeben ist.

Um den Tail des ARCH-Prozesses zu beschreiben, wird nachfolgendes Lemma benötigt. Zudem gilt in diesem Abschnitt $X = X_0$ und $Z = Z_1$.

Lemma 3.1. *(vgl. [7, S. 463])*
Für eine standard-normalverteilte Zufallsvariable Z und $\alpha_1 \in (0, 2\exp(\gamma))$, wobei $\gamma \approx 0.5772$ die Euler-Mascheroni-Konstante ist, wird definiert

$$h(u) = E(\alpha_1 Z^2)^u, \quad u \geq 0.$$

Dann ist

$$h(u) = \frac{(2\alpha_1)^u}{\sqrt{\pi}}\Gamma\Big(u + \frac{1}{2}\Big), \quad u \geq 0. \tag{3.22}$$

Die Funktion h ist strikt konvex in u und für die Gleichung $h(u) = 1$ existiert eine eindeutige Lösung $\kappa = \kappa(\alpha_1) > 0$. Weiter gilt

$$\kappa(\alpha_1) \begin{cases} > 1, & \alpha_1 \in (0,1), \\ = 1, & \alpha_1 = 1, \\ < 1, & \alpha_1 \in (1, 2\exp(\gamma)), \end{cases} \tag{3.23}$$

und

$$E((\alpha_1 Z^2)^\kappa \log(\alpha_1 Z^2)) > 0 \quad \textit{und} \quad E(\log(\alpha_1 Z^2)) < 0. \tag{3.24}$$

Beweis:
Durch Substitution erhält man, ähnlich wie in (3.19),

$$\begin{aligned} h(u) = E(\alpha_1 Z^2)^u &= \frac{2\alpha_1^u}{\sqrt{2\pi}} \int_0^\infty x^{2u} \exp(-\frac{x^2}{2}) dx \\ &= \frac{(2\alpha_1)^u}{\sqrt{\pi}}\Gamma\Big(u + \frac{1}{2}\Big), \end{aligned}$$

also (3.22).
Es gilt $h(0) = 1$ für alle α_1 und

$$h'(u) = E((\alpha_1 Z^2)^u \log(\alpha_1 Z^2)).$$

Damit ergibt sich, analog zu (3.19),

$$h'(0) = E(\log(\alpha_1 Z^2)) = \log(\alpha_1) - \log(2) - \gamma < 0, \tag{3.25}$$

wobei $0 < \alpha_1 < 2\exp(\gamma)$. Es gilt

$$h''(u) = E((\alpha_1 Z^2)^u (\log(\alpha_1 Z^2))^2) > 0 \tag{3.26}$$

und daraus folgt, dass h strikt konvex ist auf $\mathbb{R}^+$. Zudem ist

$$h(u) \geq E((\alpha_1 Z^2)^u \mathbf{1}_{\{\alpha_1 Z^2 > 2\}}) \geq 2^u P(\alpha_1 Z^2 > 2) \to \infty,$$

für $u \to \infty$. Also gibt es ein eindeutiges $\kappa > 0$, so dass $h(\kappa) = 1$ ist, da $h(0) = 1$ und h strikt konvex ist.
Wegen der strikten Konvexität, gilt $h'(\kappa) > 0$ und somit ergeben sich die Ungleichungen (3.24) mit Hilfe von (3.25) und (3.26).
(3.23) folgt, da $h(1) = \alpha_1$ gilt. Denn wenn $\alpha_1 = 1$ ist, folgt $\kappa = 1$, da $h(\kappa) = 1$ ist mit eindeutigem κ. Die anderen Fallunterscheidungen ergeben sich analog. □

Sei (X_t) ein ARCH(1)-Prozess mit festen Parametern $\alpha_0 > 0$ und $0 < \alpha_1 < 2\exp(\gamma)$ und X_0 ist unabhängig von (Z_t). Nach (3.18) ist bekannt, dass (X_t) strikt stationär ist, wenn

$$X_0^2 \stackrel{d}{=} \alpha_0 \sum_{m=1}^{\infty} Z_m^2 \prod_{j=1}^{m-1} (\alpha_1 Z_j^2).$$

Damit kann die Marginal-Verteilung gebildet werden. Denn jeder strikt stationäre ARCH(1)-Prozess (X_t) hat die Marginal-Verteilung

$$X_t \stackrel{d}{=} |X_0| r_0,$$

wobei X_0^2 der Gleichung (3.18) genügt und r_0 ist eine Bernoulli-verteilte Zufallsvariable mit $P(r_0 = \pm 1) = 0.5$, die unabhängig von $|X_0|$ ist. Da die Innovationen (Z_t) symmetrisch sind, ist auch X_t symmetrisch und damit ergibt sich $X_t \stackrel{d}{=} |X_t| r_0$ und daraus folgt die Behauptung (vgl. [7, S. 465]).

Im nächsten Theorem wird der Tail der Marginal-Verteilung beschrieben. Wenn ein $\kappa > 0$ existiert, so dass $E(\alpha_1 Z^2)^\kappa = 1$ gilt, dann hat der ARCH(1)-Prozess heavy Tails, selbst dann wenn die Innovationen thin-tailed sind.

Theorem 3.4.
Sei (X_t) ein strikt stationärer ARCH(1)-Prozess mit Parametern $\alpha_0 > 0$ und $0 < \alpha_1 < 2\exp(\gamma)$, wobei γ die Euler-Mascheroni-Konstante ist. Sei $\kappa > 0$ die eindeutige positive Lösung der Gleichung $h(u) = 1$. Dann gilt

$$P(X > x) \sim \frac{c}{2} x^{-2\kappa}, \quad x \to \infty,$$

wobei

$$c = \frac{E[((\alpha_0 + \alpha_1 X^2)^\kappa - (\alpha_1 X^2)^\kappa)(Z^2)^\kappa]}{\kappa E[(\alpha_1 Z^2)^\kappa \log(\alpha_1 Z^2)]} \in (0, \infty),$$

für eine standard-normalverteilte Zufallsvariable Z, die unabhängig ist von $X = X_0$.

In [7, S. 468ff] wird das Theorem für den ARCH(1)-Prozess geführt.
Der Parameter κ ist entscheidend für das Tail-Verhalten der Marginal-Verteilung, also auch für die Extrema eines ARCH-Prozesses.
Möchte man κ bestimmen, muss entweder die Gleichung $E(\alpha_1 Z^2)^\kappa = 1$ numerisch gelöst werden, oder die Gleichung

$$\Gamma\Big(\kappa + \frac{1}{2}\Big) = \sqrt{\pi}(2\alpha_1)^{-\kappa},$$

da die (Z_t) standard-normalverteilt sind.

α_1	0.1	0.3	0.5	0.7	0.9	0.95	0.99
κ	13.24	4.180	2.365	1.586	1.152	1.072	1.014

Tabelle 3.1: Numerische Lösungen für κ für feste α_1-Werte.

In der Tabelle 3.1 sind einige numerische Lösungen für κ aufgelistet. (Die Ergebnisse wurden von [5] übernommen). Der Parameter α_1 hat großen Einfluss auf κ, denn je größer α_1 ist, desto kleiner ist κ, d.h. desto breitere Flanken hat der ARCH-Prozess.

Doch welches extremale Verhalten hat der ARCH(1)-Prozess?
Dafür sei im Folgenden (X_t) ein ARCH(1)-Prozess mit den Parametern $\alpha_0 > 0$ und $0 < \alpha_1 < 2\exp(\gamma)$. Es gilt wie in Theorem 3.4, dass $P(X > x) \sim \frac{c}{2} x^{-2\kappa}$.

Wenn (X_t) eine Folge von unabhängig und identisch verteilten Zufallsvariablen ist, so dass $X_1 \stackrel{d}{=} X$ gilt, dann liegt X im maximalen Anziehungsbereich einer Fréchet-Verteilung. Das folgt aus Satz 2.1 aus dem zweiten Kapitel. Es gilt also

$$\lim_{n\to\infty} P(n^{-1/(2\kappa)} M_n \leq x) = \exp(-\frac{c}{2} x^{-2\kappa}). \tag{3.27}$$

Doch hier werden keine unabhängigen Zufallsvariablen betrachtet. Deswegen müssen die Bedingungen so geändert werden, dass auch der strikt stationäre ARCH-Prozess ein ähnliches extremales Verhalten hat. Der extremale Index spielt dafür eine wichtige Rolle und für den ARCH(1)-Prozess ist die $\mathcal{D}(u_n)$-Bedingung aus Definition 3.7 erfüllt, deswegen existiert auch der extremale Index für den ARCH-Prozess (Beweis siehe [5]).

Der ARCH(1)-Prozess hat also eine schwache lange Abhängigkeit. Um jetzt herzuleiten, dass der ARCH-Prozess im Anziehungsbereich einer Fréchet-Verteilung liegt, wird dieses Resultat zuerst für den quadrierten ARCH-Prozess (X_t^2) gezeigt.

Theorem 3.5.
Sei (X_t^2) ein strikt stationärer quadrierter ARCH(1)-Prozess und $M_n^{(2)} = \max(X_1^2, \ldots, X_n^2)$. Dann gilt

$$\lim_{n\to\infty} P(n^{-1/\kappa} M_n^{(2)} \leq x) = \exp(-c\theta^{(2)} x^{-\kappa}), \quad x > 0, \tag{3.28}$$

wobei κ die positive Lösung der Gleichung $E(\alpha_1 Z^2)^u = 1$ ist. c ist definiert wie in Theorem 3.4 und $\theta^{(2)}$ ist der extremale Index mit

$$\theta^{(2)} = \kappa \int_1^\infty P\Big(\max_{n\geq 1} \prod_{t=1}^n (\alpha_1 Z_t^2) \leq y^{-1}\Big) y^{-\kappa-1} dy.$$

Durch den Vergleich von (3.27) und (3.28) erkennt man, dass $\theta^{(2)}$ der extremale Index des quadrierten ARCH(1)-Prozesses ist.

Dieses Theorem ist noch zu erweitern, wenn auch Punkt- bzw. Poisson-Prozesse betrachtet werden. Ausführlich wird dies in [7, S. 219ff], [7, S. 473ff] und besonders in [5] beschrieben, dort findet sich auch ein Beweis für das Theorem.
Werden aber auch Punktprozesse betrachtet, dann kann gezeigt werden, dass der Punktprozess von Überschreitungen (point process of exceedances) gegen einen Compound-Poisson-Prozess[1] mit Intensität $c\theta^{(2)} x^{-\kappa}$ konvergiert.
Mit $\pi_k^{(2)}$ werden dann die Cluster-Wahrscheinlichkeiten des Compound-Poisson-Prozess bezeichnet. π_k ist die Wahrscheinlichkeit, dass ein Ereignis Vielfachheit k hat. Diese Cluster-Wahrscheinlichkeit kann mit Hilfe des extremalen Indexes berechnet werden. Es gilt:

$$\pi_k^{(2)} = \frac{\theta_k^{(2)} - \theta_{k+1}^{(2)}}{\theta^{(2)}}, \quad k \in \mathbb{N}, \tag{3.29}$$

wobei

$$\theta_k^{(2)} = \kappa \int_1^\infty P\Big(\#\Big\{n \in \mathbb{N} : \prod_{t=1}^n (\alpha_1 Z_t^2) > y^{-1}\Big\} = k-1\Big) y^{-\kappa-1} dy.$$

Insbesondere gilt $\theta_1^{(2)} = \theta^{(2)}$.

Das extremale Verhalten von (X_t^2) ist nun bekannt, also auch das von $(|X_t|)$. Doch welche Bedingungen gelten für die Extrema von (X_t)? Da die Innovationen (Z_t) iid und symmetrisch sind, lässt sich folgern:

$$(X_t) = (r_t |X_t|),$$

[1] Seien $(N_t)_{t\geq 0}$ Zufallsvariablen und für $0 = t_0 < t_1 < \ldots < t_n$ sind die Differenzen voneinander unabhängig und Poisson-verteilt mit Parameter $\lambda(t_i - t_{i-1})$, $1 \leq i \leq n$. Dann heißt $(N(t))_{t\geq 0}$ ein Poisson-Prozess mit Intensität λ. Sei (η_i) eine Folge von iid nicht-negativen ganzzahligen Zufallsvariablen, die unabhängig von $(N(t))_{t\geq 0}$ sind. Dann heißt $\tilde{N}(t) = \sum_{i=1}^{N(t)} \eta_i$, $t \geq 0$ der Compound-Poisson-Prozess mit Intensität λ und Cluster-Größe η_i. $\pi_k = P(\eta_1 = k)$, $k \geq 0$, sind die Cluster-Wahrscheinlichkeiten von $\tilde{N}$.

wobei (r_t) iid und unabhängig von (X_t) ist und $P(r_1 \pm 1) = 0.5$. Damit haben die Maxima von (X_t) und $(r_t|X_t|)$ dieselbe Verteilung und die bekannten Ergebnisse von $(|X_t|)$ können verwendet werden, um die Überschreitungen eines Thresholds $u > 0$ von (X_t) zu untersuchen. Dabei werden aber die Überschreitungen von $\{X_t < -u\}$ ausgeschlossen.
Eine wichtige Hilfe für das nachfolgende Theorem ist die erzeugenden Funktion der Cluster-Wahrscheinlichkeiten $(\pi_k^{(2)})$, wie sie in (3.29) definiert wurden,

$$\Pi^{(2)}(u) = \sum_{k=1}^{\infty} \pi_k^{(2)} u^k.$$

Theorem 3.6.
Sei (X_t) ein strikt stationärer ARCH(1)-Prozess. Dann gilt

$$\lim_{n\to\infty} P\Big(n^{-1/(2\kappa)} M_n \leq x\Big) = \exp\Big(-c\theta^{(2)}(1 - \Pi^{(2)}(0.5))x^{-2\kappa}\Big).$$

Der extremale Index vom Prozess (X_t) ist $\theta = 2\theta^{(2)}(1 - \Pi^{(2)}(0.5))$, wenn man das Theorem mit (3.27) vergleicht.
Der Beweis findet sich in [5] und [7, S. 476].

Auch hier kann das Theorem entsprechend erweitert werden, so dass der Punktprozess der Überschreitungen gegen den Compound-Poisson-Prozess mit Intensität $c\theta x^{-2\kappa}$ konvergiert.

Mit Theorem 3.6 wurde gezeigt, dass der ARCH(1)-Prozess ein ähnliches Grenzverhalten hat wie eine Folge von iid Zufallsvariablen, wenn sie die gleiche Marginal-Verteilung haben. So liegt die Marginal-Verteilung des ARCH(1)-Prozesses im Anziehungsbereich einer Fréchet-Verteilung (siehe auch Theorem 3.4). Der Parameter κ hat dabei eine besondere Bedeutung, denn er beschreibt das Tail-Verhalten der Marginal-Verteilung und κ ist somit nichts anderes als das Inverse des Shape-Parameters $\xi = 1/(2\kappa)$, wie er in Kapitel 2 eingeführt wurde. Somit ist die asymptotische Verteilung der Maxima des ARCH(1)-Prozesses eine Fréchet-Verteilung mit Shape-Parameter ξ (vgl. [15, S. 8]).
In Tabelle 3.2 sind verschiedene Werte für α_1 und die entsprechenden Shape-Parameter aufgelistet. Daraus lässt sich ablesen, dass der Shape-Parameter ansteigt, wenn α_1 größer wird. Man gelangt zum selben Resultat, wie in Tabelle 3.1. Denn je größer α_1, desto größer ist ξ und desto breitere Flanken hat die Verteilung des ARCH-Prozesses. Neben dem Shape-Parameter wird auch der extremale Index berechnet. $\theta^{(2)}$ und $\theta_k^{(2)}$ aus den vorherigen Theoremen, können aber leichter berechnet werden, wenn α_0 und α_1 bekannt sind:

Lemma 3.2.
Sei E_κ eine exponential-verteilte Zufallsvariable mit Parameter κ, die unabhängig

ist von dem Random Walk (S_n), der definiert wird als

$$S_0 = 0, \quad S_n = \sum_{t=1}^{n} \log(\alpha_1 Z_t^2).$$

Dann gilt

$$\theta^{(2)} = \theta_k^{(1)} = P\Big(\max_{n\geq 1} S_n \leq -E_\kappa\Big),$$
$$\theta_k^{(2)} = P\Big(\sum_{n=1}^{\infty} \mathbf{1}_{\{S_n > -E_\kappa\}} = k-1\Big), \quad k \geq 2.$$

Der Beweis findet sich in [7, S. 479].
Mit Hilfe dieses Lemmas kann man unabhängige Replikationen von (S_n) und den unabhängigen, exponential-verteilten Zufallsvariablen E_κ simulieren. Anschließend werden die Ereignisse $\{S_n > -E_\kappa\}$ für jede Replikation gezählt. Die Anzahl der Überschreitungen ist endlich, da der Random Walk einen negativen Drift hat.
Damit können Schätzer für $\theta^{(2)}$ und $\theta_k^{(2)}$ hergeleitet werden:

$$\hat{\theta}^{(2)} = 1 - \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{max_{n\geq 1} S_n^{(i)} > -E_\kappa^{(i)}\}},$$
$$\hat{\theta}_k^{(2)} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{\sum_{n=1}^{\infty} \mathbf{1}_{G_i} = k-1\}},$$

wobei

$$G_i = \Big\{S_n^{(i)} > -E_\kappa^{(i)}\Big\}$$

und $((S_n^{(i)}), E_\kappa^{(i)})$ sind N iid Replikationen von $((S_n), E_\kappa)$.

Die Tabelle 3.2 wurde von [5, S. 222] übernommen, denn dort wird eine Monte-Carlo Simulation von $((S_n), E_\kappa)$ mit $N = 1000$ durchgeführt, um den extremalen Index zu schätzen. Jede Simulation hat eine maximale Länge von $n = 1000$ für den Random Walk S_n.

