

Technische Universität München

ZENTRUM MATHEMATIK

**Eine Alternative zur  
Korrelationsfunktion - Vergleich  
verschiedener nicht-linearer Modelle**

Diplomarbeit

von

Friederike Hoyer

Themensteller/in: Prof. Dr. Claudia Klüppelberg

Betreuer/in: Dr. Vicky Fasen

Abgabetermin: 04. Dezember 2006

Hiermit erkläre ich, dass ich die Diplomarbeit selbstständig angefertigt und nur die angegebenen Quellen verwendet habe.

Garching, den 04. Dezember 2006

Friederike Hoyer

*„All’s well that ends.“*  
Murphy’s Laws on technology

## Danksagung

An dieser Stelle möchte ich mich bei Frau Prof. Dr. Claudia Klüppelberg für die interessante Themenstellung und freundliche Unterstützung bedanken. Die Gespräche mit ihr haben stets zum besseren Verständnis sowie zu neuen Ideen beigetragen. Besonderer Dank geht an Frau Dr. Vicky Fasen für die umfassende Betreuung auch über eine sehr weite räumliche Entfernung hinweg. Robert Stelzer danke ich für seine Hilfe bei meinen Fragen zu Matlab. Meinen Eltern danke ich dafür, dass sie mir das Studium an der Technischen Universität München ermöglicht, und mich stets in meinen Vorhaben bestärkt haben. Meinem Freund Alex danke ich für sein Verständnis, dass die Diplomarbeit die letzten sechs Monate bestimmt hat.

# Inhaltsverzeichnis

Abkürzungsverzeichnis . . . . .	VI
Abbildungsverzeichnis . . . . .	VIII
Tabellenverzeichnis . . . . .	X
<b>1 Einleitung</b>	<b>1</b>
1.1 Aufbau der Arbeit . . . . .	2
1.2 Präliminarien . . . . .	3
1.2.1 Tail . . . . .	3
1.2.2 Reguläre Variation . . . . .	4
1.3 Zeitreihen . . . . .	6
<b>2 Wichtige Resultate der Extremwerttheorie</b>	<b>12</b>
2.1 Hauptsatz der Extremwerttheorie . . . . .	12
2.1.1 Parameterschätzung der GEV . . . . .	15
2.1.2 Tail- und Quantilschätzung der GEV . . . . .	16
2.1.3 Anziehungsbereiche . . . . .	17
2.2 Stationäre Daten . . . . .	17
2.3 Die verallgemeinerte Pareto Verteilung . . . . .	19
2.3.1 Parameterschätzung der GPD . . . . .	20
2.3.2 Tail- und Quantilschätzung . . . . .	21
<b>3 Extremwertstatistik</b>	<b>23</b>
3.1 Schätzung des Tailindex . . . . .	23
3.2 Bestimmung der Schwelle $\mathbf{u}$ . . . . .	26
3.3 Methoden zur Tail Schätzung . . . . .	28
3.3.1 Empirische Schätzung . . . . .	28
3.3.2 Blockmaxima-Methode . . . . .	29
3.3.3 Vorgehensweise in McNeil (1998) . . . . .	30
3.3.4 POT-Methode . . . . .	32
<b>4 Nicht-lineare Modelle</b>	<b>35</b>
4.1 Stylized Facts von Finanzzeitreihen . . . . .	35
4.1.1 Distributional Stylized Facts . . . . .	35
4.1.2 Time Series Stylized Facts . . . . .	37
4.2 Nicht-lineare Zeitreihenmodelle: diskrete Zeit . . . . .	39
4.2.1 ARCH(1) Prozess . . . . .	39

4.2.2	GARCH(1,1) Prozess . . . . .	45
4.2.3	AR(1) + ARCH(1) Prozess . . . . .	52
4.3	Extremale Maße . . . . .	57
4.3.1	Anforderungen an Extremale Maße . . . . .	57
4.3.2	Theoretische Berechnung . . . . .	58
<b>5</b>	<b>Eigene Studie</b>	<b>60</b>
5.1	Simulation . . . . .	61
5.1.1	Ausgangssituation . . . . .	61
5.1.2	Analyse der Daten . . . . .	62
5.1.3	ARCH(1) Prozess . . . . .	64
5.1.4	Beurteilung für den ARCH(1) Prozess . . . . .	74
5.1.5	GARCH(1,1) Prozess . . . . .	79
5.1.6	Beurteilung für den GARCH(1,1) Prozess . . . . .	86
5.1.7	AR(1) + ARCH(1) Prozess . . . . .	87
5.1.8	Beurteilung für den AR(1) + ARCH(1) Prozess . . . . .	94
5.2	Anwendung auf einen Datensatz . . . . .	95
<b>6</b>	<b>Zusammenfassung und Ausblick</b>	<b>106</b>
<b>A</b>	<b>Abbildungen und Tabellen</b>	<b>107</b>
A.1	GARCH(1,1) Prozess . . . . .	108
A.2	AR(1) + ARCH(1) Prozess . . . . .	111
<b>B</b>	<b>Punkt-Prozesse</b>	<b>114</b>
	<b>Literaturverzeichnis</b>	<b>117</b>

# Verzeichnis der Abkürzungen und Symbole

#	Anzahl
$\approx$	approximativ
$\xrightarrow{f.s.}$	fast sichere Konvergenz
$\stackrel{D}{=}$	gleich in Verteilung
$\mathbf{1}$	Indikatorfunktion
$\xrightarrow{D}$	Konvergenz in Verteilung
$\xrightarrow{P}$	Konvergenz in Wahrscheinlichkeit
$\sim$	verteilt oder verhält sich wie
ACF	Autocorrelationfunction (Autokorrelationsfunktion)
ALL	Allianz AG
AR(1)	Autoregressiver Prozess der Ordnung $p = 1$
AR(1) + ARCH(1)	AR(1) Prozess mit ARCH(1) Prozess als Fehlerterm
ARCH(1)	Auto-Regressive Conditional Heteroskedasticity Prozess der Ordnung $p=1$
BMW	Bayerische Motoren Werke AG
bzw.	beziehungsweise
card	Kardinalität
corr( $\cdot$ )	Korrelation
cov( $\cdot$ )	Kovarianz
Def.	Definition
$\mathbb{E}(\cdot)$	Erwartungswert
et al.	et altera (und andere)
etc.	et cetera (und weitere)
f.	folgende
ff.	fort folgende
f.s.	fast sicher
$\gamma(\cdot)$	Autokovarianzfunktion
GARCH(1,1)	Generalized Auto-Regressive Conditional Heteroskedasticity Prozess der Ordnungen $p = 1$ und $q = 1$

GEV	generalised extreme value distribution (verallgemeinerte Extremwertverteilung)
GPD	generalised pareto distribution (verallgemeinerte Pareto Verteilung)
iid	identically independently distributed (unabhängig identisch verteilt)
$\mathcal{JB}$	Jarque-Bera Test
$\kappa$	Tailindex
$\mathcal{K}$	Kurtosis (Wölbungsmaß)
$\log(\cdot)$	natürlicher Logarithmus
$\mathcal{LB}$	Ljung-Box Test
$\mathcal{LM}$	Lagrange-Multiplikator-Test von Engle (1982)
MC	Monte-Carlo
MDA	Maximum Domain of Attraction (maximaler Anziehungsbereich)
MEF	Mean Excess Function (Mittlere Exzessfunktion)
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator (Schätzer)
$\mathbb{N}$	Menge der natürlichen Zahlen
$\mathbb{N}_0$	Menge der natürlichen Zahlen mit der Null
$N(\mu, \sigma^2)$	Normalverteilung mit Erwartungswert $\mu$ und Varianz $\sigma^2$
Nr.	Nummer
o.B.d.A.	ohne Beschränkung der Allgemeinheit
$\mathbb{P}(\cdot)$	Wahrscheinlichkeit
POT	Peaks over Threshold (Schwellenwertmethode)
QQ-Plot	Quantil-Quantil Plot
$\rho(\cdot)$	Autokorrelationsfunktion
$\mathbb{R}$	Menge der reellen Zahlen
$\mathbb{S}$	Einheitssphäre
$\mathcal{S}$	Skewness (Schiefemaß)
S.	Seite
sog.	so genannt
SRE	Stochastic Recurrence Equation (stochastische Rekurrenzgleichung)
supp	Support (Träger)
$t$	Student $t$ -Verteilung
u.a.	unter anderem
$\text{var}(\cdot)$	Varianz
vgl.	vergleiche
$x_R$	rechter Endpunkt
z.B.	zum Beispiel
ZV	Zufallsvariable

# Abbildungsverzeichnis

1.1	Pfad einer Brown'schen Bewegung. . . . .	10
1.2	White Noise. . . . .	11
2.1	Dichtefunktionen der GEV. . . . .	15
2.2	Dichtefunktionen der GPD. . . . .	19
3.1	Hill-Plot für einen simulierten quadrierten ARCH(1) Prozess. . . . .	25
3.2	ME-Plot der quadrierten Allianz Log>Returns. . . . .	28
3.3	QQ-Plot der Residuen eines simulierten quadrierten ARCH(1) Prozesses. . . . .	31
3.4	Daten $X_1, \dots, X_{10}$ und die entsprechenden $Y_1, \dots, Y_4$ über der Schwelle $u$ . . . . .	33
4.1	Nichtparametrische Kerndichteschätzung der Allianz Log>Returns. . . . .	36
4.2	Vergleich der Volatilitätscluster zwischen Allianz und BMW Log>Returns. . . . .	38
4.3	Simulierte ARCH(1) Prozesse. . . . .	41
4.4	Nichtparametrische Kerndichteschätzung eines simulierten ARCH(1) Prozesses. . . . .	42
4.5	Simulierte GARCH(1,1) Prozesse. . . . .	46
4.6	Nichtparametrische Kerndichteschätzung eines simulierten GARCH(1,1) Prozesses. . . . .	47
4.7	Simulierte AR(1) + ARCH(1) Prozesse. . . . .	53
5.1	Simulierte quadrierte nicht-lineare Prozesse. . . . .	63
5.2	ACF eines simulierten ARCH(1) Prozesses. . . . .	65
5.3	$\chi(k)_{ARCH}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	66
5.4	$\chi(k)_{emp\ Tails}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	67
5.5	Verschiedene $\chi(k)_{emp\ Tails}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	68
5.6	$\chi(k)_{Block}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	70
5.7	$\chi(k)_{McNeil}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	71
5.8	GPD Plots für einen simulierten quadrierten ARCH(1) Prozess. . . . .	73
5.9	$\chi(k)_{POT}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	74
5.10	Plots der Methoden für einen simulierten quadrierten ARCH(1) Prozess. . . . .	75
5.11	Absoluter Fehler der Methoden für einen simulierten quadrierten ARCH(1) Prozess. . . . .	76



5.12	Relativer Fehler der Methoden für einen simulierten quadrierten ARCH(1) Prozess. . . . .	78
5.13	ACF eines simulierten GARCH(1,1) Prozesses. . . . .	80
5.14	Hill-Plot für einen simulierten GARCH(1,1) Prozess. . . . .	81
5.15	$\chi(k)_{GARCH}$ für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	82
5.16	Plots der Methoden für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	83
5.17	QQ-Plot der Residuen eines simulierten quadrierten GARCH(1,1) Prozesses. . . . .	84
5.18	GPD Plots für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	85
5.19	Absoluter Fehler der Methoden für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	86
5.20	Relativer Fehler der Methoden für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	87
5.21	Nichtparametrische Kerndichteschätzung eines simulierten AR(1) + ARCH(1) Prozesses. . . . .	88
5.22	ACF eines simulierten AR(1) + ARCH(1) Prozesses. . . . .	88
5.23	Hill-Plot für einen simulierten AR(1) + ARCH(1) Prozess. . . . .	89
5.24	$\chi(k)_{ARARCH}$ für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	90
5.25	Plots der Methoden für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	91
5.26	QQ-Plot der Residuen eines simulierten quadrierten AR(1) + ARCH(1) Prozesses. . . . .	92
5.27	GPD Plots für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	93
5.28	Absoluter Fehler der Methoden für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	94
5.29	Relativer Fehler der Methoden für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	95
5.30	Kurse und Log>Returns der Aktienkursdaten der Allianz AG. . . . .	96
5.31	ACF der Allianz Log>Returns. . . . .	97
5.32	QQ- und Normal Probability Plot der Allianz Log>Returns. . . . .	98
5.33	Hill-Plot der angepassten Allianz Log>Returns. . . . .	99
5.34	$\chi(k)$ mit Formel (4.15) für ARCH(1), Formel (4.16) für GARCH(1,1) und Formel (4.17) für AR(1) + ARCH(1). . . . .	100
5.35	Quadrierte Allianz Log>Returns. . . . .	100
5.36	Ergebnisse der verschiedenen Schätzmethoden für die quadrierten Allianz Log>Returns. . . . .	101
5.37	QQ-Plot der Residuen der quadrierten Allianz Log>Returns. . . . .	102
5.38	Exzess Verteilung für die quadrierten Allianz Log>Returns. . . . .	103
5.39	GPD Plots für die quadrierten Allianz Log>Returns. . . . .	104
B.1	Realisierungen eines Poisson-Prozesses. . . . .	116

# Tabellenverzeichnis

3.1	Mittlere Exzessfunktion für einige Standardverteilungen. . . . .	27
4.1	Tailindex $\kappa$ für ARCH(1) Prozesse. . . . .	44
4.2	Tailindex $\kappa$ für GARCH(1,1) Prozesse. . . . .	51
4.3	Tailindex $\kappa$ für AR(1) + ARCH(1) Prozesse. . . . .	55
4.4	Extremaler Index $\hat{\theta}$ für ARCH(1) Prozesse. . . . .	59
5.1	Deskriptive Statistiken für simulierte nicht-lineare Prozesse. . . . .	62
5.2	Mittels Monte-Carlo Simulation bestimmtes $\chi(k)_{ARCH}$ . . . . .	66
5.3	$\chi(k)_{emp\ Tail}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	67
5.4	$\chi(k)_{Block}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	69
5.5	$\chi(k)_{McNeil}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	71
5.6	$\chi(k)_{POT}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	74
5.7	Absoluter Fehler für die Berechnung von $\chi(k)_{ARCH}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	77
5.8	Relativer Fehler für die Berechnung von $\chi(k)_{ARCH}$ für einen simulierten quadrierten ARCH(1) Prozess. . . . .	79
5.9	Deskriptive Statistiken der Log>Returns der Allianz Aktie. . . . .	97
5.10	Induktive Statistiken der Log>Returns der Allianz Aktie. . . . .	98
5.11	Ergebnisse der Methoden für die quadrierten Allianz Log>Returns. . . . .	105
A.1	Ergebnisse der Methoden für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	108
A.2	Absoluter Fehler für die Berechnung von $\chi(k)_{GARCH}$ für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	109
A.3	Relativer Fehler für die Berechnung von $\chi(k)_{GARCH}$ für einen simulierten quadrierten GARCH(1,1) Prozess. . . . .	110
A.4	Ergebnisse der Methoden für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	111
A.5	Absoluter Fehler für die Berechnung von $\chi(k)_{ARARCH}$ für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	112
A.6	Relativer Fehler für die Berechnung von $\chi(k)_{ARARCH}$ für einen simulierten quadrierten AR(1) + ARCH(1) Prozess. . . . .	113

# Kapitel 1

## Einleitung

Die Kapitalmärkte geraten in die Schlagzeilen, wenn sich extreme Ereignisse einstellen, wie dies z.B. nach den Terroranschlägen am 11. September 2001 in den USA der Fall war. Vor allem in der Finanzwirtschaft soll das Risiko von Verlusten berechnet bzw. vorhergesagt werden, damit sich die Öffentlichkeit, vor allem aber die Investmenthäuser, dagegen absichern können. Zunächst wurde dazu die Brown'sche Bewegung (siehe Definition 1.17), die ihren Ursprung in der Naturwissenschaft hat, herangezogen. Mit ihrer Hilfe konnten frühe erfolgreiche Modelle für Bewegungen von Marktpreisen aufgestellt werden. Für einige Anwendungen wie z.B. die Optionsbewertung nach Black-Scholes bietet sie immer noch eine genügende Näherung. In der neueren empirischen Forschung jedoch ist die Brown'sche Bewegung nicht mehr ausreichend, da sie Risiken unterschätzt und Volatilitätsschwankungen ignoriert. Die Sprünge, die am Kapitalmarkt auftreten, können daher mit der Brown'schen Bewegung nicht ausreichend abgebildet werden. Daher mussten neue Modelle gefunden werden, die diese Schwäche nicht aufweisen bzw. minimieren. Diese Modelle werden in der Extremwerttheorie verwendet. Ein generelles Ziel der Extremwerttheorie ist das Verständnis für das Verhalten extremer Beobachtungen und die Verwendung dieser Informationen um Vorhersagen für zukünftige extremale Ereignisse treffen zu können. Es gibt dazu bereits verschiedene Methoden bzw. Kennzahlen, die in dieser Arbeit teilweise zusammengefasst und ergänzt werden. Die Idee der Extremwerttheorie ist dabei, dass man nicht die gesamte Zeitreihe modelliert. Es ist ausreichend, den Teil der Zeitreihe zu betrachten, der über einer hohen Schwelle liegt, den sog. Tail. Die Wahl der Schwelle ist dabei ein zentrales Problem.

Die Korrelationsfunktion ist die klassische Methode, um lineare Abhängigkeiten in Zeitreihen zu beschreiben. Basierend auf zweiten Momenten ist sie jedoch teilweise ungeeignet, die gesamte Abhängigkeit in den Extremwerten einer Zeitreihe zu beschreiben, da zweite Momente nicht immer existieren. Die Abhängigkeit in den Extremwerten einer Zeitreihe ist aber im Risikomanagement bei Finanzinstituten von großem Interesse. Die Extremwerte, das können im Marktrisikobereich sowohl hohe Gewinne als auch sog. „Totalausfälle“ sein, stehen hier im Fokus. Eine Alternative zur Korrelationsfunktion bildet die so genannte „extreme dependence function“, die ursprünglich für räumliche Daten entwickelt wurde, aber auch für Zeitreihen verwendet werden kann. In dieser Arbeit wird

diese Funktion für verschiedene Zeitreihenmodelle konkret ausgerechnet und für simulierte und reale Finanzzeitreihen statistisch geschätzt. Unter Finanzzeitreihen versteht man die zeitlich sortierte Abfolge von Preisbeobachtungen oder Indexwerten, Rohstoffpreisen, Wechselkursen zwischen Währungen u. a., die täglich, wöchentlich, monatlich oder in einer anderen zeitlichen Einteilung ermittelt werden.

Für die Praxis ist die interessanteste Fragestellung, die das Studium von Tails beantworten kann, was die extremen Bewegungen sind, die am Finanzmarkt erwartet werden können. Sind die größten Bewegungen bereits beobachtet worden oder können noch größere erwartet werden? Gibt es theoretische Prozesse die schwere Tails, die in der empirischen Analyse auftreten, modellieren können? Antworten auf solche Fragen sind essentiell für solides Risikomanagement. Mit der Extremwerttheorie wird versucht diese Fragen zu beantworten.

## 1.1 Aufbau der Arbeit

Diese Arbeit beschäftigt sich mit dem extremen Korrelationsmaß  $\chi(\cdot)$ , das den Zusammenhang von extremen Ereignissen innerhalb einer Zeitreihe beschreiben soll. Dieses Maß wird für verschiedene nicht-lineare Modelle berechnet und geschätzt.

In **Kapitel 1** wird eine kurze Einführung in die Themenstellung und den Aufbau dieser Arbeit, sowie die mathematischen Grundlagen gegeben.

**Kapitel 2** fasst wichtige Resultate der Extremwerttheorie zusammen. Diese Resultate gehen in **Kapitel 3** ein. Dort werden verschiedene Methoden der Tailschätzung vorgestellt. In **Kapitel 4** werden zunächst Eigenschaften von Finanzzeitreihen zusammengefasst. Anschließend werden folgende nicht-lineare Prozessklassen betrachtet und deren Eigenschaften herausgearbeitet:

- **AutoRegressive Conditional Heteroskedasticity** Prozess der Ordnung 1, kurz ARCH(1) Prozess,
- **Generalized AutoRegressive Conditional Heteroskedasticity** Prozess der Ordnungen 1 und 1, kurz GARCH(1,1) Prozess,
- **AutoRegressive** Prozess der Ordnung 1 mit einem ARCH(1) Prozess als Fehlerterm, kurz AR(1) + ARCH(1) Prozess,

Abschließend werden die theoretischen Formeln zur Berechnung des extremalen Maßes  $\chi(\cdot)$  für das jeweilige nicht-lineare Modell angegeben.

In **Kapitel 5** werden die Methoden aus Kapitel 3 auf die Modelle aus Kapitel 4 sowie einen reellen Datensatz angewandt. Die empirischen Resultate werden anschließend mit den theoretischen Resultaten verglichen (Güte der Schätzung).

Im abschließenden **Kapitel 6** werden die Ergebnisse zusammengefasst und mögliche weitere Anknüpfungspunkte aufgezeigt.

## 1.2 Präliminarien

In diesem Kapitel werden einige Grundlagen für die später verwendeten Methoden und Modelle vorgestellt. Die in Kapitel 1.2.1 und 1.2.2 gegebenen Definitionen beruhen hauptsächlich auf Embrechts et al. (2003), Beirlant et al. (2004) und Bingham et al. (1987). Für weitere Informationen wird auf diese Quellen verwiesen.

### 1.2.1 Tail

Da in dieser Arbeit die Tails der Verteilungen von Zeitreihen betrachtet werden, wird hier zunächst die Definition vom sog. Tail (Schwanz) gegeben.

#### Definition 1.1 (Tail einer Verteilung/Zufallsvariable)

Für jede Zufallsvariable  $X$  mit Verteilungsfunktion  $F$  schreibt man

$$\bar{F}(x) = 1 - F(x) = P(X > x), \quad x \in \mathbb{R}$$

und nennt dies den Tail von  $X$ .

In biostatistischen Anwendungen wird  $\bar{F}$  auch als Überlebensfunktion (*survival function*) bezeichnet (Beirlant et al., 1996, S. 6). Daher werden in dieser Arbeit die Begriffe Tail und Überlebensfunktion äquivalent verwendet.

#### Definition 1.2 (Schwerer Tail, leichter Tail)

Sei  $F$  eine Verteilungsfunktion auf  $[0, \infty)$  und  $g$  eine Funktion mit  $\lim_{x \rightarrow \infty} g(x) = 0$ .

Der Tail  $\bar{F}(x)$  heißt leichter als  $g$ , wenn er schneller als  $g$  gegen 0 fällt, d.h.

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x)}{g(x)} = 0.$$

Der Tail  $\bar{F}(x)$  heißt schwerer als  $g$ , wenn er langsamer als  $g$  gegen 0 fällt, d.h.

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x)}{g(x)} = \infty.$$

$F$  hat einen schweren (polynomialen) Tail  $\bar{F}(x)$ , wenn ein  $\kappa > 0$  existiert, so dass  $\bar{F}(x) \sim cx^{-\kappa}$  für  $x \rightarrow \infty$  gilt.  $\kappa$  wird als Tailindex bezeichnet.

Im Allgemeinen haben Verteilungen, deren exponentielle Momente nicht existieren einen schweren („heavy“) Tail. Dazu gehören auch regulär variierende Verteilungen. Der Begriff der regulären Variation wird im Folgenden eingeführt. Die reguläre Variation ist bei den Modellen in Kapitel 4 von zentraler Bedeutung.

### 1.2.2 Reguläre Variation

Da Extremwerte, also die Werte, die am Rand einer Verteilungsfunktion liegen, betrachtet werden, wird zunächst der rechte Endpunkt einer Verteilungsfunktion definiert.

**Definition 1.3 (rechter Endpunkt, Embrechts et al. (2003), S. 114 f.)**

Für eine Verteilungsfunktion  $F$  ist der rechte Endpunkt  $x_R$  gegeben durch

$$\begin{aligned} x_R &= \sup\{x \in \mathbb{R} : F(x) < 1\} \\ &= \inf\{x \in \mathbb{R} : F(x) \geq 1\} \in (-\infty, \infty]. \end{aligned}$$

**Definition 1.4 (Reguläre Variation, Resnick (1987), S. 13)**

Eine Lebesgue-messbare Funktion  $U: (0, \infty) \rightarrow (0, \infty)$  ist regulär variierend in  $\infty$  mit Index  $\kappa \in \mathbb{R}$ , wenn gilt:

$$\lim_{x \rightarrow \infty} \frac{U(tx)}{U(x)} = t^\kappa, \text{ für alle } t > 0.$$

Man schreibt:  $U \in \mathcal{R}_\kappa$ .  $\kappa$  wird als Index der regulären Variation bezeichnet.

Wenn  $\kappa = 0$  ist, dann heißt  $U$  langsam variierend in  $\infty$ .

Man schreibt:  $U \in \mathcal{R}_0$ . Die Klasse aller regulär variierenden Funktionen wird mit  $\mathcal{R}$  bezeichnet.

Beispiele für regulär variierende Funktionen mit Index  $\kappa$  sind u.a.:

(i)  $x^\kappa$ :

$$\lim_{x \rightarrow \infty} \frac{(xt)^\kappa}{x^\kappa} = t^\kappa.$$

(ii)  $x^\kappa \log x$ :

$$\lim_{x \rightarrow \infty} \frac{(x^\kappa t^\kappa \log(xt))}{x^\kappa \log x} = \lim_{x \rightarrow \infty} t^\kappa \left( \frac{\log tx}{\log x} \right) \stackrel{\text{L'Hospital}}{=} \lim_{x \rightarrow \infty} t^\kappa \frac{t/(tx)}{1/x} = t^\kappa.$$

(iii)  $x^\kappa \log(1+x)$ :

$$\lim_{x \rightarrow \infty} \frac{(tx)^\kappa \log(1+tx)}{x^\kappa \log(1+x)} = \lim_{x \rightarrow \infty} t^\kappa \frac{\log(1+tx)}{\log(1+x)} = t^\kappa.$$

Regulär variierende Funktionen sind Funktionen, die sich asymptotisch wie Potenzfunktionen verhalten.

Es gibt auch Verteilungen, die regulär variierend sind. Hierzu einige Beispiele, die in Mikosch (2004, S. 216) angegeben sind.

- Pareto Verteilung:

$$\bar{F}(x) = \frac{\alpha^\kappa}{(x + \alpha)^\kappa}, \quad x \geq 0, \quad \kappa > 0, \quad \alpha > 0.$$

- Student- $t$ -Verteilung mit  $n$  Freiheitsgraden:

$$f(x) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{\pi n}} (1 + x^2/n)^{-(n+1)/2}, \quad x \in \mathbb{R}.$$

- Log-Gamma Verteilung:

$$f(x) = \frac{\kappa^\beta}{\Gamma(\beta)} (\log x)^{\beta-1} x^{-\kappa-1}, \quad x \geq 1, \quad \kappa, \beta > 0.$$

Im Folgenden wird für eine langsam variierende Funktion die Bezeichnung  $l$  verwendet.  $\mathcal{R}_0$  bezeichnet die Menge der langsam variierenden Funktionen.

Beispiele für langsam variierende Funktionen:

- (i) Konstante Funktionen sind trivialerweise langsam variierend.
- (ii) Die Logarithmusfunktion, denn

$$\lim_{x \rightarrow \infty} \frac{\log(tx)}{\log(x)} \stackrel{L'Hospital}{=} \lim_{x \rightarrow \infty} \frac{t/(tx)}{1/x} = 1.$$

- (iii)  $l(x) = \exp\{(\log x)^\beta\}$  wobei  $\beta > 0$ .

Für die Eigenschaften von langsam variierenden Funktionen wird auf Beirlant et al. (2004, Kapitel 2) verwiesen. In dieser Arbeit ist  $l$  meist eine Konstante  $c$ .

### Definition 1.5 (Tailindex)

Seien  $X$  eine Zufallsvariable (ZV),  $\bar{F}(x) = \mathbb{P}(X > x)$ ,  $x > 0$ , die Funktion, die den rechten Tail (upper Tail) der ZV beschreibt, und  $l$  eine langsam variierende Funktion in  $\infty$ , dann gilt

$$\mathbb{P}(X > x) \sim l(x)x^{-\kappa}. \tag{1.1}$$

$\bar{F}$  ist dabei regulär variierend in  $\infty$  mit Index  $-\kappa$ , denn

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = t^{-\kappa}, \quad x \rightarrow \infty.$$

Den Index  $\kappa$  bezeichnet man als Tailindex oder extreme value Index des rechten Tails der Verteilung von  $X$ . Analog lässt sich auch der Tailindex des linken Tails von  $X$  definieren.

Der Tailindex einer Verteilung beschreibt, wie schnell sich der Tail der Verteilung gegen 0 bewegt, d.h. wie viel Masse in den Enden liegt, also mit welcher Wahrscheinlichkeit extreme Ereignisse passieren, z.B. starke Preissprünge vorkommen. Je kleiner der Tailindex ist, umso langsamer fällt der Tail der Verteilung. Je größer der Tailindex ist, umso stärker fällt ceteris paribus der Tail der Verteilung.

### Bemerkung 1.6

$\mathbb{P}(X > x) = cx^{-\kappa}$  impliziert

$$\log \mathbb{P}(X > x) = \log(c) - \kappa \log(x).$$

$\kappa$  stellt das Steigungsmaß in einer einfachen linearen Beziehung zwischen dem Log-Tail der empirischen Überlebensfunktion und den Log-Extremwerten dar.

In Kapitel 4 werden nicht-lineare Zeitreihenmodelle betrachtet. Bevor diese Modelle definiert und erläutert werden, werden vorab einige wichtige Eigenschaften von Zeitreihen definiert. Als Grundlage für die folgenden Ausführungen gilt Brockwell und Davis (1991).

## 1.3 Zeitreihen

### Definition 1.7 (Stochastischer Prozess, Sato (1999), Definition 1.4)

Ein stochastischer Prozess ist eine Familie von Zufallsvariablen  $(X_t)_{t \in I}$ , definiert auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ .  $I$  ist dabei eine stetige oder diskrete Indexmenge. Meistens gilt  $I \in \{\mathbb{N}, \mathbb{R}_+\}$ .

### Definition 1.8 (Kovarianzfunktion)

$(X_t)_{t \in I}$  sei ein stochastischer Prozess mit  $\text{var}(X_t) < \infty$  für alle  $t \in I$ . Dann heißt

$$\gamma(h) = \gamma_X(h) = \text{cov}(X_{t+h}, X_t) = \mathbb{E}((X_{t+h} - \mathbb{E}(X_{t+h}))(X_t - \mathbb{E}(X_t))), \quad \forall t, h \in I \quad (1.2)$$

Kovarianzfunktion von  $(X_t)_{t \in I}$ .

### Definition 1.9 (Korrelationsfunktion)

$(X_t)_{t \in I}$  sei ein stochastischer Prozess mit  $\text{var}(X_t) < \infty$  für alle  $t \in I$ . Dann heißt

$$\rho(h) = \rho_X(h) = \text{corr}(X_{t+h}, X_t) = \frac{\text{cov}(X_{t+h}, X_t)}{\sqrt{\text{var}(X_{t+h})\text{var}(X_t)}} \quad \text{mit} \quad -1 \leq \rho_X(h) \leq 1 \quad (1.3)$$



Korrelationsfunktion von  $(X_t)_{t \in I}$ .

Hat man Realisierungen  $(X_1, \dots, X_n)$  der Datenlänge  $n$ , so wird man versuchen die empirische Autokorrelationsfunktion aufzustellen, gegeben durch

$$\hat{\rho}(h) = \frac{n^{-1} \sum_{j=1}^{n-h} (X_{j+h} - \bar{X})(X_j - \bar{X})}{n^{-1} \sum_{j=1}^n (X_j - \bar{X})^2},$$

wobei  $\bar{X} = n^{-1} \sum_{j=1}^n X_j$ .

Die Korrelation misst die lineare Abhängigkeit. Es gilt:  $|\rho(X, Y)| = 1$  genau dann, wenn  $Y = aX + b$  mit Wahrscheinlichkeit 1 für  $a \in \mathbb{R} \setminus \{0\}$  und  $b \in \mathbb{R}$ . Bei einem guten Modell sollten die empirische und theoretische Autokorrelationsfunktion nahe beieinander liegen. Typisch für Log>Returns von Finanzzeitreihen ist, dass  $\hat{\rho}(0) = 1$  und  $\hat{\rho}(h) \approx 0$  für  $h > 0$ . Es liegt also praktisch keine Korrelation zwischen den Log>Returns vor. Dennoch sind sie meist nicht unabhängig und unkorreliert. An der Autokorrelationsfunktion (ACF) der quadrierten Log>Returns  $(X_t^2)$  oder an der ACF der Absolutbeträge  $(|X_t|)$  erkennt man, dass eine weit reichende Korrelation vorliegt (siehe z.B. Abbildung 5.31). Wird für die quadrierten und die absoluten Werte der Test auf Unkorreliertheit nicht abgelehnt, macht man meist die Annahme, dass die Log>Returns iid sind.

Die Log>Returns (siehe Definition 5.1) von Finanzmarktdaten sind meist nicht unabhängig, erfüllen aber das Kriterium der Stationarität. Die Stationarität ist eine wichtige vereinfachende Annahme, welche fordert, dass der Prozess in einem besonderen Zustand von statistischem Gleichgewicht ist.

### Definition 1.10 (Strikte Stationarität)

Ein Prozess  $(X_t)_{t \in I}$  heißt strikt stationär, falls  $(X_{t_1}, \dots, X_{t_n}) \stackrel{D}{=} (X_{t_1+h}, \dots, X_{t_n+h})$  für alle  $t_1, \dots, t_n \in I$ ,  $h \geq 0$  und  $n \in \mathbb{N}$  sind.

Also ist ein stochastischer Prozess strikt stationär, wenn er invariant gegenüber Zeitverschiebungen ist.

### Definition 1.11 (Schwache Stationarität)

Ein stochastischer Prozess  $(X_t)_{t \in \mathbb{Z}}$  heißt schwach stationär, falls

$$(i) \quad \mathbb{E}(|X_t|^2) < \infty \quad \forall t \in \mathbb{Z},$$

$$(ii) \quad \mathbb{E}(X_t) = \mu \quad \text{für ein } \mu \in \mathbb{R} \text{ und } \forall t \in \mathbb{Z},$$

$$(iii) \quad \gamma_X(t, t') = \gamma_X(t+h, t'+h), \quad \forall t, t', h \in \mathbb{Z}.$$

**Bemerkung 1.12**

- (i) Im Verlauf der Arbeit wird häufig der Begriff *stationär* verwendet. Damit ist stets *strikt stationär* gemeint.
- (ii) Besitzt ein *strikt stationärer* Prozess *endliche Varianz*, dann ist er auch *schwach stationär*. Die Umkehrung gilt im Allgemeinen nicht.
- (iii) Ist der Prozess  $(X_t)_{t \in \mathbb{Z}}$  *schwach stationär*, dann gilt

$$\gamma_X(t, t') = \gamma_X(t - t', 0), \quad t, t' \in \mathbb{Z}.$$

Oft werden Daten modelliert, indem man die Normalverteilung anpasst. Die Normalverteilung ist jedoch selten die passende Verteilung. Dies wird in Kapitel 5 bei der Betrachtung von simulierten und realen Finanzmarktdaten verdeutlicht. Die Abweichung von der Normalverteilung spiegelt sich in einer Abweichung der Momente wider. So wird die Eigenschaft einer gegenüber der Normalverteilung größeren Kurtosis als Leptokurtosis bezeichnet.

**Definition 1.13 (Leptokurtosis, Embrechts et al. (2003), S. 407)**

Eine Verteilung ist *leptokurtisch*, d.h. sie besitzt eine *spitze Wölbung* und *schmale Taille* in der Umgebung des Mittelwertes, wenn für ihre Kurtosis gilt:

$$\mathcal{K}(X) := \frac{\mathbb{E}((X - \mu)^4)}{(\mathbb{E}((X - \mu)^2))^2} > 3.$$

In diesem Fall ist die Wölbung der Verteilung spitzer als die der Normalverteilung. In der Finanzmathematik wird der Begriff „leptokurtisch“ auch verwendet, um anzuzeigen, dass mehr Wahrscheinlichkeitsmasse im Ursprung und in den Tails vorhanden ist, als bei der Standardnormalverteilung, d.h. die Verteilung besitzt *schwere Tails*. Oft wird statt der Kurtosis, dem Wölbungsmaß, die *Überschußkurtosis* verwendet. Diese ist so bereinigt, dass bei einem Wert von Null Mesokurtosis vorliegt. Verteilungen mit einer *Überschusskurtosis* von Null heißen *mesokurtisch*. Ein Beispiel dafür ist die Normalverteilung. Eine alternative Definition der Leptokurtosis kann über die Quantile erfolgen. Der Vorteil dabei ist, dass  $k_p$  gegenüber  $\mathcal{K}$  stets wohldefiniert ist.

**Definition 1.14 (Embrechts et al. (2003), Definition 3.3.5, Quantilsfunktion)**

Die *verallgemeinerte Inverse* der Verteilungsfunktion  $F$

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

wird als *Quantilsfunktion* der Verteilungsfunktion  $F$  bezeichnet. Die Quantität  $x_p = F^{\leftarrow}(p)$  definiert das  $p$ -Quantil von  $F$ .

**Definition 1.15 (Ebrechts et al. (2003), S. 407)**

Eine Verteilungsfunktion  $F$  heißt leptokurtisch, falls gilt:

$$k_p = \frac{\frac{1}{2}(x_{0.75} - x_{0.25})}{x_{0.9} - x_{0.1}} < 0.263.$$

Dabei sind  $x_p$  die jeweiligen Quantile.

**Definition 1.16 (Schiefe)** Der Schiefekoeffizient ist definiert durch

$$\mathcal{S}(X) := \frac{\mathbb{E}(X - \mu)^3}{\sigma^3}.$$

In Kapitel 1 wurde bereits die Brown'sche Bewegung erwähnt, daher folgt hier die Definition.

**Definition 1.17 (Brown'sche Bewegung, Embrechts et al. (2003), Def. 2.4.1)**

Sei  $(B_t)_{t \in [0,1]}$  ein stochastischer Prozess der folgende Bedingungen erfüllt:

- (i)  $B_0 = 0$  f.s., d.h. der Prozess beginnt bei 0.
- (ii) Der Prozess hat unabhängige Zuwächse. Für alle  $0 \leq t_0 < t_1 < \dots < t_m \leq 1$  und irgendein  $m$  sind  $B_{t_1} - B_{t_0}, \dots, B_{t_m} - B_{t_{m-1}}$  unabhängig.
- (iii) Für jedes  $t \in [0,1]$  besitzt  $B_t$  eine Gauß'sche Verteilung mit Erwartungswert 0 und Varianz  $t$ .
- (iv) Die einzelnen Pfade sind f.s. stetig.

Dieser Prozess heißt Brown'sche Bewegung oder Wiener Prozess auf  $[0,1]$ .

Abbildung 1.1 zeigt beispielhaft einen Pfad einer Brown'schen Bewegung.

**Definition 1.18 (Weißes Rauschen, White Noise)**

Sei  $(Z_t)_{t \in \mathbb{Z}}$  ein stationärer stochastischer Prozess mit  $\mathbb{E}Z_t = 0$  und

$$\gamma_Z(h) = \begin{cases} \sigma^2, & \text{falls } h = 0, \\ 0, & \text{sonst.} \end{cases}$$

Dann heißt  $(Z_t)_{t \in \mathbb{Z}}$  Weißes Rauschen oder White Noise mit Erwartungswert 0 und Varianz  $\sigma^2$ , kurz  $WN(0, \sigma^2)$ .

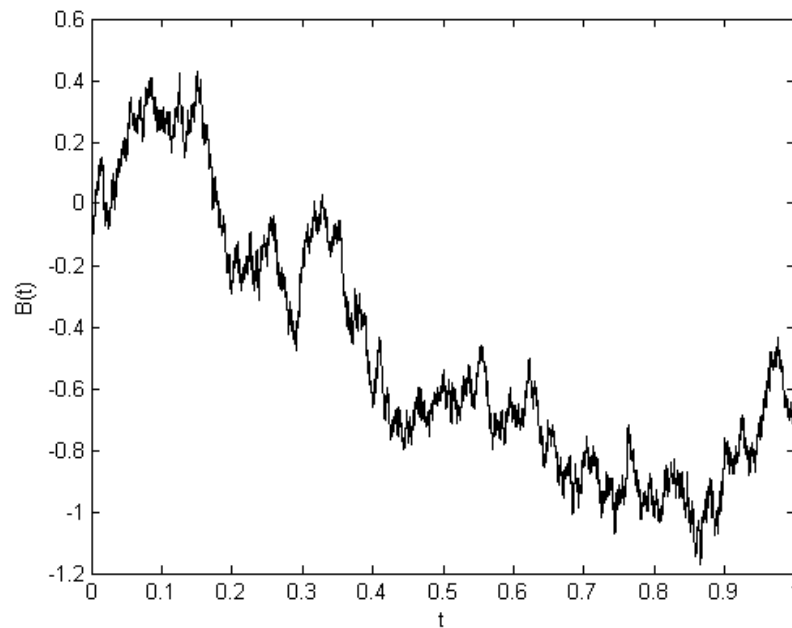
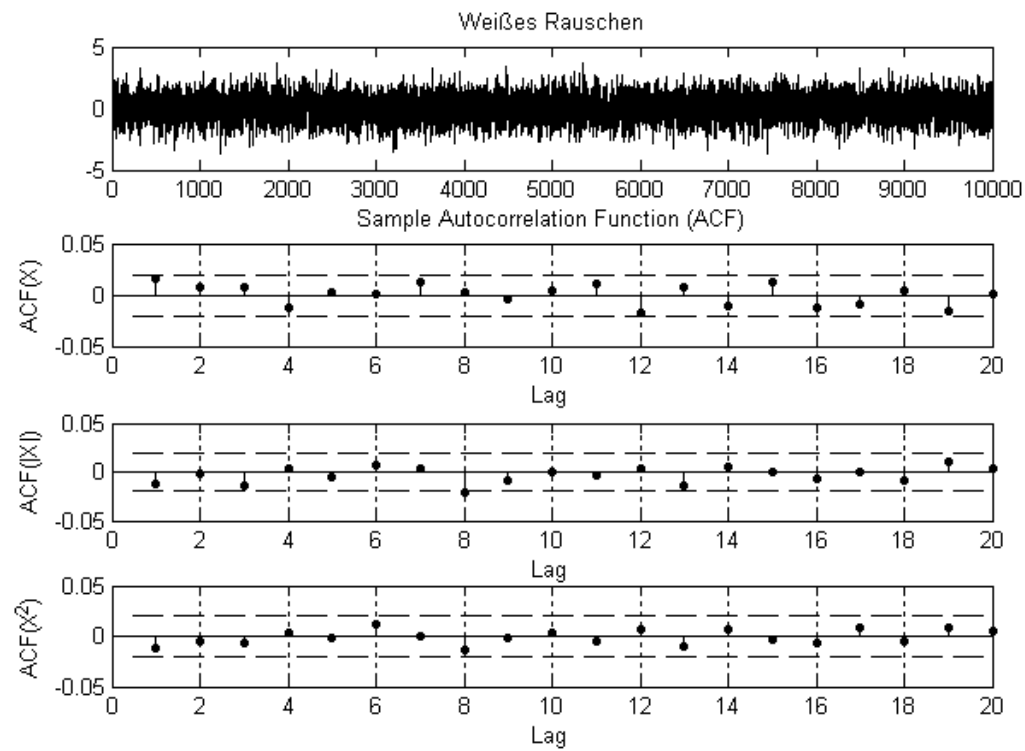


Abbildung 1.1: Ein Pfad einer Brown'schen Bewegung mit  $n = 10000$ .

In Abbildung 1.2 ist ein Weißes Rauschen,  $WN(0, 1)$ , der Länge  $n = 10000$  und die dazugehörige ACF abgebildet. Betrachtet man im Vergleich z.B. Abbildung 4.3 in Kapitel 4.2.1 erkennt man deutliche Unterschiede, aber auch den Zusammenhang. Setzt man im ARCH(1) Modell aus (4.1) den Parameter  $\alpha_1 = 0$ , dann erhält man ein Weißes Rauschen, da die Innovationen  $(Z_t)_{t \in \mathbb{N}}$  in dieser Arbeit standardnormalverteilt sind. Weitere Erläuterungen zum ARCH(1) Modell erfolgen in Kapitel 4.2.1.

Das Verhalten der ACF in Abbildung 1.2 ist typisch für iid Daten. Daher werden Daten, deren ACF der quadrierten und absoluten Werte auch Unkorreliertheit anzeigt als iid angenommen. Diese Annahme wird im Verlauf der Arbeit verwendet, um Daten auf Unabhängigkeit zu testen.

Abbildung 1.2: White Noise  $WN(0, 1)$  mit  $n = 10000$ .

# Kapitel 2

## Wichtige Resultate der Extremwerttheorie

In diesem Kapitel werden wichtige Resultate der Extremwerttheorie zusammengefasst, die im weiteren Verlauf der Arbeit verwendet werden. Sie sind Voraussetzung um ein extremes Korrelationsmaß schätzen und beurteilen zu können. Für weiterführende Informationen wird auf die jeweils zitierten Arbeiten verwiesen.

### 2.1 Hauptsatz der Extremwerttheorie

Die Normalverteilung ist die bedeutendste Grenzverteilung für Durchschnitte oder Stichprobensummen. Ein bedeutendes Werkzeug ist dabei der Zentrale Grenzwertsatz. Die Grenzverteilung von Maxima, also extremen Werten, ist damit jedoch nicht abgedeckt. Die Familie der Grenzverteilungen kann unter einer einfachen Parametrisierung dargestellt werden und ist als verallgemeinerte Extremwertverteilung (GEV) bekannt. Das Theorem von Fisher und Tippett liegt im Zentrum der Extremwerttheorie. Es beschäftigt sich mit der Verteilung von Maxima. Es seien  $X_1, X_2, \dots, X_n$ ,  $n \in \mathbb{N}$  iid Zufallsvariablen mit Verteilungsfunktion  $F$ . Im Folgenden ist die Verteilung der Zufallsvariable  $M_n = \max\{X_1, \dots, X_n\}$  von Interesse. Die Verteilung von  $M_n$  ist dann

$$\mathbb{P}(M_n \leq x) = \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n \mathbb{P}(X_i \leq x) = F^n(x).$$

In praktischen Anwendungen ist die Verteilung  $F$  meist nicht vollständig oder überhaupt nicht bekannt. Daher ist auch die exakte Verteilung des Maximums  $M_n$  nicht explizit bestimmbar. Es ist daher naheliegend das asymptotische Verhalten von  $M_n$  zu betrachten.

**Satz 2.1 (Fisher-Tippett Theorem, Embrechts et al. (2003), Theorem 3.2.3)**

*Sei  $(X_n)_{n \in \mathbb{N}}$  eine Folge von iid Zufallsvariablen.  $M_n = \bigvee_{1 \leq i \leq n} X_i = \max\{X_1, \dots, X_n\}$ .*

Ohne Schwierigkeiten kann man die exakte Verteilungsfunktion des Maximums  $M_n$  angeben:

$$\mathbb{P}(M_n \leq x) = \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) = F^n(x), \quad x \in \mathbb{R}, \quad n \in \mathbb{N}.$$

Falls es normierende Konstanten  $a_n > 0$  und  $b_n \in \mathbb{R}$ ,  $n \geq 1$  und eine nichtausgeartete Verteilungsfunktion  $H$  gibt mit

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

also  $\frac{M_n - b_n}{a_n} \xrightarrow{\mathcal{D}} H$  für  $n \rightarrow \infty$ , dann gehört  $H$  zu einer der folgenden drei Klassen von Verteilungsfunktionen:

(i) Fréchet:

$$H(x) = \Phi_\kappa(x) = \begin{cases} 0, & x < 0 \\ \exp(-x^{-\kappa}), & x \geq 0 \end{cases} \quad \text{für } \kappa > 0.$$

(ii) Weibull:

$$H(x) = \Psi_\kappa(x) = \begin{cases} \exp(-(-x)^\kappa), & x < 0 \\ 1, & x \geq 0 \end{cases} \quad \text{für } \kappa > 0.$$

(iii) Gumbel:

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

Einen Beweis des Satzes findet man in (Resnick, 1987, S. 9 f.) oder in Embrechts et al. (2003, Theorem 3.2.3).

Gilt  $F^n(a_n + b_n x) \xrightarrow{n \rightarrow \infty} H$ , wobei  $H$  zu einer der drei oben genannten Verteilungen gehört, sagt man auch  $F$  liegt im Anziehungsbereich von  $H$ , d.h.  $F \in MDA(H)$ . Die Grenzverteilung  $H$  hängt ausschließlich vom Tail der Verteilungsfunktion  $F$  ab. Dies folgt aus der Tatsache, dass die folgenden Aussagen für iid Daten äquivalent sind. Dabei sei  $u_n = a_n x + b_n$ ,  $a_n > 0$ ,  $b_n \in \mathbb{R}$  und  $X$  eine ZV mit der gleichen Verteilung wie  $X_1$ .

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

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

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

$$\lim_{u_n \rightarrow \infty} \mathbb{P} \left( \frac{X - u_n}{a(u_n)} > x \mid X > u_n \right) = \frac{\bar{F}(u_n + a(u_n)x)}{\bar{F}(u_n)} = (1 + \xi x)^{-1/\xi} = -\log H(x),$$

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

Extreme Werte treten in der Nähe des oberen Endes des Trägers der Verteilung auf. Daher hängt das asymptotische Verhalten von  $M_n$  mit dem Verhalten der Verteilungsfunktion  $F$  im rechten Tail in der Nähe des rechten Endpunktes zusammen. Das Fisher-Tippett Theorem gibt an, dass die asymptotische Verteilung der Maxima zu einer der drei oben genannten Verteilungen gehört, egal welche ursprüngliche Verteilung die beobachteten Daten besitzen. Daher kann also das Tailverhalten der Zeitreihe mittels einer dieser drei Verteilungen geschätzt werden. Die  $a_n$  werden auch als normalisierende Konstanten und die  $b_n$  als zentrierende Konstanten bezeichnet.

In der Jenkinson von Mises Darstellung sind die Verteilungen aus Satz 2.1 zur verallgemeinerten Extremwertverteilung zusammengefasst.

**Definition 2.2 (Jenkinson von Mises Darstellung: GEV)** [*Embrechts et al. (2003), Definition 3.4.1*]

Definiere die Verteilungsfunktion  $H_\xi$  durch

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

wobei  $1 + \xi x > 0$ .

$\xi = 1/\kappa$  ist der Formparameter. Dabei bezeichnet  $\kappa$  den Tailindex.

$\xi = \kappa^{-1} > 0$ entspricht der Fréchetverteilung $\Phi_\kappa$ , $\xi = 0$ entspricht der Gumbelverteilung $\Lambda$ , $\xi = -\kappa^{-1} < 0$ entspricht der Weibullverteilung $\Psi_\kappa$ .
--

Der Parameter  $\xi$  wird als Formparameter bezeichnet. Je größer der Formparameter, desto schwerere Tails hat die Verteilung. Die GEV beschreibt die Grenzverteilungen von normierten Maxima, ist also die Familie der möglichen Grenzverteilungen für normierte Maxima. Für  $\xi > 0$  ist der Tail polynomial fallend. Dies entspricht einer Verteilung mit schweren Tails.  $F$  liegt im Anziehungsbereich einer Fréchetverteilung. Diese Verteilungsklasse schließt die Pareto-,  $t$ -, Burr-, Loggamma- und Cauchyverteilung ein. Der Fall  $\xi = 0$  ist der eines exponentiell fallenden Tails. Dann ist  $F$  im Anziehungsbereich von Gumbel; Beispiele in diesem Anziehungsbereich sind die Normal-, Lognormal- und Gammaverteilung. Für  $\xi < 0$  ist  $F$  im Anziehungsbereich von Weibull. Dies ist der Fall eines endlichen oberen Endpunktes und damit leichten Tails. Zu dieser Verteilungsklasse gehören die Uniform- und die Betaverteilung. Jede univariate Extremwertverteilung



kann in eine standard Fréchetverteilung transformiert werden. Daher kann o.B.d.A. die standard Fréchetverteilung als Grenzverteilung betrachtet werden.

$H_\xi$  wird als standard GEV bezeichnet. Das Argument  $x$  kann o.B.d.A. ersetzt werden durch  $(x - \mu)/\psi$  für  $\mu \in \mathbb{R}$  und  $\psi > 0$ , um die Familie  $H_{\xi;\mu,\psi}$  zu erhalten.

$$H_{\xi;\mu,\psi}(x) = \exp \left\{ - \left( 1 + \xi \frac{x - \mu}{\psi} \right)^{-1/\xi} \right\}, \quad 1 + \xi \frac{x - \mu}{\psi} > 0. \quad (2.1)$$

Abbildung 2.1 zeigt beispielhaft die Dichtefunktionen von je einer Fréchet-, Weibull- und Gumbelverteilung.

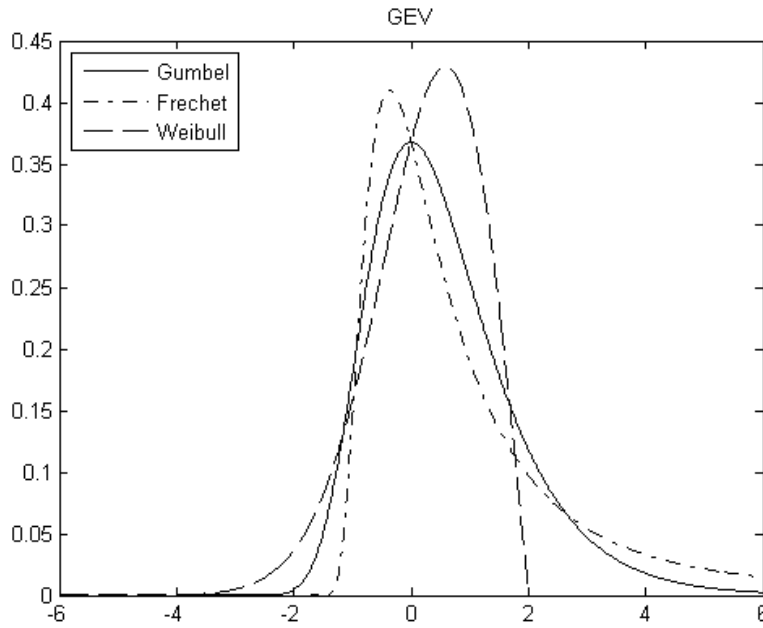


Abbildung 2.1: Dichtefunktionen der GEV mit  $\xi = 0.5$  (Fréchet) bzw.  $\xi = -0.5$  (Weibull),  $\mu = 0$  und  $\psi = 1$ .

Die bisherigen Ergebnisse gelten für Folgen von iid Daten. Was passiert jedoch wenn keine iid Daten vorliegen? Für stationäre Folgen erhält man unter gewissen Voraussetzungen

$$F_n(a_n x + b_n) = \mathbb{P}(\max(X_1, \dots, X_n) \leq a_n x + b_n) \rightarrow H^\theta(x), \quad x \in \mathbb{R}.$$

Das Verhalten der Extrema wird durch die lokale Abhängigkeit im Prozess beeinflusst, da Cluster von extremen Werten auftreten. Der Extremale Index  $\theta$  modelliert die Beziehung zwischen der Abhängigkeitsstruktur und dem extremalen Verhalten des Prozesses. Auf dieses Ergebnis wird in Abschnitt 2.2 näher eingegangen.

### 2.1.1 Parameterschätzung der GEV

Der Parameter  $\vartheta = (\xi, \mu, \psi) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$  besteht aus dem Formparameter  $\xi$ , dem Lagereparameter  $\mu$  und dem Skaleneparameter  $\psi$ . Mit  $H_{\xi,\mu,\psi}$  bzw.  $H_\vartheta$  wird die GEV bezeichnet.

net. Die Parameter  $\xi$ ,  $\mu$  und  $\psi$  sollen nun geschätzt werden. Dafür gibt es verschiedene Möglichkeiten. In dieser Arbeit wird die Maximum Likelihood Schätzung verwendet, die im Folgenden vorgestellt wird. Weitere Informationen und Schätzmethoden findet man in Embrechts et al. (2003, S. 316 ff.) und Beirlant et al. (2004).

### Maximum Likelihood Schätzer

Der Maximum Likelihood Schätzer (MLE) ist am Weitesten verbreitet und einfach in der Handhabung. Allgemein geht man davon aus, dass  $H_\vartheta$  die Dichte  $h_\vartheta$  hat. Dann gilt für die Likelihood Funktion basierend auf  $\mathbf{X} = (X_1, \dots, X_n)$

$$L(\vartheta; \mathbf{X}) = \prod_{i=1}^n h_\vartheta(X_i) \mathbf{1}_{\{1+\xi(X_i-\mu)/\psi > 0\}}.$$

Mit  $l(\vartheta, \mathbf{X}) = \log L(\vartheta, \mathbf{X})$  wird die Log-Likelihood Funktion bezeichnet. Der MLE für  $\vartheta$  ist gegeben durch

$$\hat{\vartheta}_n = \arg \max_{\vartheta \in \Theta} l(\vartheta, \mathbf{X}).$$

Im Fall  $\xi \neq 0$ , ist die Log-Likelihood Funktion für eine Stichprobe  $X_1, \dots, X_n$  von iid GEV Zufallsvariablen gegeben durch

$$\log L(\xi, \mu, \psi) = -n \log \psi - \left(\frac{1}{\xi}\right) \sum_{i=1}^n \log \left(1 + \xi \frac{X_i - \mu}{\psi}\right) - \sum_{i=1}^n \left(1 + \xi \frac{X_i - \mu}{\psi}\right)^{-\frac{1}{\xi}}$$

mit  $1 + \xi \frac{X_i - \mu}{\psi} > 0$ ,  $i = 1, \dots, n$ . Falls  $\xi = 0$  gilt, dann reduziert sich die Log-Likelihood Funktion zu

$$\log L(0, \mu, \psi) = -n \log \psi - \sum_{i=1}^n \exp\left(-\frac{X_i - \mu}{\psi}\right) - \sum_{i=1}^n \frac{X_i - \mu}{\psi}.$$

Den MLE  $(\hat{\xi}, \hat{\mu}, \hat{\psi})$  erhält man durch Maximieren der Gleichungen.

### 2.1.2 Tail- und Quantilschätzung der GEV

Im Rahmen der Extremwerttheorie wurden verschiedene Schätzverfahren für die Schätzung der Tails entwickelt, die im weiteren Verlauf der Arbeit Verwendung finden. Nach Embrechts et al. (2003, S. 323 f.) gilt für den Tailschätzer  $\overline{H}_\vartheta$ :

$$\overline{H}_\vartheta(x) = 1 - \exp \left\{ - \left( 1 + \hat{\xi} \frac{x - \hat{\mu}}{\hat{\psi}} \right)^{-\frac{1}{\hat{\xi}}} \right\}, \quad (2.2)$$

dabei kann  $\hat{\vartheta} = (\hat{\xi}, \hat{\mu}, \hat{\psi})$  z.B. mit der ML-Methode geschätzt werden. Ein natürlicher Schätzer für das  $p$ -te Quantil von  $H_\vartheta$  ist gegeben durch:

$$\hat{x}_p = H_{\hat{\vartheta}}^{\leftarrow}(p), \quad p \in (0, 1).$$

Dabei bezeichnet  $H_{\hat{\vartheta}}^{\leftarrow}(p)$  die Quantilsfunktion. Mit der Definition 1.14 von  $H_\vartheta$  folgt also

$$\hat{x}_p = \hat{\mu} - \frac{\hat{\psi}}{\hat{\xi}} \left( 1 - (-\log p)^{-\hat{\xi}} \right).$$

### 2.1.3 Anziehungsbereiche

Bisher wurde davon ausgegangen, dass die Beobachtungen iid sind und eine Extremwertverteilung  $H_\theta$  besitzen. In diesem Abschnitt wird angenommen, dass  $X_1, \dots, X_n$  iid sind mit  $F \in MDA(H_\xi)$ .  $F \in MDA(H_\xi)$  ist äquivalent zu

$$\lim_{n \rightarrow \infty} n\bar{F}(a_n x + b_n) = -\log H_\xi(x),$$

mit normierenden Folgen  $a_n > 0$  und  $b_n$ .

Verteilungen in  $MDA(H_\xi)$  für  $\xi > 0$  (der Fréchet Fall) besitzen schwere Tails (siehe Abschnitt 2.1). Es gilt dann

$$1 - F(x) = \bar{F}(x) = x^{-1/\xi} l(x) = x^{-\kappa} l(x).$$

$l(x)$  ist dabei eine langsam variierende Funktion (siehe Definition 1.4). Eine Möglichkeit ist, dass für große  $u_n = a_n x + b_n$

$$n\bar{F}(u_n) \approx \left(1 + \frac{u_n - b_n}{a_n}\right)^{-1/\xi}$$

gilt, so dass der Tailschätzer für adäquate Schätzer  $\hat{\xi}, \hat{a}_n, \hat{b}_n$  folgende Darstellung annehmen kann

$$\widehat{\bar{F}}(u_n) = \frac{1}{n} \left(1 + \hat{\xi} \frac{u_n - \hat{b}_n}{\hat{a}_n}\right)^{-1/\hat{\xi}}.$$

Nun kann das Quantil geschätzt werden durch

$$\hat{x}_p = \hat{b}_n + \frac{\hat{a}_n}{\hat{\xi}} \left( (n(1-p))^{-\hat{\xi}} - 1 \right).$$

Für die Schätzung der normierenden Konstanten  $a_n$  und  $b_n$  siehe Embrechts et al. (2003, S. 345 ff.).

## 2.2 Stationäre Daten

In der Realität tendieren extreme Ereignisse dazu in Clustern aufzutreten, d.h. auf große Ausschläge folgen weitere Ausschläge. Dies wird verursacht durch die lokale Abhängigkeit in den Daten. Große Schwankungen am Finanzmarkt werden beispielsweise verursacht durch (extreme) Ereignisse wie Naturkatastrophen, Terroranschläge etc. Finanzmarktdaten wie Wechselkurse und Aktienpreise sind dann mit diesen Ereignissen verbunden und treten nicht unabhängig auf. Stationäre Zeitreihen stellen daher realistischere Modelle für Log>Returns einer Finanzzeitreihe dar, zumindest über Zeitperioden, die nicht zu lang sind. Eine stationäre Finanzzeitreihe, d.h. die Log>Returns, weist in der Regel Cluster auf. Daher kann die Theorie aus Kapitel 2.1, die auf der Annahme von iid Daten basiert, nicht unverändert angewendet werden. Die Abhängigkeiten, die durch die Cluster dargestellt werden, müssen in der Berechnung berücksichtigt werden. Daher wird der Extremale Index eines stationären Prozesses definiert.

**Definition 2.3 (Extremaler Index Embrechts et al. (2003), Definition 8.1.2)**

Sei  $(X_n)$  eine stationäre Folge und  $\theta$  eine nicht-negative Zahl. Falls man für  $\tau \in (0, \infty)$  eine Folge  $(u_n)$  findet, so dass

$$\begin{aligned}\lim_{n \rightarrow \infty} n\mathbb{P}(X_i > u_n) &= \lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau, \\ \lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) &= \exp(-\theta\tau)\end{aligned}$$

gelten, so heißt  $\theta$  der *Extremale Index* der Folge  $(X_n)$ .

Äquivalent zu Definition 2.3 sind folgende Aussagen:

$$\begin{aligned}\lim_{n \rightarrow \infty} n\mathbb{P}(X_i > u_n) &= \lim_{n \rightarrow \infty} n(1 - F(u_n)) = \tau, \\ \lim_{n \rightarrow \infty} \mathbb{P}\{M_n \leq u_n\} &= \exp(-\theta\tau), \\ \lim_{n \rightarrow \infty} \mathbb{P}\{\tilde{M}_n \leq u_n\} &= \exp(-\tau).\end{aligned}$$

Dabei sind  $M_n = \max(X_1, \dots, X_n)$  und  $\tilde{M}_n = \max(\tilde{X}_1, \dots, \tilde{X}_n)$ , wobei  $(\tilde{X}_n)$  eine assoziierte iid Folge von ZV und  $F$  die Verteilungsfunktion von  $X$  und  $\tilde{X}$  ist.

Der Extremale Index ist eine Größe, die die Abhängigkeitsstruktur von extremen Daten beschreibt. Er beschreibt also den Unterschied zwischen der Verteilung des Maximums der Zeitreihe und des Maximums einer unabhängigen Stichprobe mit der gleichen Randverteilung  $F$ . Für große  $n$  gilt

$$\mathbb{P}\{M_n \leq u_n\} \sim \mathbb{P}^\theta\{\tilde{M}_n \leq u_n\} = F^{n\theta}(u_n). \quad (2.3)$$

$\theta$  lässt sich auch darstellen durch

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

Der Kehrwert des Extremalen Index gibt die mittlere Clustergröße einer stationären Folge an, wenn die Schwelle gegen den rechten Endpunkt  $x_R$  tendiert. Es gilt  $\theta \in [0, 1]$ . Eine stationäre Folge clustert, wenn  $\theta < 1$ .  $\theta = 1$  gilt, wenn schwache Abhängigkeit oder Unabhängigkeit vorliegt. Dies ist der Fall bei iid Daten. Der Fall  $\theta = 0$  ist ein pathologischer Fall. Je niedriger der Extremale Index, also je kleiner  $\theta$ , desto größer ist die extreme Abhängigkeit. Um einen Vergleich zwischen der Zeitreihe und den entsprechenden iid Beobachtungen zu ziehen, werden die Abhängigkeiten der Zeitreihe ignoriert. Es wird also die sog. unabhängige assoziierte Folge  $\tilde{X}_n$  betrachtet. Für  $F \in MDA(H_\xi)$  ist also die asymptotische Verteilung der normierten Maxima einer stationären Folge  $(X_n)$  mit Extremalem Index  $\theta$  auch eine Extremwertverteilung. Der Formparameter ist auch hier  $\xi$ , da nur der Lage- und Skalenparameter durch  $\theta$  beeinflusst werden.

**Lemma 2.4 (Schätzung von  $\theta$ , Embrechts et al. (2003), S. 419 f)**

Sei  $N = \#\{\text{Exzesse von } u\}$  und  $K$  die Anzahl der Blöcke, in denen ein oder mehr Exzesse

aufzutreten, d.h.  $K = \#\{\text{„Cluster-Exzesse“}\}$ .  $u$  bezeichnet dabei eine Schwelle. Dann kann man  $\theta$  durch die Block-Methode wie folgt schätzen:

$$\hat{\theta} = \frac{\text{„Anzahl der Blockmaxima größer } u\text{“}}{\text{„Anzahl der Exzesse“}} = \frac{K}{N}. \quad (2.4)$$

Zu beachten ist dabei, dass die Schätzungen für  $\theta$  stark von der Schwelle  $u$  und von der Blockgröße abhängen. Es gibt noch einige weitere Methoden zur Schätzung des Extremalen Index. Auf diese wird in dieser Arbeit nicht eingegangen, da für die Berechnungen in Kapitel 5 diese einfache Schätzung als ausreichend angenommen wird. Für weitere Informationen zur Schätzung des Extremalen Index siehe auch Regler (2004).

## 2.3 Die verallgemeinerte Pareto Verteilung

Die verallgemeinerte Pareto Verteilung spielt bei der POT-Methode in Abschnitt 3.3.4 eine wichtige Rolle und wird daher im Folgenden erörtert.

### Definition 2.5 (Verallgemeinerte Pareto Verteilung (GPD))

In Embrechts et al. (2003, S. 163 ff.) ist die GPD folgendermaßen dargestellt.

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

Dabei ist  $\beta > 0$ . Der Support ist  $x \geq 0$ , wenn  $\xi \geq 0$  und  $0 \leq x \leq -\beta/\xi$ , wenn  $\xi < 0$ .

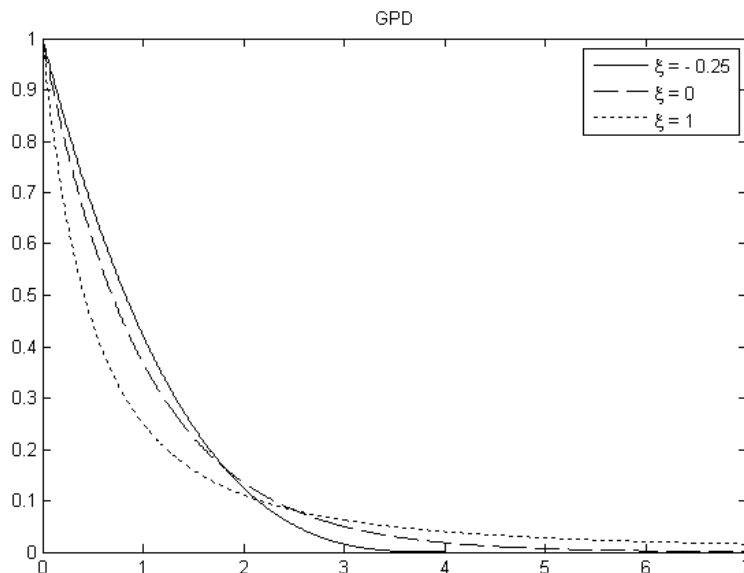


Abbildung 2.2: Dichtefunktionen der GPD mit  $\beta = 1$  für  $\xi = -0.25$ ,  $\xi = 0$  und  $\xi = 1$ .

In Abbildung 2.2 sind die Dichtefunktionen der GPD für  $\beta = 1$ ,  $\xi = -0.25$ ,  $\xi = 0$  und  $\xi = 1$  geplottet.

Zusammenfassung:

Die GEV  $H_\xi$ ,  $\xi \in \mathbb{R}$ , sind die Grenzverteilungen von normierten Maxima.  
Die GPD  $G_{\xi,\beta}$ ,  $\xi \in \mathbb{R}$  und  $\beta > 0$ , sind die Grenzverteilungen von skalierten Exzessen.

Der Formparameter  $\xi$  stimmt mit dem Formparameter der GEV überein. Die GPD ist die Grenzverteilung für Exzesse, falls die Ausgangsverteilung von  $X$  im Anziehungsbereich einer GEV liegt. Entsprechend des Wertes von  $\xi$ , dem Formparameter, kann die GPD analog zur GEV, in drei Gruppen eingeteilt werden.

- Der Fall  $\xi > 0$  entspricht Verteilungen mit schweren Tails, die regulär variierend sind. Dazu zählen u.a. die Pareto-, Student- $t$  und Fréchet-Verteilung. Der Fall der Pareto-Verteilung ist für die Betrachtung von Finanzzeitreihen besonders relevant.
- Der Fall  $\xi = 0$  entspricht der Normal-, Exponential- und Lognormalverteilung. Deren Tails fallen exponentiell.
- Die dritte Gruppe bildet der Fall  $\xi < 0$  für Verteilungen mit leichten bzw. kurzen Tails und einem endlichen rechten Endpunkt wie z.B. die Uniform- und Beta-Verteilung.