Der Shape-Parameter ξ wird immer größer mit ansteigendem α_1, dagegen wird der extremale Index immer kleiner, je größer α_1 wird. Bei einem Wert von $\alpha_1 = 0.1$ ist der extremale Index $\theta = 0.999$, das heißt, die (X_t) weisen nur eine sehr schwache Abhängigkeit auf, da der extremale Index nahe bei eins liegt. Das ändert sich aber, wenn α_1 größer wird. Bei großem α_1 besitzen die Zufallsvariablen eine starke Abhängigkeit und der Prozess hat sehr breite Tails, dies wird durch den Shape-Parameter deutlich.

α_1	ξ	$\hat{\theta}$
0.1	0.038	0.999
0.3	0.120	0.939
0.5	0.211	0.835
0.7	0.315	0.721
0.9	0.434	0.612
0.95	0.466	0.589
0.99	0.493	0.571

Tabelle 3.2: Der geschätzte extremale Index θ und der Shape-Parameter ξ von dem ARCH(1)-Prozess, die alle vom Parameter α_1 abhängen.

3.3.4 Return-Level und Return-Periode

In Kapitel 2.2.3 wurde das Return-Level für unabhängig und identisch verteilte Zufallsvariablen definiert. Das Return-Level ist eine Schwelle, die nur in einem n-Block innerhalb von k Jahren überschritten wird. Nun soll dieses Risikomaß unter der Bedingung strikter Stationarität definiert werden.

Theorem 3.3 aus Abschnitt 3.3.1 zeigt, dass die Maxima von stationären Zufallsvariablen, bis auf den extremalen Index, dasselbe Grenzverhalten haben wie iid Zufallsvariablen. Durch den Index wird die Abhängigkeit der Zufallsvariablen ausgedrückt und wenn diese GEV-verteilt sind, dann haben die stationären Zufallsvariablen zwar denselben Shape-Parameter, aber der Loc- und der Scale-Parameter ändern sich.

Die Wahrscheinlichkeit, dass in einem n-Block innerhalb von k Jahren das Return-Level $R_{n,k}$ überschritten wird, liegt bei

$$P(M_{nj} > R_{n,k}) = \frac{1}{k}, \tag{3.30}$$

wobei M_{nj} für $j \geq 1$ die Maxima sind, die anhand der Block Maxima Methode ermittelt wurden. Dann gilt nach (3.11) aus Abschnitt 3.3.1

$$P(M_{nj} \leq R_{n,k}) \approx F^{n\theta}(R_{n,k}).$$

Mit dieser Aussage und Theorem 3.3 ergibt sich dann die Definition für das Return-Level:

Definition 3.10. *Seien $M_{n1}, M_{n2}, \ldots$ die n-Block Maxima und es gilt*

$$P(M_{nj} \leq R_{n,k}) \approx P^{\theta}(\tilde{M}_{nj} \leq R_{n,k}) = F^{n\theta}(R_{n,k}), \quad j \geq 1,$$

wobei $\tilde{M}_{nj}$ die Maxima der assoziierten iid Folge sind und θ der extremale Index ist. Dann ist das k n-Block Return-Level $R_{n,k} \approx F^{-n\theta}(1 - \frac{1}{k})$.

Mit Aussage (3.11) aus Abschnitt 3.3.1 und (3.30) ergibt sich für die stationären Log-Renditen (X_n):

$$\Big(1-\frac{1}{k}\Big)^{1/n\theta} = P^{1/n\theta}(M_n \le R_{n,k}) \approx F(R_{n,k}). \tag{3.31}$$

Also ist $R_{n,k} = x_{(1-1/k)^{1/n\theta}}$, das heißt, das Return-Level ist das $(1-1/k)^{1/n\theta}$-Quantil von F (vgl. [15]).

Wenn die n-Block Maxima GEV-verteilt sind und mit H wird die GEV-Verteilung bezeichnet, dann ist das Return-Level

$$\begin{aligned} R_{n,k} &\approx 1 - H^{\theta}_{\xi,\mu,\sigma}\Big(1-\frac{1}{k}\Big) \\ &= \begin{cases} \mu - \frac{\sigma}{\xi}\Big(1-\Big(-\frac{1}{\theta}\log\Big(1-\frac{1}{k}\Big)\Big)^{-\xi}\Big) & \text{für } \xi \neq 0, \\ \mu - \sigma\log\Big(-\frac{1}{\theta}\log\Big(1-\frac{1}{k}\Big)\Big) & \text{für } \xi = 0. \end{cases} \end{aligned} \tag{3.32}$$

In Kapitel 2 wurde stets die Maximum-Likelihood Methode verwendet, um die Parameter zu schätzen. Wenn man nun die Parameter von H^θ schätzen möchten, wird die Schätzung zuerst für die GEV-Verteilung H der assoziierten iid Folge durchgeführt und anschließend kann der extremale Index mit der Block Methode geschätzt werden. Das Return-Level wird dann mit den geschätzten Parametern $(\hat{\xi}, \hat{\mu}, \hat{\sigma}, \hat{\theta})$ berechnet:

$$\hat{R}_{n,k} = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\Big(1-\Big(-\frac{1}{\hat{\theta}}\log\Big(1-\frac{1}{k}\Big)\Big)^{-\hat{\xi}}\Big) & \text{für } \hat{\xi} \neq 0, \\ \hat{\mu} - \hat{\sigma}\log\Big(-\frac{1}{\hat{\theta}}\log\Big(1-\frac{1}{k}\Big)\Big) & \text{für } \hat{\xi} = 0. \end{cases} \tag{3.33}$$

Nachdem das Return-Level definiert worden ist, soll nun auch die Return-Periode für stationäre Zufallsvariablen bestimmt werden. Die Herleitung der Return-Periode erfolgt ähnlich wie beim Return-Level mit Hilfe des extremalen Indexes und Theorem 3.3 aus Kapitel 3.3.1.

Die Return-Periode klärt, in welchem Zeitraum eine bestimmte Schwelle u überschritten wird. Dabei ist die Schwelle u vorher festzusetzten, sie sollte möglichst hoch sein, wenn man untersuchen möchte wie häufig, oder wie selten, Extremereignisse auftreten. Es ist $P(M_{nj} \le u) \approx F^{n\theta}(u)$ nach (3.11), deshalb ist die Return-Periode des Ereignisses $P(M_{nj} > u)$ auf folgende Weise definiert:

Definition 3.11. *Seien $M_{n1}, M_{n2}, \ldots$ die n-Block Maxima und es gilt*

$$P(M_{nj} \le R_{n,k}) \approx P^{\theta}(\tilde{M}_{nj} \le R_{n,k}) = F^{n\theta}(R_{n,k}), \quad j \ge 1,$$

wobei $\tilde{M}_{nj}$ die Maxima der assoziierten iid Folge sind und θ der extremale Index ist. Die Return-Periode *des Ereignisses $\{M_{nj} > u\}$ ist durch*

$$k_{n,u} \approx \frac{1}{1-F^{n\theta}(u)}$$

gegeben, wobei u eine vorgegebene Schwelle ist, $k_{n,u}$ geht gegen ∞, wenn $u \to \infty$ und $0 < \theta \leq 1$ ist der extremale Index.

Wenn die n-Block Maxima GEV-verteilt sind, dann ist

$$k_{n,u} \approx \frac{1}{1 - H^{\theta}_{\xi,\mu,\sigma}(u)} \tag{3.34}$$

die Return-Periode mit GEV-Verteilungsfunktion H^{θ}.
Genauso wie beim Return-Level werden die Parameter ξ, μ und σ der GEV-Verteilung H der iid Zufallsvariablen mit der Maximum-Likelihood Methode geschätzt. Dann kann die Return-Periode mit $(\hat{\xi}, \hat{\mu}, \hat{\sigma}, \hat{\theta})$ berechnet werden:

$$\hat{k}_{n,u} = \frac{1}{1 - H^{\hat{\theta}}_{\hat{\xi},\hat{\mu},\hat{\sigma}}(u)}.$$

3.4 Datenanalyse der Log-Renditen

Die Ergebnisse aus dem Abschnitt 3.3 sollen auch, wie im zweiten Kapitel, an einem Beispiel aus der Praxis veranschaulicht werden. Als Datensatz werden wieder die Log-Renditen des Dow Jones Indexes aus den Jahren 1960 bis 2007 verwendet, wobei die Renditen wieder negiert wurden.
Da die Log-Renditen nicht mehr als unabhängig angesehen werden, kann die GEV-Verteilung nicht mehr so einfach wie in Kapitel 2 an die Daten angepasst werden, sondern die geschätzten Parameter der GEV-Verteilung müssen mit dem extremalen Index modifiziert werden. In Abschnitt 3.4.2 erfolgt deswegen die Bestimmung des extremalen Indexes für die Log-Renditen.
Doch zuerst wird ein ARCH(1)-Prozess simuliert, denn dieser spiegelt gut die Besonderheiten von Finanzzeitreihen, wie sie die Log-Renditen des Dow Jones Indexes sind, wider. Nach der Simulation des ARCH-Prozesses kann dann überprüft werden, wie gut er die tatsächlichen Daten modelliert. Zudem wird auch für den ARCH-Prozess der extremale Index berechnet, um anschließend die GEV-Schätzung durchführen zu können. Zum Schluss werden das Return-Level und die Return-Periode berechnet. Die Vorgehensweise in diesem Kapitel richtet sich nach McNeil [15].

3.4.1 Simulation eines ARCH-Prozesses

Es wird ein ARCH(1)-Prozess simuliert, der genauso lang ist, wie die Log-Renditen des Dow Jones Indexes. Insgesamt wird ein Prozess simuliert, der aus 12081 Daten besteht. Für den Startwert wird $X_0 = 0$ gesetzt und es gilt $\alpha_0 = 0.00002$. Die Innovationen Z_t sind standard-normalverteilt und da der Parameter α_1 für das Tail-Verhalten des ARCH-Prozesses verantwortlich ist, wurden gleich vier Simulationen mit wechselnden Werten für α_1 durchgeführt.

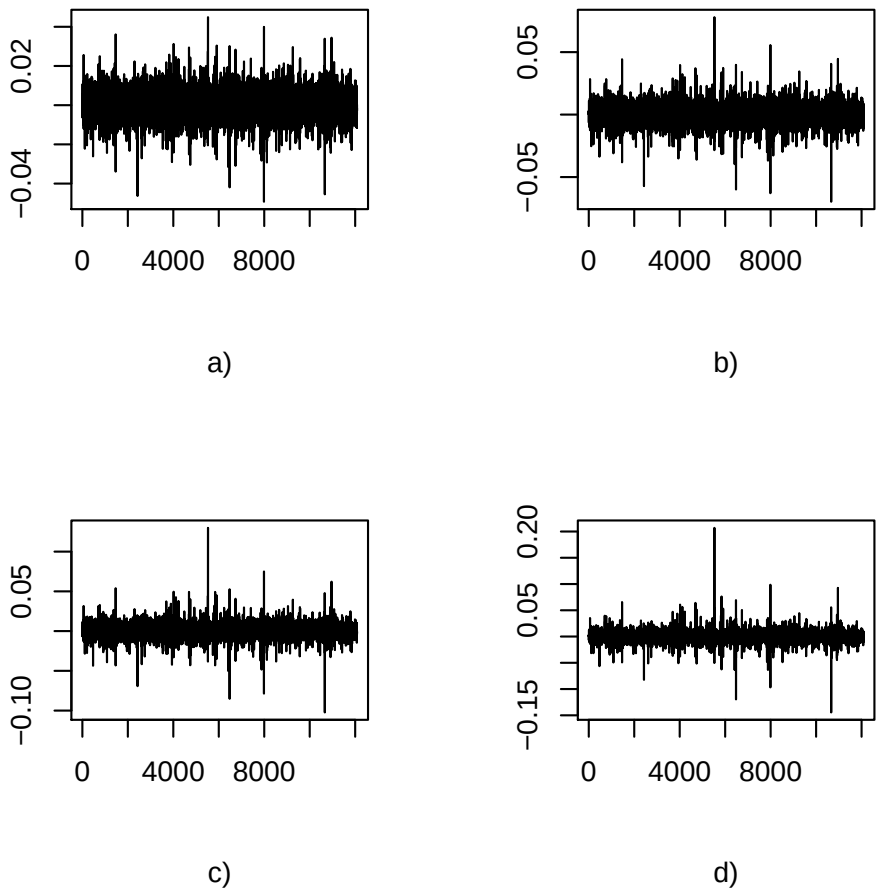

Abbildung 3.1: Simulierter ARCH(1)-Prozess mit $\alpha_0 = 0.00002$ und a) $\alpha_1 = 0.5$, b) $\alpha_1 = 0.6$, c) $\alpha_1 = 0.7$ und d) $\alpha_1 = 0.8$.

In Abbildung 3.1 ist deutlich zu erkennen, dass bei wachsendem α_1 immer mehr extrem kleine und extrem große Werte auftreten.
Im Vergleich mit den Log-Renditen ergibt sich für $\alpha_1 = 0.8$ die beste Simulation. Denn sie hat einige Gemeinsamkeiten mit den tatsächlichen Log-Renditen, wie in Abbildung 3.2 erkennbar ist. Dort wurde die x-Achse mit denselben Jahreszahlen beschriftet wie bei den Log-Renditen und auffällig ist, dass auch in der Simulation ein sehr großes Extremereignis auftritt. Allerdings ist es nicht ganz so groß wie bei den Dow Jones Daten. Insgesamt ist der ARCH-Prozess aber symmetrischer als die Log-Renditen, was auch an den standard-normalverteilten Innovationen liegt.

Der Erwartungswert des ARCH-Prozesses ist Null und da $\alpha_1 = 0.8 < 1$ ist, ist der ARCH-Prozess strikt stationär mit endlicher Varianz. Die Varianz beträgt $\alpha_0/(1-\alpha_1) = 0.0001$. Da $\alpha_1 = 0.8 > \sqrt{1/3}$, ist das vierte Moment schon unendlich. Denn es gilt $EX^4 < \infty$, wenn $\alpha_1 < \sqrt{1/3}$ ist (vgl. (3.20) aus Abschnitt 3.3.2).

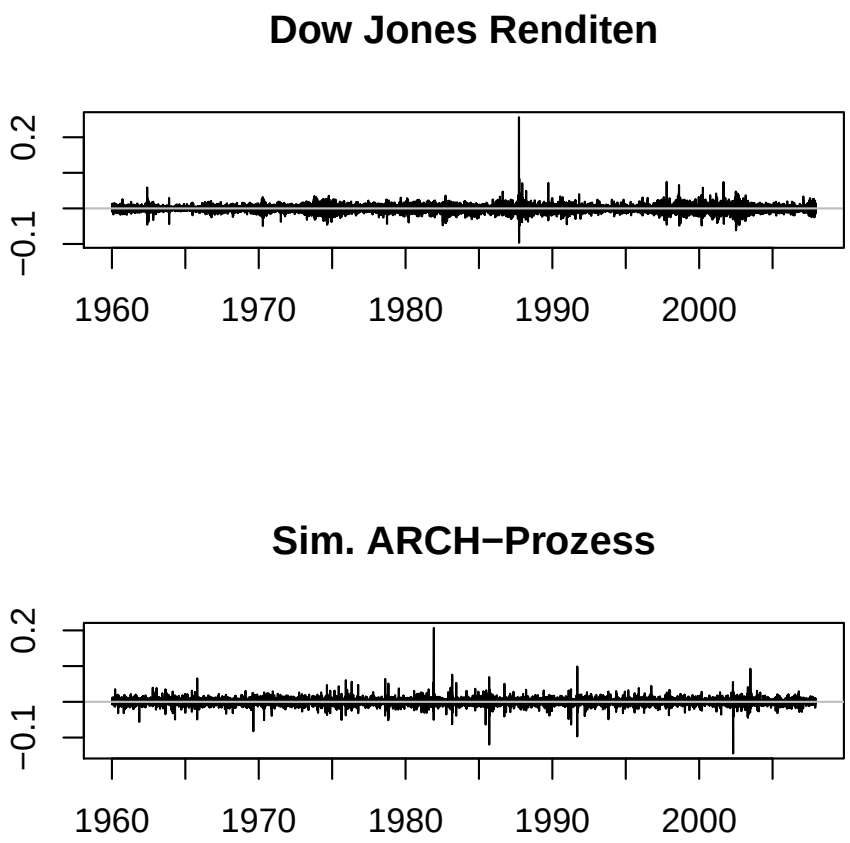

Abbildung 3.2: Die Log-Renditen des Dow Jones Indexes von 1960 bis 2007 sind in der oberen Grafik dargestellt und in der unteren Grafik der simulierte ARCH(1)-Prozess.

Der Prozess mit $\alpha_1 = 0.8$ hat also schon sehr breite Flanken, da nicht alle Momente existieren (siehe dazu das Ende von Abschnitt 2.2.1 aus Kapitel 2). Der einzige Prozess, der ein endliches viertes Moment hat, ist der Prozess in Abbildung 3.1 a). Hier beträgt die Kurtosis 9 nach (3.21) und ist größer als bei einer Normalverteilung.

Zu Beginn dieses Kapitels wurde erwähnt, dass Renditen von Finanzzeitreihen besondere statistische Eigenschaften haben, die sogenannten stilisierten Fakten. Auch bei den Dow Jones Renditen und dem simulierten ARCH-Prozess sind diese Eigenschaften zu erkennen. So ist die Kurtosis immer größer als drei, wie eben erwähnt, die Verteilung der Renditen ist somit leptokurtisch und heavy-tailed. Darüber hinaus treten große Werte vermehrt in Clustern auf (siehe Abbildung 3.2) und die Dow Jones Renditen und die Werte des ARCH-Prozesses sind unkorreliert, aber nicht unabhängig.