Im Gegensatz zur Normalverteilung, für die alle Momente existieren, besitzt eine Verteilung mit schweren Tails nicht alle Momente. Im Fall der GPD mit  $\xi > 0$  ist  $\mathbb{E}(X^k)$  unendlich für  $k \geq 1/\xi$ . Falls  $\xi = 1/2$ , dann besitzt die GPD eine unendliche Varianz und falls  $\xi = 1/4$  ist das vierte Moment unendlich.

### 2.3.1 Parameterschätzung der GPD

Wie auch bei der GEV kann man die Parameter der GPD u.a. mit der ML-Methode schätzen. Diese wird auch in dieser Arbeit verwendet.

#### Maximum Likelihood Schätzung

In Embrechts et al. (2003, S. 356) ist die ML-Schätzung der Parameter der GPD dargestellt. Die Daten  $(X_n)_{n \in \mathbb{N}} = (X_1, \dots, X_n)$  seien iid mit gemeinsamer Verteilungsfunktion  $F$ . Angenommen  $F$  ist eine GPD mit Parametern  $\xi$  und  $\beta$  wie in Definition 2.5, dann ist die Dichtefunktion  $f$  gegeben durch

$$f(x) = \frac{1}{\beta} \left(1 + \xi \frac{x}{\beta}\right)^{-\frac{1}{\xi}-1}.$$

Dabei gilt:

$$x \in \mathbb{D}(\xi, \beta), \text{ wobei } \mathbb{D}(\xi, \beta) = \begin{cases} [0, \infty) & \text{falls } \xi \geq 0, \\ \left[0, -\frac{\beta}{\xi}\right] & \text{falls } \xi < 0. \end{cases}$$

Die Log-Likelihood Funktion ist gegeben durch

$$\log L(\beta, \xi) = -n \log \beta - \left( \frac{1}{\xi} + 1 \right) \sum_{i=1}^n \log \left( 1 + \frac{\xi X_i}{\beta} \right), \quad 1 + \frac{\xi X_i}{\beta} > 0.$$

Falls  $\xi = 0$ , dann ist

$$\log L(\beta, 0) = -n \log \beta - \frac{1}{\beta} \sum_{i=1}^n X_i.$$

### 2.3.2 Tail- und Quantilschätzung

Dieser Abschnitt basiert auf Embrechts et al. (2003, S. 352 ff.) und Klüppelberg (2003, S. 112 ff.). Es wird angenommen, dass die  $(X_n)_{n \in \mathbb{N}}$  iid sind und  $X_n \stackrel{D}{=} X \stackrel{D}{=} F$  gilt. Für eine hohe Schwelle  $u$  wird

$$N_u = \text{card}\{i : X_i > u, i = 1, \dots, n\}$$

als die Anzahl der Überschreitungen der Schwelle  $u$  von  $X_1, \dots, X_n$  definiert. Die dazugehörigen Exzesse werden mit  $Y_1, \dots, Y_n$  bezeichnet. Der Tail der Exzess-Verteilungsfunktion (vgl. Definition 3.3) ist dann gegeben durch

$$\bar{F}_u(y) = \mathbb{P}(X - u > y | X > u) = \frac{1 - F(u + y)}{1 - F(u)} = \frac{\bar{F}(u + y)}{\bar{F}(u)}, \quad y \geq 0.$$

Dies ist äquivalent zu

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

Die GPD  $G_{\xi, \beta}$  mit den Parametern  $\xi \in \mathbb{R}$  und  $\beta > 0$  hat den Tail

$$\bar{G}_{\xi, \beta}(x) = \begin{cases} (1 + \xi x / \beta)^{-1/\xi} & \text{falls } \xi \neq 0, \\ \exp(-x/\beta) & \text{falls } \xi = 0. \end{cases} \quad (2.5)$$

Dabei gilt:  $x \in \mathbb{D}(\xi, \beta)$ .

Die Schätzung von  $\bar{F}(u + y)$  erfolgt mit dem empirischen Tailschätzer:

$$\widehat{\bar{F}}(u) = \frac{1}{n} \sum_{i=1}^n I(X_i > u) = \frac{N_u}{n}.$$

Approximiert man

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

und schätzt  $\xi$  und  $\beta$  durch  $\hat{\xi}$  und  $\hat{\beta}$ , dann ergeben sich der folgende Tail- und Quantilschätzer:

- Tailschätzer:

$$\widehat{\overline{F}}(u+y) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{y}{\hat{\beta}}\right)^{-1/\hat{\xi}}, \quad y \geq 0. \quad (2.6)$$

- Quantilschätzer:

$$\hat{x}_p = u + \frac{\hat{\beta}}{\hat{\xi}} \left( \left( \frac{n}{N_u} (1-p) \right)^{-\hat{\xi}} - 1 \right), \quad p \in (0, 1).$$

Eine standardisierte Vorgehensweise um  $\xi$  und  $\beta$  zu schätzen ist wieder die ML-Methode.

**Beispiel 1** (Coles, 2001, S. 77)

Für das standard Fréchet Modell,  $F(x) = \exp(-1/x)$  für  $x \geq 0$  folgt

$$\overline{F}_u(y) = \frac{1 - F(u+y)}{1 - F(u)} = \frac{1 - \exp(-(u+y)^{-1})}{1 - \exp(-u^{-1})} \approx \left(1 + \frac{y}{(-u)}\right)^{-1}$$

für  $u \rightarrow \infty$  und alle  $y > 0$ . Dies entspricht der GPD mit  $\hat{\xi} = 1$  und  $\hat{\beta} = -u$ .



# Kapitel 3

## Extremwertstatistik

### 3.1 Schätzung des Tailindex

#### Hill-Schätzer nach Hill (1975)

$X_{j,n}$ ,  $j = 1, \dots, n$  bezeichnet die Ordnungsstatistik einer iid Folge  $X_i$ ,  $i = 1, \dots, n$  von iid ZV mit gemeinsamer Extremwertverteilung, d.h.  $X_{n,n} \leq \dots \leq X_{1,n}$ . Der Hill-Schätzer für den Tailindex  $\kappa$  wird folgendermaßen definiert

$$\hat{\kappa}_{k,n}^{(H)} := \left( \frac{1}{k} \sum_{j=1}^k \log X_{j,n} - \log X_{k,n} \right)^{-1}, \quad k = 2, \dots, n. \quad (3.1)$$

Die Wahl des optimalen  $k$  erfolgt meist mit Hilfe des Hill-Plots und kann mit einer Faustregel bestimmt werden:

Wähle  $k$  nicht zu groß und so, dass  $\hat{\kappa}_{k,n}^{(H)}$  aus einem Gebiet stammt, das ungefähr horizontal ist.

Mit dem Hill-Schätzer kann man auch einen Schätzer für den Tail  $\bar{F}(x)$  erhalten, indem die deterministische Schwelle  $u$  durch eine zufällige Schwelle  $X_{k,n}$  ersetzt wird, wobei  $k$  definiert ist durch  $k = \text{card}\{i : X_{i,n} > u, i = 1, \dots, n\}$ . Für den Tailschätzer gilt

$$\widehat{\bar{F}}(x) = \frac{k}{n} \left( \frac{x}{X_{k,n}} \right)^{-\hat{\kappa}_{k,n}^{(H)}}$$

und für das  $p$ -Quantil

$$\hat{x}_p = \left( \frac{n}{k} (1-p) \right)^{-1/\hat{\kappa}_{k,n}^{(H)}} X_{k,n}.$$

### Eigenschaften des Hill-Schätzers

Im Folgenden werden einige Eigenschaften des Hill-Schätzers beschrieben. Die Grundlage dafür bilden Embrechts et al. (2003, S. 336 ff.) und Beirlant et al. (2004, S. 104 ff.).

**Satz 3.1** Sei  $(X_n)_{n \in \mathbb{N}}$  eine Folge von iid ZV mit Verteilungsfunktion  $F \in \mathcal{R}_{1/\kappa}$ ,  $\kappa > 0$ .

(a) Schwache Konsistenz:

Wenn  $k \rightarrow \infty$  und  $k/n \rightarrow 0$  für  $n \rightarrow \infty$ , dann gilt:

$$\hat{\kappa}_{k,n}^{(H)} \xrightarrow{P} \kappa.$$

(b) Starke Konsistenz:

Wenn  $k/n \rightarrow 0$  und  $k/(\log \log n) \rightarrow \infty$  für  $n \rightarrow \infty$ , dann gilt:

$$\hat{\kappa}_{k,n}^{(H)} \xrightarrow{a.s.} \kappa.$$

Weiterhin wird angenommen, dass eine Konstante  $\beta > 0$  und eine Funktion  $b \in \mathcal{R}_{-\beta}$  existieren.  $U(y) = Q(1 - 1/y)$  sei die Tailquantilsfunktion. Dann gilt

$$\frac{U(ux)}{U(x)} = u^\kappa (1 + h_{-\beta}(u)b(x) + o(b(x))) \text{ für } x \rightarrow \infty.$$

Sei  $b_{n,k} = b((n+1)/(k+1))$ :

(c) Asymptotische Normalität:

Wenn  $k, n \rightarrow \infty$ ,  $k/n \rightarrow 0$  und  $\sqrt{k} b_{n,k} \rightarrow 0$ , dann gilt

$$\sqrt{k} \left( \frac{\hat{\kappa}_{k,n}^{(H)}}{\kappa} - 1 \right) \xrightarrow{\mathcal{D}} N(0, 1).$$

(d) Der asymptotische Bias ist

$$ABias(\hat{\kappa}_{k,n}^{(H)}) \sim \frac{b_{n,k}}{1 + \beta}.$$

(e) Die asymptotische Varianz ist

$$Avar(\hat{\kappa}_{k,n}^{(H)}) \sim \frac{\kappa^2}{k}.$$

$\hat{\kappa}_{k,n}^{(H)}$  ist ein konsistenter Schätzer für  $\kappa$ .

**Bemerkung 3.2** (Embrechts et al. (2003), S. 337 ff.)

- Die Analyse, basierend auf dem Hill-Schätzer wird graphisch mit dem Hill-Plot vorgenommen (siehe Abbildung 3.1). Der Hill-Plot

$$\{(k, \hat{\kappa}_{k,n}^{(H)}) : k = 2, \dots, n\},$$

hat sich als nützliches Instrument herausgestellt, um das optimale  $k$  zu finden.

- Die asymptotische Varianz von  $\hat{\kappa}_{k,n}^{(H)}$  hängt vom unbekanntem Parameter  $\kappa$  ab, so dass für die Berechnung der asymptotischen Konfidenzintervalle ein geeigneter Schätzer von  $\kappa$  benötigt wird, typischerweise wird  $\hat{\kappa}_{k,n}^{(H)}$  verwendet.
- Der Hill-Schätzer ist sehr sensibel gegenüber Abhängigkeiten innerhalb der Daten. In diesem Fall liefert der Hill-Schätzer meistens unbrauchbare Ergebnisse.

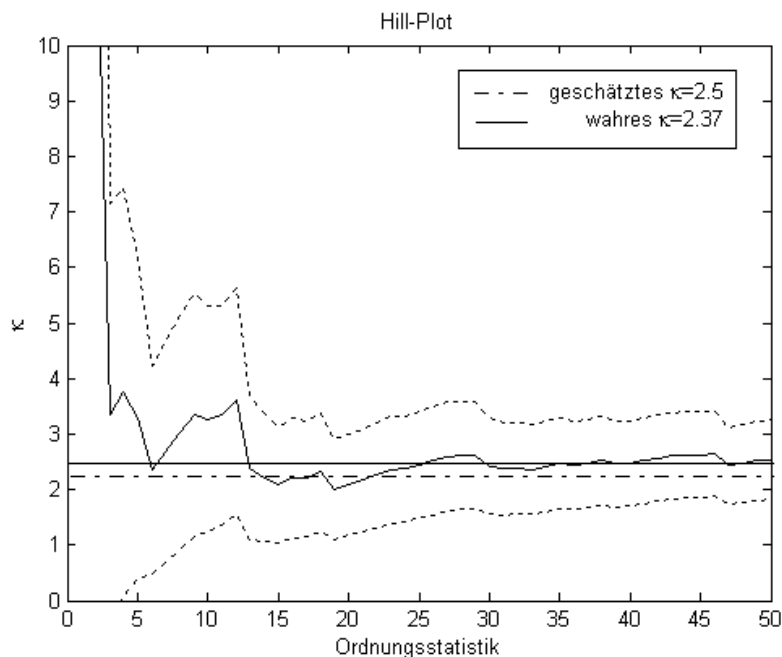


Abbildung 3.1: Hill-Plot für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

### Nachteile bei der Anwendung des Hill-Schätzers auf einen konkreten Datensatz

- (i) Die Wahl von  $k$  ist nicht eindeutig.
- (ii) Für jede Wahl von  $k$  erhält man einen anderen Schätzer für  $\kappa$ .
- (iii) Der Hill-Schätzer ist nicht invariant gegenüber einer Verschiebung der Daten. Dies kann zu einem systematischen Fehler führen.
- (iv) Oft sind die Stichproben so klein, dass die Asymptotik nicht wirkt.
- (v) Die asymptotische Varianz von  $\hat{\kappa}_{k,n}^{(H)}$  und die Konfidenzintervalle hängen von  $\kappa$  ab.

## 3.2 Bestimmung der Schwelle $u$

Da ein wesentliches Problem bei den Methoden in Abschnitt 3.3 die Bestimmung der Schwelle  $u$  ist, wird nun ein Hilfsmittel dafür angegeben.

### Die Exzessfunktion

Eine Möglichkeit die Schwelle  $u$  zu bestimmen, ist durch die Exzessfunktion gegeben.

#### Definition 3.3 (Exzess-Verteilungsfunktion, Mittlere Exzessfunktion (MEF))

[Embrechts et al. (2003), Definition 3.4.6]

Sei  $X \stackrel{\mathcal{D}}{=} F$  eine ZV mit  $x_R \leq \infty$ . Für ein festes  $u < x_R$  heißt

$$F_u(x) = \mathbb{P}(X - u \leq x | X > u) = \frac{F(x + u) - F(u)}{1 - F(u)}, \quad x \geq 0,$$

die Exzess-Verteilungsfunktion von  $X$  oder  $F$  über der Schwelle  $u$ . Die Funktion

$$e(u) = \mathbb{E}(X - u | X > u) = \int_u^{x_R} \frac{\bar{F}(t)}{\bar{F}(u)} dt, \quad u \geq 0,$$

heißt Mittlere Exzessfunktion von  $X$  oder  $F$ .

Die empirische Version der Mittleren Exzessfunktion von  $X$ , lautet

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

Für Verteilungen mit schweren Tails geht die Mittlere Exzessfunktion gegen  $\infty$ , bei leichteren Tails gegen Null (vgl. Tabelle 3.1). Im Fall der Exponentialverteilung ist sie konstant. Für die GPD ist die Mittlere Exzessfunktion eine lineare Funktion, die genau dann wächst

wenn der Parameter  $\xi$  positiv ist. Die MEF für die GPD mit  $\xi < 1$  und  $\beta$  für  $u < x_R$  sieht folgendermaßen aus:

$$e(u) = \mathbb{E}(X - u | X > u) = \frac{\beta + \xi u}{1 - \xi}, \quad \beta + \xi u > 0.$$

Die Mittlere Exzessfunktion von Verteilungen mit schweren Tails liegen also zwischen der Pareto- und der Exponentialverteilung.

Pareto	$\frac{\alpha+u}{\kappa-1}, \kappa > 1$
Loggamma	$\frac{u}{\kappa\tau-1}(1 + o(1)), \kappa\tau > 1$
Lognormal	$\frac{\sigma^2 u}{\log u - \mu}(1 + o(1))$
Weibull	$\frac{u^{1-\beta}}{c\tau}(1 + o(1))$
Exponential	$\frac{1}{\lambda}$

Tabelle 3.1: Mittlere Exzessfunktion für einige Standardverteilungen. Asymptotik gilt für  $u \rightarrow \infty$ .

#### Definition 3.4 (Mean Excess Plot)

Der Mean Excess Plot (ME-Plot) ist ein Werkzeug, um eine Zeitreihe in die drei Kategorien, „long-tailed“, „medium-tailed“ oder „short-tailed“ einzuteilen. Der ME-Plot ist wie folgt definiert:

$$\{(X_{(j)}, e_n(X_{(j)})) : j = 1, \dots, n\}.$$

Dabei repräsentiert  $e_n(\cdot)$  die empirische MEF und  $(X_{(j)})_{j \in \mathbb{N}}$  ist eine Ordnungsstatistik  $X_1 \geq \dots \geq X_n$ .

In Abbildung 3.2 ist der Mean Excess der quadrierten Allianz Log>Returns gegen die Schwellen geplottet. Die sechs größten Werte sind aus der Darstellung herausgenommen worden. Es ist bekannt, dass hier Randeffekte auftreten, die dazu führen, dass der ME-Plot gegen Null fällt. Ein Aufwärtstrend wie in der Abbildung ist ein Zeichen für schwere Tails. Die Wahl der Schwelle  $u$  kann in diesem Fall erfolgen, wenn sich das Verhalten des Graphen ändert. Bis zur Schwelle  $u = 0.002$  verläuft der Graph ungefähr linear. Dennoch bleibt es schwer eine eindeutige Wahl zu treffen.

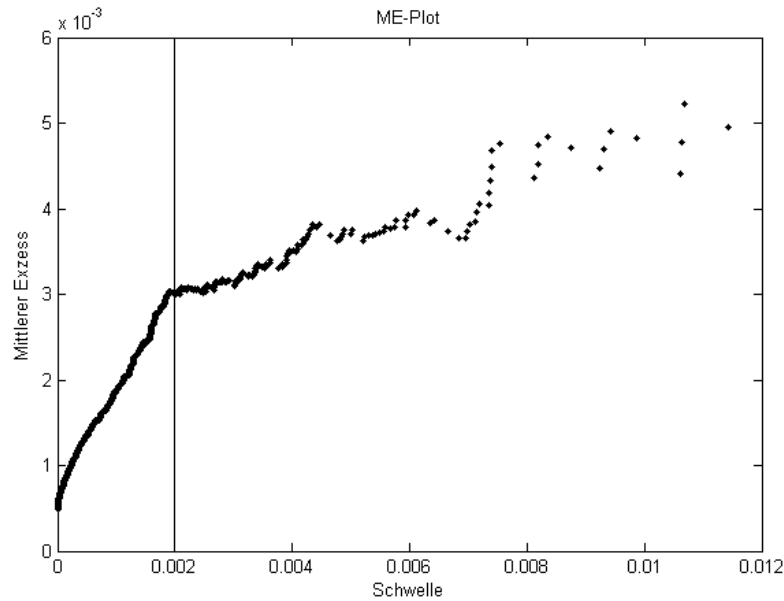


Abbildung 3.2: ME-Plot der quadrierten Allianz Log&gt;Returns.

### 3.3 Methoden zur Tail Schätzung

In diesem Kapitel werden Verfahrensweisen zur Schätzung des Tails beschrieben, die dann für die Schätzung des extremalen Abhängigkeitsmaßes  $\chi(\cdot)$  verwendet werden. Für ausführlichere Informationen zu diesen Verfahren wird auf Beirlant et al. (2004), Embrechts et al. (2003) und Coles (2001) verwiesen. Ein generelles Problem bei der Schätzung der Tails besteht darin, die Datenmenge herauszufinden, die für die Schätzung verwendet werden soll. Es handelt sich dabei stets um Daten über einer Schwelle. Die Wahl der Schwelle ist dabei nicht trivial aber essentiell.

Zunächst werden zwei Methoden beschrieben, bei denen die Abhängigkeit der Daten nicht berücksichtigt wird. In den darauf folgenden Methoden wird die Abhängigkeit berücksichtigt. Interessant ist, dass die Unterschiede der Ergebnisse in Kapitel 5 zwischen den verschiedenen Methoden meist nicht sehr groß sind.

#### 3.3.1 Empirische Schätzung

##### Empirische Schätzung

Bei dieser Schätzmethode wird die Beziehung von zwei direkt nebeneinanderliegenden Datenpunkten betrachtet. Als Erstes wird eine Schwelle  $u$  gewählt. Es wird überprüft, ob beide Datenpunkte die Schwelle  $u$  überschreiten. Liegt  $X_1$  über  $u$  wird überprüft, ob auch  $X_2$  über  $u$  liegt. Ist dies der Fall wird 1 gezählt. Dann betrachtet man  $X_2$  und  $X_3$  etc. Jedes Mal, wenn beide Datenpunkte über  $u$  liegen wird 1 gezählt. Zur Schätzung des

Tails wird dann der Quotient aus der Anzahl, dass beide Punkte über  $u$  liegen und der Anzahl aller Daten über  $u$  gebildet.

### Empirische Schätzung der Tails

Die einfachste Methode, um einen ersten Eindruck über den Tail einer Verteilung zu erhalten, ist die empirische Schätzung des Tails. Es wird gezählt, wie viele Daten die Schwelle  $u$  übersteigen.  $N$  bezeichnet die Anzahl der Überschreitungen von  $u$ . Das Verhältnis aus der Anzahl der Exzesse und der Gesamtlänge  $n$  der Daten ergibt die Schätzung.

$$\frac{N}{n} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i > u\}}.$$

Diese Wahl ist motiviert durch das Glivenko-Cantelli Theorem für stationäre ergodische Folgen  $(X_n)$ .

### Satz 3.5 (Glivenko-Cantelli Theorem, Embrechts et al. (2003), Example 2.1.4)

Mit

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}, \quad x \in \mathbb{R}$$

sei die empirische Verteilungsfunktion der iid Folge  $X_1, \dots, X_n$  bezeichnet. Nach dem starken Gesetz der großen Zahlen gilt:

$$F_n(x) \xrightarrow{f.s.} \mathbb{E}(\mathbf{1}_{\{X \leq x\}}) = F(x), \quad \forall x \in \mathbb{R}.$$

Dies kann man verschärfen zu dem Theorem von Glivenko-Cantelli:

$$\delta_n = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{f.s.} 0.$$

Die empirischen Schätzmethode stellen im Vergleich zu den anderen Methoden, die in dieser Arbeit vorgestellt werden, die wenigsten Anforderungen an die Daten, berücksichtigen aber die spezifische Abhängigkeitsstruktur nicht.

Da der Tail in den äußersten Punkten geschätzt werden soll, müssen andere Methoden herangezogen werden. Die empirische Schätzung der Tails ist nur eine grobe Schätzung und nicht geeignet, um den Tail in  $x \rightarrow \infty$  zu schätzen. Auch reicht das Glivenko-Cantelli Theorem nicht aus, die Tails in den äußersten Punkten zu betrachten.

### 3.3.2 Blockmaxima-Methode

Bei der Blockmaxima-Methode werden die Daten in Blöcke gleicher Länge aufgeteilt. Der Datensatz der Länge  $n$  wird in  $k$  Blöcke der Länge  $r$  eingeteilt, also ist  $n = rk$ . Damit keine Probleme mit der Ganzzahligkeit der Blöcke auftreten, wählt man  $r = \lfloor n/k \rfloor$ . (Für

andere Methoden der Blockbildung siehe Embrechts et al. (2003, S. 419 ff.)). Anschließend werden die Maxima der Blöcke, die so genannten Blockmaxima, betrachtet. Man nimmt an, dass die Blockmaxima unabhängig sind und einer verallgemeinerten Extremwertverteilung folgen. Es wird versucht die GEV an die empirische Verteilung der Blockmaxima anzupassen. Dazu werden die Parameter der GEV in dieser Arbeit mit Hilfe der Maximum Likelihood Methode geschätzt. Es handelt sich hier also um einen parametrischen Ansatz. Definition 2.3 ist eine Rechtfertigung für die Anpassung der GEV an die Blockmaxima einer stationären Zeitreihe, die eine Tendenz zum Clustern aufweist. Die Abhängigkeit der Daten hat die Auswirkung, dass die Konvergenz gegen eine GEV langsamer ist, da die effektive Stichprobengröße  $n\theta$  kleiner ist als  $n$ . Dies ist aus Gleichung (2.3) erkennbar. Für die Blockmaxima-Methode bedeutet dieses Ergebnis, dass auf die Wahl der Blocklänge sehr viel Aufmerksamkeit zu legen ist. Die Blöcke müssen ausreichend groß gewählt werden, damit ein GEV Modell für die Blockmaxima geeignet ist. Die Blockmaxima-Methode hat aber auch die Schwachstelle, dass nur die Maxima benutzt werden und die anderen Daten nicht verwendet werden (Informationsverlust). Daher müssen zum Einen die Blöcke ausreichend groß gewählt werden, um die Annahme der Unabhängigkeit und der GEV der Maxima zu gewährleisten. Zum Anderen dürfen die Blöcke aber nicht zu groß gewählt werden, da stets das Problem der begrenzten Datenmenge gegeben ist. Der Tail wird dann mit Formel (2.2) geschätzt.

Der Blockmaxima-Ansatz wird häufig in den Umweltwissenschaften verwendet. Zum Beispiel wird die GEV an jährliche maximale Temperaturen oder jährliche maximale Flusspegel angepasst.

Auf Abschnitt 2.1.3 basiert die folgende Vorgehensweise von McNeil (1998).

### 3.3.3 Vorgehensweise in McNeil (1998)

Bei der Blockmaxima-Methode aus Kapitel 3.3.2 wird angenommen, dass die Blockmaxima von iid Daten einer GEV folgen. Satz 2.1 beruht auf der Annahme von iid Daten. In der Realität liegen jedoch meistens Abhängigkeiten in den Daten vor. Diese sollten bei der Schätzung der Tails berücksichtigt werden, damit man eine fundierte Aussage treffen kann. McNeil (1998) hat dazu folgende Vorgehensweise gewählt:

(i) *Blockbildung:*

Zunächst wird der Datensatz analog der Blockmaxima-Methode in Blöcke  $M_{ri} = \max_{r(i-1) \leq t \leq ri} \{X_t\}$ ,  $i = 1, \dots, \lfloor n/r \rfloor$  eingeteilt. Hierbei stellt sich wiederum das Problem der Wahl der Blockgröße  $r$ .

(ii) *GEV:*

Die Blockmaxima  $(M_{ri})_{i=1, \dots, \lfloor n/r \rfloor}$  werden an eine GEV gefittet. Man geht davon aus, dass die  $(M_{ri})$  iid GEV verteilt sind. Die Annahme der Unabhängigkeit kann man z.B. mit dem Lagrange-Multiplikator ( $\mathcal{LM}$ )-Test (siehe Kapitel 5) überprüfen und gegebenenfalls die Blockgröße verändern. Aus dieser Anpassung erhält man durch ML-Schätzung die geschätzten Parameter der GEV  $(\hat{\xi}, \hat{\psi}, \hat{\mu})$ .

(iii) *QQ-Plot:*

Um die Modellanpassung zu überprüfen, werden Residuen gebildet. Residuen sind



die Abweichungen der beobachteten Werte der abhängigen Variable, von den mit einer bestimmten Modellannahme geschätzten Werten. Sind die Residuen von einander abhängig, ist das gewählte Modell nicht das Richtige, da es nicht alle Abhängigkeiten modelliert. Dazu werden Residuen  $W_{ri} = (1 + \hat{\xi} \frac{M_{ri} - \hat{\mu}}{\hat{\psi}})^{-1/\hat{\xi}}$  gebildet und in einem QQ-Plot gegen die Quantile der Exponentialverteilung geplottet. Die  $W_{ri}$  sind exponentialverteilt, da

$$\begin{aligned} \mathbb{P}(W_{ri} \leq x) &= \mathbb{P}\left(\left(1 + \hat{\xi} \frac{M_{ri} - \hat{\mu}}{\hat{\psi}}\right)^{-1/\hat{\xi}} \leq x\right) = \mathbb{P}\left(M_{ri} \leq \left[(x^{-\hat{\xi}} - 1) \frac{\hat{\psi}}{\hat{\xi}}\right] + \hat{\mu}\right) \\ &= H_{\xi, \mu, \psi}\left(\left[(x^{-\hat{\xi}} - 1) \frac{\hat{\psi}}{\hat{\xi}}\right] + \hat{\mu}\right) = \exp(-x). \end{aligned}$$

Daher sollte man eine Gerade erhalten. In Abbildung 3.3 erkennt man, dass die Residuen fast auf der Geraden liegen, also unabhängig sind. Die Modellwahl ist also für die verwendeten Daten geeignet.

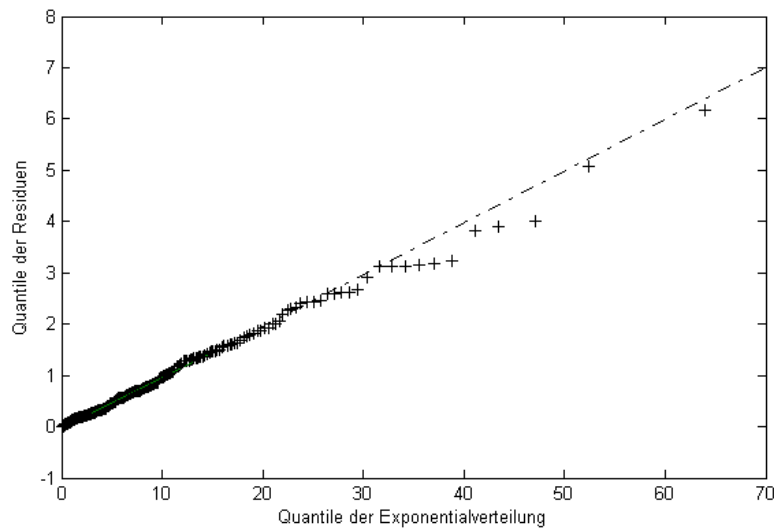


Abbildung 3.3: QQ-Plot der Residuen eines simulierten quadrierten ARCH(1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ , standardnormalverteilter Noise-Variable,  $n = 10000$  und die Quantile der Exponentialverteilung.

(iv) *Schätzung des Quantils:*

Es wird das  $1 - \frac{mr}{n}$ -te Quantil der Verteilung  $F$  geschätzt. Dabei wird zunächst  $m \in \mathbb{N}$  gewählt, z.B.  $m = 3$ . Bei der Wahl von  $m$  muss darauf geachtet werden, dass  $m$  so gewählt wird, dass noch ausreichend Daten vorhanden sind. Aus den Daten wird der  $mr$ -größte Wert  $\tilde{R}_{n,mr}$  zur Schätzung des Quantils gewählt. Es wird  $c_{\frac{n}{mr}} := F^{\leftarrow}\left(1 - \frac{mr}{n}\right) \approx \tilde{R}_{n,mr}$  gesetzt.  $\tilde{R}_{n,mr}$  wird auch als Return-Level bezeichnet.

(v) *Bestimmung des Extremalen Index nach Lemma 3.1:*

Es wird die Schwelle  $u$  bestimmt. Anschließend wird  $\hat{\theta}$  mit der Block-Methode (2.4) geschätzt. Statt die Schwelle  $u$  direkt zu wählen, kann man auch die Anzahl der Daten, die über  $u$  liegen sollen, wählen. Hierbei ist zu beachten, dass die Schwelle  $u$  so gewählt wird, dass nicht zu viele Blöcke darüber liegen.

(vi) *Schätzung des Tails:* Mit

$$\bar{F}(x) = \frac{mr}{\hat{\theta}n} \left( 1 + \hat{\xi} \frac{c_{(n/(mr))}^{-1} x - \hat{\mu}}{\hat{\psi}} \right)^{-1/\hat{\xi}} \quad (3.2)$$

wird der Tail geschätzt.

Im Unterschied zu der Berechnung des Tails bei der Blockmaxima-Methode in Kapitel 3.3.2 geht hier der Extremale Index sowie das Quantil in die Berechnung ein. Dadurch wird der Abhängigkeit der Daten Rechnung getragen. Diese Vorgehensweise wird im Folgenden auch als „Methode von McNeil“ bezeichnet.

### 3.3.4 POT-Methode

Die wohl bekannteste Methode, um Tails oder Quantile zu schätzen, ist die Peaks over Threshold Methode (POT). Sie wurde von Hydrologen entwickelt und ist mathematisch fundiert. Diese Methode basiert scheinbar auf ad hoc statistischen Modellen, ist aber flexibel genug, um saisonale Schwankungen und Abhängigkeiten zu meistern. Bei der POT-Methode betrachtet man die  $k$ -größten Ordnungsstatistiken oder man legt eine hohe Schwelle  $u$  fest und betrachtet die Exzesse über dieser Schwelle, die sog. Peaks. Ein Peak ist definiert als der maximale Wert, der während einer Folge von aufeinander folgenden Beobachtungen über der Schwelle erreicht wird. Die Zeitpunkte an denen die Peaks auftreten, nennt man *Peak Times* und die dazugehörigen Exzesse *Peak Values* (vgl. Smith (1984)). Die POT-Methode ist eine Schätzmethode für einen Tail oder ein Quantil und basiert auf extremen Beobachtungen.

Eine Schwierigkeit bei der POT-Methode liegt in der Bestimmung der Schwelle. Zum Einen muss die Schwelle so hoch sein, dass die asymptotische Approximation verwendbar ist. Zum Anderen darf sie aber nicht zu hoch sein, damit man nicht zu wenige Überschreitungen hat. Wenn  $u$  zu groß ist, ist die Anzahl der Überschreitungen klein und die Varianz damit hoch. Die Schwellen Approximation basiert auf der Verteilung der Exzesse über der Schwelle  $u$ . Als Verteilung wird dazu die GPD verwendet.

Bei stationären Zeitreihen bleibt die Vorgehensweise der POT-Methode wie oben beschrieben. Jedoch hat man meist keine iid verteilten Zufallsvariablen mehr zur Verfügung. Daher wird eine assoziierte Folge von iid Zufallsvariablen gebildet.

Auch bei der POT-Methode besteht das Problem, dass in der Theorie von iid Daten ausgegangen wird, die in der Realität nicht vorliegen. In diesem Fall werden häufig Residuen aus einem an die Log>Returns angepassten Modells verwendet, da diese häufig die Annahme der Unabhängigkeit erfüllen. Dies wird anhand einer graphischen Darstellung wie z.B. dem QQ-Plot überprüft. Es gibt auch verschiedene Tests wie z.B. den Ljung-Box Test

(siehe Kapitel 4.1.2), mit deren Hilfe man die Annahme von iid Daten überprüfen kann. Die POT-Methode kann daher folgendermaßen zusammengefasst werden (Embrechts et al. (2003, S. 366)):

- (1) Für eine hohe Schwelle  $u$  werden die Beobachtungen der Stichprobe markiert, die  $u$  übersteigen. In Abbildung 3.4 sind das die Beobachtungen 1, 4, 6, 10. Die Exzesse einer iid oder stationären Folge über einer hohen Schwelle  $u$  folgen einem Poisson-Prozess (siehe Definition B.4), d.h. die Zeiten zwischen den Exzessen sind exponentialverteilt.
- (2) Die entsprechenden Exzesse über  $u$  sind unabhängig und haben eine GPD. Man interessiert sich nicht nur dafür wann und wie oft die Überschreitungen auftreten, sondern auch wie groß die Exzesse sind. In Abbildung 3.4 sind die Exzesse mit  $Y_1, \dots, Y_4$  bezeichnet.
- (3) Die Exzesse und die Zeitpunkte der Überschreitungen sind voneinander unabhängig.