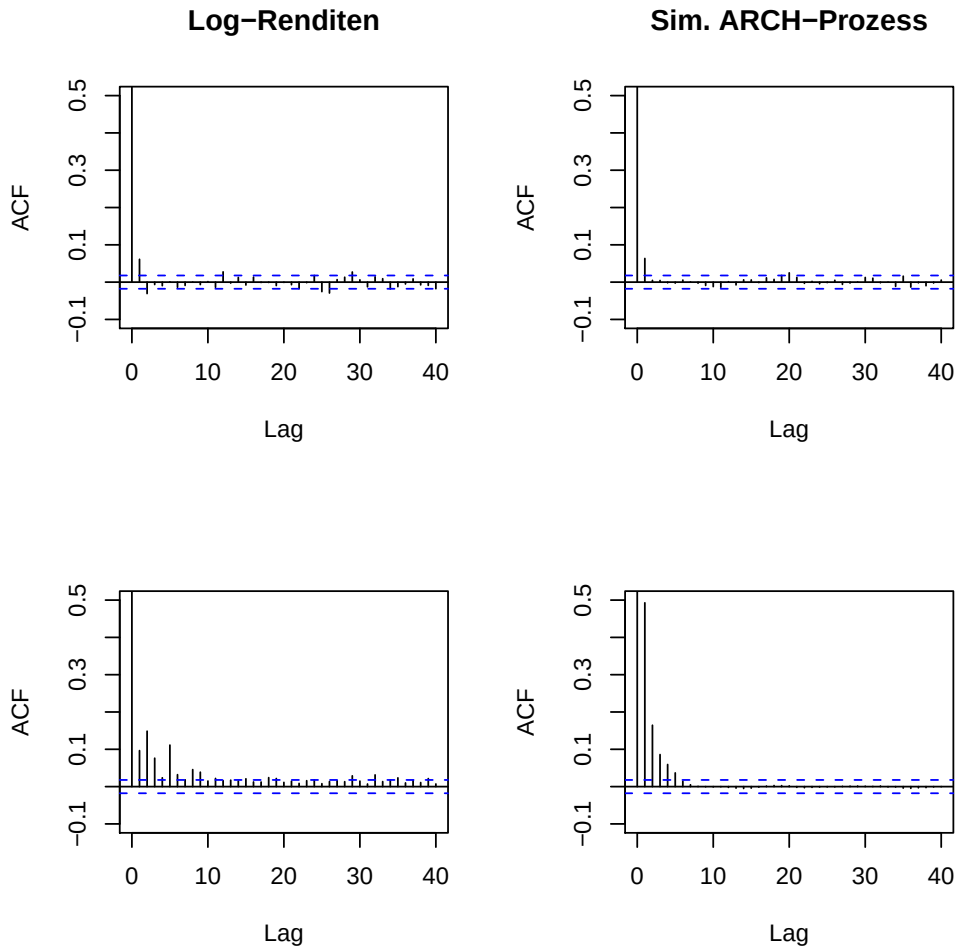

Abbildung 3.3: Empirische ACF-Funktionen der Log-Renditen und des ARCH-Prozesses (oben) und ACF-Funktionen der quadrierten Werte (unten), schwarz-gestrichelt sind die Konfidenzintervalle.

Diese Eigenschaft kann mit der empirischen Autokorrelationsfunktion überprüft werden. Denn die Autokorrelationsfunktion misst den linearen Zusammenhang zwischen zwei Zeitpunkten und durch den Lag wird angegeben, wie weit diese Zeitpunkte voneinander entfernt sind.

In Abbildung 3.3 sind die empirischen Autokorrelationen für Zeitdifferenzen von 1 bis 40 dargestellt. In den oberen beiden Grafiken ist zu erkennen, dass die Annahme der Unkorreliertheit gerechtfertigt ist. Denn das Konfidenzintervall wird kaum überschritten. Wenn es zu einer Überschreitung kommt, ist diese auch nur sehr gering. Dies gilt sowohl für die Dow Jones Daten, als auch für den ARCH-Prozess. In den unteren beiden Grafiken wurde die empirische Autokorrelationsfunktion für die quadrierten Werte geplottet. Bei den Log-Renditen kann nicht von Unabhängigkeit ausgegangen werden, da das Konfidenzintervall immer wieder überschritten wird. Allerdings werden die Überschreitungen immer kleiner und mit ansteigendem Lag

immer seltener. So wird das Konfidenzintervall im ACF-Plot der Log-Renditen bis zu einem Lag von 10 immer überschritten. Ein ähnliches Verhalten ergibt sich auch bei der Autokorrelationsfunktion des ARCH-Prozesses. Dort gibt es aber insgesamt nicht ganz so viele Überschreitungen, doch auch hier wird das Konfidenzintervall bis zu einem Lag von 10 sehr deutlich überschritten.

Abschließend ist noch zu bemerken, dass die Parameter des ARCH(1)-Prozesses sehr „willkürlich“ festgesetzt wurden. Eine andere Möglichkeit wäre, die Parameter mittels Maximum-Likelihood Methode aus den vorhandenen Log-Renditen zu schätzen. Aber dafür ist der ARCH(1)-Prozess mit standard-normalverteilten Innovationen doch recht einfach, man müsste dann einen Prozess mit einer höheren Ordnung wählen. Dies hätte aber zur Folge, dass das Modell komplizierter wird. Außerdem würde sich durch eine Maximum-Likelihood Schätzung ein sehr kleiner Wert für α_1 ergeben und demzufolge auch nur ein sehr kleiner Shape-Parameter. Dadurch wird das Risiko eines Kursabsturzes aber unterschätzt und da hier eine besonders vorsichtige Methode angewendet werden soll, werden die Parameter des ARCH-Prozesses festgesetzt und der ARCH(1)-Prozess wird simuliert.
Allerdings wird in Abschnitt 3.5 dieses Problem noch einmal aufgegriffen und eine Alternative aufgezeigt.

3.4.2 Extremaler Index

Da sowohl die Log-Renditen als auch der simulierte ARCH-Prozess nicht unabhängig ist, wird nun der extremale Index berechnet, um ein Maß für die Abhängigkeit der Daten zu haben. Der extremale Index wird mit der Block Methode berechnet und die einzelnen Blockgrößen entsprechen der Anzahl von Handelstagen in einem Jahr, Halbjahr, Quartal und Monat.
Wie schon in Kapitel 3.3.1 erwähnt, wird in dieser Arbeit angenommen, dass der extremale Index immer existiert. Das bedeutet, die $D(u_n)$-Bedingung ist erfüllt und auch in der Abbildung 3.3 wird in den unteren beiden Grafiken die schwache lang andauernde Abhängigkeit der Log-Renditen und des ARCH-Prozesses sichtbar.

In den Tabellen 3.3 und 3.4 sind die Ergebnisse des extremalen Indexes für den simulierten ARCH-Prozess und für die Log-Renditen aufgelistet. θ habe ich mit der Block Methode berechnet und dafür Formel (3.12) verwendet. Die Thresholds werden so gesetzt, dass ich Überschreitungen von 200 bis 15 habe. Dies entspricht bei dem ARCH-Prozess einen Threshold von 0.019 bis 0.054. Bei den Log-Renditen liegen die Thresholds im Bereich von 0.02 bis 0.042. Der extremale Index wird dann für alle vier Blockgrößen berechnet.

In Tabelle 3.3 sind die unterschiedlichen Ergebnisse für den extremalen Index aufgelistet und es wird deutlich, dass die Blockgröße einen großen Einfluss auf die Ergebnisse hat. Denn bei den Monatsblöcken ist der extremale Index sehr hoch, das entspricht einer geringen Abhängigkeit in den Daten. Dagegen ist der extrema-

(m,n)	N_u	200	100	50	40	30	25	20	15
Monat	K_u	132	70	37	29	22	18	16	11
(576,21)	$\hat{\theta}$	0.74	0.74	0.76	0.74	0.75	0.73	0.81	0.74
Quartal	K_u	104	62	33	26	19	17	15	10
(192,63)	$\hat{\theta}$	0.74	0.75	0.72	0.70	0.67	0.71	0.78	0.68
Halbjahr	K_u	73	52	28	23	17	15	13	10
(96,126)	$\hat{\theta}$	0.68	0.75	0.66	0.66	0.62	0.65	0.70	0.70
Jahr	K_u	45	38	22	19	15	13	12	10
(48,252)	$\hat{\theta}$	0.66	0.75	0.59	0.60	0.60	0.61	0.69	0.75

Tabelle 3.3: Der geschätzte extremale Index für den ARCH-Prozess bei unterschiedlichen Blockgrößen. N_u ist die Anzahl aller Überschreitungen des Thresholds und K_u ist die Anzahl der Blöcke, in denen der Threshold überschritten wird.

le Index kleiner, wenn die Unterteilung der Blöcke nach Jahren stattfindet. Auch innerhalb der verschiedenen Blockgrößen variieren die Ergebnisse sehr.
Vorsicht ist jedoch bei den Monats- und Jahresblöcken geboten. Denn nach der Schätzformel (3.12) müssen sowohl m als auch n hinreichend groß sein und bei der Aufteilung in Monate und Jahre ist n bzw. m sehr klein. Deswegen liefern die Quartals- und Halbjahresblöcke zuverlässigere Ergebnisse.

Um nur einen Schätzer pro Blockgröße zu haben, wurde in der Tabelle 3.5 der durchschnittliche extremale Index über die verschiedenen Thresholds gebildet. Zusätzlich sind in der Tabelle die durchschnittlichen Cluster-Größen eingezeichnet.
In der Tabelle 3.2 aus Abschnitt 3.3.3 wurde der extremale Index mit Hilfe einer Monte Carlo Simulation geschätzt und nach dieser Simulation liegt der Wert des extremalen Indexes im Bereich von 0.612 bis 0.721, wenn $\alpha_1 = 0.8$ ist. Im Vergleich mit der Tabelle 3.5 liegen alle Schätzwerte, bis auf die Monatsblöcke, innerhalb dieses Intervalls. $\hat{\theta}$ hat nach der Monte Carlo Simulation ungefähr einen Wert von 0.667. Danach ist der Schätzwert der Halbjahresblöcke am besten. Insgesamt lässt sich sagen, dass die Schätzung gut gelungen ist.

Die Ergebnisse für die Log-Renditen sind in Tabelle 3.4 zu sehen. Auch hier ist der extremale Index bei den Monatsblöcken größer als bei der Aufteilung nach Jahren, wo der extremale Index sogar einen Wert von 0.26 annimmt. Besonders auffällig ist, dass die Ergebnisse sehr großen Schwankungen unterliegen, vor allem innerhalb der einzelnen Blockgrößen. Das tritt bei dem ARCH-Prozess nicht so deutlich hervor.
Der größte Wert ist 0.81, der kleinste 0.26. Das entspricht einer durchschnittlichen Cluster-Größe von 1.23 bis 3.85. Das heißt, in einem Cluster liegt die Anzahl extremer Renditen zwischen eins und vier.
Aber auch hier ist zu beachten, dass bei den Monats- und Jahresblöcken n bzw. m sehr klein sind. Dadurch entsteht ein sehr kleiner extremaler Index, wenn man den Threshold so setzt, dass es 200 Überschreitungen gibt, aber insgesamt nur 48 Blöcke.

(m,n)	N_u	200	100	50	40	30	25	20	15
Monat	K_u	118	67	37	31	23	19	16	12
(576,21)	$\hat{\theta}$	0.65	0.71	0.76	0.80	0.78	0.77	0.81	0.81
Quartal	K_u	76	47	28	25	19	16	14	11
(192,63)	$\hat{\theta}$	0.48	0.54	0.60	0.67	0.67	0.67	0.73	0.75
Halbjahr	K_u	49	34	25	23	17	14	13	10
(96,126)	$\hat{\theta}$	0.34	0.42	0.58	0.66	0.62	0.60	0.70	0.70
Jahr	K_u	32	26	21	19	14	11	11	10
(48,252)	$\hat{\theta}$	0.26	0.37	0.55	0.60	0.55	0.50	0.62	0.74

Tabelle 3.4: Der geschätzte extremale Index für die Log-Renditen bei unterschiedlichen Blockgrößen. N_u ist die Anzahl aller Überschreitungen des Thresholds und K_u ist die Anzahl der Blöcke, in denen der Threshold überschritten wird.

In Tabelle 3.5 stehen die Durchschnittswerte des extremalen Indexes für die verschiedenen Blockgrößen und auch hier ist noch einmal der große Unterschied zwischen den einzelnen Blockgrößen zu erkennen.
Die Aufteilung in Halbjahresblöcke ist bei dem ARCH-Prozess am besten, deswegen wird im Weiteren davon ausgegangen, dass die Renditen des ARCH-Prozesses einen extremalen Index von 0.68 haben. Für die Log-Renditen des Dow Jones Index wird demzufolge ein extremaler Index von 0.58 verwendet. Also nehme ich auch hier das Ergebnis der Halbjahresblöcke. Zusätzlich zum durchschnittlichen extremalen Index sind auch die durchschnittlichen Cluster-Größen in der Tabelle zu finden. Bei den Log-Renditen beträgt die Cluster-Größe zwischen 1.32 und 1.89, beim ARCH-Prozess sind sie nicht ganz so groß, denn sie liegen zwischen 1.33 und 1.52.

		Monat	Quartal	Halbjahr	Jahr
ARCH-Prozess	$\hat{\theta}$	0.75	0.72	0.68	0.66
	$\hat{\theta}^{-1}$	1.33	1.39	1.47	1.52
Log-Renditen	$\hat{\theta}$	0.76	0.64	0.58	0.53
	$\hat{\theta}^{-1}$	1.32	1.56	1.72	1.89

Tabelle 3.5: Der durchschnittliche extremale Index und die durchschnittliche Cluster-Größe der Log-Renditen und des ARCH-Prozesses.

Wie schon festgestellt, ändert sich der extremale Index zum Teil erheblich, wenn der Threshold geändert wird. Diese Veränderungen sind in Abbildung 3.4 zu sehen. Dort wurde der extremale Index für die Halbjahresblöcke geplottet, einmal für die Log-Renditen und einmal für den ARCH-Prozess. Für die Log-Renditen gilt, je kleiner der Threshold desto kleiner der extremale Index. Dagegen streuen die Werte für θ beim ARCH-Prozess nicht so stark wie bei den Log-Renditen.
Zusammenfassend lässt sich sagen, dass bei der Bestimmung des extremalen Indexes immer ein gewisser Spielraum bleibt. Den „richtigen Wert“ gibt es nicht.

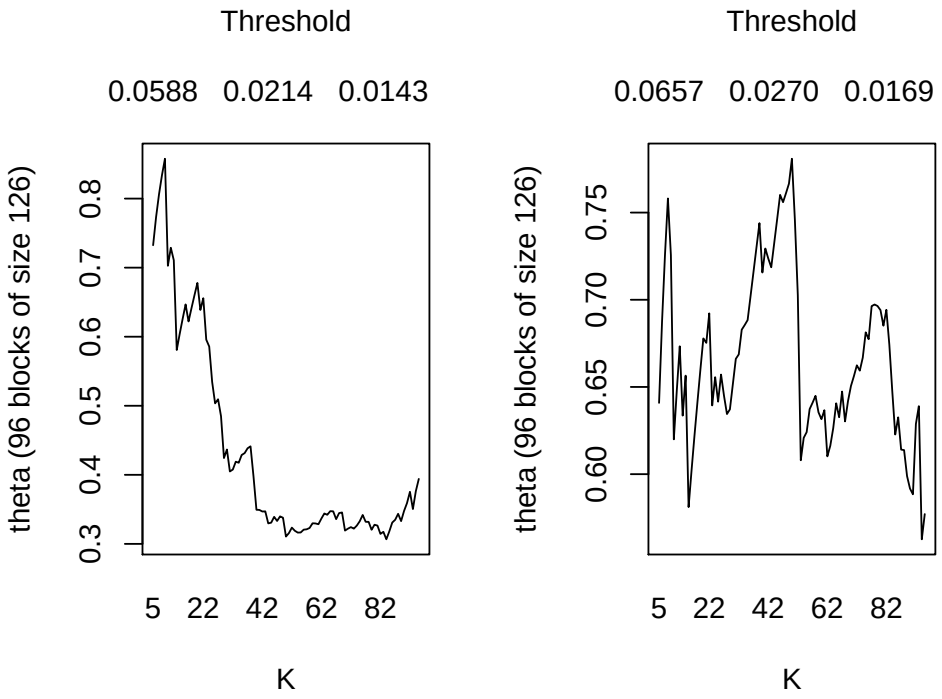

Abbildung 3.4: Veränderung des extremalen Indexes bei absteigendem Threshold für die Aufteilung nach Halbjahren (links: Log-Renditen, rechts: ARCH-Prozess).

3.4.3 GEV-Schätzung

Nachdem der extremale Index, sowohl für den ARCH-Prozess als auch für die Daten des Dow Jones Indexes, geschätzt wurde, soll nun die GEV-Verteilung an die Daten angepasst werden. Denn nach Theorem 3.3 in Abschnitt 3.3.1 ist bekannt, dass auch die skalierten Maxima gegen eine GEV-Verteilung konvergieren, wenn die zugrunde liegenden (X_n) strikt stationär sind.
Die Aufteilung der Blockgrößen erfolgt wie in Kapitel 2, es werden also wieder Jahres-, Halbjahres-, Quartals- und Monatsmaxima betrachtet.

Als erstes wird die GEV-Schätzung für den ARCH-Prozess durchgeführt. Dazu nehme ich an, dass die einzelnen Blöcke der Maxima groß genug sind, damit die Maxima als unabhängig angesehen werden können. Dann kann ich eine Maximum-Likelihood Schätzung durchführen, wie in Kapitel 2.
In der Tabelle 3.6 stehen die Ergebnisse der Schätzung, dabei wurde der extremale Index noch nicht weiter beachtet.