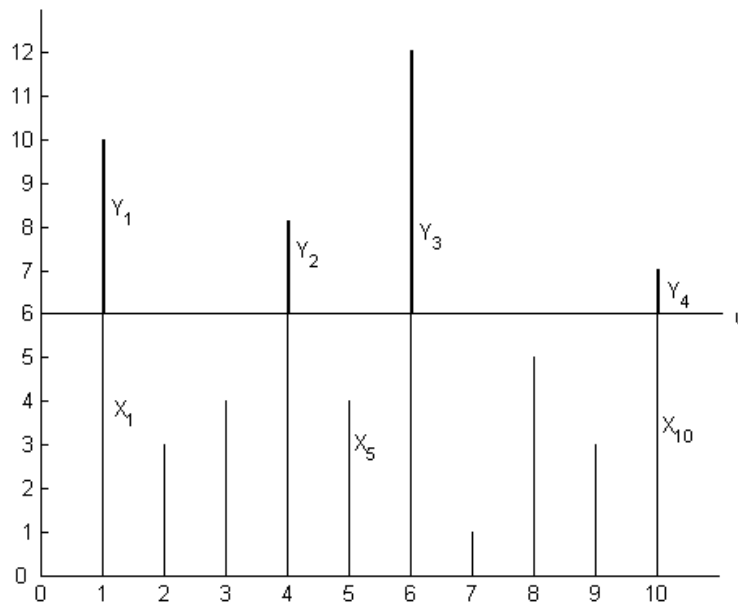


Abbildung 3.4: Daten  $X_1, \dots, X_{10}$  und die entsprechenden  $Y_1, \dots, Y_4$  über der Schwelle  $u$ .

Für die Schätzung des Tails wird Formel (2.6) verwendet.

Möchte man der Abhängigkeit, die bei stationären Daten vorliegt, Rechnung tragen kann auch bei der POT-Methode der Extremale Index bei der Tailschätzung berücksichtigt werden. Dies wurde in Borkovec und Klüppelberg (2000) gezeigt. Der Vollständigkeit halber wird der Tailschätzer

$$\widehat{\overline{F}}(u+y) = \frac{N_u^c}{n\hat{\theta}} \left( 1 + \hat{\xi} \frac{y}{\hat{\beta}} \right)^{-1/\hat{\xi}} \quad (3.3)$$

und der Quantilschätzer

$$\hat{x}_p = u + \frac{\hat{\beta}}{\hat{\xi}} \left( \left( \frac{\hat{n}\theta}{N_u^c} (1-p) \right)^{-\hat{\xi}} - 1 \right)$$

angegeben. Dabei bezeichnet  $N_u^c$  die Anzahl der Cluster von Exzessen von  $u$ . Man kann dazu die Daten z.B. in Blöcke einteilen. Es zählt dann jeder Block, in dem ein oder mehrere Datenpunkte die Schwelle übersteigen zu  $N_u^c$ . Cluster können auf verschiedene Weise gebildet werden. Nähere Informationen dazu findet man in Embrechts et al. (2003) und Regler (2004). Die Ergebnisse in Kapitel 5 verändern sich jedoch kaum, wenn man Formel (3.3) statt (2.6) verwendet. Daher wird Formel (3.3) in dieser Arbeit nicht verwendet.

# Kapitel 4

## Nicht-lineare Modelle

Nicht-lineare Zeitreihenmodelle haben sehr viel Interesse auf sich gezogen, denn bei diesen Modellen ist es möglich, dass die Varianz von Informationen aus der Vergangenheit abhängig ist. Man spricht dann auch von bedingter Varianz.

Zunächst werden die Eigenschaften von Finanzzeitreihen, die sog. Stylized Facts, vorgestellt. Anschließend werden einige nicht-lineare Modelle in diskreter Zeit vorgestellt. Es zeigt sich, dass diese Modelle über viele der Stylized Facts verfügen und daher geeignet sind, Finanzzeitreihen zu modellieren.

### 4.1 Stylized Facts von Finanzzeitreihen

Erst die Verfügbarkeit langer Finanzzeitreihen und die ständige Weiterentwicklung der Software zur Analyse dieser Daten macht es möglich, fundierte Aussagen über die statistischen Eigenschaften von Finanzzeitreihen zu treffen. Finanzzeitreihen haben besondere Eigenschaften, die sie von anderen Zeitreihen wie z.B. Versicherungsschäden, unterscheiden. Diese Eigenschaften werden mit dem Begriff „Stylized Facts“ bezeichnet. Das Verständnis dieser Eigenschaften und der dadurch verursachten Probleme und Möglichkeiten bei der Modellierung ist von besonderem Interesse. Essentiell für die Anpassung eines Modells an die verschiedenen Zeitreihen ist die Kenntnis der Verteilung und ihrer Momente in den zugrunde liegenden Daten. Finanzzeitreihen sind im Allgemeinen nicht iid verteilt. Im Folgenden werden einige dieser Stylized Facts näher erläutert. Die Stylized Facts werden dazu in zwei Gruppen aufgeteilt: Besonderheiten hinsichtlich der Verteilung (Distributional Stylized Facts) und Besonderheiten im Zeitablauf (Time Series Stylized Facts). Für weitere Informationen siehe u.a. Cont (2001).

#### 4.1.1 Distributional Stylized Facts

Um einen ersten Eindruck über die Verteilung zu bekommen wird eine nichtparametrische Kerndichteschätzung der Log>Returns der Finanzzeitreihen durchgeführt und mit der

Dichtefunktion einer angepassten Normalverteilung verglichen. Dies ist in Abbildung 4.1 am Beispiel der Log>Returns der Allianz Aktie dargestellt.

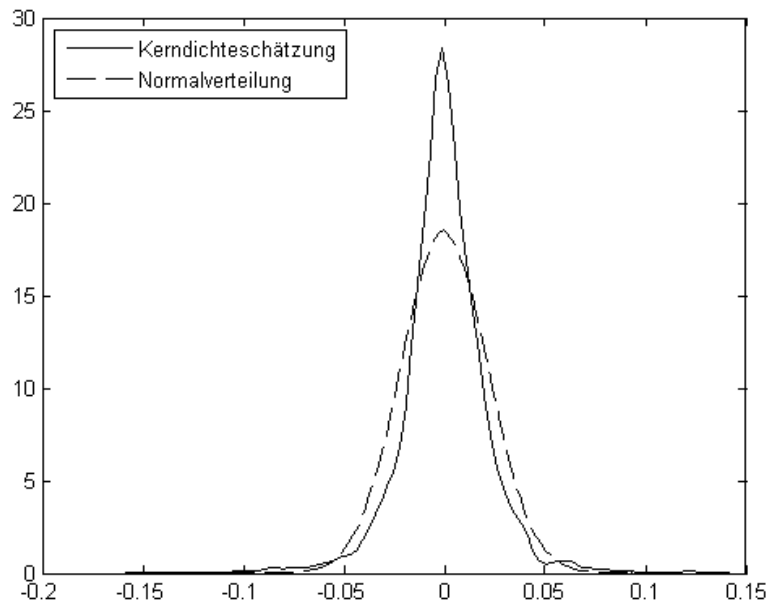


Abbildung 4.1: Nichtparametrische Kerndichteschätzung der Allianz Log>Returns im Vergleich mit der Dichte der angepassten Normalverteilung.

#### (1) Nichtnormalität

Die Renditen von Finanzmarktdaten sind nicht normalverteilt. Optisch ist diese Abweichung deutlich bei der Kerndichteschätzung im Vergleich mit der an die Momente angepassten Normalverteilung zu erkennen. Am Beispiel der Log>Returns der Allianz Aktie in Abbildung 4.1 ist im Vergleich zur Normalverteilung auffallend, dass die empirische Dichtefunktion Spitzgipfligkeit (High-Peakedness), schwere Tails (Heavy Tails), einen schlankeren Mittelbereich (Thin Waists) und eine leichte Schiefe (Skewness) aufweist. Das bedeutet, dass im Vergleich zur Normalverteilung mehr Wahrscheinlichkeitsmasse in den Tails der Verteilung zu finden ist. Ein Modell, das auf der Normalverteilung basiert, würde daher die Eintrittswahrscheinlichkeit von extremen Werten unterschätzen. Dies kann besonders beim Risikomanagement unerwünschte Auswirkungen haben. Finanzzeitreihen sind geprägt von Spitzgipfligkeit, schweren Rändern und schlanken Mittelbereichen. Dementsprechend treten extreme Veränderungen öfter auf, als dies bei der Normalverteilung der Fall ist. Eine weitere Möglichkeit die Nichtnormalität festzustellen, beruht auf Testverfahren, die auf höheren Momenten, in der Regel auf dem dritten und vierten Moment, basieren. Der wohl bekannteste Vertreter ist der Test auf Normalverteilung von Jarque und Bera (1987) ( $\mathcal{JB}$ -Test). Der  $\mathcal{JB}$ -Test stützt sich auf die Schiefe- und des Kurtosiskoeffizienten. Es wird die Hypothese getestet, ob die standardisierten Residuen unabhängig normalverteilt sind.

**(2) Asymmetrie**

Die Verteilung ist nicht symmetrisch, sondern leicht negativ verzerrt. Asymmetrie wird in der Regel deskriptiv mit dem dritten standardisierten Moment, der Schiefe bestimmt. Rechtsschiefe Verteilungen weisen ein positives Schiefemaß auf, während Linksschiefe im Falle eines negativen Schiefemaßes festzustellen ist. Eine Schiefe von Null bedeutet Symmetrie.

**(3) Aggregational Gaussianity (Aggregierte Normalität)**

In der Regel wird davon ausgegangen, dass bei steigendem Stichprobenumfang der zentrale Grenzwertsatz greift und deswegen die Kurtosis tendenziell sinkt. Der zentrale Grenzwertsatz besagt, dass sich die Verteilung der empirischen Daten für eine Anzahl  $n \rightarrow \infty$  der Normalverteilung annähert (Aggregational Gaussianity).

**(4) Existenz niedriger Momente**

Wie viele Momente bei Finanzzeitreihen tatsächlich existieren, ist umstritten. Für tägliche Returns ist mittlerweile die allgemeine Auffassung, dass diese eine endliche Varianz besitzen, wohingegen die Existenz des dritten und vierten Momentes sich immer noch im wissenschaftlichen Disput befindet. Das vierte und das fünfte Moment existieren häufig nicht.

Zusammenfassend kann man festhalten, dass die Unterstellung einer Normalverteilung für die Log>Returns schwer wiegende Folgen hat. Das Auftreten extremer Ereignisse wird unterschätzt, da die Normalverteilung nicht genügend Masse in den Tails hat und damit Extrema nicht ausreichend berücksichtigt.

### 4.1.2 Time Series Stylized Facts

Ein Begriff, der bei der Betrachtung von Finanzzeitreihen häufig auftritt, ist die Volatilität. Eine einheitliche Definition dafür ist schwer zu finden. Der Begriff wird eher intuitiv gebraucht. Ursprünglich versteht man darunter die Standardabweichung oder auch die Varianz. Wenn man die Daten betrachtet und die Volatilität in den Log>Returns erkennen will, achtet man darauf, wie sich die momentane Streuung und die Größe der Ausschläge verändert. Im Allgemeinen ist die Volatilität aber eine unbeobachtete Größe, die erst geschätzt werden muss. Volatilität gibt die Schwankung des Basiswertes, zum Beispiel einer Aktie, um seinen Mittelwert, über einen festgelegten Zeitraum an und gilt als Kriterium zur Beurteilung des Risikos des Basiswertes. Je höher die Volatilität ist, desto größer ist das damit verbundene Risiko, aber auch die Chance, höhere Gewinne zu realisieren.

**(1) Volatilitätsschwankungen**

Die Volatilität ändert sich in der Zeit, d.h. sie ist nicht konstant. Daten, die diese Eigenschaft besitzen, werden auch als (bedingt) heteroskedastisch bezeichnet. Die Volatilität ist zufällig.

**(2) Tailverhalten und Leptokurtosis**

Die Volatilität und die Log>Returns haben schwere Tails. Die Returns weisen Verteilungen auf, die in den Enden mehr Masse haben als die Normalverteilung. Dafür

verwendet man auch den Ausdruck *fat* oder *heavy tailed*. Die Daten verfügen häufig über eine Kurtosis  $\mathcal{K} > 3$ , d.h. sie sind leptokurtisch. Zwischen Volatilitätsclustern und Leptokurtosis besteht ein enger Zusammenhang.

### (3) Volatilitätscluster

Die Volatilität der Log>Returns weist Clusterverhalten auf, d.h. große Ausschläge treten in Gruppen auf. Betrachtet man die Log>Returns im Zeitablauf, so sind im univariaten Fall turbulentere Perioden mit höheren Ausschlägen und ruhigere mit niedrigeren Ausschlägen erkennbar. Dieses Phänomen wird als Volatilitätscluster bezeichnet. Betrachtet man die beiden Zeitreihen der Allianz und der BMW Log>Returns in Abbildung 4.2, so erkennt man, dass die Volatilitätscluster bei den Zeitreihen oftmals simultan verlaufen.

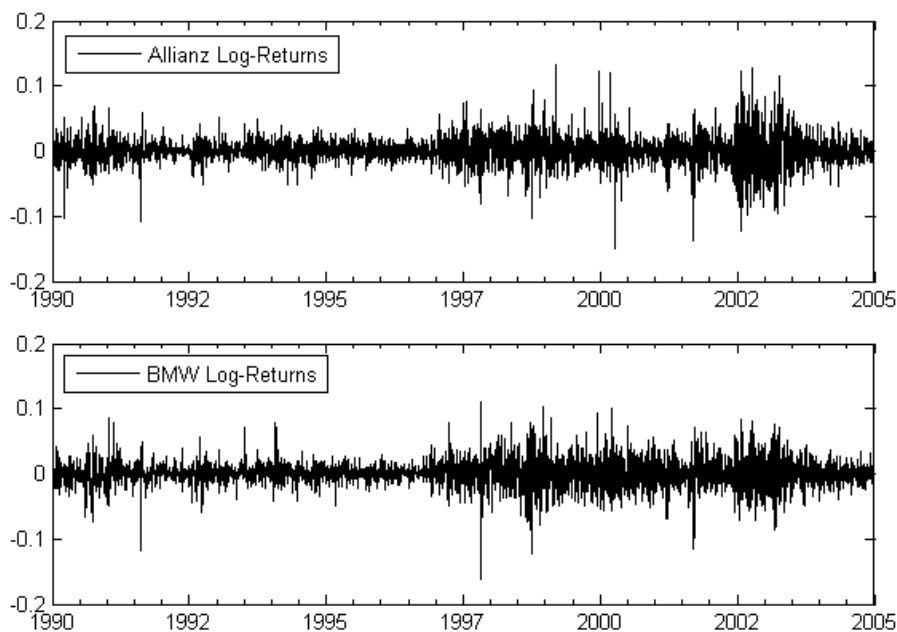


Abbildung 4.2: Vergleich der Volatilitätscluster zwischen Allianz Log>Returns und BMW.

### (4) Autokorrelation und Long-Memory-Eigenschaft

Autokorrelation bedeutet, dass die Daten zeitlich voneinander abhängen. Diese Abhängigkeitsstruktur kann über mehrere Perioden, die so genannten Lags, in die Vergangenheit zurückreichen. Bei Finanzzeitreihen kann man in der Regel keine Abhängigkeit in den Log>Returns erkennen. Die Log>Returns sind meist unkorreliert, aber nicht unabhängig, wie man z.B. an der Autokorrelationsfunktion der quadrierten Log>Returns oder ihrer Absolutbeträge erkennen kann (siehe z.B. Abbildung 5.31). Bei den einfachen Log>Returns bricht die Autokorrelationsfunktion schnell (bei Lag 1) ab, während bei den quadrierten Log>Returns die Lags signifikante Abhängigkeiten zeigen. Diese Abhängigkeiten sind meist über sehr lange



Perioden zu beobachten, so dass auch oft von längerfristiger Abhängigkeit (Long Range Dependence, Long-Memory-Eigenschaft) die Rede ist. Entscheidende Ereignisse beeinflussen in diesem Fall den Markt über längere Zeit hinweg. Die Daten besitzen die sog. Long-Memory-Eigenschaft, d.h. für die Korrelationsfunktion  $\rho(\cdot)$  gilt  $\sum_{h=0}^{\infty} |\rho(h)| = \infty$ .

Der bekannteste Test auf Autokorrelation ist für den univariaten Fall der Test von Ljung und Box (1978) ( $\mathcal{LB}$ -Test), der auch als Portmanteau Statistik bekannt ist. Der  $\mathcal{LB}$ -Test betrachtet die geschätzte Autokorrelationsfunktion. Sind die Residuen unkorreliert, müsste diese etwa gleich Null sein. Weicht die geschätzte Autokorrelationsfunktion zu stark ab, wird die Nullhypothese unkorrelierter Residuen verworfen. Statt den Residuen können auch die Log>Returns mit dem  $\mathcal{LB}$ -Test auf Autokorrelation getestet werden.

#### (5) Leverage Effekt

Der Leverage Effekt beschreibt ein asymmetrisches Verhalten des Kursverlaufs einer Aktie auf Nachrichten. Der Markt reagiert asymmetrisch auf gute und schlechte Nachrichten. Schlechte Nachrichten werden oft überbewertet und führen zu einer höheren Volatilität, während bei guten Nachrichten die Volatilität geringer ausfällt.

## 4.2 Nicht-lineare Zeitreihenmodelle: diskrete Zeit

Besonders für die Modellierung von Finanzzeitreihen benutzt man Modelle der Form

$$X_t = \mu_t + \sigma_t Z_t, \quad t \in \mathbb{N}.$$

Dabei ist  $(Z_t)_{t \in \mathbb{N}}$  ein Weißes Rauschen und unabhängig von  $\sigma_t$ .  $\mu_t$  ist der bedingte Erwartungswert von  $X_t$  gegeben  $X_{t-1}, X_{t-2}, \dots$  und  $\sigma_t$  ist die bedingte Standardabweichung von  $X_t$  gegeben  $X_{t-1}, X_{t-2}, \dots$ .  $(Z_t)_{t \in \mathbb{N}}$  wird als Noise-Variable oder Innovation bezeichnet. Des Weiteren ist  $(X_t)_{t \in \mathbb{N}}$  in der Regel durch die vorangegangenen  $X_{t-1}, X_{t-2}, \dots$  mitbestimmt, so dass zu Beginn einige Zufallsvariablen, die so genannten Start-ZV, gegeben werden müssen. Im Folgenden werden einige heteroskedastische Modelle erläutert. Heteroskedastisch bedeutet, dass die Varianz nicht konstant ist, wohingegen der Begriff homoskedastisch für konstante Varianz steht.

### 4.2.1 ARCH(1) Prozess

1982 führte Robert Engle (1982) ARCH( $p$ ) Prozesse (Auto Regressive Conditionally Heteroscedastic Processes of order  $p$ ) für  $p \in \mathbb{N}$  ein. Sein ARCH Modell (d.h. autoregressiv mit nicht konstanter bedingter Varianz) beschreibt die Entwicklung der Volatilität. Es geht von der Annahme aus, dass die Varianz der zufälligen Modellfehler vom realisierten Zufallsfehler der Vorperiode abhängig ist, so dass große und kleine Fehler dazu tendieren, in Gruppen aufzutreten (Clusterverhalten). Der bedingte Erwartungswert  $\mu_t$  ist in diesem Modell 0. ARCH Modelle sind weit verbreitet, um zeitvariierende Volatilitäten und die Persistenz von Veränderungen auf die Volatilität zu modellieren. Im Folgenden wird

der ARCH Prozess der Ordnung 1 betrachtet. Die Aussagen in diesem Kapitel basieren hauptsächlich auf Engle (1982) und Bollerslev et al. (1994) sowie Embrechts et al. (2003).

**Definition 4.1 (ARCH(1) Prozess, Engle (1982), S. 988)**

Durch

$$X_t = \sigma_t Z_t \text{ mit } Z_t \sim N(0, 1), \text{ wobei} \quad (4.1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2, \quad t \in \mathbb{N} \quad (4.2)$$

wird ein ARCH(1) Prozess definiert. Dabei ist  $X_0$  unabhängig von der symmetrischen iid Folge  $(Z_t)_{t \in \mathbb{N}}$  mit  $\mathbb{E}Z_1^2 < \infty$  und  $\alpha_0, \alpha_1 > 0$ .

**Bemerkung 4.2 (Embrechts et al. (2003), S. 466)**

$\alpha_1$  ist der entscheidende Parameter des ARCH(1) Prozesses:

- für  $\alpha_1 = 0$  ist  $(X_t)_{t \in \mathbb{N}}$  iid normalverteilt,
- für  $\alpha_1 \in (0, 1)$  ist  $(X_t)_{t \in \mathbb{N}}$  stationär mit endlicher Varianz,
- für  $1 \leq \alpha_1 \leq 2e^\gamma \approx 3.56856$  ist  $(X_t)_{t \in \mathbb{N}}$  stationär mit unendlicher Varianz.

In Abbildung 4.3 sind verschiedene ARCH(1) Prozesse simuliert. Je kleiner  $\alpha_1$  ist, desto mehr ähnelt der Prozess dem Weißen Rauschen in Abbildung 1.2 und umso weniger stark geht die Vorperiode ein. Je größer  $\alpha_1$  wird, umso deutlicher sind starke Ausschläge und Clusterbildung zu erkennen. Die Parameter des ARCH(1) Modells können mit Hilfe der Maximum Likelihood Methode geschätzt werden.  $\alpha_0$  und  $\alpha_1$  sind positive Parameter, um die Stationarität zu sichern.  $X_{t-1}$  und  $Z_t$  sind unabhängig für jedes  $t \in \mathbb{N}$  (Embrechts et al. (2003)).

Definition (4.2) lässt sich umschreiben zu

$$X_t = (\alpha_0 + \alpha_1 X_{t-1}^2)^{1/2} Z_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} Z_t \text{ für } t \in \mathbb{N}. \quad (4.3)$$

**Bemerkung 4.3 (Stochastische Rekurrenzgleichung)**

$(X_t)_{t \in \mathbb{N}}$  erfüllt die stochastische Rekurrenzgleichung (SRE)

$$X_t^2 = A_t X_{t-1}^2 + B_t, \text{ wobei } (A_t, B_t) = (\alpha_1 Z_t^2, \alpha_0 Z_t^2).$$

Eine mögliche Interpretation der stochastischen Rekurrenzgleichung ist folgende:  $A_t$  kann als die Veränderung des Zinsfaktors und  $B_t$  als die Einlage und somit die ganze stochastische Differentialgleichung als die Bilanz eines Sparkontos interpretiert werden (Gomes et al. (2004)).

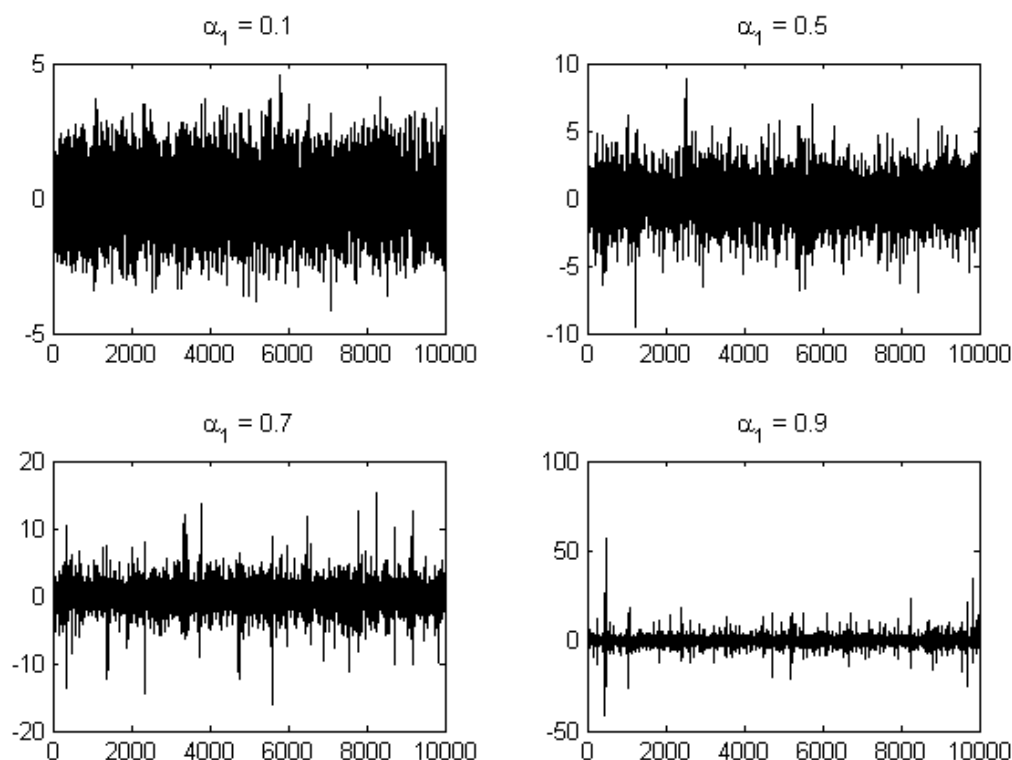
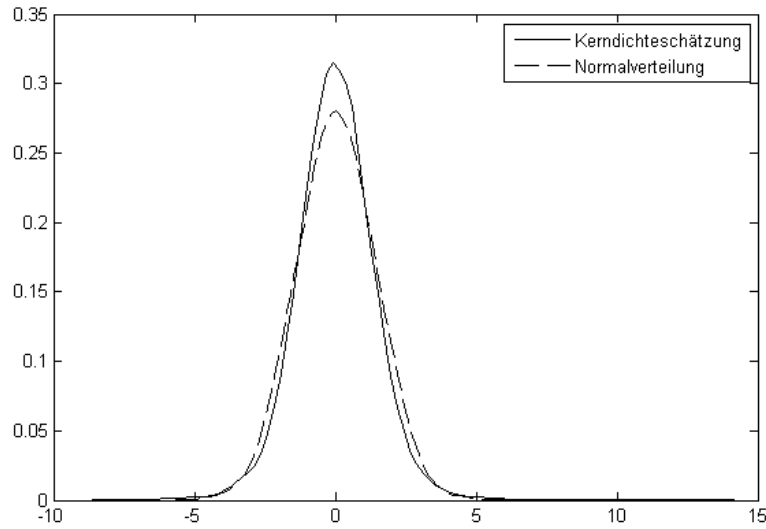


Abbildung 4.3: Simulierte ARCH(1) Prozesse mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1, 0.5, 0.7, 0.9$ ,  $n = 10000$  und normalverteilter Noise-Variable.

### Eigenschaften des ARCH(1) Prozesses

Der unbedingte Erwartungswert des ARCH(1) Prozesses ist Null. Für die unbedingte Varianz gilt:  $\sigma^2 = \mathbb{E}(X_t^2) = \alpha_0 / (1 - \alpha_1)$ . Des Weiteren ist der ARCH(1) Prozess unter bestimmten Voraussetzungen stationär. Außerdem handelt es sich bei ihm auch um einen Markov Prozess. Der ARCH(1) Prozess generiert Daten mit schwereren Tails als die Dichte der Normalverteilung. Dies erkennt man auch in Abbildung 4.4 am Beispiel eines simulierten ARCH(1) Prozesses sowie daran, dass die Kurtosis in Definition 4.4 unter bestimmten Voraussetzungen größer ist als drei.



Abbildungung 4.4: Nichtparametrische Kerndichteschätzung eines simulierten ARCH(1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ , standarnormalverteilter Noise-Variable,  $n = 10000$  und Dichte der an die Momente des ARCH(1) Prozesses angepassten Normalverteilung.

Für die Kurtosis des ARCH(1) Prozesses gilt:

**Satz 4.4 (Kurtosis des ARCH(1) Prozesses)**

$$\mathcal{K}(X_t) = \frac{\mathbb{E}(X_t^4)}{(\mathbb{E}(X_t^2))^2} = \frac{3(1 - \alpha_1^2)}{(1 - 3\alpha_1^2)} = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2} \text{ mit } 0 \leq \alpha_1 < \frac{1}{\sqrt{3}}.$$

Für das vierte Moment des ARCH(1) Prozesses gilt:  $\mathbb{E}(X_t^4) = c < \infty$ . Dann ist

$$\mathbb{E}[X_t^4] = \frac{3\alpha_0^2}{(1 - \alpha_1)^2} \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2}$$

mit  $3\alpha_1^2 < 1$ . Für die Korrelation gilt dann  $\text{corr}(X_t^2, X_0^2) = \alpha_1^t$  für alle  $t \in \mathbb{N}$ .

**Beweis:** Setze

$$\begin{aligned} c &= \mathbb{E}[X_t^4] = \mathbb{E}[\sigma_t^4 Z_t^4] = \mathbb{E}[\sigma_t^4] \mathbb{E}[Z_t^4] = \mathbb{E}[Z_t^4] \mathbb{E}[(\alpha_0 + \alpha_1 X_{t-1}^2)^2] \\ &= 3(\alpha_0^2 + 2\alpha_0\alpha_1 \mathbb{E}[X_{t-1}^2] + \alpha_1^2 \mathbb{E}[X_{t-1}^4]). \end{aligned}$$

Es ist  $\mathbb{E}[X_{t-1}^2] = \alpha_0/(1 - \alpha_1)$  und  $\mathbb{E}[X_{t-1}^4] = c$ . Dann folgt die Behauptung durch Einsetzen und Umformen:

$$\begin{aligned} c &= 3(\alpha_0^2 + 2\alpha_0\alpha_1\alpha_0/(1 - \alpha_1) + \alpha_1^2 c) \\ c - 3\alpha_1^2 c &= \frac{3\alpha_0^2(1 - \alpha_1) + 6\alpha_0^2\alpha_1}{1 - \alpha_1} \\ c &= \frac{3\alpha_0^2 + 3\alpha_0^2\alpha_1}{(1 - \alpha_1)(1 - 3\alpha_1^2)}. \end{aligned}$$

Iterieren der ARCH(1) Gleichung (4.1) führt zu

$$\begin{aligned}\mathbb{E}(X_t^2 X_0^2) &= \mathbb{E}(\sigma_t^2 Z_t^2 \sigma_0^2 Z_0^2) = \mathbb{E}((\alpha_0 + \alpha_1 X_{t-1}^2) Z_t^2 (\alpha_0 + \alpha_1 X_0^2) Z_0^2) \\ &= \dots = \frac{\alpha_0^2}{1 - \alpha_1} \sum_{k=0}^{t-1} \alpha_1^k + \alpha_1^t \mathbb{E}(X_t^4) = (\mathbb{E}(X_t^2))^2 + \alpha_1^t \text{var}(X_t^2).\end{aligned}$$

Weiter folgt mit

$$\sigma_t^2 = \mathbb{E}(X_t^2) = \frac{\alpha_0}{1 - \alpha_1}$$

das Ergebnis. □

Für den Grenzfall  $\alpha_1 = 0$  und normalverteilte Innovationen ist  $\mathcal{K}(X_t) = 3$ , während für  $\alpha_1 > 0$   $\mathcal{K}(X_t) > 3$  gilt. Die unbedingte Verteilung ist also bei bedingter Heteroskedastie leptokurtisch, d.h. die Wölbung im Zentrum der Verteilung ist stärker und die Tails sind dicker als bei der Normalverteilung, was häufig bei Finanzmarktdaten zu beobachten ist. Die Dicke der Tails und damit die Existenz von Momenten und der ACF hängt von den Parametern des ARCH(1) Modells ab. Das vierte Moment im Falle von normalverteilten Innovationen existiert, falls  $3\alpha_1^2 < 1$ . Für die Herleitung der Momente des ARCH(1) Prozesses siehe Bollerslev (1986).

Die folgenden Ergebnisse sind Embrechts et al. (2003) entnommen.

**Lemma 4.5 (Embrechts et al. (2003), Lemma 8.4.6)**

Für eine standardnormalverteilte ZV  $Z_1$  und  $\alpha_1 \in (0, 2e^\gamma)$ , wobei  $\gamma$  die Euler Konstante mit dem näherungsweisen Wert 0.5772 ist, wird definiert:

$$h(u) = \mathbb{E}(\alpha_1 Z_1^2)^u, \quad u \geq 0.$$

Dann folgt

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

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

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

und

$$\mathbb{E}[(\alpha_1 Z_1^2)^\kappa \log(\alpha_1 Z_1^2)] > 0 \quad \text{und} \quad \mathbb{E}[\log(\alpha_1 Z_1^2)] < 0.$$

**Beweis:**  $h(0) = 1$  für alle  $\alpha_1$ . Des Weiteren besitzt  $h$  Ableitungen von jeder Ordnung. Im Besonderen gilt

$$\begin{aligned} h'(u) &= \mathbb{E}((\alpha_1 Z_1^2)^u \log(\alpha_1 Z_1^2)) \\ h''(u) &= \mathbb{E}((\alpha_1 Z_1^2)^u (\log(\alpha_1 Z_1^2))^2) > 0. \end{aligned}$$

Daraus folgt

$$\begin{aligned} h'(0) &= \mathbb{E}(\log(\alpha_1 Z_1^2)) = \log(\alpha_1) + \mathbb{E}(\log(Z_1^2)) \\ &= \log(2\alpha_1) + \Gamma'(1/2)/\Gamma(1/2) = \log(2\alpha_1) - \gamma - 2 \log 2 \\ &= \log(\alpha_1) - \log 2 - \gamma < 0 \end{aligned}$$

für  $0 < \alpha_1 < 2e^\gamma$ , wobei  $\gamma$  die Eulerkonstante ist und  $\Gamma$  die Gamma-Funktion bezeichnet.  $h''(u) > 0$  impliziert, dass  $h$  streng konvex auf  $\mathbb{R}_+$  ist. Mit der Symmetrie der Dichte der Normalverteilung und mit partieller Integration erhält man

$$\begin{aligned} h(u) = \mathbb{E}(\alpha_1 Z^2)^u &= \frac{\alpha_1^u}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2u} e^{-x^2/2} dx \\ &= \frac{\alpha_1^u}{\sqrt{2\pi}} 2 \int_0^{\infty} (2y)^{u-1/2} e^{-y} dy = \frac{2\alpha_1^u}{\sqrt{2}\sqrt{\pi}} 2^{u-1/2} \int_0^{\infty} y^{u-1/2} e^{-y} dy \\ &= \frac{(2\alpha_1)^u}{\sqrt{\pi}} \Gamma(u + 1/2). \end{aligned}$$

Außerdem gilt

$$h(u) \geq \mathbb{E}((\alpha_1 Z_1^2)^u \mathbf{1}_{\{\alpha_1 Z_1^2 > 2\}}) \geq 2^u \mathbb{P}(\alpha_1 Z_1^2 > 2) \rightarrow \infty, \quad u \rightarrow \infty.$$

Zusammen mit  $h(0) = 1$  und der Konvexität von  $h$  impliziert dies, dass ein eindeutiges  $\kappa > 0$  existiert, so dass  $h(\kappa) = 1$  gilt. Mit  $h'(\kappa)$  folgen dann

$$\mathbb{E}[(\alpha_1 Z_1^2)^\kappa \log(\alpha_1 Z_1^2)] > 0 \text{ und } \mathbb{E}[\log(\alpha_1 Z_1^2)] < 0.$$

Da  $h(1) = \alpha_1$  gilt, folgt

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

mit der Monotonie. □

Der Wert von  $\kappa = \kappa(\alpha_1)$  ist ausschlaggebend für die Existenz der Momente und das extremale Verhalten des ARCH(1) Prozesses. Die Gleichung  $h(u) = 1$  kann nicht explizit gelöst werden. Numerische Lösungen werden in Tabelle 4.1 angegeben. Die Tabelle ist an de Haan et al. (1989, S. 222) angelehnt.

$\alpha_1$	0.1	0.3	0.5	0.7	0.9	0.95	0.99
$\kappa$	13.24	4.18	2.37	1.59	1.15	1.07	1.01

Tabelle 4.1: Tailindex  $\kappa$  für ARCH(1) Prozesse.