Der geschätzte Shape-Parameter ist für alle vier Blockgrößen größer als Null, also kann davon ausgegangen werden, dass die zugrundeliegende Marginal-Verteilung heavy-tailed ist. Allerdings gibt es zwischen den Shape-Parametern der einzelnen Blockgrößen große Unterschiede.

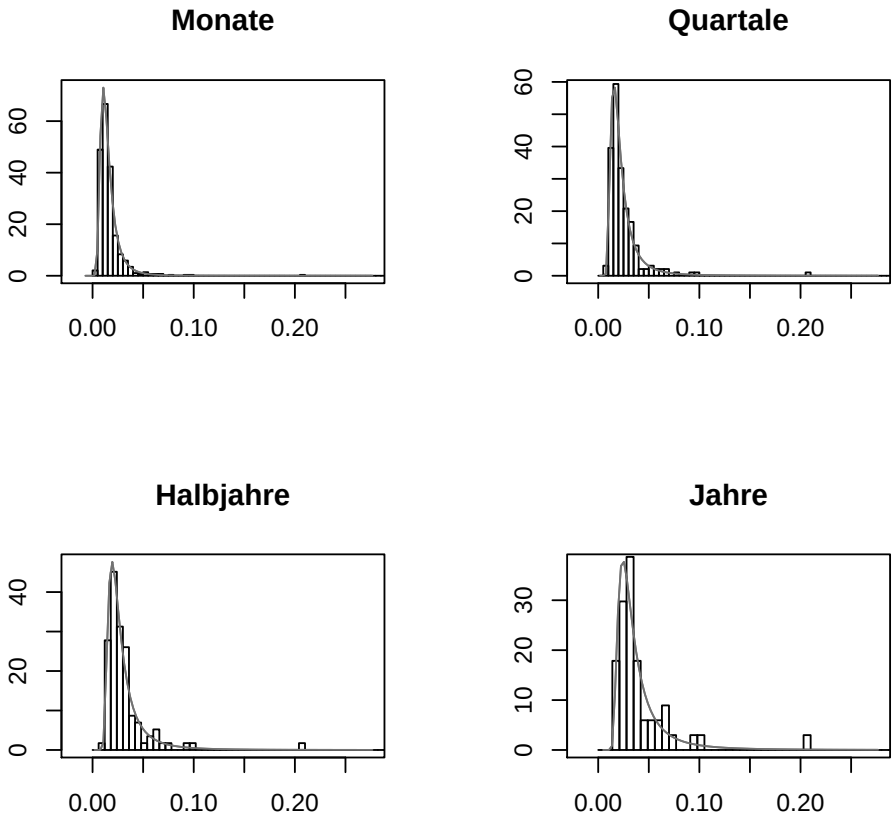

Abbildung 3.5: GEV-Anpassung an die Maxima des ARCH-Prozesses, grau eingezeichnet ist die Dichte-Funktion der GEV-Verteilung mit den geschätzten Parametern.

	$\approx n$	m	ξ	μ	σ
Monat	21	576	**0.2452** (0.1868-0.3036)	**0.0116** (0.0112-0.0121)	**0.0052** (0.0050 -0.0054)
Quartal	63	192	**0.3746** (0.2401-0.5090)	**0.0174** (0.0163 -0.0184)	**0.0066** (0.0059 -0.0073)
Halbjahr	126	96	**0.3972** (0.2051 -0.5894)	**0.0218** (0.0199 -0.0236)	**0.0083** (0.0068 -0.0097)
Jahr	252	48	**0.4375** (0.1571- 0.7179)	**0.0278** (0.0244-0.0312)	**0.0105** (0.0077 -0.0133)

Tabelle 3.6: Parameter-Schätzer der GEV-Verteilung, angepasst an die Block Maxima des ARCH-Prozesses (95% Konfidenzintervalle in Klammern)

In Abbildung 3.5 ist die GEV-Verteilung mit den geschätzten Parametern in einem Histogramm eingezeichnet und es zeigt sich, dass die Anpassung gut gelungen ist.

In der Tabelle 3.2 aus Abschnitt 3.3.3 wurde ermittelt, dass der ARCH(1)-Prozess für $\alpha_1 = 0.8$ einen Shape-Parameter hat, der im Bereich von 0.315 bis 0.434 liegt. Also ist ξ ungefähr 0.3745 für $\alpha_1 = 0.8$. Im Vergleich mit den Werten aus der Tabelle 3.6 ist bei den Quartalsmaxima der geschätzte Shape-Parameter fast identisch und bei den Jahres- und Halbjahresmaxima liegt der „wahre" Wert von ξ zumindest noch im Konfidenzintervall. Nur bei den Monatsmaxima ist der geschätzte Shape-Parameter zu klein. Doch insgesamt stimmt das Modell mit den Ergebnissen aus Kapitel 3.3.3 gut überein.

Simulationsstudie

Um zu zeigen, dass die Ergebnisse aus der Tabelle 3.6 nicht „rein zufällig" zu der Theorie passen, habe ich 500 ARCH-Prozesse simuliert. Danach habe ich für jede Simulation die Maxima mittels Block Maxima Methode ausgewählt und dann die Shape-Parameter für die einzelnen Blockgrößen geschätzt.

In Abbildung 3.6 sind die geschätzten Shape-Parameter der Simulationen eingezeichnet. Außerdem ist der „wahre" Wert des Shape-Parameters von 0.3745 grau markiert, der mit Hilfe von der Tabelle 3.2 aus Abschnitt 3.3.3 ermittelt wurde. So ist zu erkennen, dass die Monatsmaxima in den meisten Fällen einen zu niedrigen Wert liefern. Denn der durchschnittliche Shape-Parameter liegt bei 0.2982 (siehe Tabelle 3.7). Dagegen „passen" die Ergebnisse der Jahresmaxima besser, allerdings haben sie eine hohe Standardabweichung. So gibt es sogar negative Shape-Parameter. Die besten Ergebnisse liefern die Quartals- und Halbjahresmaxima. Hier stimmen die Erwartunswerte fast mit dem „wahren" Wert überein und auch die Standardabweichung und der mittlere quadratische Fehler sind bei beiden Blockgrößen recht klein.

	Monat	Quartal	Halbjahr	Jahr
(n,m)	(21,576)	(63,192)	(126,96)	(252,48)
Erwartungswert	0.2982	0.3569	0.3661	0.3664
Standardabweichung	0.0427	0.0664	0.1002	0.1502
Root mean squared error	0.0874	0.0686	0.1004	0.1503

Tabelle 3.7: Auswertung der 500 Simulationen des ARCH-Prozesses. Für jeden Prozess wurde der Shape-Parameter geschätzt.

Insgesamt ist zu sagen, dass die Ergebnisse aus der GEV-Schätzung in Tabelle 3.6 nicht besonders zufällig sind, sondern einem normalen ARCH-Prozess entsprechen.

Da die Zufallsvariablen nicht unabhängig sind, müssen die Loc- und Scale-Parameter aus Tabelle 3.6 noch mit dem extremalen Index geändert werden, wie es im Theorem 3.3 in Abschnitt 3.3.1 beschrieben wurde. Der extremale Index wurde im vorherigen Abschnitt geschätzt und beträgt 0.68. Allerdings hat der extremale Index keine

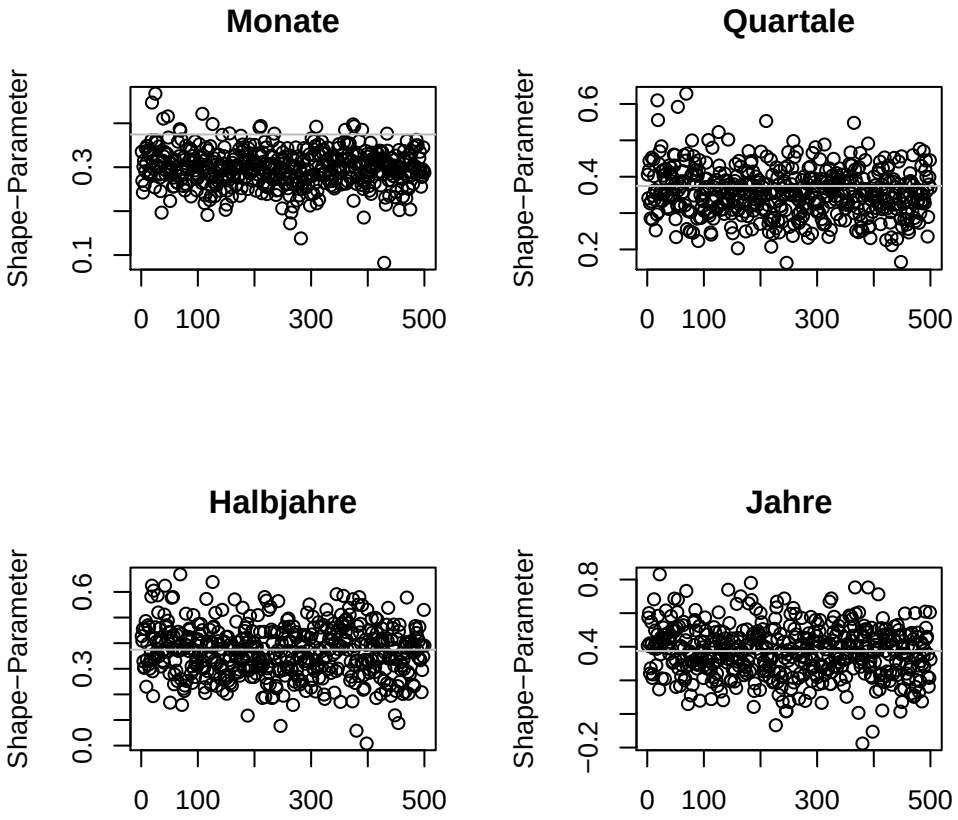

Abbildung 3.6: Geschätzte Shape-Parameter von 500 Simulationen des ARCH-Prozesses. Grau eingezeichnet der Shape-Parameter mit dem Wert 0.3745.

Auswirkungen auf den Shape-Parameter, also auf das Tail-Verhalten der zugrundeliegenden Marginal-Verteilung. Die Änderungen mittles θ für den ARCH-Prozess sind in der Tabelle 3.8 zu sehen.
Unter der Berücksichtigung des extremalen Indexes verkleinern sich die Loc- und Scale-Parameter. Wird dann die Dichtefunktion betrachtet, wie in Abbildung 3.5 dargestellt, dann verschiebt sich die Dichtefunktion etwas nach links und wird mehr gestaucht, das heißt sie wird „spitziger“.

Wie im Anschluss zu Theorem 3.3 erklärt, ist die Konvergenzgeschwindigkeit im stationären Fall langsamer als im unabhängigen. Deswegen ist es wichtig, dass die einzelnen Blöcke groß genug sind. Dies führt bei den Monatsmaxima zu Problemen, denn die tatsächliche Blockgröße liegt dann nicht bei 21, sondern bei $21 * 0.68 \approx 14$ und das ist schon sehr klein. Demzufolge ist diese Blockgröße mit Vorsicht zu behandeln.

	$\approx n$	m	θ	ξ	μ	σ
Monat	21	576	0.68	0.2452	0.0097	0.0047
Quartal	63	192	0.68	0.3746	0.0150	0.0057
Halbjahr	126	96	0.68	0.3972	0.0188	0.0071
Jahr	252	48	0.68	0.4375	0.0241	0.0089

Tabelle 3.8: Parameter der GEV-Schätzung des ARCH-Prozesses, die mit dem extremalen Index modifiziert wurden.

Die GEV-Schätzung für die Log-Renditen wurde schon in Kapitel 2 durchgeführt. Nun werden aber auch hier die Loc- und Scale-Parameter mit dem extremalen Index geändert. Im vorherigen Abschnitt wurde ein Wert von 0.58 für den extremalen Index berechnet. Die Ergebnisse für die „neuen“ Parameter finden sich in Tabelle 3.9. Genau wie beim ARCH-Prozess verkleinern sich auch hier die Loc- und Scale-Parameter.

	$\approx n$	m	θ	ξ	μ	σ
Monat	21	576	0.58	0.1877	0.0088	0.0050
Quartal	63	192	0.58	0.3106	0.0126	0.0053
Halbjahr	126	96	0.58	0.4731	0.0153	0.0051
Jahr	252	48	0.58	0.5024	0.0184	0.0069

Tabelle 3.9: Parameter der GEV-Schätzung der Log-Renditen, die mit dem extremalen Index modifiziert wurden.

In der Abbildung 3.7 sind die Maxima der Log-Renditen[2] dargestellt. Zusätzlich wurden dazu die GEV-Dichtefunktionen eingezeichnet. Die schwarz eingezeichnete Funktion ist die Funktion, die sich mit den Parametern des ARCH-Prozesses ergibt (siehe Tabelle 3.8). Die Dichtefunktion mit den geänderten Parametern der Log-Renditen aus der Tabelle 3.9 ist gestrichelt gekennzeichnet. Insgesamt passen sich die Dichtefunktionen der Log-Renditen, als auch die Dichtefunktionen des ARCH-Prozesses, gut an die Maxima der Log-Renditen an. Nur bei den Halbjahres- und Jahresmaxima gibt es kleine Unterschiede. Es lässt sich aber feststellen, dass die Maxima der Log-Renditen auch durch eine GEV-Verteilung eines simulierten ARCH-Prozesses angepasst werden können.

3.4.4 Return-Level und Return-Periode

Nachdem für jede Blockgröße eine GEV-Verteilung vorliegt, die die Maxima gut beschreibt, ist der nächste Schritt das Return-Level zu berechnen. Die besondere Aufmerksamkeit liegt vor allem darauf, Quantile aller Log-Renditen zu betrachten

[2]Um die Grafik etwas übersichtlicher zu gestalten, wurde das Fenster verkleinert. Dies hat zur Folge, dass der größte Wert der Maxima (das Extremereignis von 1987) nicht mehr abgebildet ist.

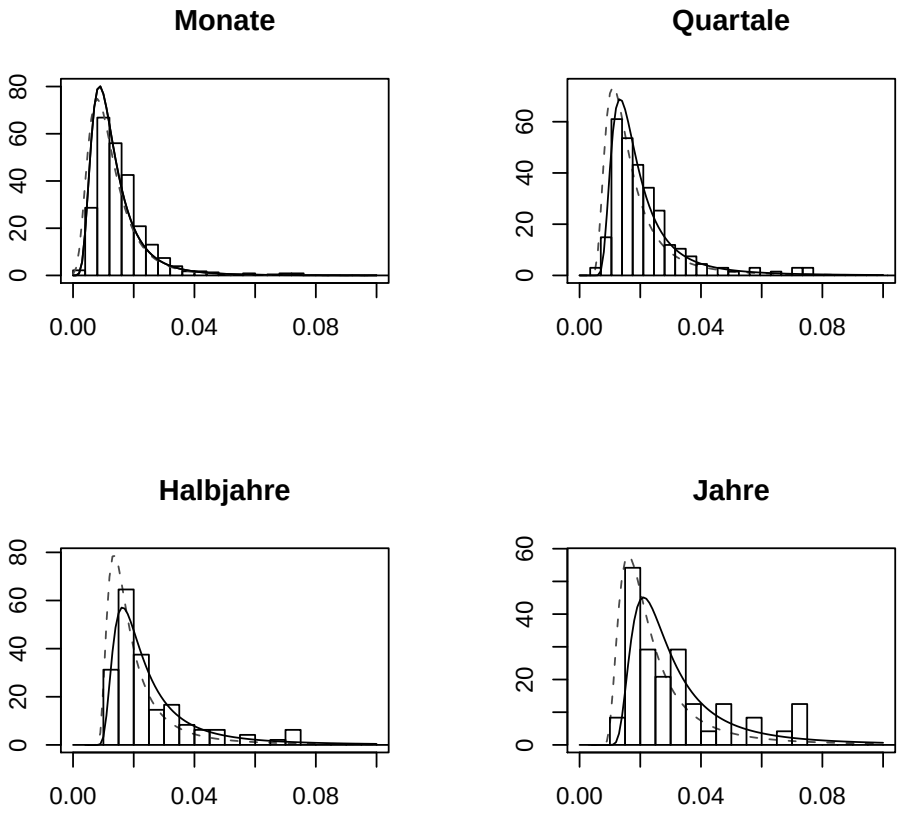

Abbildung 3.7: In den Histogrammen der Maxima der Log-Renditen wurden die Dichtefunktionen der Log-Renditen (gestrichelt) und die Dichtefunktionen des ARCH-Prozesses (schwarz) eingezeichnet.

und nicht nur die Maxima. Außerdem stellt sich die Frage, wie sich das Return-Level verändert, wenn der extremale Index mit berücksichtigt wird. Das heißt zu welchen Schlussfolgerungen gelangt man, wenn keine Unabhängigkeit mehr vorliegt? Zum Schluss wird noch kurz auf die Return-Periode eingegangen.

Ein k-Jahr Return-Level ist eine Schwelle, welche durchschnittlich nur in einem n-Block im Zeitraum von k Jahren überschritten wird. In diesem Abschnitt ist $k \in \{5, 10, 25, 50\}$.
Zunächst wird das Return-Level für den ARCH-Prozess berechnet, dabei wird der extremale Index noch nicht berücksichtigt. Das Return-Level und die dazugehörigen 95% Konfidenzintervalle werden somit wie in Kapitel 2 berechnet. Die Ergebnisse stehen in der Tabelle 3.10.