### Tailverhalten des ARCH(1) Prozesses

Stationäre ARCH(1) Prozesse sind regulär variierend, wie folgender Satz zeigt:

#### Satz 4.6 (Embrechts et al. (2003), Theorem 8.4.12)

Es sei  $(X_t)$  ein stationärer ARCH(1) Prozess mit den Parametern  $\alpha_0 > 0$  und  $\alpha_1 \in (0, 2e^\gamma)$ , wobei  $\gamma$  die Euler Konstante bezeichnet. Sei  $\kappa > 0$  die eindeutige positive Lösung der Gleichung  $h(u) = 1$ , dann gilt

$$\mathbb{P}(X_1 > x) \sim \frac{c}{2} x^{-2\kappa}, \quad x \rightarrow \infty,$$

wobei

$$c = \frac{\mathbb{E} [((\alpha_0 + \alpha_1 X_1^2)^\kappa - (\alpha_1 X_1^2)^\kappa)(Z_1^2)^\kappa]}{\kappa \mathbb{E} [(\alpha_1 Z_1^2)^\kappa \log(\alpha_1 Z_1^2)]} \in (0, \infty) \text{ ist.}$$

Dabei ist  $Z_1$  eine standardnormalverteilte Zufallsvariable, die unabhängig von  $X_1$  ist.  $c$  ist eine konstante langsam variierende Funktion.

Für den Beweis wird auf (Embrechts et al., 2003, S. 470 f.) verwiesen. Es besteht hier die Beziehung  $2\kappa = \frac{1}{\xi}$ . Dabei ist  $\kappa$  der Tailindex, der u.a. mit dem Hill-Schätzer (3.1) ermittelt werden kann.

### 4.2.2 GARCH(1,1) Prozess

In Bollerslev (1986) führt Bollerslev eine Verallgemeinerung von ARCH Prozessen auf die so genannten GARCH Modelle (generalized autoregressive conditional heteroscedasticity) durch. Bei diesen Prozessen hängt die Varianz nicht nur von der Vergangenheit der Zeitreihe ab, sondern auch von ihrer eigenen Vergangenheit. Dadurch besitzt diese Prozessklasse eine flexiblere Abhängigkeitsstruktur, um Log>Returns (siehe Definition 5.1) zu beschreiben. In der Finanzwirtschaft werden GARCH Modelle verwendet, um z.B. Zinsgewinne oder Gewinne bei Währungsgeschäften zu modellieren (siehe auch Brockwell et al. (2005)). Auch in diesem Modell ist der bedingte Erwartungswert  $\mu_t = 0$ . Die Klasse der GARCH Modelle beinhaltet als Spezialfall die ARCH Modelle.

Der einfachste, aber oft sehr nützliche GARCH(1,1) Prozess ist folgendermaßen definiert:

#### Definition 4.7 (GARCH(1,1) Prozess)

$$X_t = \sigma_t Z_t, \text{ wobei} \tag{4.4}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta) \sigma_{t-1}^2, \text{ für } t \in \mathbb{N}. \tag{4.5}$$

Dabei ist  $(Z_t)_{t \in \mathbb{N}_0}$  eine symmetrische iid Folge von Zufallsvariablen mit  $\mathbb{E}(Z_1^2) = 1$  und der Volatilitätsprozess  $(\sigma_t^2)_{t \in \mathbb{N}_0}$  ist die Lösung der stochastischen Rekurrenzgleichung (4.5).

Die Parameter  $\alpha_0$ ,  $\alpha_1$  und  $\beta$  sind nicht-negativ. Im Fall von  $\beta = 0$  reduziert sich das GARCH(1,1) Modell auf ein ARCH(1) Modell. Eine typische Situation bei Finanzmarktdaten ist gegeben durch einen GARCH(1,1) Prozess mit kleinem  $\alpha_1$  und großem  $\beta$ .  $\alpha_0$  wird meistens klein gewählt. In Abbildung 4.5 sind verschiedene GARCH(1,1) Modelle abgebildet.

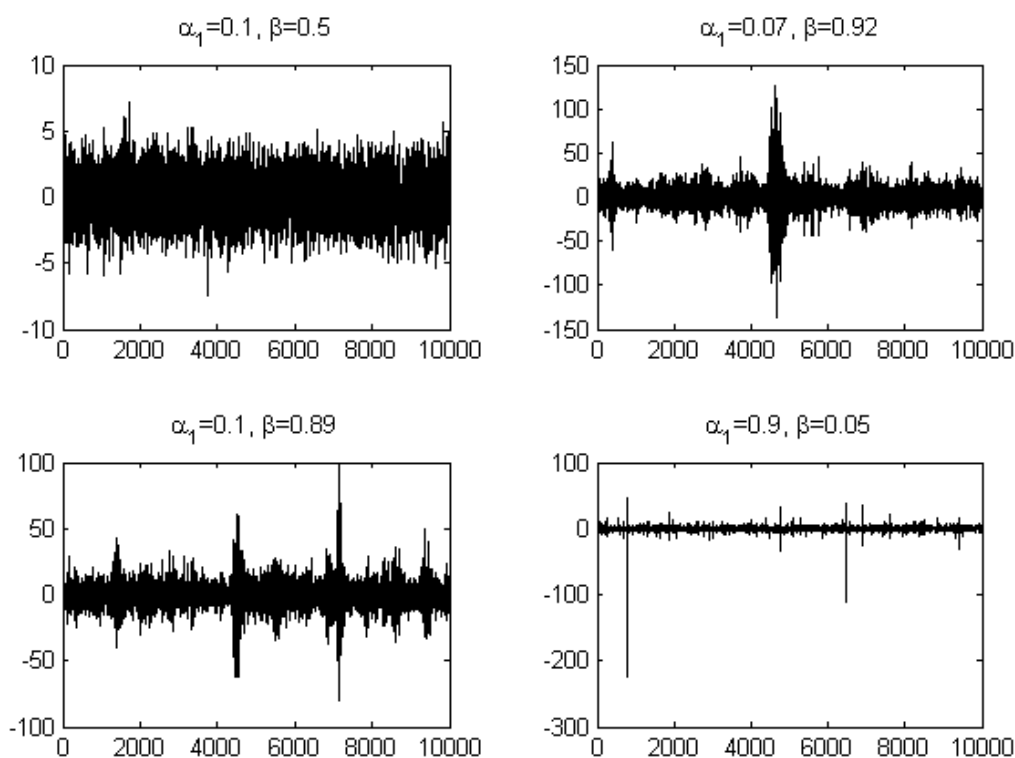


Abbildung 4.5: Simulierte GARCH(1,1) Prozesse mit verschiedenen  $\alpha_0 = 1$ ,  $\alpha_1$ ,  $\beta$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable.

Trotz seiner Einfachheit (nur drei Parameter) verfügt der GARCH(1,1) Prozess über viele der Stylized Facts, wie schwere Tails, Volatilitätscluster etc. (vgl. Mikosch und Stărică (2000) und Kapitel 4.1).

In Abbildung 4.6 wird die nichtparametrische Kerndichteschätzung des GARCH(1,1) Prozesses sowie die Normalverteilung mit dem Erwartungswert und der Standardabweichung des GARCH(1,1) Prozesses geplottet. Man erkennt deutlich, dass keine Normalverteilung vorliegt und, dass die Tails schwerer sind als bei der Normalverteilung.



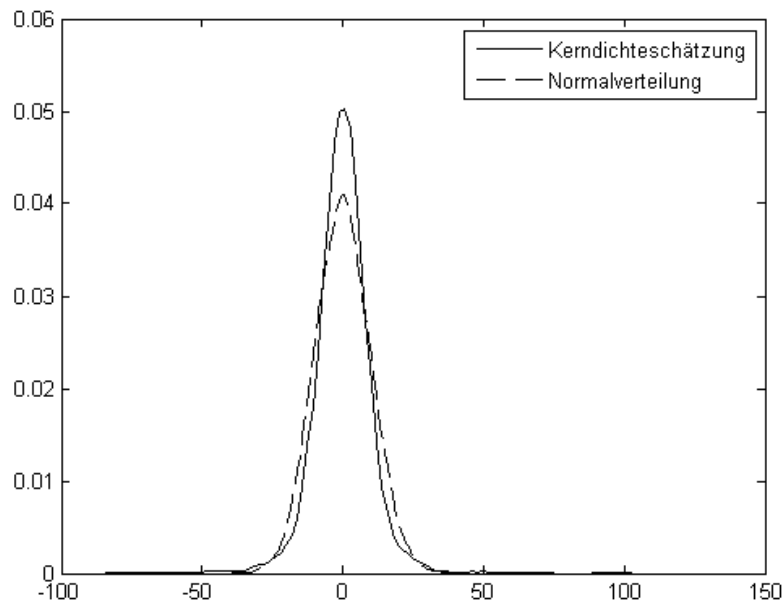


Abbildung 4.6: Nichtparametrische Kerndichteschätzung eines simulierten GARCH(1,1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable,  $n = 10000$  und Dichte der an die Momente angepassten Normalverteilung.

#### Bemerkung 4.8 (Stochastische Rekurrenzgleichung)

Setze  $t \in \mathbb{N}$ ,  $B_t = \alpha_0$  und  $A_t = \alpha_1 Z_{t-1}^2 + \beta$ , dann ist

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 Z_{t-1}^2 + \beta \sigma_{t-1}^2 = \underbrace{(\alpha_1 Z_{t-1}^2 + \beta)}_{=: A_t} \sigma_{t-1}^2 + \underbrace{\alpha_0}_{=: B_t} = A_t \sigma_{t-1}^2 + B_t.$$

Für die stochastische Rekurrenzgleichung folgt dann:

$$\mathbf{X}_t = \begin{pmatrix} X_t^2 \\ \sigma_t^2 \end{pmatrix}, \quad \mathbf{A}_t = \begin{pmatrix} \alpha_1 Z_t^2 & \beta Z_t^2 \\ \alpha_1 & \beta \end{pmatrix} \quad \text{und} \quad \mathbf{B}_t = \begin{pmatrix} \alpha_0 Z_t^2 \\ \alpha_0 \end{pmatrix},$$

also

$$\mathbf{X}_t = \mathbf{A}_t \mathbf{X}_{t-1} + \mathbf{B}_t. \tag{4.6}$$

Der GARCH(1,1) Prozess kann also als Lösung der stochastischen Rekurrenzgleichung (4.6) betrachtet werden, wobei  $(\mathbf{A}_t, \mathbf{B}_t)$  eine iid Folge ist. Für jedes  $t$  ist der Vektor  $\mathbf{X}_{t-1}$  unabhängig von  $(\mathbf{A}_t, \mathbf{B}_t)$ . Die  $(\mathbf{A}_t)_{t \in \mathbb{N}_0}$  sind eine Folge von iid  $d \times d$  Matrizen und die  $\mathbf{B}_t$  sind iid  $d$ -dimensionale Zufallsvektoren.

### Eigenschaften des GARCH(1,1) Prozesses

Der GARCH(1,1) Prozess besitzt ebenso wie der ARCH(1) Prozess die Short-Memory-Eigenschaft, d.h. die Autokorrelationen fallen exponentiell und die Summe der Autokorrelationen ist endlich (siehe Abbildungen 5.2 und 5.13). Die meisten Zeitreihen, vor allem Finanzzeitreihen, besitzen aber die Long-Memory-Eigenschaft, da entscheidende Ereignisse den Markt über längere Zeit hinweg beeinflussen. Die Autokorrelationen klingen also mit wachsendem Lag nur langsam ab, so dass die Summe der Autokorrelationen unendlich ist. Bei der Modellwahl ist auf diese Eigenschaft zu achten.

Der unbedingte Erwartungswert des GARCH(1,1) Prozesses ist Null. Für die unbedingte Varianz gilt:  $\sigma^2 = \mathbb{E}(X_t^2) = \alpha_0 / (1 - \alpha_1 - \beta)$ . Dass die Normalverteilung für den GARCH(1,1) Prozess nicht passend ist, bestätigt auch die Kurtosis.

#### Satz 4.9 (Kurtosis, Bollerslev (1986), S. 312 f.)

Sei  $X_t$  ein GARCH(1,1) Prozess mit  $\text{var}(X_t) = \sigma^2 < \infty$  und  $Z_t \sim N(0, 1)$ . Dann ist  $\mathbb{E}(X_t^4) < \infty$  genau dann, wenn  $3\alpha_1^2 + 2\alpha_1\beta + \beta^2 < 1$ . Die Kurtosis  $\mathcal{K}(X_t)$  ist dann gegeben durch

$$\mathcal{K}(X_t) = \frac{\mathbb{E}(X_t^4)}{(\mathbb{E}(X_t^2))^2} = 3 + \frac{6\alpha_1^2}{1 - \beta^2 - 2\alpha_1\beta - 3\alpha_1^2}.$$

**Beweis:** Das Ergebnis folgt nach einigen Rechnungen aus

$$\begin{aligned} \mathbb{E}(X_t^4) &= \mathbb{E}(Z_t^4)\mathbb{E}(\sigma_t^4) = \mathbb{E}(\sigma_t^4) \frac{\mathbb{E}(Z_t^4)}{1} = \mathbb{E}(\sigma_t^4) \frac{\mathbb{E}(Z_t^4)}{(\mathbb{E}(Z_t^2))^2} \\ &= \mathcal{K}(Z_t)\mathbb{E}(\sigma_t^4) = 3\mathbb{E}(\sigma_t^4) = 3\mathbb{E}[(\alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2)^2] \end{aligned}$$

und der Stationarität von  $X_t$ . Dabei ist

$$\begin{aligned} \mathbb{E}(\sigma_t^4) &= \mathbb{E}[(\alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2)^2] \\ &= \mathbb{E}[\alpha_0^2 + 2\alpha_0\alpha_1 X_{t-1}^2 + 2\alpha_0\beta \sigma_{t-1}^2 + \alpha_1^2 X_{t-1}^4 + 2\alpha_1\beta X_{t-1}^2 \sigma_{t-1}^2 + \beta^2 \sigma_{t-1}^4] \\ &= \alpha_0^2 + 2\alpha_0\alpha_1 \mathbb{E}(X_{t-1}^2) + 2\alpha_0\beta \mathbb{E}(\sigma_{t-1}^2) + \alpha_1^2 \mathbb{E}(\sigma_{t-1}^4 Z_{t-1}^4) \\ &\quad + 2\alpha_1\beta \mathbb{E}(\sigma_{t-1}^4 Z_{t-1}^2) + \beta^2 \mathbb{E}(\sigma_{t-1}^4) \\ &= \frac{\alpha_0^2(1 + \alpha_1 + \beta)}{(1 - \alpha_1 - \beta)} + \mathbb{E}((\alpha_1 Z_{t-1}^2 + \beta)^2) \mathbb{E}(\sigma_{t-1}^4) \\ &= \frac{\alpha_0^2}{1 - \mathbb{E}((\alpha_1 Z_1^2 + \beta)^2)} \frac{1 + \alpha_1 + \beta}{1 - \alpha_1 - \beta} \\ &= \frac{\alpha_0^2(1 + \alpha_1 + \beta)}{(1 - \alpha_1^2 \mathbb{E}(Z_1^4) - 2\alpha_1\beta \mathbb{E}(Z_1^2) - \beta^2)(1 - \alpha_1 - \beta)} \\ &= \frac{\alpha_0^2(1 + \alpha_1 + \beta)}{(1 - 3\alpha_1^2 - 2\alpha_1\beta - \beta^2)(1 - \alpha_1 - \beta)}. \end{aligned}$$

Mit

$$\sigma^2 = \mathbb{E}(X_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta}$$

folgt das Ergebnis.  $\square$

Gelten die Annahmen aus Satz 4.9 so erkennt man, dass für alle Werte  $\alpha_1 > 0$  für die Kurtosis  $\mathcal{K} > 3$  gilt. Damit ist die Verteilung von  $(X_t)_{t \in \mathbb{N}}$  leptokurtisch.

Notwendige und hinreichende Bedingungen für die Stationarität eines GARCH(1,1) Prozesses ergeben sich aus folgendem Satz:

**Satz 4.10 (Mikosch (2004), Theorem 5.3.1)**

*Es existiert eine f.s. eindeutige nicht verschwindende strikt stationäre kausale Lösung der stochastischen Rekurrenzgleichung (4.6) mit  $B_t \equiv \alpha_0$  genau dann, wenn  $\alpha_0 > 0$  und  $\mathbb{E} \log(\alpha_1 Z_1^2 + \beta) < 0$ .*

Im Fall von  $\alpha_0 = 0$  ist  $X_t = 0$  immer die stationäre Lösung. Es wurde jedoch empirisch gezeigt, dass einige Finanzzeitreihen (z.B. Wechselkurse und Zinssätze) typischerweise Parameter aufweisen, die nicht in dem schwach stationären Gebiet liegen (Engle und Bollerslev, 1986). Obwohl diese Zeitreihen nicht quadrat-integrierbar sind, sind sie strikt stationär.

**Tailverhalten des GARCH(1,1) Prozesses**

Stationäre GARCH(1,1) Prozesse sind regulär variierend, d.h. sie besitzen schwere Tails. Dies impliziert, dass der GARCH(1,1) Prozess nie über Momente jeder Ordnung verfügt.

**Satz 4.11 (Mikosch und Stărică (2000), Theorem 2.1)**

*Sei  $A_t = \alpha_1 Z_{t-1}^2 + \beta$ ,  $t \in \mathbb{Z}$ . Angenommen die Verteilung von  $\log A_1$  ist nicht arithmetisch,  $\mathbb{E} \log A_1 < 0$ ,  $\mathbb{P}(A_1 > 1) > 0$  und es gibt  $h_0 \leq \infty$ , so dass  $\mathbb{E} A_1^h < \infty$  für alle  $h < h_0$  und  $\mathbb{E} A_1^{h_0} = \infty$ . Dann gelten folgende Bedingungen:*

(a) Die Gleichung

$$\mathbb{E} A_1^{\kappa/2} = \mathbb{E}(\alpha_1 Z_1^2 + \beta)^{\kappa/2} = 1 \quad (4.7)$$

hat eine eindeutige positive Lösung.

(b) Angenommen  $\alpha_0 > 0$  und  $\kappa$  erfüllen (4.7). Dann existiert eine stationäre Lösung  $(\sigma^2)$  für  $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta$ . Für unabhängige  $A_1, \sigma^2$  mit  $\sigma^2 \stackrel{D}{=} \sigma_1^2$  gibt es eine positive Konstante

$$c = \frac{\mathbb{E}[(\alpha_0 + A_1 \sigma^2)^{\kappa/2} - (A_1 \sigma^2)^{\kappa/2}]}{[(\kappa/2) \mathbb{E} A_1^{\kappa/2} \log A_1]},$$

so dass

$$\mathbb{P}(\sigma > x) \sim cx^{-\kappa}$$

und

$$\mathbb{P}(|X| > x) \sim \mathbb{E}|Z|^\kappa \mathbb{P}(\sigma > x) \text{ für } x \rightarrow \infty.$$

Des Weiteren ist der Vektor  $(X, \sigma)$  regulär variierend mit Index  $\kappa$  und das Spektralmaß auf der Einheitskugel  $\mathbb{S}^1$  ist gegeben durch

$$\mathbb{P}(\theta \in \cdot) = \frac{\mathbb{E}|(Z, 1)|^\kappa \mathbf{1}_{\{|(Z, 1)| \in \cdot\}}}{\mathbb{E}|(Z, 1)|^\kappa}.$$

Für den Beweis wird folgender Satz verwendet:

**Satz 4.12 (Breimans Lemma, Breiman (1965))**

Seien  $X, Y \geq 0$  unabhängige Zufallsvariablen so, dass

$$\mathbb{P}(X > x) \sim cx^{-\kappa}, \quad x \rightarrow \infty,$$

für irgendein  $c, \kappa > 0$ . Weiterhin wird angenommen, dass  $\mathbb{E}(Y^{\kappa+\epsilon}) < \infty$  für ein  $\epsilon > 0$  und  $\mathbb{P}(Y > 0) > 0$  gilt. Dann gilt

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

**Beweis:**

$$\mathbb{P}(XY > x) = \mathbb{P}(XY > x, Y > x) + \mathbb{P}(XY > x, Y \leq x) =: p_1 + p_2.$$

Dann folgt

$$\frac{p_1}{\mathbb{P}(X > x)} \leq \frac{\mathbb{P}(Y > x)}{\mathbb{P}(X > x)} = \frac{\mathbb{E}(\mathbf{1}_{\{Y > x\}})}{\mathbb{P}(X > x)} \leq \frac{\mathbb{E}((Y/x)^\kappa \mathbf{1}_{\{Y > x\}})}{\mathbb{P}(X > x)} \sim \frac{1}{c} \mathbb{E}(Y^\kappa \mathbf{1}_{\{Y > x\}}) \rightarrow 0, \quad x \rightarrow \infty.$$

Mit der Unabhängigkeit von  $X$  und  $Y$  und dem Satz von Fubini folgt für  $p_2$ :

$$\frac{p_2}{\mathbb{P}(X > x)} = \int_0^x \frac{\mathbb{P}(Xy > x)}{\mathbb{P}(X > x)} d\mathbb{P}(Y \leq y).$$

Mit dem Satz von Lebesgue von der majorisierten Konvergenz folgt dann unter Beachtung von  $y \leq x \Rightarrow x/y \geq 1$

$$\begin{aligned} \lim_{x \rightarrow \infty} \int_0^x \frac{\mathbb{P}(X > x/y)}{\mathbb{P}(X > x)} d\mathbb{P}(Y \leq y) &= \int_0^\infty \lim_{x \rightarrow \infty} \frac{\mathbb{P}(X > x/y)}{\mathbb{P}(X > x)} d\mathbb{P}(Y \leq y) \\ &= \int_0^\infty \lim_{x \rightarrow \infty} \frac{c(x/y)^{-\kappa}}{cx^{-\kappa}} d\mathbb{P}(Y \leq y) = \int_0^\infty y^\kappa d\mathbb{P}(Y \leq y) = \mathbb{E}(Y^\kappa) < \infty. \end{aligned}$$

Es folgt:  $\mathbb{P}(XY > x) \sim 0 + \mathbb{E}(Y^\kappa)\mathbb{P}(X > x)$ .  $\square$

**Beweis von Satz 4.11:** Die Funktion  $\mathbb{E}A_1^h$  ist konvex und stetig in  $h$ . Da  $\mathbb{E} \log A_1 < 0$  werden Werte  $< 1$  in der Umgebung des Ursprungs angenommen. Für ausreichend große  $h$  gilt  $\mathbb{E}A_1^h \geq 1$ , da  $\mathbb{P}(A_1 > 1) > 0$  und  $\mathbb{E}A_1^{h_0} = \infty$ . Daher hat  $\mathbb{E}A_1^{\kappa/2} = 1$  eine eindeutige positive Lösung. Wenn  $\alpha_0 > 0$  und  $\mathbb{E} \log(\alpha_1 Z_t^2 + \beta) < 0$  gilt, dann existiert nach Satz 4.10 eine stationäre Lösung ( $\sigma_t^2$ ). Die Annahmen bestärken, dass  $\mathbb{E}A_1^{\kappa/2+c} < \infty$  für kleine  $c > 0$  und alle Bedingungen aus dem Theorem 4.1 von Goldie (1991) erfüllt werden. Das letzte Ergebnis liefert  $\mathbb{P}(\sigma > x) \sim cx^{-\kappa}$ .  $\mathbb{P}(|X| > x) \sim \mathbb{E}|Z|^\kappa \mathbb{P}(\sigma > x)$  ist die Konsequenz aus einem Ergebnis von Breiman (1965) (siehe Satz 4.12). Angenommen  $\xi$  und  $\eta$  sind unabhängige positive Zufallsvariablen, so dass  $\mathbb{P}(\xi > x) = l(x)x^{-\kappa}$  für eine langsam variierende Funktion  $l$  gilt und  $\mathbb{E}\eta^{\kappa+\epsilon} < \infty$  für irgendein  $\epsilon > 0$ . Dann gilt

$$\mathbb{P}(\eta\xi > x) \sim \mathbb{E}\eta^\kappa \mathbb{P}(\xi > x), \quad x \rightarrow \infty.$$

Eine andere Anwendung der Ergebnisse von Breiman (1965) führt für eine beliebige Borelmenge  $B \subset \mathbb{S}^1$  zu

$$\begin{aligned} \mathbb{P}(|(X, \sigma)| > xt, (X, \sigma)/|(X, \sigma)| \in B) &= \mathbb{P}(\sigma|(Z, 1)|\mathbf{1}_{\{(Z, 1)/|(Z, 1)| \in B\}} > xt) \\ &\sim \mathbb{E}|(Z, 1)|^\kappa \mathbf{1}_{\{(Z, 1)/|(Z, 1)| \in B\}} x^{-\kappa} \mathbb{P}(\sigma > t), \\ \mathbb{P}(|(X, \sigma)| > t) = \mathbb{P}(\sigma|(Z, 1)| > t) &\sim \mathbb{E}|(Z, 1)|^\kappa \mathbb{P}(\sigma > t). \end{aligned}$$

$\square$

Aus Mikosch (2004) ist Tabelle 4.2 entnommen.  $\kappa$  bezeichnet den Tailindex, der für standardnormalverteilte Innovationen mittels Monte-Carlo Simulation bestimmt wurde. Es handelt sich um den Tailindex des einfachen Prozesses, also  $\mathbb{P}(X > x) \sim cx^{-\kappa}$ . Der Tailindex des quadrierten Prozesses ist daher  $\kappa/2$ . Für alle Werte wurde  $\alpha_1 = 0.1$  gewählt. Der Tailindex hängt von der Verteilung der Noise-Variable  $(Z_t)_{t \in \mathbb{N}}$  ab (siehe dazu Kiefersbeck (1999)). In dieser Arbeit ist  $(Z_t)_{t \in \mathbb{N}}$  stets standardnormalverteilt.

$\beta$	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
$\kappa$	2	12.4	16.40	19.00	21.00	23.00	24.80	26.40	28.20
$\beta$	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89
$\kappa$	12.00	11.40	10.80	10.00	9.00	8.00	7.00	5.60	3.80

Tabelle 4.2: Tailindex  $\kappa$  für GARCH(1,1) Prozesse mit  $\alpha_1 = 0.1$

Für  $\alpha_1 + \beta = 1$  gilt  $\kappa = 2$ . Dies erkennt man an Gleichung (4.7). Es gilt für standardnormalverteilte Noise-Variablen bei  $\kappa = 2$

$$\mathbb{E}(\alpha_1 Z_1^2 + \beta)^{\kappa/2} = \alpha_1 \mathbb{E}(Z_1^2) + \beta = \alpha_1 + \beta.$$

Der Tailindex kann auch mit dem Hill-Schätzer aus Kapitel 3.1 geschätzt werden. Um allerdings eine einigermaßen sichere Aussage über den Tailindex einer Datenreihe zu erhalten, ist es unabdingbar, verschiedene Schätzmethode miteinander zu vergleichen.

### 4.2.3 AR(1) + ARCH(1) Prozess

Für diesen Prozess werden ein AR(1) (autoregressiver) Prozess und ein ARCH(1) Prozess kombiniert. Der ARCH(1) Prozess stellt dabei den Fehlerterm dar. Der AR(1) Prozess ist wie folgt definiert:

#### Definition 4.13 (AR(1) Prozess)

$(X_t)_{t \in \mathbb{N}}$  heißt autoregressiv der Ordnung 1, falls

$$X_t = \lambda X_{t-1} + Z_t = \lambda^2 X_{t-2} + \lambda Z_{t-1} + Z_t = \dots = \lambda^m X_{t-m} + \sum_{j=0}^{m-1} \lambda^j Z_{t-j}.$$

Dabei ist  $Z_t$  iid mit  $\mathbb{E}(Z_t) = 0$  der Fehlerterm bzw. die Noise-Variable.

Der AR(1) Prozess ist stationär und kausal  $\Leftrightarrow |\lambda| < 1$ .

Gemeinsam mit dem ARCH(1) Prozess aus (4.1) ist der AR(1) + ARCH(1) Prozess folgendermaßen definiert:

#### Definition 4.14 (AR(1) + ARCH(1) Prozess)

$$X_t = \lambda X_{t-1} + \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} Z_t \text{ für } t \in \mathbb{N}, \quad (4.8)$$

wobei  $(Z_t)_{t \in \mathbb{N}}$  iid symmetrische ZV sind,  $\lambda \in \mathbb{R}, \alpha_0, \alpha_1 > 0$  und  $X_0$  ist unabhängig von  $(Z_t)_{t \in \mathbb{N}}$ .

Der bedingte Erwartungswert ist gegeben durch  $\mu_t = \lambda X_{t-1}$ . Dieses Modell kombiniert die Vorteile eines AR(1) Prozesses, der sich auf den bedingten Erwartungswert von  $(X_t)_{t \in \mathbb{N}}$ , gegeben die Vergangenheit, konzentriert, und die Vorteile eines ARCH(1) Modells, das sich auf die bedingte Varianz von  $(X_t)_{t \in \mathbb{N}}$  gegeben die Vergangenheit konzentriert.

#### Bemerkung 4.15 (Stochastische Rekurrenzgleichung)

$$\begin{aligned} X_t^2 &= (\lambda X_{t-1} + \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} Z_t)^2 = \lambda^2 X_{t-1}^2 + 2\lambda X_{t-1} Z_t \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} + \alpha_0 Z_t^2 \\ &\quad + \alpha_1 X_{t-1}^2 Z_t^2 \\ &= \lambda^2 X_{t-1}^2 + \alpha_1 X_{t-1}^2 Z_t^2 + 2\sqrt{\alpha_1} \lambda X_{t-1}^2 Z_t - 2\sqrt{\alpha_1} \lambda X_{t-1}^2 Z_t + 2\lambda X_{t-1} Z_t \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} \\ &\quad + \alpha_0 Z_t^2 \\ &= (\lambda + \sqrt{\alpha_1} Z_t)^2 X_{t-1}^2 + 2\lambda X_{t-1} Z_t (\sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} - \sqrt{\alpha_1} X_{t-1}) + \alpha_0 Z_t^2 \\ &=: A_t X_{t-1}^2 + B_t, \end{aligned}$$

wobei

$$A_t = (\lambda + \sqrt{\alpha_1} Z_t)^2 \text{ und } B_t = 2\lambda X_{t-1} Z_t (\sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} - \sqrt{\alpha_1} X_{t-1}) + \alpha_0 Z_t^2.$$

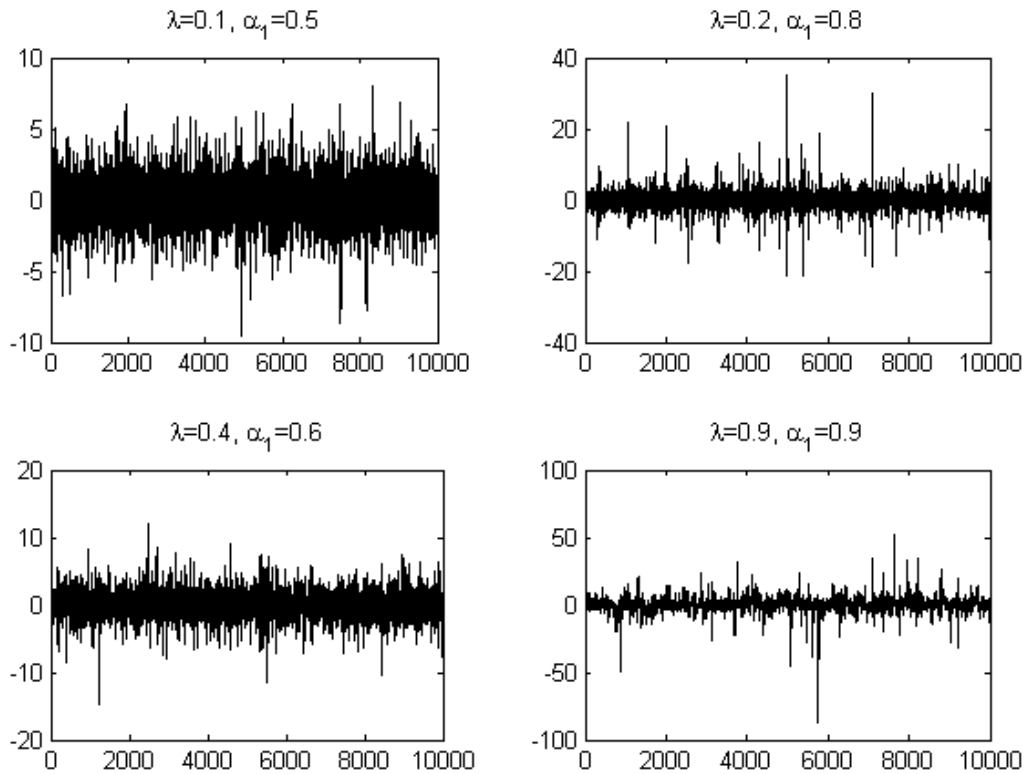


Abbildung 4.7: Simulierte AR(1) + ARCH(1) Prozesse mit  $\alpha_0 = 1$ , verschiedenen  $\lambda$  und  $\alpha_1$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

In Abbildung 4.7 erkennt man, dass  $\alpha_1$  und  $\lambda$  für das extreme Verhalten des Prozesses entscheidend sind.  $\alpha_0$  ändert lediglich die Skalierung.

### Eigenschaften des AR(1) + ARCH(1) Prozesses

Die folgenden Aussagen basieren auf Klüppelberg (2003) sowie Borkovec und Klüppelberg (2001).

Für  $\alpha_1 = 0$  ist der Prozess (4.8) ein AR(1) Prozess, dessen stationäre Verteilung durch die Innovationen  $(Z_t)_{t \in \mathbb{N}}$  bestimmt wird.

Es sei  $Z$  eine generische Zufallsvariable mit der gleichen Verteilungsfunktion  $H$  wie  $Z_t$ .  $Z_t$  sei wie in (4.8) und  $Z \stackrel{D}{=} Z_1$ . Im Folgenden nehmen wir o.B.d.A an, dass  $\lambda \geq 0$  ist und die folgenden generellen Bedingungen gelten:

$$\begin{aligned}
 &\text{Der Träger von } Z \text{ sei ganz } \mathbb{R} : \text{supp}(Z) = \mathbb{R}. \\
 &Z \text{ besitze eine um 0 symmetrische Lebesgue Dichte } p \text{ und} \\
 &Z \text{ besitze ein endliches zweites Moment: } \mathbb{E}(Z^2) < \infty.
 \end{aligned} \tag{4.9}$$

Bei  $(X_t)_{t \in \mathbb{N}}$  handelt es sich um eine homogene Markovkette mit Zustandsraum  $\mathbb{R}$  und mit der Borel- $\sigma$ -Algebra, denn für alle  $t \in \mathbb{N}$  und alle  $B \in \mathcal{B}(\mathbb{R})$  gilt:

$$\begin{aligned} \mathbb{P}(X_t \in B | X_{t-1} = x_1, \dots, X_0 = x_t) &= \mathbb{P}(X_{t-1} + \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} Z_t | X_{t-1} = x_1, \dots, X_0 = x_t) \\ &= \mathbb{P}(x_1 + \sqrt{\alpha_0 + \alpha_1 x_1^2} Z_t \in B), \end{aligned}$$

da  $Z_t$  unabhängig von  $X_{t-1}, X_{t-2}, \dots$  ist. Für  $X$  gilt die Fixpunktgleichung

$$X \stackrel{D}{=} \lambda X + \sqrt{\alpha_0 + \alpha_1 X^2} Z,$$

wobei  $Z$  unabhängig von  $X$  ist. Zusätzlich sollen folgende drei Annahmen gelten:

(D1)  $p(x) \geq p(x'), \forall x' \geq x > 0$ .

(D2) Für jedes beliebige feste  $c > 0$  existieren eine Konstante  $q = q(c) \in (0, 1)$  und Funktionen  $f_+(c, \cdot)$  bzw.  $f_-(c, \cdot)$  mit  $f_+(c, x) \xrightarrow{x \rightarrow \infty} 1$  bzw.  $f_-(c, x) \xrightarrow{x \rightarrow \infty} 1$ , so dass  $\forall x > 0$  und  $t > x^q$

$$\begin{aligned} p\left(\frac{x+c+\lambda t}{\sqrt{\alpha_0 + \alpha_1 t^2}}\right) &\geq p\left(\frac{x+\lambda t}{\sqrt{\alpha_0 + \alpha_1 t^2}}\right) f_+(c, x), \\ p\left(\frac{x+c-\lambda t}{\sqrt{\alpha_0 + \alpha_1 t^2}}\right) &\geq p\left(\frac{x-\lambda t}{\sqrt{\alpha_0 + \alpha_1 t^2}}\right) f_-(c, x). \end{aligned}$$

(D3) Es existiere eine Konstante  $\eta > 0$ , so dass mit  $q$  aus (D2) und  $N := \inf\{u \geq 0; \mathbb{E}(|\sqrt{\alpha_1} \varepsilon|^u > 2)\}$  gilt:

$$\lim_{x \rightarrow \infty} p(x) x^{(N+1+\eta+3q)/(1-q)} = 0.$$

Die Normal-, die  $t$ - und die Laplaceverteilung erfüllen die Bedingungen (D1) – (D3). Die folgende Proposition beschreibt die Eigenschaft des Erwartungswertes  $\mathbb{E}(|\lambda + \sqrt{\alpha_1} Z|^u)$  in Abhängigkeit von der Potenz  $u$ . Dieser Erwartungswert spielt eine Rolle für die Stationarität des AR(1) + ARCH(1) Prozesses.

**Proposition 4.16 (Borkovec und Klüppelberg (2001), Proposition 2)**

*Z sei eine ZV mit Wahrscheinlichkeitsdichte  $p$ , die die generellen Bedingungen (4.9) und (D1) – (D3) erfüllt. Sei  $h : \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}$  definiert durch*

$$h(\lambda, \alpha_1, u) = \mathbb{E}(|\lambda + \sqrt{\alpha_1} Z|^u), \quad u \geq 0.$$

*Für feste  $(\lambda, \alpha_1)$  werde die Funktion  $h(\lambda, \alpha_1, \cdot)$  als  $h_{\lambda, \alpha_1}(\cdot)$  bezeichnet.*

*Es gilt: Die Funktion  $h_{\lambda, \alpha_1}$  ist in  $[0, T)$  strikt konvex, wobei  $T := \inf\{u \geq 0; \mathbb{E}(|\sqrt{\alpha_1} Z|^u) = \infty\}$ . Falls die Parameter  $\lambda$  und  $\alpha_1$  so gewählt werden, dass*

$$\frac{d}{du} h(\lambda, \alpha_1, u) \Big|_{u=0} = h'_{\lambda, \alpha_1}(0) = \mathbb{E}(\log |\lambda + \sqrt{\alpha_1} Z|) < 0, \quad (4.10)$$



dann existiert eine eindeutige Lösung  $\kappa = \kappa(\lambda, \alpha_1) > 0$  für die Gleichung  $h_{\lambda, \alpha_1}(u) = 1$ .  
Für  $h'_{\lambda, \alpha_1}(0) < 0$  gilt

$$\kappa(\lambda, \alpha_1) \begin{cases} > 2, & \lambda^2 + \alpha_1 \mathbb{E}(Z^2) < 1, \\ = 2, & \lambda^2 + \alpha_1 \mathbb{E}(Z^2) = 1, \\ < 2, & \lambda^2 + \alpha_1 \mathbb{E}(Z^2) > 1. \end{cases}$$

**Beweis (vgl. Kiefersbeck (1999)):** Der Beweis erfolgt ähnlich zu dem Beweis für Satz 4.5. Für die ersten Ableitungen von  $h_{\lambda, \alpha_1}(u)$  gilt:

$$\begin{aligned} h'_{\lambda, \alpha_1}(u) &= \mathbb{E}(|\lambda + \sqrt{\alpha_1}Z|^u \log |\lambda + \sqrt{\alpha_1}Z|), \\ h''_{\lambda, \alpha_1}(u) &= \mathbb{E}(|\lambda + \sqrt{\alpha_1}Z|^u (\log |\lambda + \sqrt{\alpha_1}Z|)^2) > 0, \quad \forall u \in [0, T]. \end{aligned} \quad (4.11)$$

Wegen  $\text{supp}(Z) = \mathbb{R}$  ist der Erwartungswert in (4.11) echt positiv. Aus der ersten Ableitung ergibt sich die Gleichheit in (4.10) und aus (4.11) folgt die strikte Konvexität von  $h_{\lambda, \alpha_1}$  in  $[0, T)$ .

O.B.d.A. gelte  $\lambda \geq 0$ . Wegen der Symmetrie von  $Z$  folgt

$$h_{-\lambda, \alpha_1}(u) = \mathbb{E}(|-\lambda + \sqrt{\alpha_1}Z|^u) = \mathbb{E}(|-\lambda - \sqrt{\alpha_1}Z|^u) = \mathbb{E}(|\lambda + \sqrt{\alpha_1}Z|^u) = h_{\lambda, \alpha_1}(u).$$

Es gilt  $h_{\lambda, \alpha_1}(0) = \mathbb{E}(|\lambda + \sqrt{\alpha_1}Z|^0) = 1$ .

Außerdem gilt:

$$h_{\lambda, \alpha_1}(u) \geq \mathbb{E}(|\lambda + \sqrt{\alpha_1}Z|^u \mathbf{1}_{\{\sqrt{\alpha_1}Z \geq 2\}}) \geq 2^u \mathbb{P}(\sqrt{\alpha_1}Z \geq 2).$$

Zusammen mit der Konvexität und  $h_{\lambda, \alpha_1}(0) = 1$  impliziert dies, dass ein eindeutiges  $\kappa > 0$  mit  $h_{\lambda, \alpha_1}(\kappa) = 1$  existiert. Angenommen es gelte  $0 < \kappa_1 < \kappa_2$  und  $h_{\lambda, \alpha_1}(\kappa_1) = h_{\lambda, \alpha_1}(\kappa_2) = 1$ , dann gilt

$$h_{\lambda, \alpha_1}(0) = h_{\lambda, \alpha_1}(\kappa_1) = h_{\lambda, \alpha_1}(\kappa_2)$$

im Widerspruch zur Konvexität.

Weiter gilt

$$\begin{aligned} h_{\lambda, \alpha_1}(2) &= \mathbb{E}(|\lambda + \sqrt{\alpha_1}Z|^2) = \mathbb{E}((\lambda + \sqrt{\alpha_1}Z)^2) \\ &= \mathbb{E}(\lambda^2 + 2\sqrt{\alpha_1}\lambda Z + \alpha_1 Z^2) \\ &= \lambda^2 + \alpha_1 \mathbb{E}(Z^2), \quad \text{da } \mathbb{E}(Z) = 0. \end{aligned}$$

□

Die Werte in Tabelle 4.3 für den Tailindex  $\kappa$  für AR(1) + ARCH(1) Prozesse mit standardnormalverteilter Noise-Variable sind Kiefersbeck (1999) entnommen.

$\alpha_1$	0.2	0.4	0.6	0.8
$\lambda$	0.2	0.2	0.2	0.2
$\kappa$	11.00	5.50	3.54	2.52

Tabelle 4.3: Tailindex  $\kappa$  für AR(1) + ARCH(1) Prozesse.

**Bemerkung 4.17**

- (1)  $\kappa$  ist eine Funktion von  $\lambda$  und  $\alpha_1$ . Da  $Z$  symmetrisch ist, hängt  $\kappa$  nicht vom Vorzeichen von  $\lambda$  ab. Für ein festes  $\lambda$  und  $Z \sim N(0, 1)$  kann man zeigen, dass  $\kappa$  fallend ist in  $|\lambda|$ .
- (2) Es kann gezeigt werden, dass für ein festes  $\alpha_1$  der Exponent  $\kappa$  in  $|\lambda|$  fallend ist. Das bedeutet, dass die Verteilung von  $X$  schwerere Tails hat, wenn  $|\lambda|$  ansteigt. Im Besonderen hat der AR(1) Prozess mit ARCH(1) Fehlern für  $\lambda \neq 0$  schwerere Tails als der ARCH(1) Prozess. Siehe dazu auch Tabelle 3 in Borkovec und Klüppelberg (2000).
- (3) Generell ist es nicht möglich explizit zu bestimmen, welche Parameter  $\lambda$  und  $\alpha_1$   $h'_{\lambda, \alpha_1}(0) = \mathbb{E}(\log |\lambda + \sqrt{\alpha_1} Z|) < 0$  erfüllen. Falls  $\lambda = 0$  (das ist der ARCH(1) Fall) und  $Z \sim N(0, 1)$ , dann ist die Gleichung für  $h'_{\lambda, \alpha_1}(0)$  genau dann erfüllt, wenn  $\alpha_1 \in (0, 2e^\gamma)$ , wobei  $\gamma$  die Euler Konstante ist.
- (4) Jensens Ungleichung  $\lambda^2 + \alpha_1 \mathbb{E}(Z^2) < 1$  impliziert  $h'_{\lambda, \alpha_1}(0) < 0$ .

**Tailverhalten eines AR(1) + ARCH(1) Prozesses**

Es wird erwartet, dass  $\alpha_1$  und  $\lambda$  den Tailindex einer stationären Verteilung des AR(1) + ARCH(1) Prozesses bestimmen.

**Satz 4.18 (Borkovec und Klüppelberg (2001), Theorem 3)**

Sei  $(X_t)_{t \in \mathbb{N}}$  gegeben durch (4.8) mit  $(Z_t)_{t \in \mathbb{N}}$ , das die generellen Bedingungen (4.9) und (D1)-(D3) erfüllt und mit den Parametern  $\alpha_1$  und  $\lambda$ , die  $h'_{\lambda, \alpha_1}(0) = \mathbb{E}(\log |\lambda + \sqrt{\alpha_1} Z|) < 0$  erfüllen. Sei  $\bar{F}(x) = \mathbb{P}(X > x)$ ,  $x \geq 0$  der rechte Tail der stationären Verteilungsfunktion, so ist

$$\mathbb{P}(X > x) \sim cx^{-\kappa}, \quad x \rightarrow \infty,$$

wobei

$$c = \frac{1}{2\kappa} \frac{\mathbb{E}(|\lambda|X| + \sqrt{\alpha_0 + \alpha_1 X^2} Z|^\kappa - |(\lambda + \sqrt{\alpha_1} Z)|X|^\kappa)}{\mathbb{E}(|\lambda + \sqrt{\alpha_1} Z|^\kappa \log |\lambda + \sqrt{\alpha_1} Z|)}$$

und  $\kappa$  die eindeutige positive Lösung von

$$\mathbb{E}(|\lambda + \sqrt{\alpha_1} Z|^\kappa) = 1 \tag{4.12}$$

ist.  $c$  entspricht auch hier einer konstanten langsam variierenden Funktion.

## 4.3 Extremale Maße

### 4.3.1 Anforderungen an Extremale Maße

Mari und Kotz (2001, S. 150) haben folgende acht allgemeine Anforderungen an Abhängigkeitsmaße definiert. Diese treffen bis auf (5) und (6) auch auf die Maße, die in diesem Kapitel angegeben werden, zu.

- (1) Standardisierung  
Eine Standardisierung der Werte soll zwischen 0 und 1 stattfinden.
- (2) Unabhängigkeit  
Bei Unabhängigkeit zwischen zwei Zufallsvariablen  $X$  und  $Y$  soll das Maß 0 sein.
- (3) Funktionale Abhängigkeit  
Bei einer funktionalen Abhängigkeit zwischen beiden Variablen soll das Maß den Wert 1 haben.
- (4) Äquivalenz  
Wenn die Abhängigkeit steigt, soll auch der Wert des Maßes steigen.
- (5) Persistenz/Invarianz  
Auch nach linearer Transformation der Zufallsvariablen soll das Maß gleich bleiben.
- (6) Symmetrie  
Wenn die Variablen vertauscht werden, sollte der Wert des Maßes symmetrisch sein.
- (7) Beziehung zu anders skalierten Variablen  
Wenn das Maß sowohl für ordinale, als auch für stetige Variablen definiert ist, so sollte zwischen beiden Werten des Maßes ein starker Zusammenhang zu erkennen sein.
- (8) Interpretierbarkeit  
Der Wert des Maßes soll eine qualitative Aussage möglich machen.

Das im Folgenden vorgestellte  $\chi(\cdot)$  ist aufgrund des Definitionsbereiches  $[0, 1]$  nicht symmetrisch. Bei linearer Transformation der Daten verändern sich zudem die Ergebnisse für das Maß. Die Berechnung von einem extremalen Abhängigkeitsmaß  $\chi(\cdot)$  basiert auf der „extreme dependence function“, die ursprünglich für räumliche Daten entwickelt wurde. Für die Beziehung zwischen  $\chi(\cdot)$  und der „extreme dependence function“ wird auf Beirlant et al. (2004) verwiesen.

Es sollen die extremalen Maße für die verschiedenen Modelle, die in Abschnitt 4.2 dargestellt wurden, berechnet werden. Die extremalen Maße, die im Folgenden vorgestellt werden sind für den quadrierten Prozess hergeleitet. Durch Variation des Tailindex  $\kappa$  kann man sie aber auch als Alternative für die Korrelationsfunktion der einfachen bzw. der absoluten Werte verwenden.

### 4.3.2 Theoretische Berechnung

Dieser Abschnitt basiert auf Fasen und Klüppelberg (2006).

#### Definition 4.19 (Extremale Maße)

Sei  $(X_t)_{t \in I}$  ein stochastischer Prozess.  $H$  ist eine Extremwertverteilung und  $(X_t)_{t \in I}$  liegt im Anziehungsbereich von  $H$ .  $I$  bezeichnet eine stetige oder diskrete Indexmenge. O.B.d.A. gilt  $\inf I = 0$ .

Für  $k \in \mathbb{N}$  seien  $t_1 < \dots < t_k$  Zeitpunkte in  $I$  und  $x_1, \dots, x_k \in \text{supp}(H)$ . Dann wird die Extremale Koeffizienten Funktion (extremal coefficient function) definiert durch

$$\chi(t) = \lim_{x \uparrow x_R} \mathbb{P}(X_t > x \mid X_0 > x) \text{ für } t \in I. \quad (4.13)$$

#### Bemerkung 4.20

Abhängigkeiten in den Extremen eines stationären stochastischen Prozesses werden mit dem Extremalen Index  $\theta$  gemessen.

$$\theta = \lim_{n \rightarrow \infty} \lim_{x \uparrow x_R} \mathbb{P}(\max_{i=1, \dots, n} X_i \leq x \mid X_0 > x). \quad (4.14)$$

Dieser ist, wie bereits in Kapitel 2.2 erwähnt, ein Maß für das Clusterverhalten von Extremen über hohen Schwellen.

#### Das extremale Maß für das ARCH(1) Modell

**Satz 4.21** Der stationäre ARCH(1) Prozess  $(X_t)_{t \in \mathbb{N}}$  aus (4.1) erfülle die Bedingungen aus Satz 4.6. Das extremale Maß für den quadrierten ARCH(1) Prozess  $(X_t^2)_{t \in \mathbb{N}}$  ist dann

$$\chi(k) = \mathbb{E} \min \left\{ 1, \left( \alpha_1^{\kappa/2} \right)^k \prod_{i=1}^k |Z_i^2|^{\kappa/2} \right\}, \text{ für } k \in \mathbb{N}. \quad (4.15)$$

Der Extremale Index ist durch

$$\theta = \mathbb{E} \left( 1 - \bigvee_{k=1}^{\infty} \left( \alpha_1^{\kappa/2} \right)^k \prod_{i=1}^k |Z_i^2|^{\kappa/2} \right)^+$$

gegeben.

Für den Beweis wird die SRE aus Bemerkung 4.3 sowie Ergebnisse zur multivariaten regulären Variation verwendet. Siehe dazu Vervaat (1979) und Goldie (1991). Der Beweis für  $\theta$  folgt aus de Haan et al. (1989).

In Embrechts et al. (2003) sind in Tabelle 8.4.23 für verschiedene Werte von  $\alpha_1$  Berechnungen für den Extremalen Index angegeben, die in Tabelle 4.4 dargestellt sind.

$\alpha_1$	0.1	0.3	0.5	0.7	0.9	0.95	0.99
$\hat{\theta}$	0.999	0.939	0.835	0.721	0.612	0.589	0.571

Tabelle 4.4: Extremaler Index  $\hat{\theta}$  für ARCH(1) Prozesse.

### Das extreme Maß für das GARCH(1,1) Modell

**Satz 4.22** *Der stationäre GARCH(1,1) Prozess  $(X_t)_{t \in \mathbb{N}}$  aus (4.4) erfülle die Bedingungen aus Satz 4.11. Das extreme Maß für den quadrierten GARCH(1,1) Prozess  $(X_t^2)_{t \in \mathbb{N}}$  ist*

$$\chi(k) = \frac{\mathbb{E} \min \left\{ |Z_0|^\kappa, |Z_k|^\kappa \prod_{i=1}^k (\alpha_1 Z_{i-1}^2 + \beta)^{\kappa/2} \right\}}{\mathbb{E} |Z_0|^\kappa}, \text{ für } k \in \mathbb{N}. \quad (4.16)$$

Der Extreme Index ist dann

$$\theta = \frac{\mathbb{E} \left( |Z_0|^\kappa - |Z_k|^\kappa \prod_{i=1}^k (\alpha_1 Z_{i-1}^2 + \beta)^{\kappa/2} \right)^+}{\mathbb{E} |Z_0|^\kappa}.$$

Der Beweis erfolgt mit Satz 4.11, sowie mit Ergebnissen zur multivariaten regulären Variation (siehe dazu Angaben zu Satz 4.21).

Mit Theorem 2.3 in Mikosch und Stărică (2000) erfolgt der Beweis für  $\theta$ .

### Das extreme Maß für das AR(1) + ARCH(1) Modell

Der stationäre AR(1)+ ARCH(1) Prozess  $(X_t)_{t \in \mathbb{N}}$  aus (4.8) erfülle die Bedingungen aus Satz 4.18. Das extreme Maß für den quadrierten AR(1) + ARCH(1) Prozess  $(X_t^2)_{t \in \mathbb{N}}$  ist

$$\chi(k) = \mathbb{E} \min \left\{ 1, \left( \prod_{i=1}^k (\lambda + \sqrt{\alpha_1} Z_i)^2 \right)^{\kappa/2} \right\}, \text{ für } k \in \mathbb{N}. \quad (4.17)$$

Der Extreme Index  $\theta$  ist gegeben durch

$$\theta = \mathbb{E} \left( 1 - \prod_{k=1}^{\infty} \prod_{i=1}^k |\lambda + \sqrt{\alpha_1} Z_i|^\kappa \right)^+.$$

Der Beweis für das Extreme Maß erfolgt wie im ARCH(1) und GARCH(1,1) Fall. Für den Beweis des Extremalen Index siehe Borkovec (2000).

# Kapitel 5

## Eigene Studie

Für den empirischen Teil der Arbeit werden zunächst simulierte Prozesse und anschließend Finanzmarktdaten verwendet. Bei den Finanzmarktdaten handelt es sich um die Kurse der Allianz AG. Finanzzeitreihen sind in der Regel korreliert und nicht iid. Aufeinanderfolgende Preise weisen häufig eine starke Abhängigkeit sowie einen Trend auf. Folglich sind Preise in der Regel nicht stationär. Daher werden nicht Preise sondern Preisänderungen, die sog. Log>Returns betrachtet, die das Kriterium der Stationarität erfüllen. Aktienkurse, Aktienindizes, Wechselkurse etc. wachsen exponentiell im Zeitablauf. Daher können die logarithmierten Daten als stationär angesehen werden. Man trifft die Annahme, dass die Returns einem stationären Zeitreihenmodell mit stochastischer Volatilitätsstruktur folgen. Das Vorkommen von stochastischer Volatilität impliziert aber, dass die Log>Returns im Zeitverlauf nicht unabhängig sind. Log>Returns sind folgendermaßen definiert:

### Definition 5.1

Sei  $(Y_t)_{t \in \mathbb{N}_0}$  ein Preisprozess, insbesondere sind alle  $Y_t > 0$ . Dann heißt  $(\log Y_t)_{t \geq 0}$  der Logpreisprozess. Des Weiteren setzt man

$$X_t := \log Y_t - \log Y_{t-1} = \log(Y_t/Y_{t-1}), \quad t \in \mathbb{N}$$

und nennt die  $X_t$  Log>Returns.

Diese Definition kann man auch auf die Berechnung der Log>Returns von Wechselkursen etc. übertragen. Mit der Formel von Taylor folgt:

$$\log \frac{Y_t}{Y_{t-1}} = \log \left( 1 + \frac{Y_t - Y_{t-1}}{Y_{t-1}} \right) \approx \frac{Y_t - Y_{t-1}}{Y_{t-1}}.$$

Ein Vorteil der Transformation der Daten in Log>Returns ist, dass die Prozesse unabhängig von den zugrunde liegenden Währungen sind und daher verschiedene Renditeprozesse verglichen werden können. Nur die relative Veränderung im Zeitablauf ist von Bedeutung.

## 5.1 Simulation

Das Abhängigkeitsmaß  $\chi(\cdot)$  soll mit den Formeln aus Kapitel 4.3.2 bestimmt und anschließend mit den Methoden aus Kapitel 3.3 geschätzt werden. Bei den Formeln wird darauf verzichtet den Wert für den Lag 0 zu berechnen, da dieser in jedem Fall gleich eins ist, wie es auch bei der ACF der Fall ist. Mit  $\chi(\cdot)$  kann man zum Einen eine Aussage über die Abhängigkeit innerhalb der Daten machen, zum Anderen kann man eine Aussage darüber machen, ob ein bestimmtes Modell für einen Datensatz in Frage kommt. Es werden die nicht-linearen Modelle aus Kapitel 4 betrachtet. Mit Hilfe von Matlab werden Daten simuliert. Mit den Formeln aus Kapitel 4.3.2 für die extremalen Maße der verschiedenen nicht-linearen Modelle wird das Extremale Maß für die simulierten Daten mittels Monte-Carlo Simulation bestimmt. Die Werte für  $\kappa$  aus den Tabellen 4.1, 4.2 und 4.3 werden verwendet. Anschließend wird  $\chi(\cdot)$  geschätzt, indem der Tail für die simulierten Daten und der Tail für einen zweiten Datensatz, der aus den simulierten Daten konstruiert wird, mit den empirischen Schätzungen aus 3.3.1, Blockmaxima-Methode aus 3.3.2, der Vorgehensweise aus 3.3.3 (Methode von McNeil) und der POT-Methode aus 3.3.4 geschätzt wird. Es interessiert vorerst, ob mit den Methoden aus Kapitel 3.3 eine gute Schätzung von  $\chi(\cdot)$  möglich ist. Zunächst wird dieses Vorgehen für die Modelle aus Kapitel 4 durchgeführt. Die Formeln in Kapitel 4.3.2 sind stets für den quadrierten Prozess  $(X_t^2)_{t \in \mathbb{N}}$ . Daher werden bei der Anwendung der verschiedenen Methoden aus Kapitel 3.3 quadrierte Daten verwendet. Bei der Verwendung von  $\kappa$  muss darauf geachtet werden, ob es sich um das  $\kappa$  für den quadrierten oder für den einfachen Prozess handelt. Die Formeln für die extremalen Maße in Kapitel 4.3.2 sind so aufgebaut, dass stets das  $\kappa$  des einfachen Prozesses verwendet, jedoch  $\chi(\cdot)$  für den quadrierten Prozess berechnet wird.

Eine Schwierigkeit bei den Methoden aus Kapitel 3.3 stellt die Wahl der Schwelle  $u$  dar. Der ME-Plot bietet zwar ein Hilfsmittel, ist aber, wie in Kapitel 3.2 beschrieben, nicht eindeutig. Daher wird im Folgenden für die Schwelle  $u = z\%$  stets der Wert herangezogen, den  $z\%$  der Daten überschreiten.  $z$  wird dabei klein gewählt, damit die betrachtete Datenmenge möglichst auf die Extremwerte eingeschränkt wird.

### 5.1.1 Ausgangssituation

Gegeben ist eine stationäre Zeitreihe  $(X_n)_{n \in \mathbb{N}}$ . Ziel ist es  $\chi(\cdot)$  als nicht-lineares Abhängigkeitsmaß und Alternative zur linearen Autokorrelationsfunktion zu schätzen. Dabei gilt:

$$\chi(k) = \lim_{x \rightarrow \infty} \mathbb{P}(X_k > x \mid X_0 > x) \quad (5.1)$$

$$\begin{aligned} \mathbb{P}(X_k > x \mid X_0 > x) &= \frac{\mathbb{P}(X_k > x, X_0 > x)}{\mathbb{P}(X_0 > x)} = \frac{\mathbb{P}(\min\{X_k, X_0\} > x)}{\mathbb{P}(X_0 > x)} \\ &= \frac{\mathbb{P}(Y_k > x)}{\mathbb{P}(X_0 > x)} = \frac{1 - F_{Y_k}(x)}{1 - F_{X_0}(x)}. \end{aligned} \quad (5.2)$$

Das Problem reduziert sich auf die Schätzung zweier Tails. Zum Einen die Schätzung von  $\bar{F}_{X_0}(x) = 1 - F_{X_0}(x)$ , zum Anderen die Schätzung von  $\bar{F}_{Y_k}(x) = 1 - F_{Y_k}(x)$ ,  $k \in \mathbb{N}$ . Dabei wurde für diese Arbeit  $Y_{k,j}$  folgendermaßen gewählt:

$$Y_{k,j} = \min\{X_j, X_{j+k}\}, \text{ für } j = 1, \dots, n - k, k = 1, \dots, 20, n \in \mathbb{N}. \quad (5.3)$$

Nun hat man eine Zeitreihe  $X_n$  und eine (kürzere) Zeitreihe  $Y_{k,j}$  zur Verfügung, deren Tails man schätzen kann. Dadurch erhält man dann eine Aussage über die Abhängigkeiten innerhalb der Zeitreihe. Die Methoden aus Kapitel 3.3 werden auf  $X_n$  und  $Y_{k,j}$  angewandt.  $Y_{k,j}$  verändert sich dabei in der Regel für jedes  $k$ . In Kapitel 4.3.1 wurde bereits festgehalten, dass  $\chi \in [0, 1]$  liegt. Im Fall  $\chi(\cdot) = 1$  liegt Abhängigkeit der Daten vor, im Fall  $\chi(\cdot) = 0$  Unabhängigkeit.

Es wurde bei den Berechnungen mit den Matlab Versionen 6.5 und 7.1 gearbeitet. Das verwendete Paket EVIM ist ein Matlab-Paket zur Extremwertanalyse, das von Gençay et al. (2001) erstellt wurde. EVIM entspricht dem Paket EVIS von Alexander McNeil für S-Plus.

### 5.1.2 Analyse der Daten

Zunächst werden die Daten analysiert und beschrieben. Es stellt sich, vor allem bei der Anwendung auf einen konkreten Datensatz aus der Finanzwirtschaft die Frage, mit welchem Modell die Daten am besten dargestellt werden können. Zur Analyse der Daten werden deskriptive und induktive Statistiken sowie verschiedene graphische Darstellungen herangezogen, wie z.B. der QQ-Plot. In Tabelle 5.1 sind beispielhaft für eine Simulation von 10000 Daten jeweils die deskriptiven Statistiken für ein simuliertes ARCH(1), ein GARCH(1,1) und ein AR(1) + ARCH(1) Modell angegeben. Diese verändern sich bei jeder Simulation etwas, der Eindruck bleibt jedoch gleich. Das Ergebnis für die Kurtosis  $\mathcal{K}$  für den ARCH(1) und den GARCH(1,1) Fall wird bereits aufgrund von Satz 4.4 und Satz 4.9 erwartet, ebenso die Ergebnisse für  $\mu$  und  $s^2$ . Aufgrund der Abbildungen 4.4, 4.6 und 5.21 sowie der deskriptiven Statistiken wird deutlich, dass die Prozesse nicht normalverteilt sind.

	$n$	Min	Max	$\mu$	$x_{0.5}$	$s$	$s^2$	$\mathcal{S}$	$\mathcal{K}$
Modell									
ARCH(1)	10000	-10.2575	13.8145	0.0037	-0.0020	1.4109	1.9906	0.1842	6.9126
GARCH(1,1)	10000	-80.2297	98.6320	-0.0130	0.0502	9.8430	96.8853	0.0115	9.1413
AR(1)+ARCH(1)	10000	-35.9025	22.9402	-0.0032	0.0046	2.0811	4.3311	-0.1860	23.5159

$n$  = Stichprobenumfang.

Min = Kleinste Rendite in der jeweiligen Zeitreihe.

Max = Größte Rendite in der jeweiligen Zeitreihe.

$\mu$  = empirischer Mittelwert.

$x_{0.5}$  = Median.

$s$  = Standardabweichung.

$s^2$  = unbedingte Varianz.

$\mathcal{S}$  = Schiefemaß.

$\mathcal{K}$  = Kurtosis.

Tabelle 5.1: Deskriptive Statistiken für einen simulierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ , GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$  und AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ .

Die Parameter für die betrachteten Prozesse werden stets so gewählt, dass die Annah-



men der Stationarität und für das Tailverhalten erfüllt sind. Die Noise-Variable  $(Z_t)_{t \in \mathbb{N}}$  ist stets standardnormalverteilt. Da die Formeln in Kapitel 4.3.2 für die quadrierten Prozesse hergeleitet sind, werden die Methoden aus Kapitel 3.3 auch auf quadrierte Daten angewendet. Alle Plots bezüglich der GEV und GPD wurden sowohl für  $X_n$  als auch für  $Y_{k,j}$  durchgeführt. Beispielhaft sind stets die Plots für  $X_n$  dargestellt.

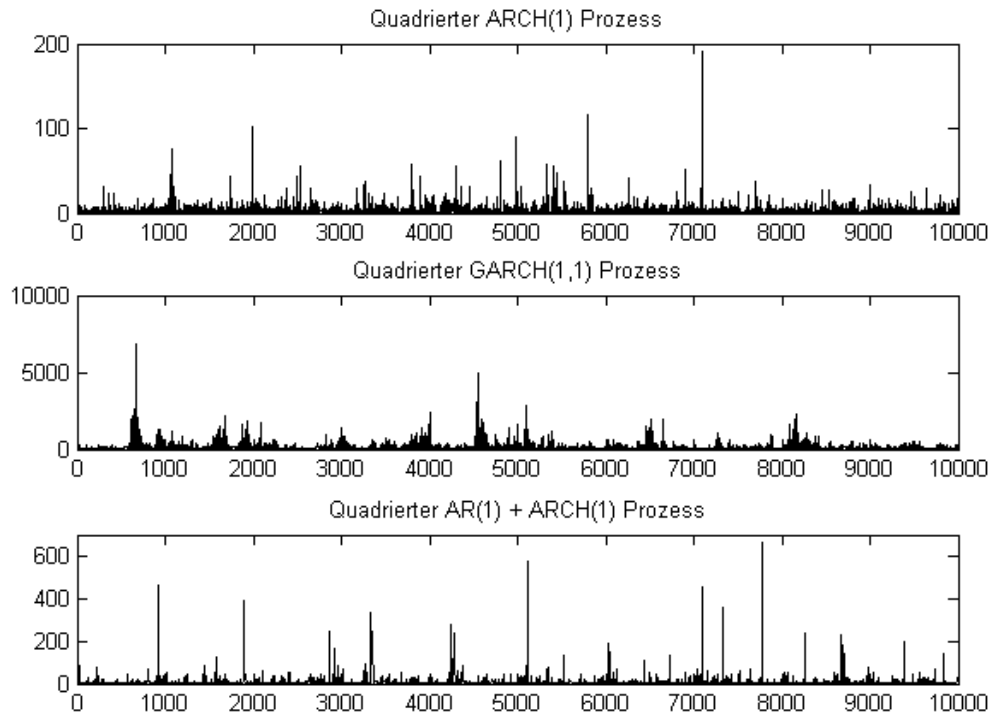


Abbildung 5.1: Oben: Simulierter quadrierter ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ , standardnormalverteilter Noise-Variable und  $n = 10000$ . Mitte: Simulierter quadrierter GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ . Unten: Simulierter quadrierter AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

In Abbildung 5.1 sind ein simulierter quadrierter ARCH(1) Prozess, ein simulierter quadrierter GARCH(1,1) Prozess und ein simulierter quadrierter AR(1) + ARCH(1) Prozess abgebildet.

Diese Bezeichnungen werden im Folgenden verwendet:  $\chi(\cdot)_{emp}$  steht für die empirische Schätzung aus Kapitel 3.3.1 von  $\chi(\cdot)$ .  $\chi(\cdot)_{emp\ Tail}$  bezeichnet die Schätzung von  $\chi(\cdot)$ , in der die Tails von  $X_n$  und  $Y_{k,j}$  empirisch geschätzt werden.  $\chi(\cdot)_{Block}$  bezeichnet die Schätzung von  $\chi(\cdot)$  mit der Blockmaxima-Methode aus Kapitel 3.3.2 und  $\chi(\cdot)_{McNeil}$  die Schätzung mit der Methode von McNeil aus Kapitel 3.3.3. Schließlich steht  $\chi(\cdot)_{POT}$  für die Schätzung von  $\chi(\cdot)$  mit der POT-Methode aus Kapitel 3.3.4.  $\chi(\cdot)_{Prozess}$  bezeichnet die Bestimmung von  $\chi(\cdot)$  mit MC Simulation für den jeweiligen nicht-linearen Prozess.

Im Anschluss an die Betrachtung jedes Modells erfolgt eine Beurteilung der Methoden. Für die Beurteilung der Schätzer werden der absolute und der relative Fehler betrachtet, die wie folgt gebildet werden:

$$\begin{aligned} \text{Absoluter Fehler: } & \chi^{(k)}_{\text{Prozess}} - \chi^{(k)}_{\text{Methode}}, \\ \text{Relativer Fehler: } & \frac{\chi^{(k)}_{\text{Prozess}} - \chi^{(k)}_{\text{Methode}}}{\chi^{(k)}_{\text{Prozess}}}. \end{aligned}$$

Ein absoluter Fehler von  $|0.1|$  wird als akzeptabel angesehen. Für den relativen Fehler gilt ein Wert von  $|1|$  als Fehler von 100%.

### 5.1.3 ARCH(1) Prozess

Als erstes wird ein ARCH(1) Prozess der Länge  $n = 10000$  wie in (4.2) simuliert. Die Stichprobenlänge  $n = 10000$  wurde gewählt, damit für die Extremwerttheorie, die hier angewandt wird, genügend Daten zur Verfügung stehen. Für diesen ARCH(1) Prozess wird dann  $\chi^{(k)}_{\text{ARCH}}$  mittels Monte-Carlo Simulation mit 10000 Wiederholungen für  $k = 1, \dots, 20$  bestimmt, sowie mit den verschiedenen Methoden aus Kapitel 3.3 geschätzt. In der Literatur wird  $\alpha_0$  oft sehr klein gewählt. Allerdings ändern sich die Ergebnisse in dieser Arbeit nur minimal, wenn  $\alpha_0 = 1$  gewählt wird. Auf die Bestimmung von  $\chi^{(k)}_{\text{ARCH}}$  mit Monte-Carlo Simulation hat  $\alpha_0$ , wie man in Formel (4.15) erkennt, keinen Einfluss.  $\alpha_0$  ist der Skalierungsparameter und von geringer Bedeutung, solange es sich um extreme Eigenschaften handelt.  $\alpha_1 = 0.5$  wird gewählt, da der Prozess dann stationär ist und einige weitere Kennzahlen, wie z.B. der Extremale Index  $\theta$  bekannt sind, sowie das 2. standardisierte Moment existiert. In Abbildung 4.3 rechts oben wurde ein simulierter ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  bereits abgebildet. Man erkennt deutlich, dass großen Ausschlägen weitere Ausschläge folgen. Die Daten clustern.