Im Vergleich mit den Ergebnissen der Log-Renditen aus Kapitel 2, ist das Return-Level für den ARCH-Prozess immer größer, auch bei den Konfidenzintervallen. Besonders deutlich wird das bei den Quartalsmaxima, aber hier ist auch der Shape-

	5 Jahre	10 Jahre	25 Jahre	50 Jahre
Monat	**0.0480** (0.0430-0.0548)	**0.0588** (0.0514-0.0691)	**0.0760** (0.0643-0.0929)	**0.0919** (0.0758-0.1158)
Quartal	**0.0531** (0.0451-0.0662)	**0.0693** (0.0562-0.0931)	**0.0980** (0.0739-0.1457)	**0.1272** (0.0897-0.2025)
Halb-jahr	**0.0518** (0.0437-0.0662)	**0.0687** (0.0550-0.0973)	**0.0990** (0.0725-0.1633)	**0.1304** (0.0882-0.2562)
Jahr	**0.0501** (0.0419-0.0651)	**0.0680** (0.0539-0.1027)	**0.1011** (0.0724-0.1932)	**0.1361** (0.0869-0.3177)

Tabelle 3.10: Return-Level des ARCH-Prozesses für die verschiedenen Blockgrößen, 95% Konfidenzintervalle in Klammern.

Parameter größer als bei den Log-Renditen. Außer für $k = 50$ sind die Intervalle der Jahres- und Halbjahresmaxima etwas kleiner. Insgesamt werden auch hier Kursabstürze gut modelliert. Denn ein Ereignis wie 1987 tritt einmal innerhalb von 50 Jahren auf, wenn die Jahres- oder Halbjahresmaxima betrachtet werden, und dass obwohl eine so große Rendite wie 1987, in den simulierten Daten des ARCH-Prozesses gar nicht auftaucht. Es ist möglich, dass die Kurse sehr tief abstürzen können, selbst dann, wenn solche Extremereignisse in der Vergangenheit noch nie stattgefunden haben. Doch auch für kleinere k ergeben sich gute Ergebnisse.
In Abbildung 3.8 sind die Halbjahresmaxima des ARCH-Prozesses zu sehen. Zusätzlich ist das 20-Halbjahr Return-Level mit einer schwarzen Linie eingezeichnet und die entsprechenden Konfidenzintervalle sind schwarz-gestrichelt. In 20 Halbjahren soll es, nach der Definition des Return-Levels, zu einer Stress-Periode kommen, wo die Grenze in einem Halbjahresblock überschritten wird. Insgesamt kommt es beim ARCH-Prozess in 50 Jahren zu fünf Überschreitungen, wobei nur zwei Werte über die obere Grenze des Konfidenzintervalls hinausgehen. Das Return-Level ist gut geschätzt worden.

Nun werden wieder die Log-Renditen betrachtet, aber um auch für sie das Return-Level zu berechnen, wird Formel (3.33) Kapitel 3.3.4 verwendet. Dort wird das Return-Level unter Berücksichtigung des extremalen Indexes berechnet. Da die Maxima als unabhängig angesehen werden, können die Konfidenzintervalle mit der Profile Log-Likelihood Methode berechnet werden, wie das schon in Kapitel 2 beschrieben wurde.
Die Ergebnisse, sowohl für die Log-Renditen also auch für den ARCH-Prozess, stehen in der Tabelle 3.11. Allerdings wird das Return-Level für die Monatsmaxima nicht mehr berechnet, da die Blockgröße zu klein ist, wie in Abschnitt 3.4.3 erwähnt.

Wenn der extremale Index, also ein Maß für die Abhängigkeit, mit in die Berechnungen einbezogen wird, dann werden das Return-Level und die entsprechenden Konfidenzintervalle kleiner. So gibt es ein Extremereignis, bei dem die Rendite einen Wert von 10% annimmt, nicht mehr alle 10 Jahre. Um zu wissen, in welchem Zeitraum eine so große Rendite auftritt, muss für k ein noch größerer Wert gewählt

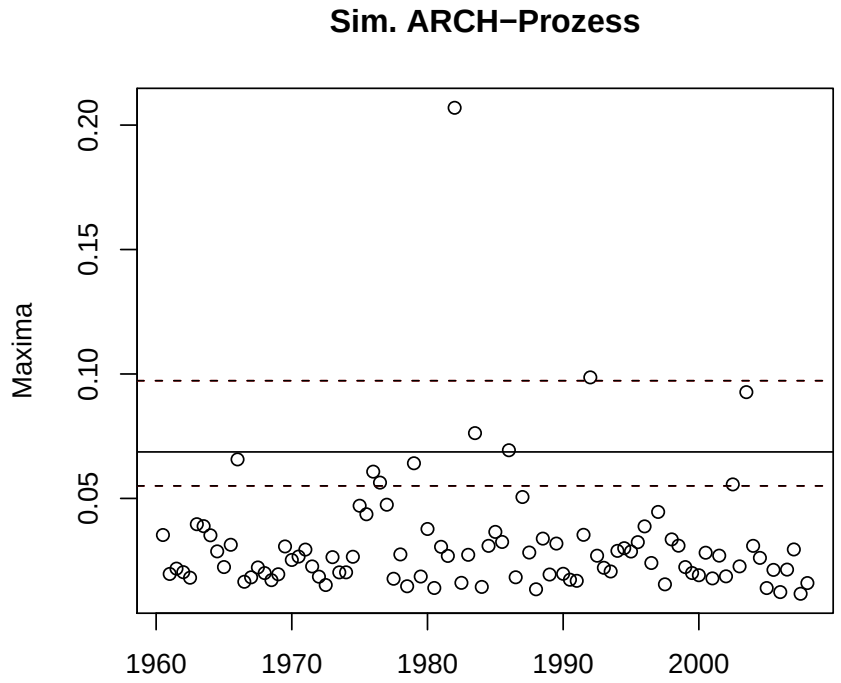

Abbildung 3.8: Zusätzlich zu den Halbjahresmaxima des ARCH-Prozesses sind das 20-Halbjahr Return-Level (schwarze Linie) und die 95% Konfidenzintervalle (schwarz-gestrichelt) zu sehen.

	Quartal		**Halbjahr**		**Jahr**	
	Log	ARCH	Log	ARCH	Log	ARCH
5	**3.85** (3.42-4.48)	**4.60** (4.00-5.52)	**3.59** (3.11-4.37)	**4.46** (3.85-5.44)	**3.39** (2.88-4.16)	**4.29** (3.66-5.27)
10	**4.90** (4.23-6.01)	**5.99** (4.99-7.71)	**4.86** (3.98-6.52)	**5.91** (4.87-7.90)	**4.73** (3.82-6.59)	**5.81** (4.75-8.06)
25	**6.68** (5.45-8.88)	**8.48** (6.59-12.08)	**7.30** (5.45-11.53)	**8.51** (6.50-13.19)	**7.33** (5.40-12.73)	**8.60** (6.38-14.85)
50	**8.40** (6.55-11.91)	**11.01** (8.06-16.92)	**9.98** (6.92-17.96)	**11.20** (7.93-19.54)	**10.24** (6.75-21.60)	**11.57** (7.86-24.17)

Tabelle 3.11: k-Jahre Return-Level für die Log-Renditen und für den ARCH-Prozess unter Berücksichtigung des extremalen Indexes, multipliziert mit 100, wobei $k \in \{5, 10, 25, 50\}$. (95% Konfidenzintervalle in Klammern)

werden.
Da nun auf die Abhängigkeitsstruktur eingegangen wird, werden die Daten genauer modelliert. Gleichzeitig ist die Berechnung des Return-Levels aber auch unsicherer, weil es nun mit vier geschätzten Parametern berechnet wird. Welchen Einfluss θ

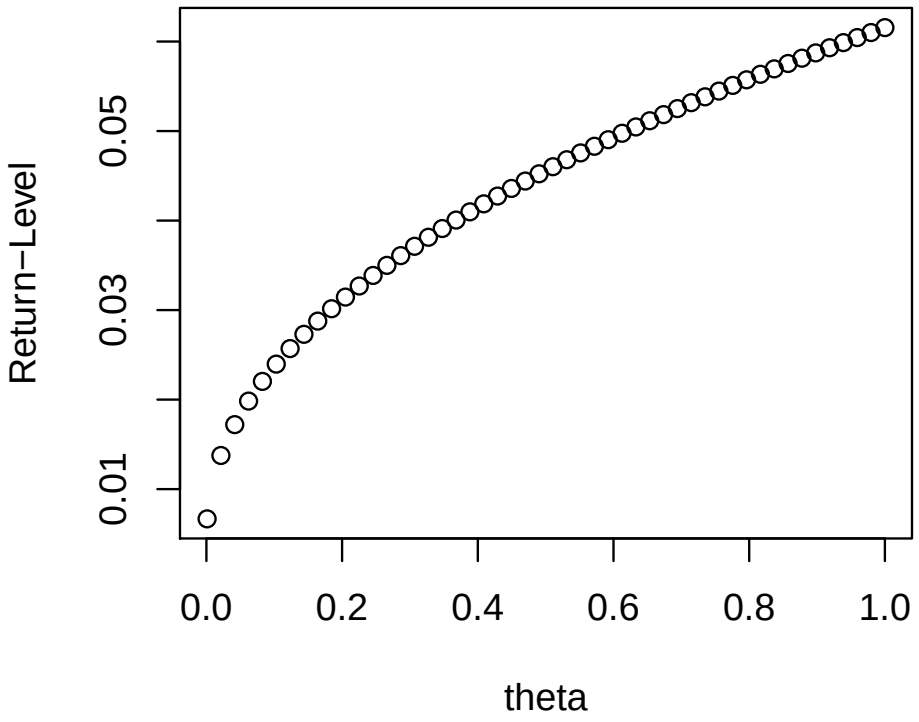

Abbildung 3.9: Veränderung des 20-Halbjahr Return-Level der Log-Renditen in Abhängigkeit von θ.

auf $R_{n,k}$ hat, wird in Abbildung 3.9 dargestellt.

Je größer der extremale Index, desto größer das Return-Level. Bei dem 20-Halbjahr Return-Level nimmt $R_{n,k}$ Werte zwischen 0.01 und 0.06 an, wenn ein kleiner bzw. ein großer extremaler Index gewählt wird. Möchte man also eine besonders vorsichtige Schätzung haben, sollte der extremale Index nicht zu klein sein. Das „konservativste“ Ergebnis erhält man somit für $\theta = 1$. Dabei sind die zugrundeliegenden Daten nicht unbedingt unabhängig, auch wenn sie einen extremalen Index von eins haben.

Wenn auch die Maxima der Log-Renditen unabhängig sind, die Log-Renditen selbst sind es nicht. Denn ein extremaler Index von 0.58 bedeutet eine Cluster-Größe von 1.72. Das heißt, extreme Renditen treten durchschnittlich in Clustern von 2 Tagen auf.
Um jetzt ein Quantil der Marginal-Verteilung der Log-Renditen zu berechnen, sollte der extreamle Index berücksichtigt werden.
Das Quantil berechnet sich mit (3.31) aus Kapitel 3.3.4. Für $k = 20$ und $n = 126$

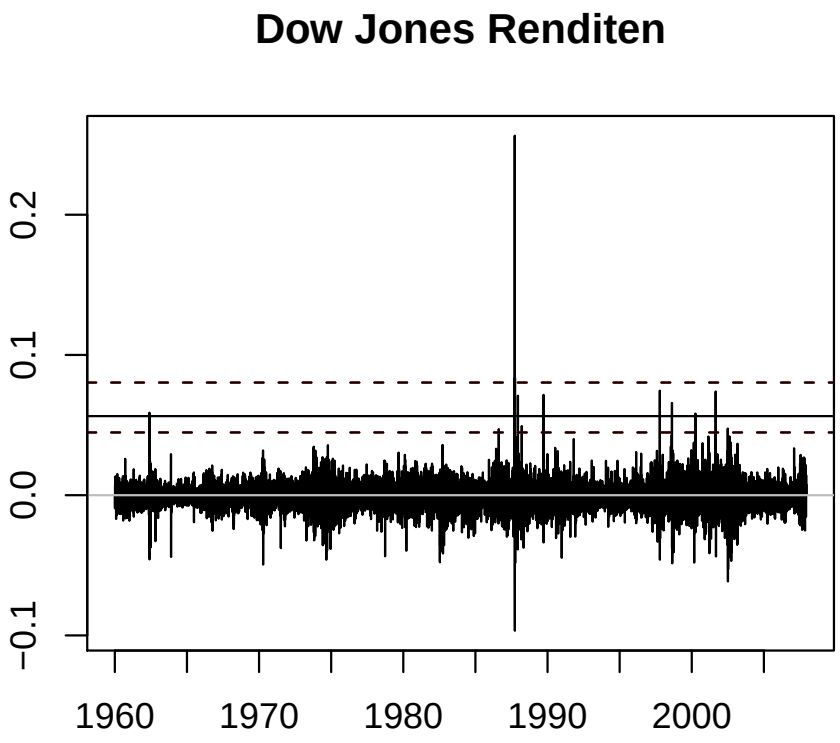

Abbildung 3.10: Die Log-Renditen des Dow Jones Indexes. Die schwarze Linie ist das 28-Halbjahr Return-Level mit den dazugehörigen 95% Konfidenzintervallen (schwarz-gestrichelt).

ergibt sich

$$F(R_{n,k}) \approx P^{1/n\theta}(M_n \leq R_{n,k}) = \left(1 - \frac{1}{k}\right)^{1/n\theta} = 0.9992984.$$

Das 20-Halbjahr Return-Level ist somit das 0.9992 Quantil der Marginal-Verteilung. Die Marginal-Wahrscheinlichkeit, das Return-Level zu überschreiten, beträgt 0.0007.

Aus der Tabelle 3.11 geht hervor, dass man, bei der Betrachtung der Halbjahresmaxima, in 10 Jahren eine Stress-Periode erwarten kann, in der die Renditen einen Wert von fast 7% annehmen können. In dieser Stress-Periode treten dann Cluster mit einer durchschnittlichen Größe von 1.7 auf. Da es in zehn Jahren ungefähr $252 * 10 = 2520$ Handelstage gibt, wird sich an $0.0007 * 2520 = 1.764$ Tagen eine größere Rendite ereignen als 0.07.

In Kapitel 2.3.2 wurde ebenfalls das Quantil berechnet und dort ist das 20-Halbjahr Return-Level sogar das 0.9995 Quantil der Marginal-Verteilung. Da nun der extremale Index berücksichtigt wird, ist hier das 0.9995 Quantil das 28-Halbjahr Return-

Level. Dieses Return-Level hat einen Wert von 0.0564 mit 95%-Konfidenzbereich von (0.0447-0.080).
Das 28-Halbjahr Return-Level ist in Abbildung 3.10 eingezeichnet. Insgesamt gibt es 13 Überschreitungen der unteren Grenze des Konfidenzintervalls. Das Return-Level selbst wird neun Mal überschritten und nur zwei Renditen sind größer als das Konfidenzintervall. (In der Abbildung ist das zum Teil nicht zu erkennen, da die Extremwerte sehr nah beieinander liegen.) Da die Log-Renditen 12081 Handelstage haben, liegt die erwartete Anzahl der Renditen, die das 0.9995 Quantil überschreiten, ungefähr bei $12081 * 0.0005 = 6.0405$. Da neun Renditen das Return-Level überschreiten, ist die erwartete Anzahl an Überschreitungen etwas zu niedrig. Unter Einbeziehung der Konfidenzintervalle wird das Risiko einer großen Rendite aber gut eingeschätzt, denn es treten nur zwei Renditen auf, die die obere Grenze des Konfidenzintervalls überschreiten. Die untere Grenze des Konfidenzintervalls ist jedoch eindeutig zu niedrig. Dies zeigt einmal mehr, wie wichtig es ist auch die Konfidenzintervalle zu betrachten. Denn gerade die obere Grenze zeigt, dass die Renditen durchaus höhere Werte annehmen können, als es das Return-Level vermuten lässt.

Auf die gleiche Art und Weise lässt sich auch das Quantil der Marginal-Verteilung des ARCH-Prozesses berechnen. Hier ist das 20-Halbjahr Return-Level das 0.9994 Quantil der Marginal-Verteilung bei einem extremalen Index von 0.68. Das Return-Level hat einen Wert von 0.059 mit Konfidenzbereich von (0.0487-0.0790) (siehe Tabelle 3.11). Da auch der ARCH-Prozess aus 12081 Daten besteht, liegt hier die erwartete Anzahl der Renditen, die das 0.9994 Quantil überschreiten ungefähr bei $12081 * 0.0006 = 7.2486$. Auch hier wird das Return-Level öfter überschritten, denn es gibt insgesamt 11 Überschreitungen. Die untere Grenze des Konfidenzintervalls wird sogar 18 Mal überschritten. Nur für die obere Grenze passt es, denn dort gibt es vier Überschreitungen. Um ein größeres Return-Level zu erhalten, kann entweder k größer gewählt werden, oder der extremale Index sollte höher sein. Denn wenn der extremale Index höher ist, dann vergrößert sich das Return-Level und es gibt weniger Überschreitungen.

Nun wird die Frage andersherum gestellt, das heißt, in welchem Zeitraum wird eine vorher festgesetzte Schwelle überschritten. Dazu wird die Return-Periode berechnet, wobei zunächst Thresholds bestimmt werden, die als extrem angesehen werden. Die Schwellenwerte werden genauso festgesetzt wie im zweiten Kapitel.
Zunächst werden die Return-Perioden für den ARCH-Prozess mit den dazugehörigen 95% Konfidenzintervallen berechnet. Die Ergebnisse befinden sich in Tabelle 3.12, dabei wurde der extremale Index noch nicht berücksichtigt.