In Abbildung 5.2 erkennt man, dass die ACF des ARCH(1) Prozesses  $(X_t)_{t \in \mathbb{N}}$  keine großen Ausschläge aufweist. Betrachtet man jedoch die ACF der quadrierten und absoluten Daten, erkennt man, dass die Autokorrelationen der absoluten und quadrierten Werte teilweise stark von Null abweichen. Dies ist ein Anzeichen für weit reichende Abhängigkeiten innerhalb der Daten, wie es bereits in Kapitel 4.1.2 erwähnt wurde. Es wird in dieser Arbeit stets die ACF mit 20 Lags dargestellt, da auch  $\chi^{(k)}$  für  $k = 1, \dots, 20$  bestimmt wird und dadurch ein Vergleich möglich ist.

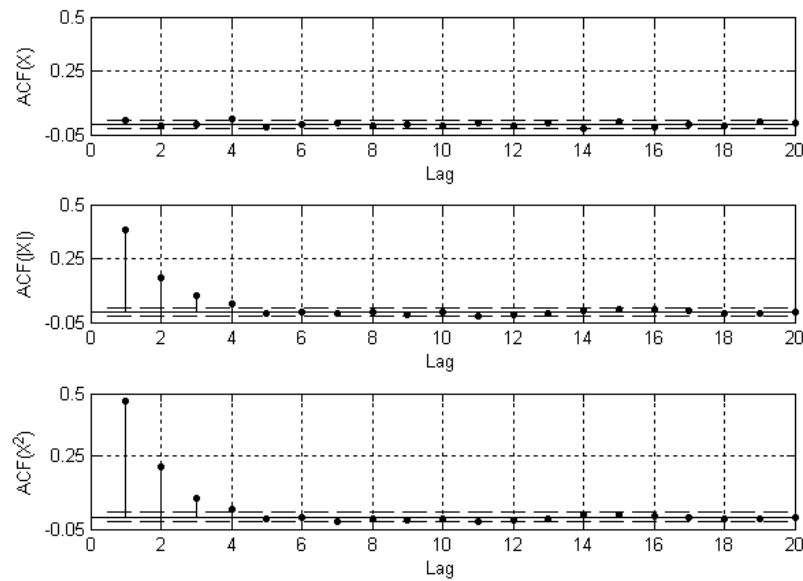


Abbildung 5.2: Oben: ACF eines simulierten ARCH(1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ , standardnormalverteilter Noise-Variable und  $n = 10000$ . Mitte: ACF der absoluten Werte. Unten: ACF der quadrierten Werte.

### Bestimmung mit Monte-Carlo (MC) Simulation

Für die Bestimmung von  $\chi(k)_{ARCH}$  mit MC Simulation und Formel (4.15) wird neben den Parametern des ARCH(1) Prozesses auch der Tailindex benötigt. Ist  $\alpha_1 = 0.5$ , dann ist der Tailindex  $\kappa$  aus Tabelle 4.1 gleich 2.37. Der Tailindex kann auch mittels Hill-Plot geschätzt werden. Allerdings kann es zu Abweichungen kommen, da der Hill-Plot oft nicht eindeutig ist. Der Hill-Plot gibt die Veränderungen des Formparameters in Bezug auf die Schwelle wieder. In Abbildung 3.1 wurde der Hill-Plot für einen simulierten ARCH(1) Prozess dargestellt. Man wählt, wie in Kapitel 3.1 erwähnt, für  $\kappa$  den Wert, bei dem der Plot ungefähr horizontal verläuft. In diesem Fall bietet sich z.B.  $\kappa = 2.5$  an. Das Ergebnis der Bestimmung von  $\chi(k)_{ARCH}$  mit MC Simulation ist in Abbildung 5.3 dargestellt. Ähnlich wie bei einer ACF der quadrierten Werte ist der erste Wert groß und fällt dann gegen Null. Dieses Ergebnis wurde erwartet, da die Abhängigkeit zwischen zwei Werten zunächst vorliegt und je mehr Werte dazwischen liegen, desto geringere Abhängigkeit ist vorhanden. Dies ist eine Folge aus der Definition des ARCH(1) Prozesses. In jede Berechnung von  $(X_t)_{t \in \mathbb{N}}$  geht immer der vorherige Wert mit ein. Je kleiner  $\alpha_1$  nun ist, desto weniger gehen die vorhergehenden Daten ein und umgekehrt. Beim Vergleich des unteren Bildes in Abbildung 5.2 und dem Bild links oben aus Abbildung 5.3 fällt eine große Ähnlichkeit auf. Jedoch sind die Werte der ACF etwas größer als die Werte von  $\chi(k)_{ARCH}$ . Es wird auch deutlich, dass mit kleinerem  $\kappa$  die Werte steigen. Auch dies entspricht dem Wesen des Tailindex, da der Tail umso langsamer fällt je kleiner  $\kappa$  ist.

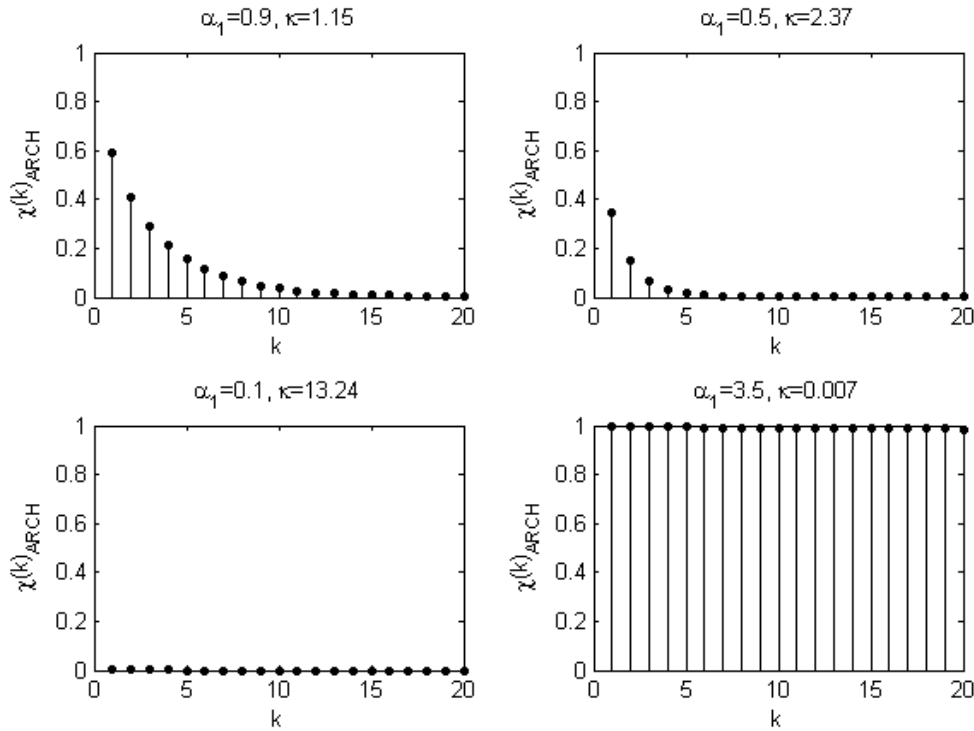


Abbildung 5.3:  $\chi(k)_{ARCH}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ , und verschiedenen  $\alpha_1$  und  $\kappa$ ,  $n = 10000$  sowie standardnormalverteilter Noise-Variable.

In Tabelle 5.2 sind die Ergebnisse der MC Simulation für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $\kappa = 2.37$ ,  $n = 10000$  sowie standardnormalverteilter Noise-Variable zusammengefasst.

$k$	1	2	3	4	5	6	7	8	9	10
$\chi(k)_{ARCH}$	0.3387	0.1534	0.0693	0.0314	0.0153	0.0075	0.0033	0.0016	0.0009	0.0005
$k$	11	12	13	14	15	16	17	18	19	20
$\chi(k)_{ARCH}$	0.0004	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Tabelle 5.2:  $\chi(k)_{ARCH}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ ,  $\kappa = 2.37$  und standardnormalverteilter Noise-Variable.

### Empirische Schätzung und empirische Schätzung der Tails

Um einen ersten Eindruck über mögliche Schätzergebnisse für  $\chi(k)_{ARCH}$  zu erhalten, wird  $\chi(k)_{ARCH}$  empirisch geschätzt. Dabei werden die Abhängigkeiten innerhalb der Daten nicht beachtet.

Da das Ergebnis für  $\chi(k)_{emp}$  und  $\chi(k)_{emp\ Tails}$  übereinstimmt, ist es in Abbildung 5.4 oben und in Tabelle 5.3 gemeinsam dargestellt.

Es wird, wie es in Kapitel 3.3.1 beschrieben wurde, vorgegangen. Dazu wird zuerst eine Schwelle  $u$  bestimmt. Für die empirische Schätzung der Tails werden anschließend die Datenpunkte gezählt, die über dieser Schwelle liegen. Es wird eine Schwelle gewählt, die sowohl für die Stichprobe  $X_n$  als auch für  $Y_{k,j}$  gleich ist. Die Bezeichnung  $u = 1\%$  bedeutet, dass 1% der Daten über der Schwelle liegen. Bei einer Datenlänge von  $n = 10000$  werden also die 100 Daten betrachtet, die über  $u$  liegen.  $\chi(k)_{emp\ Tails}$  ist dann der Quotient aus dem empirischen Tail von  $Y_{k,j}$  für  $k = 1, \dots, 20$ , dividiert durch den empirischen Tail von  $X_n$ . Das Ergebnis ist dann vergleichbar mit  $\chi(k)_{ARCH}$ .

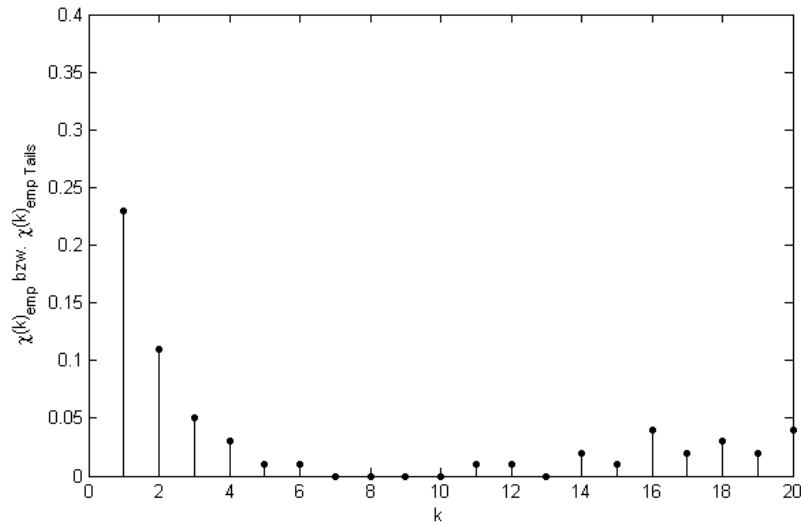


Abbildung 5.4:  $\chi(k)_{emp\ Tails}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $u = 1\%$ .

$k$	1	2	3	4	5	6	7	8	9	10
$\chi(k)_{emp\ Tails}$	0.2300	0.1100	0.0500	0.0300	0.0100	0.0100	0.0000	0.0000	0.0000	0.0000
$k$	11	12	13	14	15	16	17	18	19	20
$\chi(k)_{emp\ Tails}$	0.0100	0.0100	0.0000	0.0200	0.0100	0.0401	0.0200	0.0301	0.0200	0.0401

Tabelle 5.3:  $\chi(k)_{emp\ Tails}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $u = 1\%$ .

In Tabelle 5.3 sind die Werte für die obere Grafik aus Abbildung 5.4 dargestellt.

Die Wahl von  $u$  ist bei allen Methoden, die hier verwendet werden von großer Bedeutung. Mit dem geschätzten  $\chi(\cdot)_{Methode}$  soll eine Aussage darüber möglich sein, ob ein Modell für einen Datensatz passt. Je größer  $u$  bei der empirischen Schätzung der Tails gewählt wird, desto näher sind die Werte von  $\chi(k)_{emp\ Tails}$  den Werten von  $\chi(k)_{ARCH}$ . Der Unterschied ist allerdings, dass  $\chi(k)_{emp\ Tails}$  langsamer gegen Null fällt. Dies erkennt man in Abbildung 5.5.

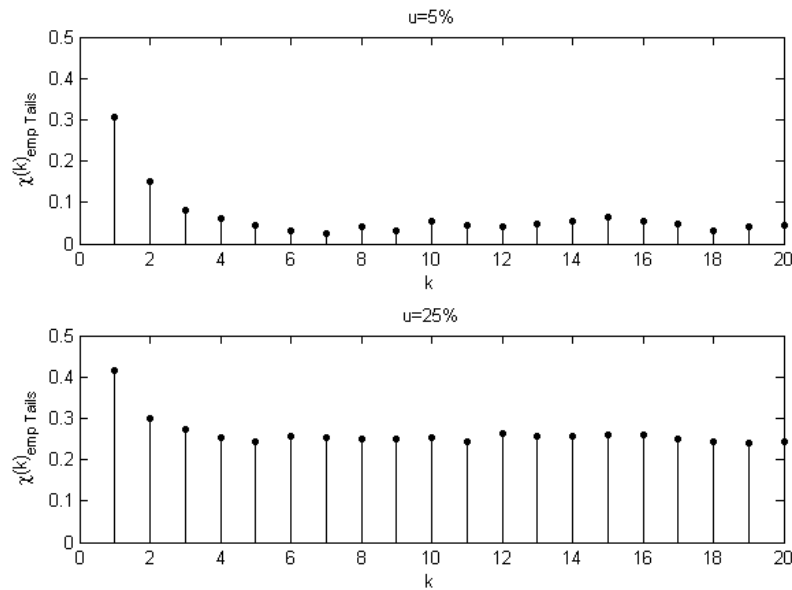


Abbildung 5.5:  $\chi(k)_{emp\ Tails}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable.

### Blockmaxima-Methode

Das Theorem von Fisher-Tippett (Satz 2.1) basiert auf der Annahme von iid Daten. Die GEV, die bei der Blockmaxima-Methode aus Kapitel 3.3.2 für die Berechnung des Tails mit Formel (2.2) verwendet wird, basiert also auch auf der Annahme von iid Daten. In unserem Fall handelt es sich aber nicht um iid Daten sondern „nur“ um stationäre, aber abhängige Daten. In der Literatur wird aber dennoch häufig diese Methode angewandt, da durch die Betrachtung der Blockmaxima die Abhängigkeitsstruktur durchbrochen wird, wenn die Blocklänge geeignet gewählt ist. Daher wurde die Blockmaxima-Methode für verschiedene Blocklängen  $r$  berechnet, um herauszufinden welche Wahl von  $r$  geeignet ist. Es wird angenommen, dass die Blockmaxima unabhängig sind. Daher werden die Blockmaxima mit dem  $\mathcal{LB}$ - und dem Lagrange-Multiplikator ( $\mathcal{LM}$ )-Test auf Autokorrelation bzw. auf ARCH-Effekte getestet. Der Ljung-Box Test ist zum Standardtest auf das Vorliegen von Autokorrelation vor allem in Residuen geworden. Es handelt sich dabei um einen Test auf keine Autokorrelation, also White Noise (Weißes Rauschen). Aus einer Ablehnung der Nullhypothese wird gefolgert, dass in den Daten eine Autokorrelation vorhanden ist. Der  $\mathcal{LM}$ -Test testet auf die Nullhypothese, dass keine ARCH-Effekte, also Homoskedastie in den Daten vorhanden ist. Dies bedeutet im ARCH(1) Fall, dass  $\alpha_1 = 0$  ist. Für weitere Informationen siehe Engle (1982, S. 999 ff.). Da  $(Z_t)_{t \in \mathbb{N}}$  in dieser Arbeit iid  $N(0, 1)$  gewählt sind können die Daten als unabhängig angesehen werden, wenn die Nullhypothesen des  $\mathcal{LM}$ -Test und des  $\mathcal{LB}$ -Test nicht verworfen werden.

Bei einer Blockgrößenwahl von  $r = 50$  sind die Blockmaxima laut der beiden Tests unkorreliert und weisen keine ARCH-Effekte auf, daher wird angenommen, dass die Blockmaxima unabhängig sind. Die Anzahl der Blockmaxima beträgt dann bei 10000 Daten

200. Das Ergebnis für  $\chi(k)_{Block}$  ist in Abbildung 5.6 und Tabelle 5.4 dargestellt.  $\chi(k)_{Block}$  fällt rasch ab und findet ab Lag 3 keine Abhängigkeiten mehr. Die Wahl der Blockgröße ist ein besonderes Problem bei der Blockmaxima-Methode und entspricht der Wahl der Schwelle  $u$ . Bei größeren Blöcken geht  $\chi(k)_{Block}$  schneller gegen Null bzw. die Berechnung ist nicht mehr möglich, da zuwenig Daten für die Anpassung der GEV vorhanden sind. Dies ist ein generelles Problem der Extremwerttheorie.  $\chi(k)_{Block}$  wird bestimmt als Quotient der Tails der Stichprobe  $Y_{k,j}$  und  $X_n$ . Für die Tailberechnung wird in Formel (2.2) für  $x$  das Maximum von  $X_n$  gewählt. Die Anpassung der GEV wurde überprüft, indem ein Abgleich zwischen  $\kappa$  und dem geschätzten Parameter  $\hat{\xi}$  durchgeführt wurde. Es galt im Schnitt  $1/\hat{\xi} \approx \kappa$  mit. Dies bestätigt wiederum die Annahme, dass die Blockmaxima an eine GEV angepasst werden können. Da der ARCH(1) Prozess, wie in Satz 4.6 festgestellt wurde, schwere Tails besitzt ist auch  $\hat{\xi}$  positiv. Des Weiteren wurde mit einem QQ-Plot der Residuen gegen die Quantile der Exponentialverteilung, wie beispielhaft in Abbildung 3.3 dargestellt, überprüft, ob die GEV das passende Modell ist. Dies wurde auch für jede Stichprobe  $Y_{k,j}$  wiederholt. Die Quantile der Residuen lagen auch für jedes  $Y_{k,j}$  zum Großteil auf einer Gerade im QQ-Plot mit exponentialverteilten Quantilen. Auch für  $\hat{\xi}$  galt im Mittel  $\hat{\xi} \approx 1/\kappa$ .  $\hat{\xi}$  sinkt mit wachsendem  $k$ .

$k$	1	2	3	4	5	6	7	8	9	10
$\chi(k)_{Block}$	0.1855	0.031	0.0014	0.0009	0.0001	0.0001	0.0000	0.0000	0.0000	0.0009
$k$	11	12	13	14	15	16	17	18	19	20
$\chi(k)_{Block}$	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0003	0.0000	0.0006	0.0008

Tabelle 5.4:  $\chi(k)_{Block}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $r = 50$ .

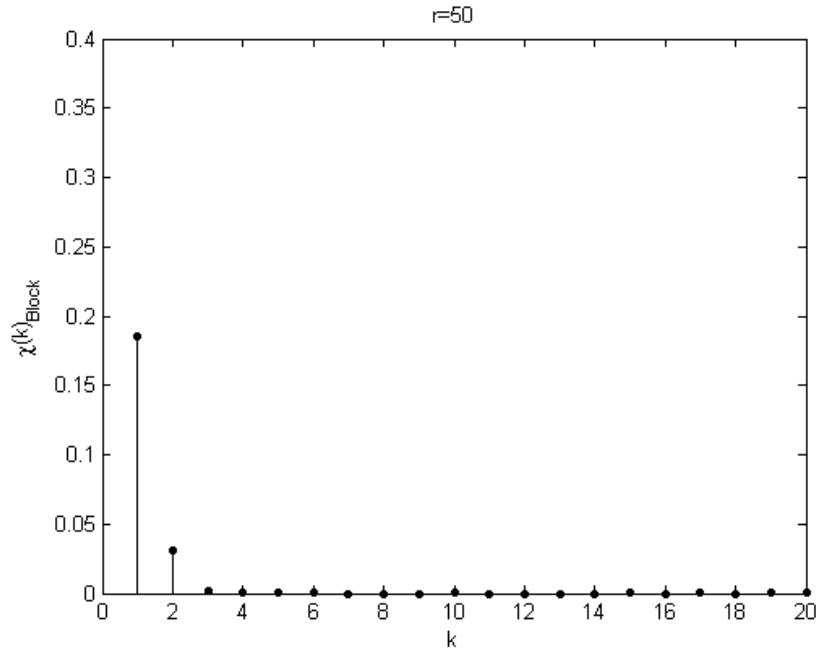


Abbildung 5.6:  $\chi(k)_{Block}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $r = 50$ .

Da die Daten in dieser Arbeit nicht iid, aber stationär sind, wird im nächsten Schritt die Methode von McNeil aus Kapitel 3.3.3 für stationäre Daten angewendet.

### Methode von McNeil

Es werden die Punkte (i)-(vi) aus Kapitel 3.3.3 für  $X_n$  und  $Y_{k,j}$ , das mit (5.3) bestimmt wurde, durchgeführt. Anschließend wird der Quotient aus den beiden Tails, die mit Formel 3.2 bestimmt wurden, gebildet. Als Blockgröße wird  $r = 50$  gewählt, da bereits aus der Anwendung der Blockmaxima-Methode bekannt ist, dass in diesem Fall die Blockmaxima als unabhängig angenommen werden können. In Abbildung 3.3 erkennt man, dass die Residuen im QQ-Plot gegen die Quantile der Exponentialverteilung fast auf einer Geraden liegen. Die Modellwahl ist also für die verwendeten Daten geeignet. Für die Schätzung des Quantils wurde  $m = 3$  gewählt. Man schätzt dann ungefähr das 98.5%-Quantil. Als Schwelle  $u$  wird 1% gewählt. Anschließend wird der Extremale Index  $\theta$  mit der Methode aus (2.4) geschätzt. Bei der Berechnung des Tails mit Formel (3.2) wird für  $x$  das Maximum von  $X_n$  gewählt. Wendet man diese Methode auf einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable an, erhält man für  $\hat{\theta} \approx 0.7$ . Es liegen zwischen 67-75 Blöcke über der Schwelle. Da der wahre Wert für  $\theta$  aus Tabelle 4.4 bekannt ist, wird mit diesem gerechnet:  $\theta = 0.835$ . Die Ergebnisse sind in Abbildung 5.7 und Tabelle 5.5 dargestellt.  $\chi(k)_{McNeil}$  fällt, wie  $\chi(k)_{Block}$  schnell ab und findet ab Lag 3 keine Abhängigkeiten mehr. Hier wird der Zusammenhang zur Blockmaxima-Methode deutlich. Des Weiteren ist bei dieser Methode, wie auch bei der Blockmaxima- und der POT-Methode, aufgefallen, dass der Wert  $\alpha_0$



einen Einfluss auf das Ergebnis hat. Je kleiner  $\alpha_0$  gewählt wird, um so schneller gehen die Ergebnisse für  $\chi(k)_{McNeil}$  gegen 0. Wählt man  $\alpha_0$  dagegen sehr groß, ist der Unterschied zwischen den verschiedenen Schätzungen von  $\chi(k)_{McNeil}$  nicht so bedeutend. Interessant ist hierbei auch die Beobachtung, dass eine Änderung von  $\alpha_0$  bei der Bestimmung mittels MC Simulation und der empirischen Schätzung von  $\chi(k)$  kaum einen Einfluss auf die Ergebnisse hat.

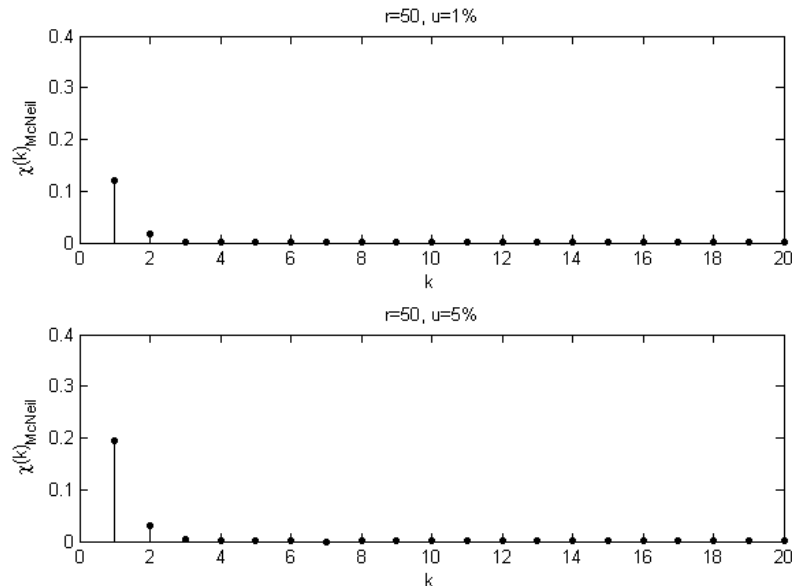


Abbildung 5.7:  $\chi(k)_{McNeil}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $r = 50$ .

$k$	1	2	3	4	5	6	7	8	9	10
$\chi(k)_{McNeil}$	0.1213	0.0168	0.0017	0.0009	0.0002	0.0002	0.0000	0.0000	0.0000	0.0006
$k$	11	12	13	14	15	16	17	18	19	20
$\chi(k)_{McNeil}$	0.0001	0.0001	0.0001	0.0002	0.0004	0.0003	0.0004	0.0001	0.0005	0.0007

Tabelle 5.5:  $\chi(k)_{McNeil}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $r = 50$ .

### POT-Methode

Eine weitere Methode zur Schätzung von  $\chi(k)_{ARCH}$  ist die POT-Methode aus Kapitel 3.3.4. Diese Methode ist sowohl für iid Daten als auch für stationäre Zeitreihen geeignet. Folgende Annahmen müssen überprüft werden:

- (1) Die Exzesse einer iid oder stationären Folge über einer hohen Schwelle  $u$  folgen einem Poisson-Prozess (siehe Definition B.4), d.h. die Zeiten zwischen den Exzessen

sind exponentialverteilt.

- (2) Die entsprechenden Exzesse über  $u$  sind unabhängig und haben eine GPD.
- (3) Die Exzesse und die Zeitpunkte der Überschreitungen sind voneinander unabhängig.

Die Zeiten zwischen zwei Exzessen sollen einer Exponentialverteilung folgen. Dies kann mit einem QQ-Plot überprüft werden. Es wurden sowohl für  $X_n$  als auch für  $Y_{k,j}$  diese QQ-Plots untersucht. Diese Plots haben die Annahme der exponentialverteilten Zwischenankunftszeiten in der Regel bestätigt. Beispielhaft ist ein QQ-Plot von Zwischenankunftszeiten (hier:  $X_{309}$  bis  $X_{363}$ ) von  $X_n$  in Abbildung 5.8 dargestellt. Um die Unabhängigkeit der Exzesse zu überprüfen werden der  $\mathcal{LM}$ - und  $\mathcal{LB}$ -Test verwendet. Für die Schwelle  $u = 1\%$  sind die Exzesse größtenteils nicht korreliert und weisen keine ARCH-Effekte auf. Die Verteilungsannahme wird mit verschiedenen Plots bezüglich der GPD überprüft. Im Plot der Exzess-Verteilung ist die kumulierte Dichtefunktion der geschätzten GPD als Kurve geplottet. Die Daten über der Schwelle  $u = 1\%$  folgen dieser Kurve. Auch mit dieser Methode wurden verschiedene  $\chi(k)_{POT}$  berechnet, indem die einzelnen Komponenten der Methode ( $u, Y_{k,j}$ ) variiert wurden. Verwendbare Ergebnisse haben sich ergeben wenn  $Y_{k,j}$  mit Variante (5.3) bestimmt wurde. Wegen dem Zusammenhang zwischen GEV und GPD gilt, dass der geschätzte Parameter  $\hat{\xi}$  gleich ist. Dies wurde überprüft und traf mit kleinen Abweichungen zu. In Abbildung 5.8 ist ein simulierter quadrierter ARCH(1) Prozess gegen die Quantile der GPD mit  $\hat{\xi} = 1/\kappa$  abgebildet. Diese Plots sind eine Rechtfertigung für die Anpassung der GPD an die Exzesse des simulierten quadrierten ARCH(1) Prozesses. Der Tail für  $X_n$  und  $Y_{k,j}$  wird mit Formel (2.5) berechnet. Für  $y$  wird das Maximum von  $X_n$  gewählt.

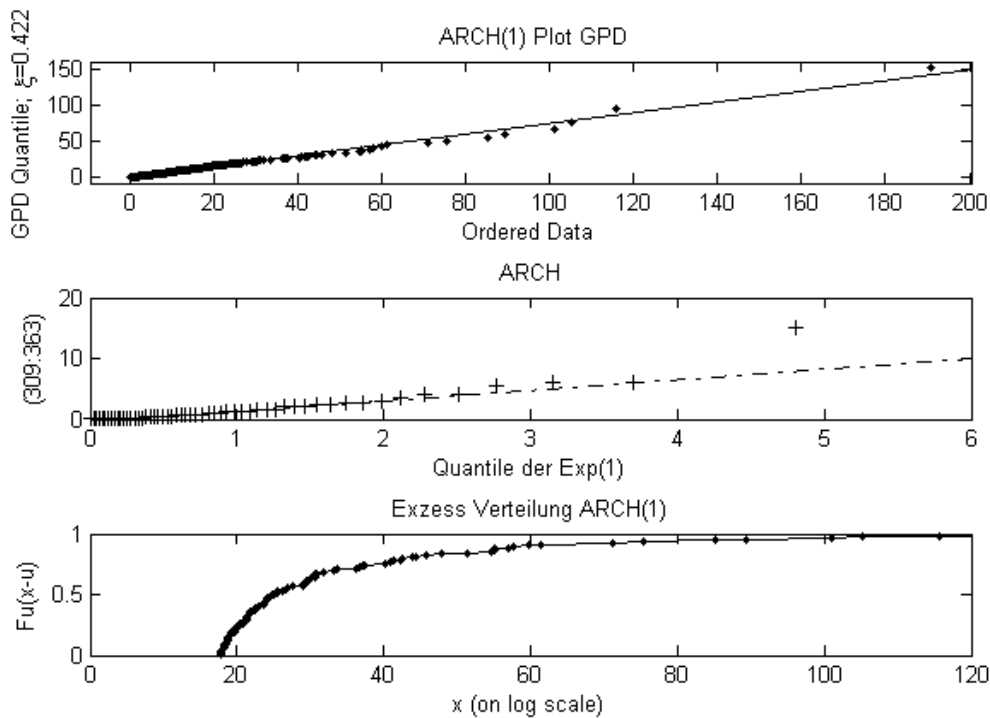


Abbildung 5.8: Oben: GPD Quantile für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable. Mitte: QQ-Plot der Quantile der Exponentialverteilung und der Zwischenankunftszeiten  $X_{309} : X_{363}$ . Unten: Exzess-Verteilung.

In Abbildung 5.9 sind für verschiedene Schwellen die Ergebnisse mit der POT-Methode aus Kapitel 3.3.4 abgebildet.  $u = 1\%$  bedeutet, dass für  $X_n$  und  $Y_{k,j}$  die Schwelle so gewählt wurde, dass jeweils 1% der Daten über der Schwelle liegen. Die Ergebnisse für  $u = 1\%$  sind in Tabelle 5.6 dargestellt. Entgegen der Erwartung, dass diese Methode für die Schätzung von  $\chi(k)_{ARCH}$  sehr gut geeignet ist, fällt auf, dass der Verlauf von  $\chi(k)_{POT}$  von dem Verlauf von  $\chi(k)_{ARCH}$  abweicht.  $\chi(k)_{POT}$  fällt bei Lag 2 fast auf Null und steigt dann noch einmal an. Bei einer Senkung der Schwelle auf  $u = 2\%$  wird die Unterschätzung von  $\chi(k)_{ARCH}$  noch größer.

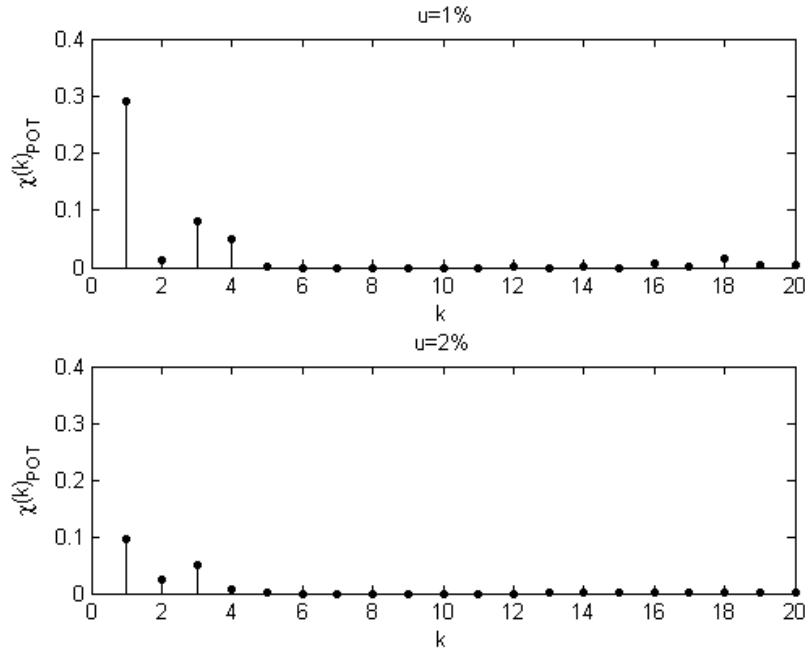


Abbildung 5.9:  $\chi^{(k)}_{POT}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $u = 1\%$  bzw.  $u = 2\%$ .

$k$	1	2	3	4	5	6	7	8	9	10
$\chi^{(k)}_{POT}$	0.2917	0.0123	0.0812	0.0501	0.0015	0.0000	0.0000	0.0000	0.0000	0.0000
$k$	11	12	13	14	15	16	17	18	19	20
$\chi^{(k)}_{POT}$	0.0000	0.0014	0.0000	0.0003	0.0000	0.0068	0.0008	0.0146	0.0030	0.0031

Tabelle 5.6:  $\chi^{(k)}_{POT}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$ , standardnormalverteilter Noise-Variable und  $u = 1\%$ .

#### 5.1.4 Beurteilung für den ARCH(1) Prozess

In Abbildung 5.10 sind die Ergebnisse aus der vorhergehenden Betrachtung des ARCH(1) Prozesses zusammengefasst. Alle Methoden unterschätzen die Bestimmung mit MC Simulation. Es fällt auf, dass die empirischen Schätzungen  $\chi^{(k)}_{emp}$  und  $\chi^{(k)}_{emp\ Tail}$  den Verlauf von  $\chi^{(k)}_{ARCH}$  am besten wiedergeben, jedoch auf niedrigerem Niveau. Auch die Blockmaxima-Methode liefert ein ähnliches Ergebnis. Die Vorgehensweise von McNeil zeigt ein schlechteres Ergebnis. Bei den ersten beiden Lags erhält man noch von Null signifikante Abweichungen, die allerdings bereits sehr klein sind. Diese Methode unterschätzt  $\chi^{(k)}_{ARCH}$ . Die Ergebnisse der POT-Methode schwanken stärker und länger im Vergleich zur MC Simulation. Bei der Betrachtung von  $\chi^{(k)}_{POT}$  fällt Lag 2 ins Auge. Ohne diesen Ausreißer nach unten würde der Verlauf von  $\chi^{(k)}_{POT}$  der MC Simulation am besten entsprechen.