Ähnlich wie bei den Berechnungen des Return-Levels entsprechen auch hier die Ergebnisse fast denen der Log-Renditen. Bis auf die Ergebnisse des 0.2563 hohen Thresholds, sind die Return-Perioden aber immer etwas niedriger. So ist eine Rendite von 0.0745 ein 12-Jahr Ereignis, bei den Log-Renditen tritt eine Rendite von 0.0745 aber nur innerhalb von 15 Jahren auf. Eine extreme Rendite von 0.2563

	0.0745	0.0838	0.12	0.2563
Halbjahr	**12.2** (6.3-28.3)	**16.4** (7.8-43.0)	**40.6** (14.7-163.1)	**275.6** (66.8-3299.0)
Jahr	**12.3** (6.3-29.3)	**16.2** (7.4-44.0)	**37.3** (12.9-169.4)	**217.3** (51.0-4095.2)

Tabelle 3.12: In der ersten Zeile stehen die vorgegebenen Thresholds. Dazu werden die entsprechenden Return-Perioden des ARCH-Prozesses berechnet. Die Angabe entspricht Jahren. In Klammern stehen die 95% Konfidenzintervalle.

tritt alle 50 Jahre auf, wenn die Konfidenzintervalle betrachtet werden. Dabei ist bei der Return-Periode die untere Grenze des Konfidenzintervalles die Wichtigere, denn die obere Grenze liefert für hohe Thresholds keine brauchbaren Ergebnisse.

Mit der Formel (3.34) aus Kapitel 3.3.4 kann die Return-Periode auch unter Berücksichtigung des extremalen Indexes berechnet werden. Die Ergebnisse, sowohl für die Log-Renditen als auch für den ARCH-Prozess, stehen in Tabelle 3.13. Die Konfidenzintervalle werden mit der Profile Log-Likelihood Methode berechnet, da die Maxima unabhängig sind.

	Halbjahr		**Jahr**	
	Log	ARCH	Log	ARCH
0.0745	**26.1** (12.6-65.9)	**17.9** (9.0-41.5)	**26.5** (12.5-71.2)	**17.9** (8.8-42.8)
0.0838	**34.0** (15.2-95.5)	**24.1** (11.2-62.9)	**33.0** (14.5-99.4)	**23.6** (10.6-64.6)
0.12	**74.9** (26.6-299.4)	**59.4** (22.1-239.4)	**69.3** (23.6-323.9)	**54.6** (18.7-249.0)
0.2563	**387.7** (87.7-3555.0)	**405.2** (138.3-4851.7)	**325.4** (89.8-4716.5)	**319.3** (81.4-6048.7)

Tabelle 3.13: Return-Periode für die Log-Renditen und den ARCH-Prozess unter der Berücksichtigung des extremalen Indexes. Die Angabe entspricht Jahren.

Im Vergleich mit dem iid Fall ergeben sich hier höhere Werte, oft sind sie fast doppelt so hoch. So ist eine Rendite von 0.0745 ein 26-Jahr Ereignis. Werden die Konfidenzintervalle hinzugegzogen, dann ist eine extreme Rendite von 0.2563 „nur“ noch ein 90-Jahr Ereignis. Aber auch hier gilt, wie bei dem Return-Level, dass alle Ergebnisse auf geschätzte Parameter beruhen, einschließlich des extremalen Indexes. Somit unterliegen die Werte einer gewissen Unsicherheit.

Welchen Einfluss der extremale Index auf die Return-Periode hat, ist in Abbildung 3.11 zu erkennen.
Je größer der extremale Index, desto kleiner die Return-Periode. So kann eine Rendite von 0.0838 ein 25- oder ein 150-Jahr Ereignis sein, je nachdem wie groß θ ist. Es gilt also, genau wie beim Return-Level, wenn man eine besonders vorsichtige

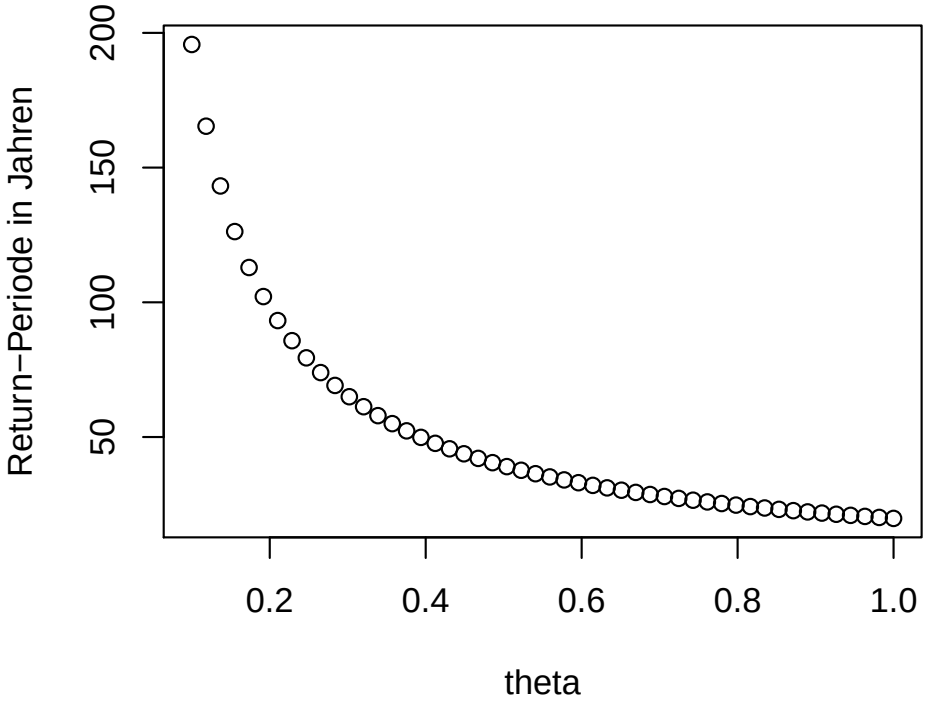

Abbildung 3.11: Veränderung der Return-Periode der Halbjahresmaxima von den Log-Renditen bei einem Threshold von 0.0838 in Abhängigkeit von θ.

Schätzung haben will, dann sollte ein hoher extremaler Index bevorzugt werden.

Auch wenn die Log-Renditen nicht unabhängig sind, können die Maxima gut durch eine GEV-Verteilung angepasst werden. Selbst dann, wenn man die Parameter der GEV-Verteilung durch einen simulierten ARCH-Prozess erhalten hat.
Durch das Return-Level und die Return-Periode kann gemessen werden, wie tief die Kurse abstürzen können und in welchen Zeiträumen Extremereignisse zu erwarten sind. Allerdings verkleinern sich Return-Level und Return-Periode, wenn der extremale Index mit in die Berechnungen einbezogen wird.
Wie groß k beim Return-Level und wie groß der Treshold u bei der Return-Periode gewählt, und welche Blockgröße beim extremalen Index bevorzugt werden, bleibt letztendlich eine Frage des Risikomanagement.

3.5 GARCH-Prozess

Im vorherigen Abschnitt wurde gezeigt, dass die Log-Renditen durch den ARCH(1)-Prozess dargestellt werden können. Man erhält auch einen hohen Shape-Parameter, wenn die GEV-Verteilung an die Maxima angepasst wird. Die Vorgehensweise richtete sich nach McNeil (vgl. [15]). Allerdings hat diese Methode einen großen Nachteil, der zu Beginn von Abschnitt 3.4 schon kurz erwähnt wurde, denn die Parameter des ARCH(1)-Prozesses wurden willkürlich festgesetzt. Der einzige Anhaltspunkt, den man bei der Festsetzung der Parameter hatte, war, dass der ARCH(1)-Prozess mehr „Ausreißer“ produziert, je größer α_1 ist.
In diesem Abschnitt wird nun ein anderer Weg gewählt. Denn nun sollen die Parameter aus den vorhandenen Log-Renditen geschätzt werden, dafür ist das Modell des ARCH(1)-Prozesses mit standard-normalverteilten Innovationen aber zu einfach. Dafür müsste schon eine höhere Ordnung für den ARCH-Prozess gewählt werden. Dies hätte aber zur Folge, dass unter Umständen sehr viele Parameter geschätzt werden müssen. Deswegen wird im Folgenden ein GARCH(1,1)-Prozess an die Log-Renditen angepasst, denn mit diesem Prozess bekommt man eine gute Datenanpassung schon mit wenigen Parametern. Doch zuvor werden kurz die wichtigsten Eigenschaften eines GARCH-Prozesses vorgestellt.

GARCH(1,1)-Prozess

Die nachfolgende Definition beschreibt einen GARCH(p,q)-Prozess. Danach wird aber nur noch ein GARCH(1,1)-Prozess betrachtet. Ausführliche Darstellungen des GARCH(p,q)-Prozesses findet sich vor allem in [17, S. 145ff], das auch hier als Literaturquelle dient.

Definition 3.12.
Sei $(Z_t)_{t\in\mathbb{Z}}$ ein SWN(0,1). Der Prozess $(X_t)_{t\in\mathbb{Z}}$ ist ein GARCH(p,q)-Prozess *(Generalized Autoregressive Conditional Heteroskedasticity), wenn er strikt stationär ist und er für alle $t \in \mathbb{Z}$ und positive $(\sigma_t)_{t\in\mathbb{Z}}$, den folgenden Gleichungen genügt*

$$\begin{aligned} X_t &= \sigma_t Z_t && \text{für alle } t \in \mathbb{Z}, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 && \text{für alle } t \in \mathbb{Z}, \end{aligned} \tag{3.35}$$

wobei $\alpha_0 > 0$, $\alpha_i \geq 0$ für $i = 1, \ldots, p$ und $\beta_j \geq 0$, $j = 1, \ldots, q$.

Der Unterschied zwischen ARCH- und GARCH-Prozess ist, dass die quadrierte Volatilität des GARCH-Prozesses von den vorherigen Volatilitäten und von den vorherigen quadrierten Werten des Prozesses abhängt. Daher wird $|X_t|$ groß, wenn entweder $|X_{t-i}|$ oder $|\sigma_{t-j}|$ groß ist.

Die Eigenschaften eines GARCH(1,1)-Prozesses können aus den Eigenschaften eines ARCH(1)-Prozesses hergeleitet werden. So ist der GARCH(1,1)-Prozess strikt stationär, wenn die Bedingung $E(\log(\alpha_1 Z_t^2 + \beta)) < 0$ erfüllt ist (vgl. Abschnitt 3.3.2). Außerdem gilt folgende Äquivalenz:
Der GARCH(1,1)-Prozess ist schwach-stationär genau dann, wenn $\alpha_1 + \beta < 1$ ist. Die Varianz des schwach-stationären Prozesses ist dann durch $\alpha_0/(1 - \alpha_1 + \beta)$ gegeben.
Genau wie beim ARCH-Prozess treffen auch auf den GARCH-Prozess die stilisierten Fakten zu, deswegen ist der GARCH(1,1)-Prozess auch unkorreliert. Auf die Berechnung der Kurtosis wird an dieser Stelle verzichtet. Aber die Kurtosis eines GARCH(1,1)-Prozesses ist größer als die einer Normalverteilung, auch dann wenn angenommen wird, dass die Innovationen standard-normalverteilt sind.
Der GARCH(1,1)-Prozess hat, genau wie der ARCH(1)-Prozess, heavy Tails. Deswegen gilt auch für den GARCH(1,1)-Prozess ein ähnliches Resultat, wie Theorem 3.4 in Kapitel 3.3.3 (siehe dazu [4]).

Parameter-Schätzung

Nach dieser kurzen Einführung wird ein GARCH(1,1)-Prozess für die Log-Renditen des Dow Jones Indexes von 1960 bis 2007 angepasst. Dazu werden die Parameter des GARCH-Prozesses geschätzt. Für diese Schätzung wird die Maximum-Likelihood Methode verwendet. Dazu wird angenommen, dass die Innovationen iid und Student-t-verteilt sind. Denn die t-Verteilung hat breitere Flanken als die Normalverteilung und diese Tatsache spiegelt sich nachher auch im GARCH(1,1)-Prozess wider.
Nähere Informationen, wie die Maximum-Likelihood Methode angewendet wird, findet sich in [17, S. 150ff] und [21, S. 180].

	Schätzer	Standardfehler	Signifikanzcode
α_0	5.975×10^{-7}	1.052×10^{-7}	1.34×10^{-8} $***$
α_1	0.05538	0.004351	2×10^{-16} $***$
β	0.9380	0.004642	2×10^{-16} $***$
Freiheitsgrad	7.995	0.5200	2×10^{-16} $***$

Tabelle 3.14: Ergebnisse der Maximum-Likelihood Schätzung unter der Annahme von t-verteilten Innovationen.

In Tabelle 3.14 stehen die Ergebnisse der Maximum-Likelihood Schätzung für den GARCH(1,1)-Prozess mit bedingter t-Verteilung. Insgesamt ist die Schätzung gut gelungen, denn die Standardfehler sind für alle Schätzer sehr klein, vor allem für $\hat{\beta}$. Außerdem ist der Signifikanzcode so niedrig, dass R drei Sterne ausgibt, was die beste Bewertung ist. Der Signifikanzcode gibt an, dass alle geschätzten Parameter signifikant in das Modell eingehen.
Der geschätzte Freiheitsgrad der t-Verteilung ist 7.995. Mit $\alpha_1 = 0.05538$ und

$\beta = 0.9380$ ergibt sich $0.05538 + 0.9380 = 0.99338 < 1$. Damit ist ein GARCH(1,1)-Prozess mit diesen Parametern schwach stationär und hat eine endliche Varianz.

Überprüfung der Anpassung

Nachdem die Parameter des GARCH(1,1)-Prozesses geschätzt wurden, wird nun überprüft, wie gut das Modell die Log-Renditen tatsächlich beschreibt. Dafür ist es wichtig die Residuen, das sind die „geschätzten Innovationen", genauer zu betrachten und zu untersuchen, ob diese t-verteilt sind. Die Residuen ergeben sich durch (s. [23, S. 109])

$$\hat{z}_t = \frac{x_t}{\hat{\sigma}_t},$$

wobei

$$\hat{\sigma}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1 x_{t-1}^2 + \hat{\beta}\hat{\sigma}_{t-1}^2$$

und (x_t) sind die 12081 Log-Renditen des Dow Jones Indexes.

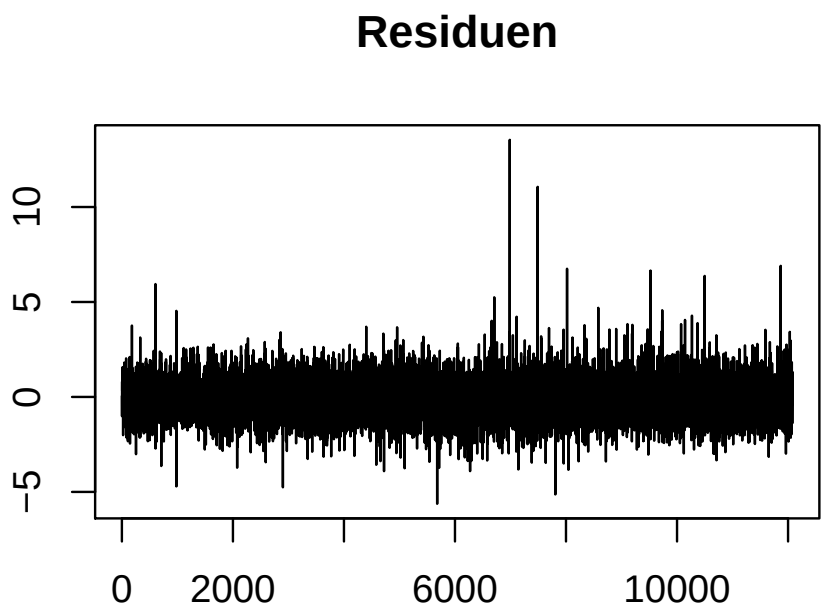

Abbildung 3.12: Die Residuen des GARCH(1,1)-Prozesses.

In Abbildung 3.12 sind die Residuen des GARCH(1,1)-Prozesses zu sehen. Es ist zu erkennen, dass sich die Extremereignisse der Log-Renditen auch in den Residuen widerspiegeln. Allerdings treten sie nicht ganz so deutlich hervor.

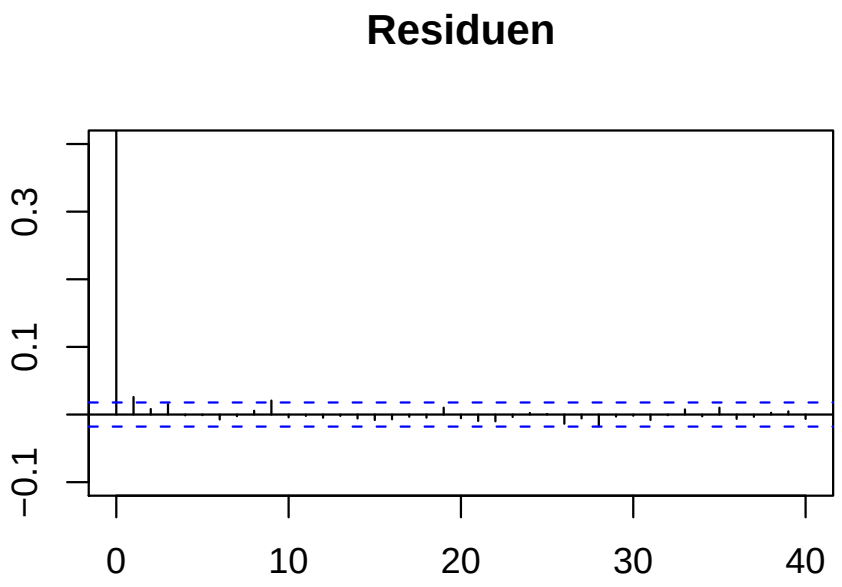

Abbildung 3.13: ACF-Plot der quadrierten Residuen, schwarz-gestrichelt ist das 95% Konfidenzintervall.

Die empirische Autokorrelationsfunktion der quadrierten Residuen ist in Abbildung 3.13 zu sehen. Es ist von unabhängigen Residuen auszugehen, da das Konfidenzintervall so gut wie gar nicht überschritten wird.

Bei der Schätzung für die Parameter des GARCH-Prozesses wurde eine skalierte t-Verteilung zugrunde gelegt. Diese Skalierung bewirkt, dass neben dem Freiheitsgrad noch ein Loc- und ein Scale-Parameter geschätzt werden müssen, denn die skalierte t-Verteilung mit υ Freiheitsgraden hat folgende Dichtefunktion

$$f_{\upsilon,\mu,\sigma}(x) = \frac{\Gamma(\frac{\upsilon+1}{2})}{\Gamma(\frac{\upsilon}{2})\sqrt{\pi\upsilon\sigma^2}}\Big(1+\frac{(x-\mu)^2}{\upsilon\sigma^2}\Big)^{-\frac{\upsilon+1}{2}}.$$

Die Parameter können mit der Maximum-Likelihood Methode geschätzt werden, da die Residuen unabhängig sind. Damit ergeben sich folgende Ergebnisse:

$$\hat{\mu} = -0.0356, \quad \hat{\sigma} = 0.8643, \quad \hat{\upsilon} = 7.9234. \tag{3.36}$$

Für den geschätzten Freiheitsgrad-Parameter $\hat{\upsilon}$ ergibt sich ein etwas kleinerer Wert als in Tabelle 3.14, aber da dort ein Standardfehler von 0.52 angegeben ist, kann dieser Unterschied ignoriert werden. In den nachfolgenden Berechnungen werden die geschätzten Parameter (3.36) verwendet.

Nun wird überprüft, ob die Residuen t-verteilt sind. Dazu werden zuerst zwei graphische Methoden benutzt. Als erstes wird die Dichtefunktion der geschätzten t-Verteilung in ein Histogramm der Residuen eingezeichnet. Das Ergebnis ist in Abbildung 3.14 in der linken Grafik zu sehen. In der rechten Grafik ist ein QQ-Plot der Residuen des GARCH(1,1)-Prozesses gegen die t-Verteilung abgebildet.

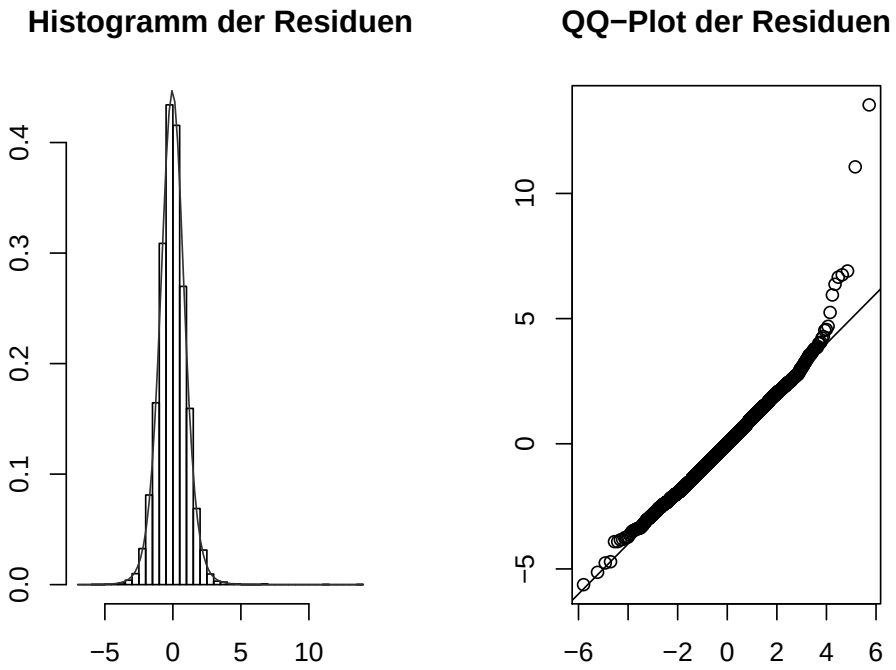

Abbildung 3.14: In der linken Grafik ist die t-Verteilung mit den geschätzten Parametern in das Histogramm der Residuen eingezeichnet (grau). Rechts ist der QQ-Plot der Residuen gegen die t-Verteilung mit den geschätzten Parametern abgebildet.

Die Anpassung im Histogramm ist gut gelungen, so dass man von t-verteilten Residuen ausgehen könnte. Im QQ-Plot kommt es im rechten Tail aber zu einer Abweichung. Um zu klären, ob diese Abweichung so gravierend ist, dass die Annahme der t-verteilten Residuen nicht mehr gerechtfertigt ist, soll nun noch ein χ^2-Anpassungstest durchgeführt werden.

Beim χ^2-Test ist die Nullhypothese, dass die Residuen t-verteilt sind. Die Teststatistik ist durch

$$\sum_{i=1}^{K} \frac{(Y_i - E_i)^2}{E_i} \sim \chi^2(K-1) \tag{3.37}$$

gegeben, wobei die Beobachtungen in K Intervalle unterteilt werden und die Inter-

valle werden mit $Y_1, \ldots, Y_K$ bezeichnet. E_i wird definiert als

$$E_i = n \int_{t_{i-1}}^{t_i} f(x)dx,$$

wobei $f(x)$ die Dichte der t-Verteilung ist.

Die Residuen wurden in sechs Intervalle unterteilt, und mit den Parametern aus (3.37) ergibt sich ein p-Wert von 0.67. Die Nullhypothese wird nicht abgelehnt, da der p-Wert größer ist als das 5%-Signifikanzniveau. Damit ist die Annahme t-verteilter Residuen gerechtfertigt.

Ein GARCH(1,1)-Prozess mit t-verteilten Innovationen beschreibt die Log-Renditen des Dow Jones Indexes gut. Dies ist auch in Abbildung 3.15 zu erkennen. Dort sind die Log-Renditen abgebildet und zusätzlich ist ein simulierter GARCH(1,1)-Prozess mit t-verteilten Innovationen über die Log-Renditen gelegt.

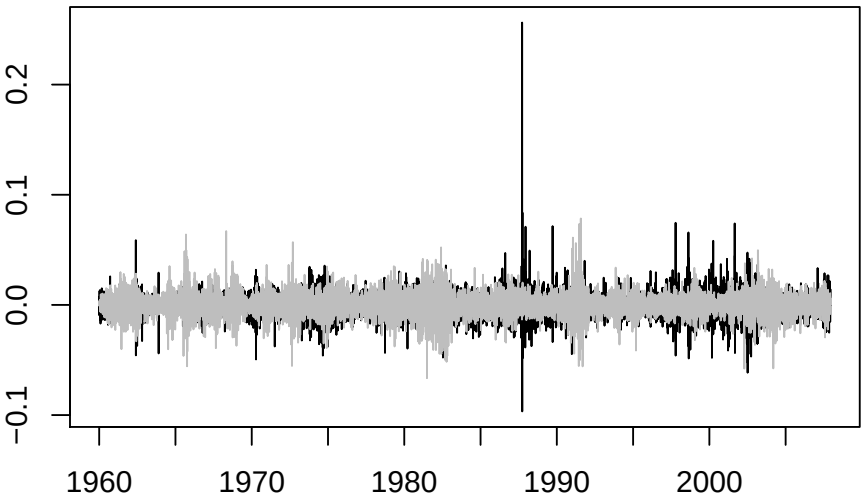

Abbildung 3.15: Über die Log-Renditen des Dow Jones Indexes von 1960 bis 2007 (schwarz) ist ein simulierter GARCH(1,1)-Prozess mit t-verteilten Innovationen gelegt (grau).

Deutlich zu erkennen ist die Clusterbildung des simulierten Prozesses. Außerdem hat auch der GARCH-Prozess einige „Ausreißer“, allerdings kein Extremereignis wie es sich 1987 beim Dow Jones Index ereignet hat.
Da die Anpassung mit einem GARCH(1,1)-Prozess gut gelungen ist, wird der simulierte GARCH-Prozess aus Abbildung 3.15 im Folgenden auf sein extremales Verhalten untersucht.

Extremaler Index

Der GARCH(1,1)-Prozess mit den geschätzten Parametern aus Tabelle 3.14 ist strikt stationär. Da nun eine GEV-Verteilung an die Maxima des GARCH-Prozesses angepasst werden soll, müssen die Parameter mit dem extremalen Index modifiziert werden. Deswegen wird zuerst der extremale Index mit der Block Methode geschätzt.

	N_u	200	100	50	40	30	25	20	15	$\emptyset$
Monat	$\hat{\theta}$	0.68	0.77	0.81	0.80	0.82	0.81	0.81	0.88	0.79
Quartal	$\hat{\theta}$	0.46	0.59	0.63	0.58	0.59	0.62	0.67	0.75	0.61
Halbjahr	$\hat{\theta}$	0.38	0.52	0.55	0.47	0.43	0.47	0.47	0.63	0.49
Jahr	$\hat{\theta}$	0.33	0.47	0.42	0.34	0.29	0.35	0.44	0.58	0.40

Tabelle 3.15: Der geschätzte extremale Index des simulierten GARCH(1,1)-Prozesses. N_u ist die Anzahl aller Überschreitungen des Thresholds, in der letzten Spalte stehen die Durchschnittswerte.

In der Tabelle 3.15 stehen die Ergebnisse der Schätzung, wobei wieder mit 15 bis 200 Überschreitungen des Thresholds gerechnet wurde. Auf die Auflistung von K_u, also die Anzahl der Blöcke, in denen der Threshold überschritten wird, wurde dieses Mal verzichtet.
Die Wahl der Blockgröße hat einen großen Einfluss auf den extremalen Index. Dieses Verhalten war auch schon bei der Berechnung des extremalen Indexes für die Log-Renditen aufgefallen. Im Vergleich mit den Ergebnissen der Log-Renditen (siehe Tabellen 3.4 und 3.5 in Kapitel 3.4.2) ergeben sich für die Monats- und Quartalsblöcke ähnliche Ergebnisse. Dagegen ist der extremale Index des GARCH-Prozesses für die Halbjahres- und Jahresblöcke etwas kleiner. Dies deutet auf eine etwas stärkere Abhängigkeit hin.

Um zu überprüfen, ob diese Ergebnisse „rein zufällig“ sind, werden 200 GARCH(1,1)-Prozesse simuliert und für jeden Prozess und jede Blockgröße wird der durchschnittliche extremale Index berechnet. Die Ergebnisse sind in Abbildung 3.16 zu sehen. Mit der schwarzen Linie eingezeichnet ist der Mittelwert der Simulation (für die Monatsblöcke ergibt sich ein Mittelwert von 0.76, für die Quartale 0.59, für die Halbjahre 0.50 und für die Jahre 0.44). Die Werte variieren zum Teil sehr stark. Aber insgesamt sind die Ergebnisse aus der Tabelle 3.15 nicht zufällig, sondern entsprechen einem „typischen“ GARCH-Prozess mit t-verteilten Innovationen.

GEV-Schätzung

Nachdem der extremale Index berechnet wurde, wird nun eine GEV-Verteilung an die Maxima des simulierten GARCH-Prozesses angepasst. Das Vorgehen gleicht dem aus Kapitel 3.4.3. Es wird angenommen, dass die Maxima unabhängig sind,

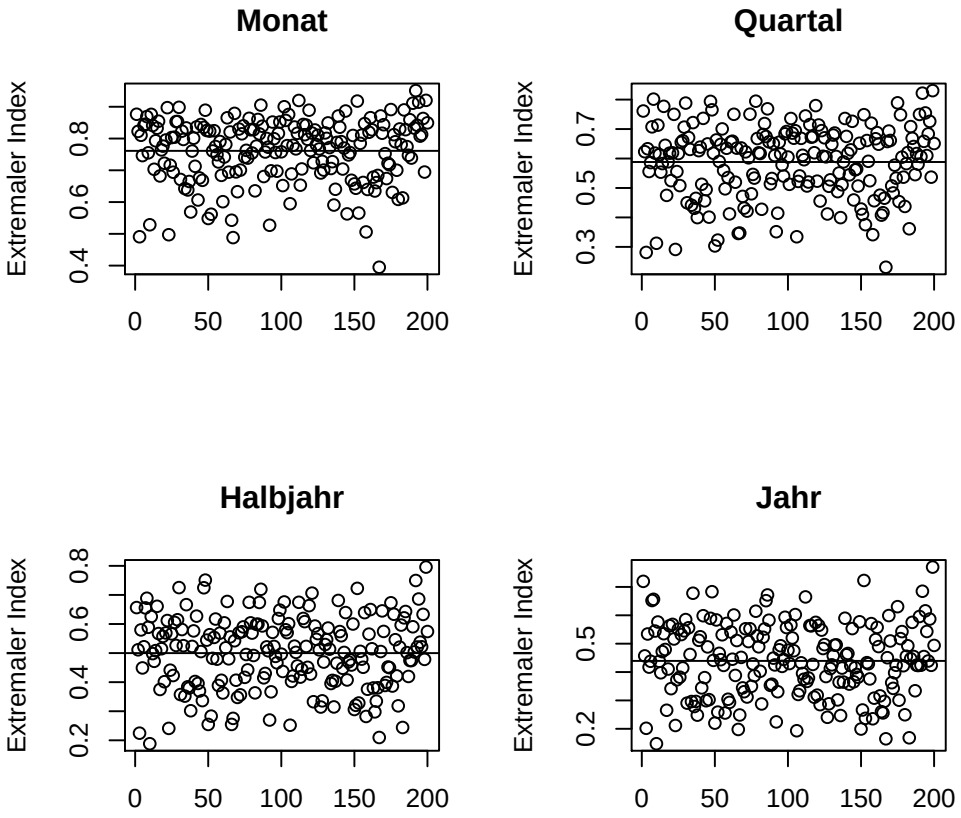

Abbildung 3.16: Von 200 Simulationen des GARCH(1,1)-Prozesses wurde der durchschnittliche extremale Index für vier verschiedene Blockgrößen berechnet. Mit der schwarzen Linie ist der Durchschnittswert aller Simulationsergebnisse eingezeichnet.

allerdings werden keine Monatsmaxima mehr betrachtet, da diese Blockgröße zu klein ist, wenn der extremale Index berücksichtigt wird.

(n,m)	Quartal (63,192)	Halbjahr (126,96)	Jahr (252,48)
ξ	**0.2033**	**0.1740**	**0.2998**
	(0.0844-0.3222)	(0.0181-0.3299)	(-0.0105-0.6100)

Tabelle 3.16: Ergebnisse der Maximum-Likelihood Schätzung für den Shape-Parameter des simulierten GARCH(1,1)-Prozesses. In Klammern stehen die 95% Konfidenzintervalle.

In der Tabelle 3.16 stehen die Ergebnisse der Schätzung für den Shape-Parameter der GEV-Verteilung. Auffällig ist, dass der Shape-Parameter für alle Blockgrößen kleiner ist als bei den Log-Renditen des Dow Jones Indexes (siehe Abschnitt 3.4.3). Bei den Jahresmaxima ist sogar ein negativer Shape-Parameter möglich. Besonders

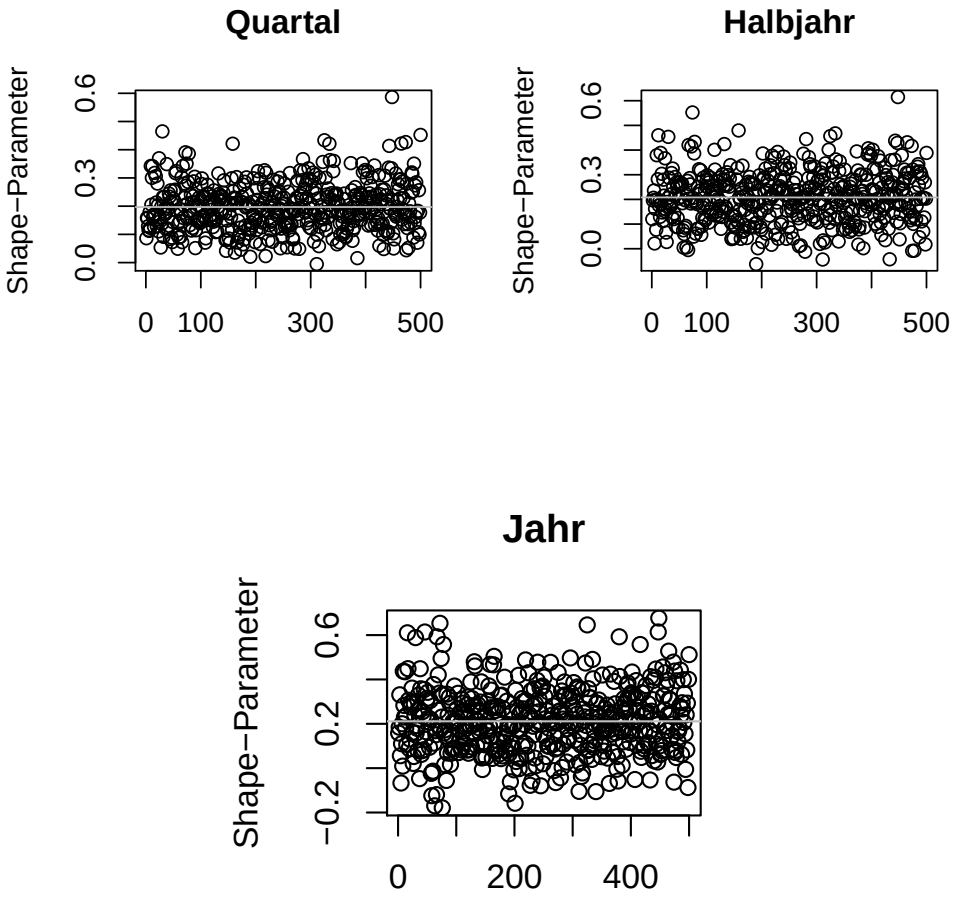

Abbildung 3.17: Geschätzte Shape-Parameter von 500 Simulationen des GARCH-Prozesses. Grau eingezeichnet ist der Mittelwert.

bemerkenswert ist aber, dass der Shape-Parameter der Halbjahresmaxima kleiner ist als der Parameter der Quartalsmaxima. Die Ursache für die kleinen Werte für ξ ergeben sich, da der simulierte GARCH(1,1)-Prozess nicht ganz so große Extremereignisse hat, wie die Log-Renditen des Dow Jones Indexes.
Dass man mit dem GARCH(1,1)-Prozess aber generell nur einen kleinen Shape-Parameter bekommt, zeigt sich auch in Abbildung 3.17. Dort ist das Ergebnis von 500 simulierten GARCH-Prozessen zu sehen, wobei für jeden Prozess ξ berechnet wurde. Der Mittelwert ist grau markiert, er beträgt für die Quartalsmaxima 0.1969, für die Halbjahresmaxima 0.2081 und für die Jahresmaxima 0.2112. Das Ergebnis aus Tabelle 3.16 entspricht somit einem normalen GARCH-Prozess, nur der Wert der Jahresmaxima ist etwas höher. Dies ist aber nicht ungewöhnlich, denn in der Abbildung sind noch deutlich höhere Werte für ξ zu erkennen.

In der Tabelle 3.17 sind alle Parameter aufgelistet, also auch die Loc- und Scale-Parameter. Diese wurden mit dem extremalen Index modifiziert. Es wurde $\theta = 0.61$ gewählt, das ist der durchschnittliche extremale Index der Quartalsblöcke. Dieser Wert wurde gewählt, da bei einer vorsichtigen Schätzung ein hoher extremaler Index zu bevorzugen ist. Die Parameter μ und σ aus der Tabelle weisen keine Besonderheiten auf.

	$\approx n$	m	θ	ξ	μ	σ
Quartal	63	192	0.61	0.2033	0.0139	0.0060
Halbjahr	126	96	0.61	0.1740	0.0168	0.0071
Jahr	252	48	0.61	0.2998	0.0211	0.0063

Tabelle 3.17: Parameter der GEV-Schätzung des GARCH-Prozesses, die mit dem extremalen Index modifiziert wurden.

Auch hier ist die Güte der Anpassung überprüft worden, indem die Dichtefunktion in ein Histogramm der Maxima eingezeichnet wurde. Auf eine ausführliche Darstellung wurde an dieser Stelle verzichtet. Aber mit den geschätzten Parametern kann jetzt das Return-Level berechnet werden.

Return-Level und Return-Periode

Das k-Jahr Return-Level wird für $k \in \{10, 25, 50\}$ berechnet. Die Ergebnisse stehen in Tabelle 3.18. Verglichen mit den Ergebnissen der Log-Renditen sind die Ergebnisse hier vor allem für $k = 50$ sehr klein. Die Ursache dafür sind die relativ kleinen Werte für ξ. Denn der Shape-Parameter hat einen großen Einfluss auf das Return-Level. Möchte man eine besonders vorsichtige Schätzung haben, sollte der Shape-Parameter nicht zu klein sein.

	10 Jahre	25 Jahre	50 Jahre
Quartal	**0.0468** (0.0412-0.0561)	**0.0597** (0.0504-0.0769)	**0.0712** (0.0577-0.0973)
Halbjahr	**0.0446** (0.0393-0.0535)	**0.0567** (0.0478-0.0742)	**0.0671** (0.0547-0.0948)
Jahr	**0.0413** (0.0358-0.0522)	**0.0548** (0.0445-0.0845)	**0.0677** (0.0516-0.1257)

Tabelle 3.18: Return-Level des GARCH-Prozesses für die verschiedenen Blockgrößen, 95% Konfidenzintervalle in Klammern.

Gleiches gilt auch für die Return-Periode, die hier für Schwellen von 0.05 bis 0.745 berechnet wurde. Hier bewirkt der Shape-Parameter, dass die Ergebnisse wesentlich größer werden. Deswegen sind in der Tabelle 3.19 auch kleinere Schwellenwerte berechnet worden als bei den vorherigen Berechnungen.

	0.05	0.06	0.0745
Halbjahr	**15.3** (8.3-31.7)	**31.5** (13.9-89.8)	**78.1** (26.6-386.4)
Jahr	**18.5** (9.5-43.8)	**33.6** (13.3-118.3)	**68.8** (19.8-468.0)

Tabelle 3.19: In der ersten Zeile stehen die vorgegebenen Thresholds. Dazu werden die entsprechenden Return-Perioden des GARCH-Prozesses berechnet. Die Angabe entspricht Jahren. In Klammern stehen die 95% Konfidenzintervalle.

Insgesamt lässt sich aber feststellen, dass der simulierte GARCH(1,1)-Prozess für kleine k-Werte bei dem Return-Level und für kleine Schwellenwerte bei der Return-Periode gute Ergebnisse liefert. Denn eine Schwelle von 0.0745 ist bei den Jahresmaxima ein 20-Jahr Ereignis, wenn das Konfidenzintervall hinzugezogen wird. Werden nun die Log-Renditen des Dow Jones Indexes betrachtet, dann tritt dort in den 60 Jahren drei Mal eine Rendite auf, die größer ist als 0.0745. Es ist also auch hier ein 20-Jahr Ereignis.

Zusammenfassung

Im Gegensatz zum ARCH(1)-Prozess wurden die Parameter des GARCH(1,1)-Prozesses mit Hilfe der vorhandenen Log-Renditen geschätzt und es zeigte sich, dass der GARCH(1,1)-Prozess mit t-verteilten Innovationen eine gute Anpassung für die Log-Renditen des Dow Jones Indexes ist. Das wurde durch graphische Methoden und dem Anpassungstest überprüft. Bei einem simulierten GARCH-Prozess kommt es auch zu einer Clusterbildung, genau wie bei den Log-Renditen. Allerdings treten im GARCH-Modell meistens nicht so große Extremereignisse auf. Dies hat zur Folge, dass der Shape-Parameter der GEV-Verteilung zu klein ist, im Gegensatz zu den Log-Renditen. Der Grund für den kleinen Shape-Parameter wird deutlich, wenn Mittelwert und Standardabweichung der Maxima betrachtet werden. So liegt der Mittelwert der Halbjahresmaxima des simulierten GARCH-Prozesses bei 0.027, die Standardabweichung bei 0.013. Die Halbjahresmaxima der Log-Renditen haben einen Mittelwert von 0.028, einen deutlichen Unterschied gibt es aber bei der Standardabweichung, die ist mit 0.027 deutlich größer als beim GARCH-Modell. Die „Ausreißer“ der Log-Renditen verursachen den hohen Shape-Parameter. Denn insgesamt sind von den 96 Halbjahresmaxima 51 Maxima des GARCH-Prozesses größer als die entsprechenden Maxima der Log-Renditen.

4 Fazit

In dieser Arbeit standen Extremereignisse im Mittelpunkt. Das Ziel war es sich damit auseinander zu setzen, wie häufig Extrema auftreten und welche Auswirkungen sie haben. Dazu wurde das Grenzverhalten von Maxima, die mit der Block Maxima Methode ermittelt wurden, untersucht. Es wurde festgestellt, dass skalierte Maxima gegen eine Extremwertverteilung konvergieren. Dabei spielt es keine Rolle, ob die zugrundeliegenden Zufallsvariablen unabhängig sind oder nicht. Denn dieses Resultat lässt sich auch auf stationäre Folgen von Zufallsvariablen übertragen. Wie in Kapitel 3.3.1 gesehen, ändert sich dabei noch nicht einmal das Tail-Verhalten der angepassten Extremwertverteilung, sondern nur die Lage und Skalierung. Diese Veränderung wird durch den extremalen Index ausgedrückt. Mit dem extremalen Index lassen sich Rückschlüsse auf die zugrundeliegenden Zufallsvariablen ziehen, denn ein hoher extremaler Index bedeutet eine geringe Abhängigkeit in den Daten. Mit den geschätzten Parametern der Extremwertverteilung lassen sich Risikomaße, wie das Return-Level und die Return-Periode, berechnen. Sie geben an, wie tief ein Kurs abstürzen kann und wie häufig so ein Kursabsturz eintritt. Wie in den Abschnitten 2.3.2 und 3.4.4 dargestellt, ist eine extreme Rendite, wie sie sich 1987 ereignet hat, kein Jahrhundertereignis, sondern kann durchaus alle 50 Jahre eintreten. Selbst mit einem Sturz von über 8%, den es auch in der Finanzmarktkrise 2008/2009 gab, kann alle 10 Jahre gerechnet werden.

Wird der extremale Index berücksichtigt und Loc- und Scale-Parameter entsprechend modifiziert, dann wird die Abhängigkeitsstruktur in die Berechnungen mit einbezogen. Dadurch werden die Daten genauer modelliert. Dies hat aber zur Folge, dass Return-Level und Return-Periode kleiner bzw. größer werden als im unabhängigen Fall. Je kleiner der extremale Index ist, desto unterschiedlicher sind die Ergebnisse. Bei einer sehr vorsichtigen Herangehensweise sollte der extremale Index nicht zu klein sein.
Doch auch unter der strengen Annahme der Unabhängigkeit kann das extremale Verhalten der Maxima gut untersucht werden und liefert sinnvolle Ergebnisse.

Die Log-Renditen haben nicht ein rein zufälliges Verhalten, sondern können durch Zeitreihen modelliert werden. Dieses hat den Vorteil, dass durch Simulation beliebig viele und beliebig lange Zeitreihen zur Verfügung stehen. Auch dann, wenn der ursprüngliche Datensatz nicht so viele Informationen liefert. Das ist vor allem bei der Anwendung der Block Maxima Methode nützlich, denn bei dieser Methode werden viele Daten benötigt.
Um zu verdeutlichen, dass die Log-Renditen durch ein Zeitreihenmodell beschrie-

ben werden können, wurde zuerst in Kapitel 3.4 ein ARCH(1)-Prozess simuliert. An die Maxima des simulierten Prozesses wurde eine GEV-Verteilung angepasst. Durch die Schätzung ergaben sich hohe Werte für den Shape-Parameter, der für das Tailverhalten wichtig ist.
Da die Parameter des ARCH-Modells festgesetzt wurden, wurde in Kapitel 3.5 der GARCH(1,1)-Prozess eingeführt. Damit war es möglich, die Parameter des GARCH-Prozesses direkt aus den vorhandenen Log-Renditen zu schätzen. Der GARCH(1,1)-Prozess mit den geschätzten Parametern und t-verteilten Innovationen lieferte eine gute Anpassung an die Log-Renditen.

Extremereignisse können mit den hier vorgestellten Risikomaßen, wie dem Return-Level und der Return-Periode, nicht mit absoluter Sicherheit vorhergesagt werden. Aber ihr Eintreten kann gut geschätzt werden. So kann auch ein „Worst-Case"-Szenario modelliert werden, dazu gibt es zwei Möglichkeiten. Entweder wird ein Zeitraum betrachtet, um dann die möglichen Renditen im ungünstigsten Fall zu berechnen. Oder es wird eine Schwelle festgesetzt und dann der Zeitraum berechnet, wann die Schwelle von den Renditen überschritten wird.
Die Beobachtungen, die für diese Berechnungen zugrunde gelegt werden, können sowohl tatsächliche Beobachtungen aus der Vergangenheit sein, oder die Werte können mittels eines ARCH- oder GARCH-Prozesses simuliert werden.

Statistikprogramm R

Alle Berechnungen, aus den Kapiteln 2.3, 3.4 und 3.5, sind mit dem Statistikprogramm R durchgeführt worden. R ist ein frei zugängliches Programm, das unter *http://www.r-project.org/* erhältlich ist.
Die Dow Jones Daten, an denen die Berechnungen veranschaulicht wurden, stammen von der Website *http://de.finance.yahoo.com/*, dort finden sich auch viele weitere Aktienindizes.
In R sind viele Pakete implementiert, die für diese Arbeit hilfreich waren. Speziell für die Extremwerttheorie zählen dazu die Pakete `evd`, `ismev`, `fExtremes` und besonders `evir`. In dem zuletzt genannten Paket befinden sich unter den Befehlen *gev*, *rlevel.gev* und *exindex* Möglichkeiten, um die GEV-Schätzung durchzuführen, das Return-Level und den extremalen Index zu berechnen. Mit *qgev*, *dgev* und *pgev* lassen sich die Quantile, die Dichten und die Wahrscheinlichkeiten der GEV-Verteilung bestimmen. Letzteres ist für die Return-Periode hilfreich.
In den Paketen `tseries` und `fGarch` stehen Befehle zu Verfügung, um die Parameter eines ARCH- oder GARCH-Prozesses zu schätzen. So ermöglicht es die Funktion *garchFit* aus dem Paket `fGarch` eine Maximum-Likelihood Schätzung mit verschiedenen bedingten Verteilungen durchzuführen.
Das Paket `stats` enthält viele wichtige Befehle für statistische Berechnungen. So kann man mit *ks.test* einen Kolmogorov-Smirnov-Test anwenden und *optim* bietet eine Möglichkeit, um Maximum-Likelihood Schätzer zu finden. Im Gegensatz zu dem Befehl *gev*, der nur für die GEV-Verteilung geeignet ist, kann hier jede beliebige Dichtefunktion berechnet werden. Außerdem beinhaltet dieses Paket mit *acf* und *plot.ecdf* Befehle, um einen ACF-Plot bzw. empirische Verteilungsfunktionen zu zeichnen. Für ausgewählte Dichtefunktionen empfiehlt sich die Funktion *fitdistr* aus dem Paket `MASS` für eine Maximum-Likelihood Schätzung.

Literaturverzeichnis

[1] Alsmeyer, Gerold: *Wahrscheinlichkeitstheorie.* Münster : Skripten zur Mathematischen Statistik, Nr. 30, 2005

[2] Beirlant, Jan: *Statistics of extremes.* Chichester : Wiley, 2004

[3] Coles, Stuart: *An Introduction to Statistical Modeling of Extreme Values.* London : Springer Verlag, 2007

[4] Davis, Richard A. ; Mikosch, Thomas: *Extreme Value Theory for GARCH Processes.* In: Andersen, Torben G. ; Davis, Richard A. ; Kreiss, Jens-Peter ; Mikosch, Thomas (Hrsg.): *Handbook of financial time series, 187-200.* Berlin : Springer Verlag, 2009

[5] deHaan, Laurens ; Resnick, Sidney I. ; Rootzén, Holger ; deVries, Casper G.: *Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes.* Stochastic Processes and their Applications 32, 213-224, 1989

[6] Drees, Holger: *Versicherungsmathematik III: Modellierung von Finanzzeitreihen.* Vorlesungsskript : Universität Hamburg, 2005

[7] Embrechts, Paul ; Klüppelberg, Claudia ; Mikosch, Thomas: *Modelling extremal events for insurance and finance.* Berlin : Springer Verlag, 2003

[8] Franke, Jürgen ; Härdle, Wolfgang ; Hafner, Christian M.: *Statistics of Financial Markets.* Berlin : Springer Verlag, 2004

[9] Gumbel, Emil J.: *Statistics of extremes.* New York : Columbia University Press, 1958

[10] Horváth, Zsuzsanna ; Johnston, Ryan: *GARCH Time Series Process.* University of Utah : working paper, o.J.

[11] Jondeau, Eric ; Poon, Ser-Huang ; Rockinger, Michael: *Financial modeling under non-Gaussian distributions.* London : Springer Verlag, 2007

[12] Leadbetter, M.R. ; Lindgren, Georg ; Rootzén, Holger: *Extremes and Related Properties of Random Sequences and Processes.* New York : Springer Verlag, 1983

[13] Löwe, Matthias: *Extremwerttheorie.* Vorlesungsskript : Universität Münster, 2008

[14] McNeil, Alexander J.: *On Extremes and Crashes.* ETH Zürich : working paper, 1997

[15] McNeil, Alexander J.: *Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory.* ETH Zürich : working paper, 1998

[16] McNeil, Alexander J.: *Extreme Value Theory for Riskmanager.* ETH Zürich : working paper, 1999

[17] McNeil, Alexander J. ; Frey, Rüdiger ; Embrechts, Paul: *Quantitative Risk Management.* Princeton : Princeton University Press, 2005

[18] Pruscha, Helmut: *Vorlesungen über Mathematische Statistik.* Stuttgart : Teubner Verlag, 2000

[19] Reiss, Rolf-Dieter ; Thomas, Michael: *Statistical Analysis of Extreme Values.* Basel : Birkhäuser Verlag, 1997

[20] Ricci, Vito: *Fitting Distributions with R.* working paper, 2005

[21] Schmid, Friedrich ; Trede, Mark: *Finanzmarktstatistik.* Berlin : Springer Verlag, 2006

[22] Smith, Richard L.: *Maximum likelihood estimation in a class of nonregular cases.* Biometrika 72, 76-90, 1985

[23] Tsay, Ruey S.: *Analysis of Financial Time Series.* New York : Wiley, 2005

[24] Zivot, Eric ; Wang, Jiahui: *Modeling Financial Time Series with S-PLUS.* New York : Spinger Verlag, 1997

Eidesstattliche Erklärung

Gemäß § 21 (6) der Diplomprüfungsordnung für den Studiengang Mathematik der Westfälischen Wilhelms-Universität Münster vom 15. Juli 1998 versichere ich, dass ich die vorliegende Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Münster, 10. November 2010 Birgit Woeste