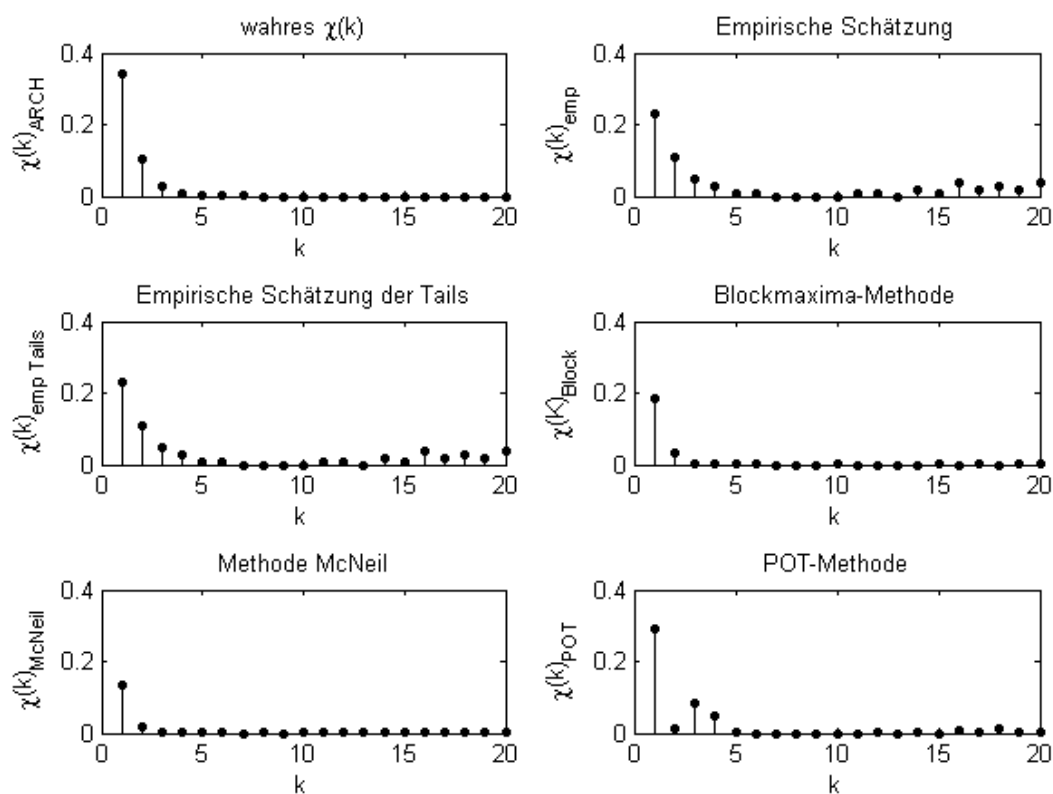


Abbildung 5.10: Plots der Methoden für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable.

Für die Ergebnisse aus Abbildung 5.10 wird in Abbildung 5.11 und in Tabelle 5.7 der jeweilige absolute Fehler dargestellt. Bis auf die POT-Methode weisen die absoluten Fehler für alle Methoden ein ähnliches Verhalten auf. Eine Abweichung von  $|0.1|$  wird als akzeptabel betrachtet, auch wenn die Ergebnisse an sich nicht sehr hoch sind. Daher sind nach der Betrachtung des absoluten Fehlers die empirischen Schätzungen am besten geeignet.

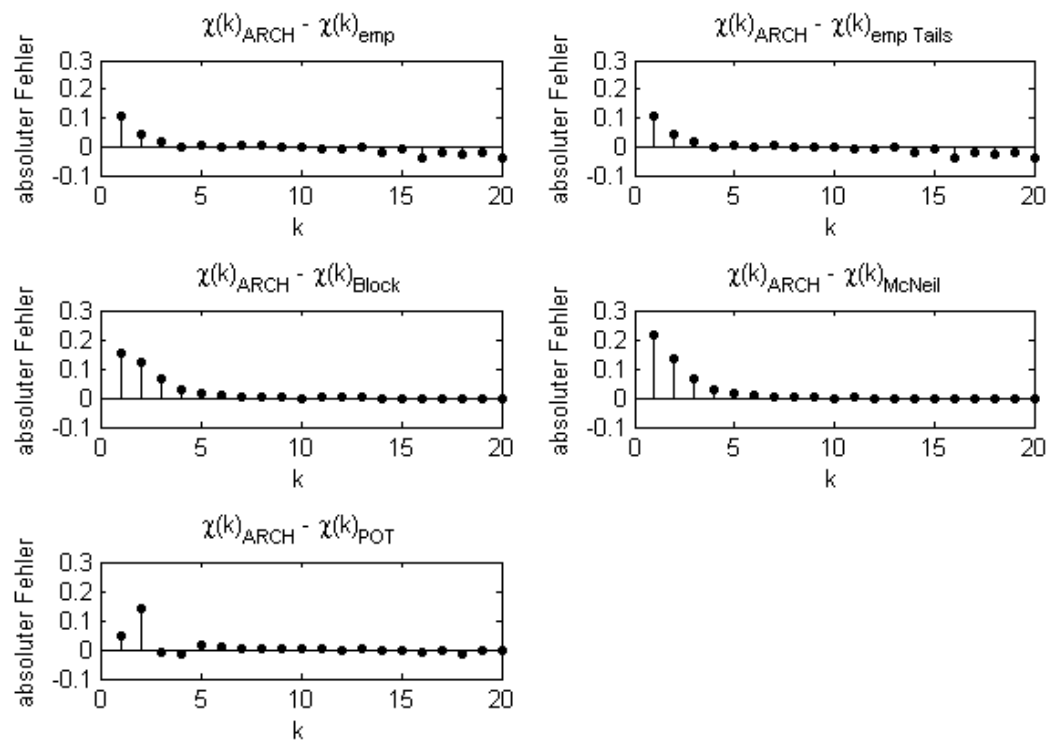


Abbildung 5.11: Absoluter Fehler der Methoden für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable.

Methode Lag $k$	Emp.	Emp. Tails	Block	McNeil	POT
1	0.1087	0.1087	0.1532	0.2174	0.0470
2	0.0434	0.0434	0.1224	0.1366	0.1411
3	0.0193	0.0193	0.0679	0.0676	-0.0119
4	0.0014	0.0014	0.0305	0.0305	-0.0187
5	0.0053	0.0053	0.0152	0.0151	0.0138
6	-0.0025	-0.0025	0.0074	0.0073	0.0075
7	0.0033	0.0033	0.0033	0.0033	0.0033
8	0.0016	0.0016	0.0016	0.0016	0.0016
9	0.0009	0.0009	0.0009	0.0009	0.0009
10	0.0005	0.0005	-0.0004	-0.0001	0.0005
11	-0.0096	-0.0096	0.0004	0.0003	0.0004
12	-0.0099	-0.0099	0.0001	0.0000	-0.0013
13	0.0001	0.0001	0.0001	0.0000	0.0001
14	-0.0200	-0.0200	0.0000	-0.0002	-0.0003
15	-0.0100	-0.0100	-0.0002	-0.0004	0.0000
16	-0.0401	-0.0401	0.0000	-0.0003	-0.0068
17	-0.0200	-0.0200	-0.0003	-0.0004	-0.0008
18	-0.0301	-0.0301	0.0000	-0.0001	-0.0146
19	-0.0200	-0.0200	-0.0006	-0.0005	-0.0030
20	-0.0401	-0.0401	-0.0008	-0.0007	-0.0031

Tabelle 5.7: Absoluter Fehler für die Berechnung von  $\chi(k)_{ARCH}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable.

In Abbildung 5.12 sind die relativen Fehler der jeweiligen Methoden abgebildet. Die relativen Fehler wurden nur im Intervall  $[-1, 1]$  dargestellt, damit die Fehler bis  $|1|$  erkennbar sind. Ein Fehler von  $|1|$  entspricht 100% Abweichung. Die entsprechenden Werte sind in Tabelle 5.8 zusammengestellt. Die relativen Fehler haben teilweise sehr hohe negative Ausschläge (man betrachte dazu z.B. den relativen Fehler der empirischen Schätzung in Lag 20). Dies liegt daran, dass  $\chi(k)_{ARCH}$  gegen Null geht.

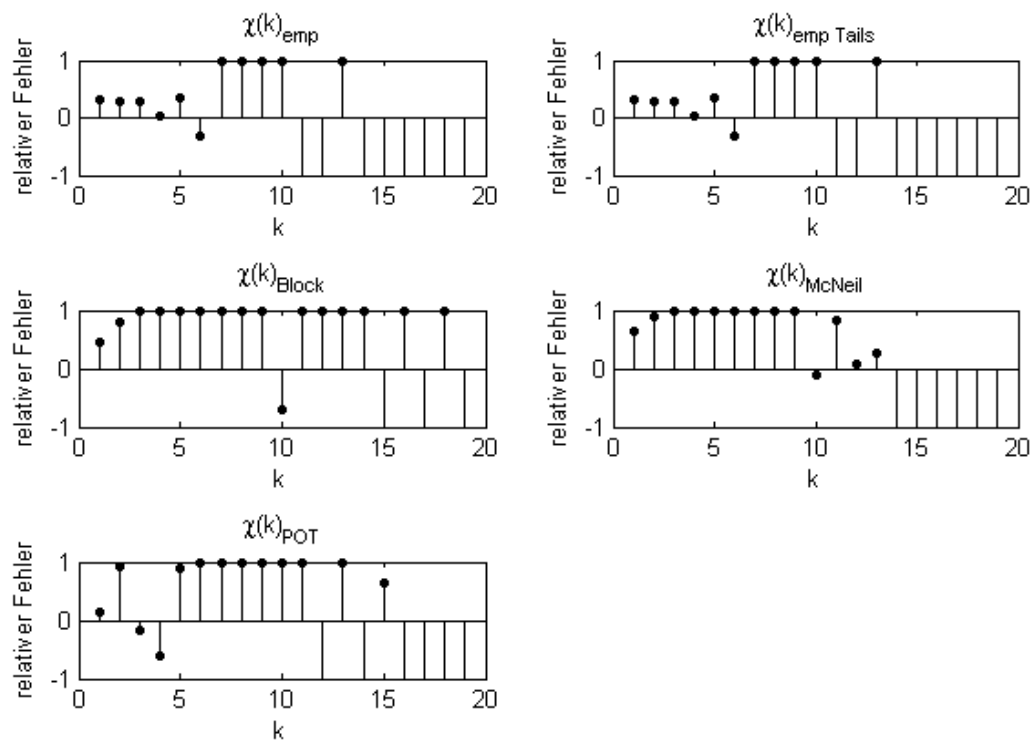


Abbildung 5.12: Relativer Fehler der Methoden für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable.



Methode Lag $k$	Emp.	Emp. Tails	Block	McNeil	POT
1	0.3208	0.3208	0.4523	0.6419	0.1387
2	0.2830	0.2830	0.7980	0.8903	0.9198
3	0.2784	0.2784	0.9798	0.9749	-0.1720
4	0.0436	0.0436	0.9713	0.9724	-0.5953
5	0.3454	0.3454	0.9935	0.9871	0.9007
6	-0.3330	-0.3330	0.9867	0.9741	0.9993
7	1.0000	1.0000	1.0000	0.9996	0.9993
8	1.0000	1.0000	1.0000	0.9762	0.9850
9	1.0000	1.0000	1.0000	0.9709	1.0000
10	1.0000	1.0000	-0.6971	-0.1150	1.0000
11	-24.0264	-24.0264	1.0000	0.8164	1.0000
12	-69.7097	-69.7097	1.0000	0.0848	-9.0605
13	1.0000	1.0000	1.0000	0.2508	0.9924
14	-6239.4704	-6239.4704	1.0000	-74.8403	-89.8028
15	-17412.8433	-17412.8433	-346.7552	-734.1198	0.6262
16	-105252.9085	-105252.9085	1.0000	-701.0780	-17818.6569
17	-178839.0325	-178839.0325	-2677.0478	-3893.5064	-7082.6150
18	-571657.4534	-571657.4534	1.0000	-2088.0813	-278562.1879
19	-1545415.0249	-1545415.0249	-46273.5591	-35955.8749	-234287.0927
20	-10765910.0710	-10765910.0710	-214887.4445	-187583.1815	-837526.7126

Tabelle 5.8: Relativer Fehler für die Berechnung von  $\chi(k)_{ARCH}$  für einen simulierten quadrierten ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.5$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable.

Die Ergebnisse für den GARCH(1,1) Prozess und den AR(1) + ARCH(1) Prozess sind denen des ARCH(1) Prozesses bis auf Besonderheiten bezüglich des jeweiligen Prozesses sehr ähnlich. Daher werden im Folgenden nur die Besonderheiten bei den jeweiligen Methoden aufgezeigt. Weitere Abbildungen und Tabellen sind in Anhang A zusammengestellt.

### 5.1.5 GARCH(1,1) Prozess

In diesem Abschnitt wird die Analyse für den GARCH(1,1) Prozess durchgeführt. Es wird ein simulierter quadriertes GARCH(1,1) Prozess der Länge  $n = 10000$  mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$  und standardnormalverteilter Noise-Variable verwendet. Dieser ist in Abbildung 5.1 in der Mitte abgebildet. Wie in Tabelle 5.1 und Abbildung 4.6 erkennbar ist, ist der GARCH(1,1) Prozess nicht normalverteilt. Dies wurde bereits in Kapitel 4.2.2 gezeigt.

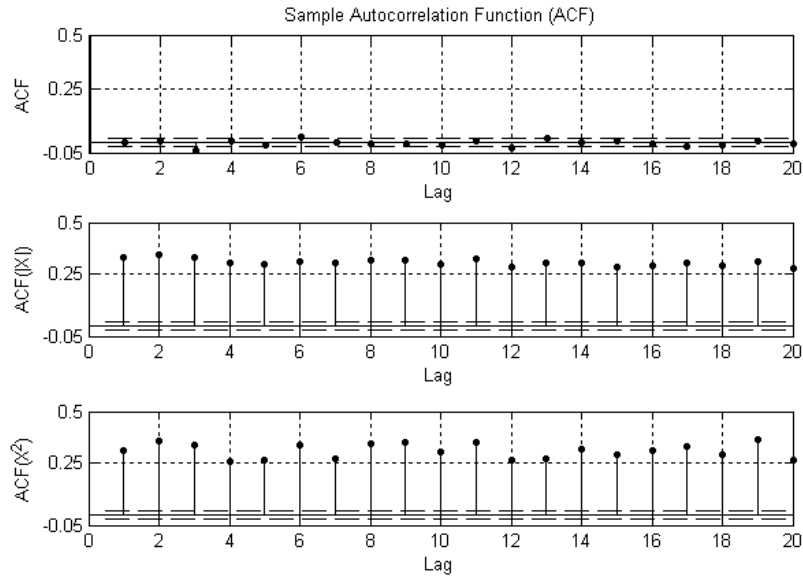


Abbildung 5.13: Oben: ACF eines simulierten GARCH(1,1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ . Mitte: ACF der absoluten Werte. Unten: ACF der quadrierten Werte.

Wie man in Abbildung 5.13 in der Mitte und unten erkennen kann, weist der GARCH(1,1) Prozess bei der ACF für die absoluten und quadrierten Werte lang anhaltende Korrelationen auf und die Werte sind fast konstant. Der GARCH(1,1) Prozess erfüllt viele der Stylized Facts aus Kapitel 4.1.

### Bestimmung mit MC Simulation

Es wird  $\chi(k)_{GARCH}$  mit Formel (4.16) und MC Simulation mit 10000 Wiederholungen für einen quadrierten GARCH(1,1) Prozess bestimmt. Für den Tailindex wird der zu  $\alpha_1$  und  $\beta$  gehörende Wert aus Tabelle 4.2 verwendet. Die Parameter wurden so gewählt, dass die Voraussetzungen (Stationarität, etc.) für  $\chi(k)_{GARCH}$  erfüllt sind. Abbildung 5.14 zeigt beispielhaft eine Schätzung von  $\kappa$  mit dem Hill-Plot und mögliche Abweichungen.

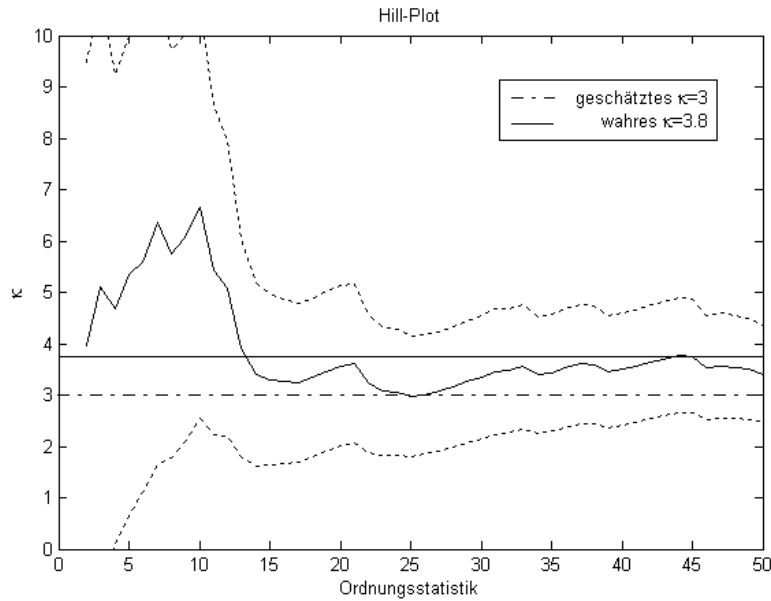


Abbildung 5.14: Hill-Plot für einen simulierten GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

In Abbildung 5.15 sind die Ergebnisse für  $\chi(k)_{GARCH}$  mit Monte-Carlo Simulation und unter Verwendung von Formel (4.16) dargestellt. Im Vergleich zum ARCH(1) Prozess sind die Korrelationen hier lang anhaltender. Das hängt mit der stärkeren Modellierung der Persistenz der Volatilität zusammen. Bei GARCH Modellen hängt die bedingte Varianz von der Vergangenheit des Prozesses und zusätzlich von der eigenen Vergangenheit der Volatilität ab. Im GARCH(1,1) Prozess geht also die Volatilität der Vorperiode mit ein, d.h. die Daten sind nicht nur von den Daten der Vorperiode abhängig. Die Persistenz wird im GARCH(1,1) Fall durch die Summe von  $\alpha_1$  und  $\beta$  gemessen und ist in dem verwendeten Beispiel nahe bei 1. Dieser Fall bildet sehr gut das Verhalten von Finanzzeitreihen ab und wurde daher so gewählt.  $\beta$  bestimmt wie schnell das Ergebnis für  $\chi(k)_{GARCH}$  gegen Null fällt. Je größer  $\beta$  ist, umso langsamer fällt  $\chi(k)_{GARCH}$ . Sehr deutlich ist auch die Ähnlichkeit zur ACF des quadrierten Prozesses in Abbildung 5.13 zu erkennen. Auch für die anderen Prozesse, die in dieser Arbeit verwendet werden, wurde dies festgestellt und bestätigt  $\chi(k)$  als Alternative zur Korrelationsfunktion. Betrachtet man  $\chi(k)_{GARCH}$  für kleine  $\beta$  und große  $\alpha_1$  ist die Beziehung zwischen ARCH(1) und GARCH(1,1) erkennbar: Für  $\beta = 0$  ist  $\chi(k)_{GARCH}$  gleich  $\chi(k)_{ARCH}$ .

Tabelle A.1 im Anhang zeigt das Ergebnis für  $\chi(k)_{GARCH}$ .

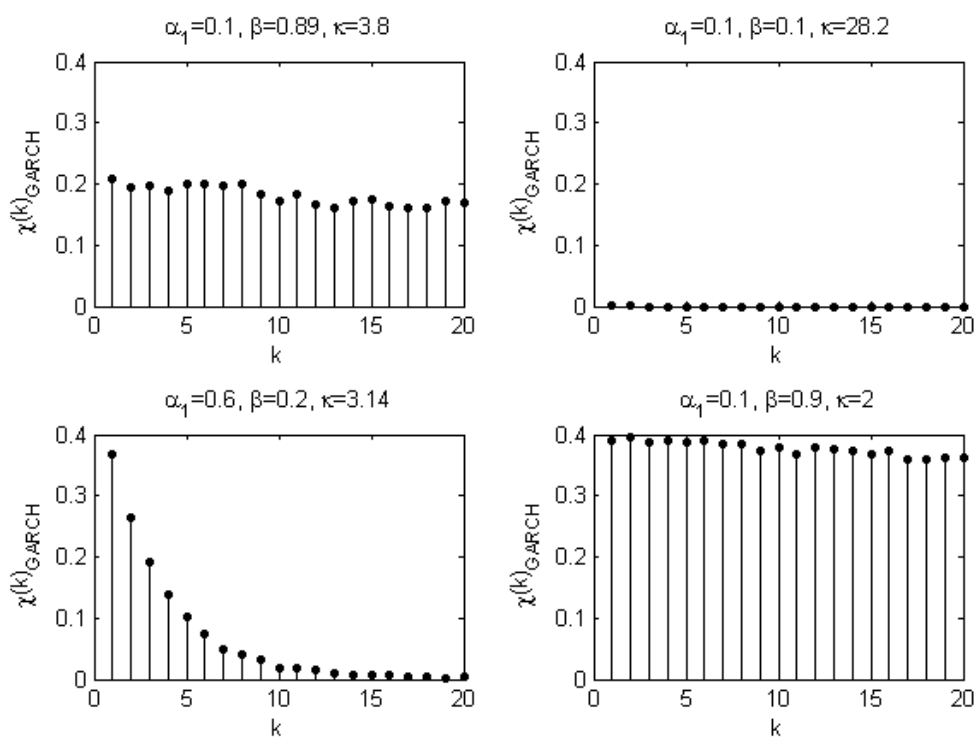


Abbildung 5.15:  $\chi(k)_{GARCH}$  für einen simulierten quadrierten GARCH(1,1) Prozess mit verschiedenen  $\alpha_0, \alpha_1, \beta$  und  $\kappa$  für eine standardnormalverteilte Noise-Variable.

In Abbildung 5.16 sind die Ergebnisse der verschiedenen Methoden im Vergleich zur MC Simulation dargestellt. Keine der verwendeten Methoden bildet das Ergebnis mit MC Simulation exakt ab. Die empirische Schätzung und die empirische Schätzung der Tails bilden das Verhalten annähernd wieder, allerdings auf niedrigerem Level.

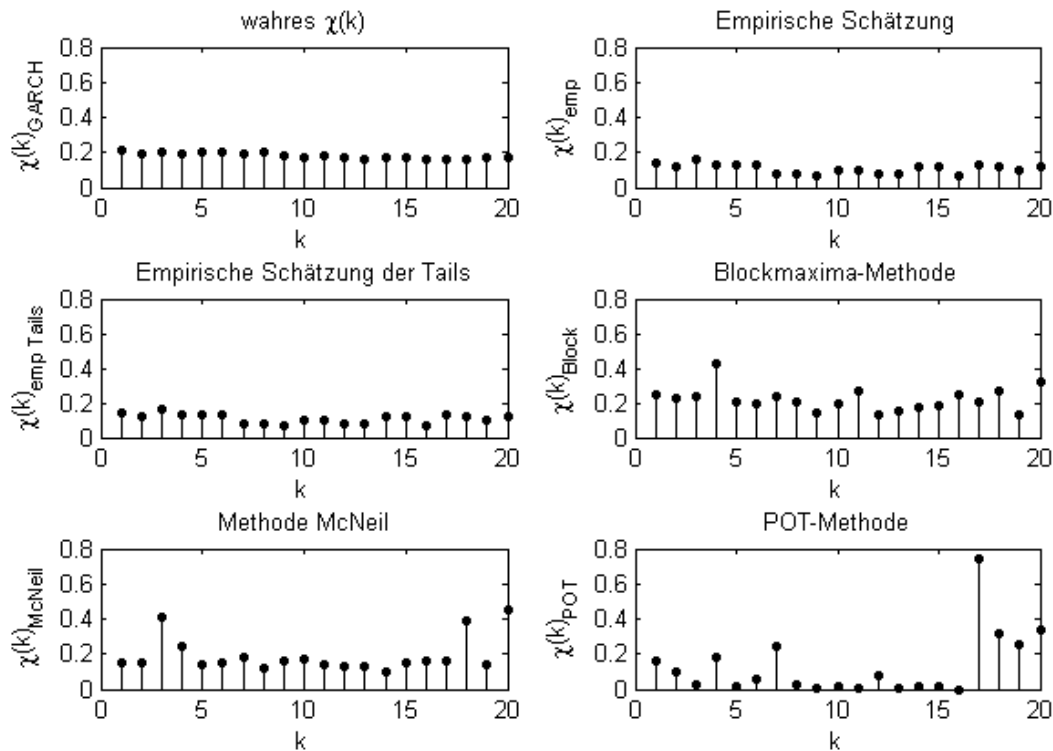


Abbildung 5.16: Ergebnisse für  $\chi(k)$  mit den verschiedenen Methoden für einen simulierten quadrierten GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

### Empirische Schätzung und empirische Schätzung der Tails

In Abbildung 5.16 rechts oben bzw. in der Mitte links und Tabelle A.1 sind die Ergebnisse für  $\chi(k)_{emp}$  und  $\chi(k)_{emp\ Tails}$  für einen simulierten quadrierten GARCH(1,1) Prozess für  $u = 1\%$  dargestellt.  $\chi(k)_{emp}$  und  $\chi(k)_{emp\ Tails}$  stimmen bis auf die vierte Nachkommastelle überein. Beide unterschätzen jedoch  $\chi(k)_{GARCH}$ .

Wie beim ARCH(1) Prozess hat die Schwellenwahl auf  $\chi(k)_{emp}$  und  $\chi(k)_{emp\ Tails}$  einen großen Einfluss. Um die Größenordnung von  $\chi(k)_{GARCH}$  zu erreichen, muss die Schwelle niedriger gewählt werden. Dies widerspricht aber der Extremwerttheorie, da die Tails und nicht der ganze Datensatz betrachtet werden sollen.

### Blockmaxima-Methode

Bei der Blockmaxima-Methode aus Kapitel 3.3.2 ergibt sich bei der Anwendung auf einen simulierten quadrierten GARCH(1,1) Prozess ein Problem bei der Wahl von  $\alpha_0$ . Wird  $\alpha_0$  sehr klein gewählt (z.B.  $\alpha_0 = 10^{-7}$ ), erhält man ein unbrauchbares Ergebnis (imaginäre Anteile treten auf, da die ML Schätzung für den Parameter  $\hat{\xi}$  teilweise für  $Y_{k,j}$  negativ

wird). Das Ergebnis mit  $\alpha_0 = 1$  ist in Abbildung 5.16 in der Mitte rechts dargestellt. Für  $r = 50$  sind die Blockmaxima von  $Y_{k,j}$  laut  $\mathcal{LM}$ - und  $\mathcal{LB}$ -Test nicht unkorreliert und weisen ARCH-Effekte auf. Dies lässt auf Abhängigkeit der Blockmaxima schließen. Für  $r = 100$  werden die Nullhypothesen der beiden Tests für die Blockmaxima nicht abgelehnt. Daher wird  $r = 100$  verwendet. Im Vergleich zum ARCH(1) Fall ist dies plausibel, da GARCH(1,1) Daten eine höhere Abhängigkeit aufweisen und dies auch bei den Extremwerten der Fall sein sollte. Die Ergebnisse sind in Abbildung 5.16 in der Mitte rechts und Tabelle A.1 zusammengefasst. Die Blockmaxima-Methode überschätzt  $\chi(k)_{GARCH}$ . Auch hier wurde wieder der Zusammenhang zwischen  $\kappa$  und  $\hat{\xi}$  überprüft. Jedoch weicht  $\hat{\xi} = 0.6388$  für  $X_n$  sehr stark von  $1/\kappa = 0.26$  ab. Die Schätzung von  $\hat{\xi}$  für  $Y_{k,j}$  ist für jedes  $k$  in der Größenordnung 0.6. Mit einer Veränderung der Blockgröße konnte dies nicht verbessert werden. In Abbildung 5.17 erkennt man, dass die Residuen im QQ-Plot mit den Quantilen der Exponentialverteilung überwiegend auf einer Geraden liegen. Die ist eine Bestätigung für die Anpassung der GEV.

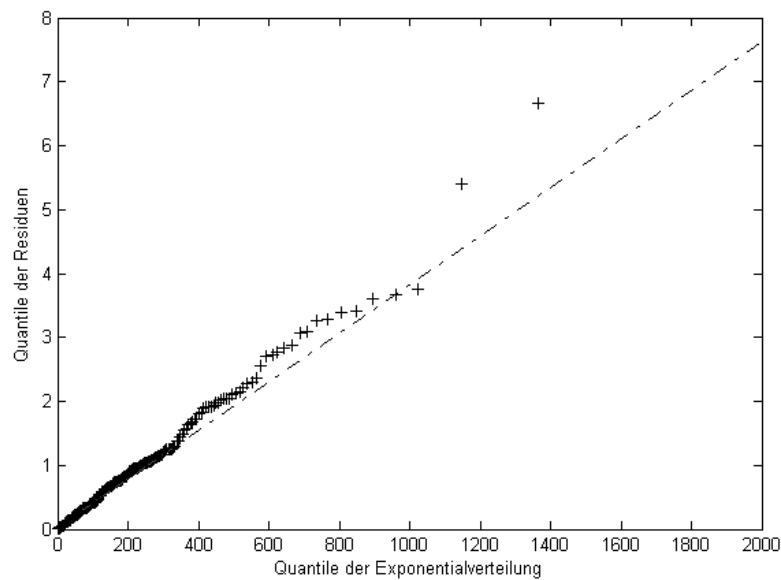


Abbildung 5.17: QQ-Plot der Residuen eines simulierten quadrierten GARCH(1,1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise Variable und  $n = 10000$ .

### Methode von McNeil

Das Vorgehen erfolgt analog zum ARCH(1) Fall. Als Blockgröße wird  $r = 100$  und  $m = 2$  gewählt, d.h. es wird ungefähr das 98%-Quantil geschätzt. Die Überprüfung der Modellwahl erfolgt mit dem QQ-Plot. In Abbildung 5.17 sind zur Überprüfung der Modellwahl die Quantile der Residuen gegen die Quantile der Exponentialverteilung geplottet. Man erkennt, dass die Modellwahl gerechtfertigt ist. Es liegen zwischen 26-30 Blöcke über der

Schwelle  $u = 1\%$ . Für  $\hat{\theta}$  ergibt sich mit (2.4) ungefähr 0.3.  $\hat{\theta}$  ist eine Kennzahl für die Abhängigkeit. Mit einem Wert von 0.3 ist diese sehr hoch.

In Abbildung 5.16 links unten und Tabelle A.1 sind die Ergebnisse dieser Vorgehensweise abgebildet. Es fällt auf, dass die Ergebnisse stark schwanken und auch für größere Lags signifikante Werte liefern.

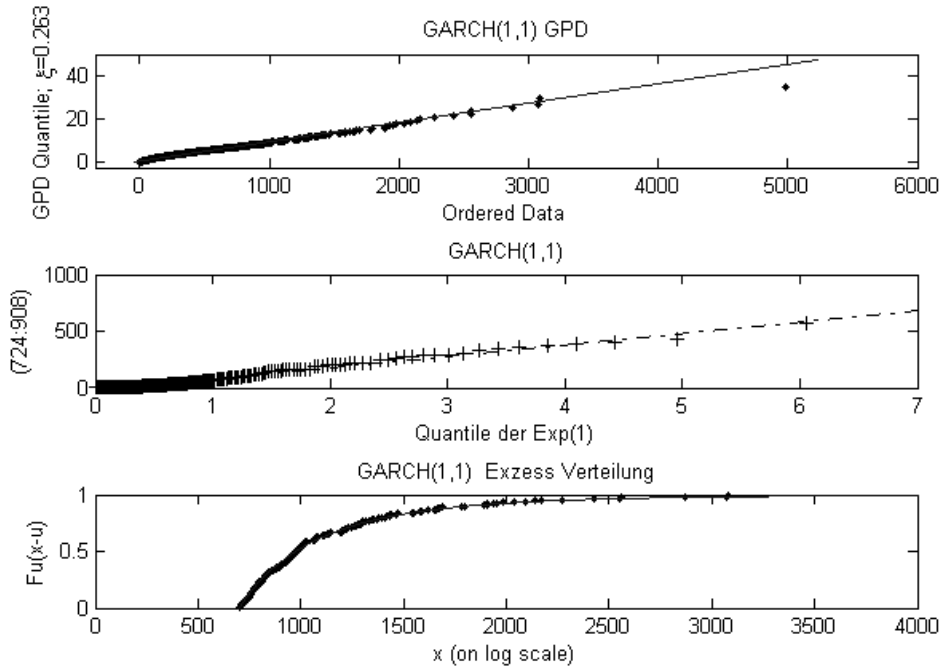


Abbildung 5.18: Oben: GPD Quantile für einen simulierten quadrierten GARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable. Mitte: Quantile der Exponentialverteilung und Zwischenankunftszeiten ( $X_{724} : X_{908}$ ). Unten: Exzess-Verteilung.

### POT-Methode

Wählt man als Schwelle bei der POT-Methode  $u = 1\%$ , ergeben sich für  $\chi^{(k)}_{POT}$  Werte mit imaginärem Anteil, da für die Schätzung des Parameters  $\hat{\xi}$  der GPD negative Werte auftreten. Für die Schwellenwahl  $u = 2\%$  werden die Nullhypothesen des  $\mathcal{LM}$ - und  $\mathcal{LB}$ -Tests nicht verworfen, daher werden die Exzesse als unabhängig angenommen. Es wird daher als Schwelle  $u = 2\%$  verwendet. Es ergibt sich folgendes Ergebnis, das in Abbildung 5.16 rechts unten und Tabelle A.1 dargestellt ist. Der Parameter  $\hat{\xi} = 0.2156$  für  $X_n$  entspricht nicht dem  $\hat{\xi}$  aus der Blockmaxima-Methode. Für  $Y_{k,j}$  ergeben sich Werte, die in der Größenordnung von  $\hat{\xi}$  von  $X_n$  sind. Da das  $\hat{\xi}$  der POT-Methode sehr nah an  $\kappa$ , liegt scheint es bei der Blockmaxima-Methode bei der Schätzung von  $\hat{\xi}$  ein Problem zu geben. Dieses ist allerdings nur bei der Betrachtung des GARCH(1,1) Modells aufgetreten. In Abbildung

5.18 sind, analog zum ARCH(1) Fall, Plots des GARCH(1,1) Modells bezüglich der GPD dargestellt. Auch hier ist die Annahme von exponentialverteilten Zwischenankunftszeiten haltbar.

### 5.1.6 Beurteilung für den GARCH(1,1) Prozess

Betrachtet man die Ergebnisse für den GARCH(1,1) Fall in Abbildung 5.16 fällt das Ergebnis der POT-Methode auf. Jedoch konnte auch durch eine Änderung in der Berechnung (für  $Y_{k,j}$  wurde u.a. eine feste Schwelle gewählt und nicht mehr für jedes  $k$  eine Neue) keine signifikante Veränderung im Ergebnis erzielt werden. Dieses Ergebnis der starken Ausreißer in den hinteren Lags scheint mit der speziellen Struktur des GARCH(1,1) Prozesses zusammenzuhängen.

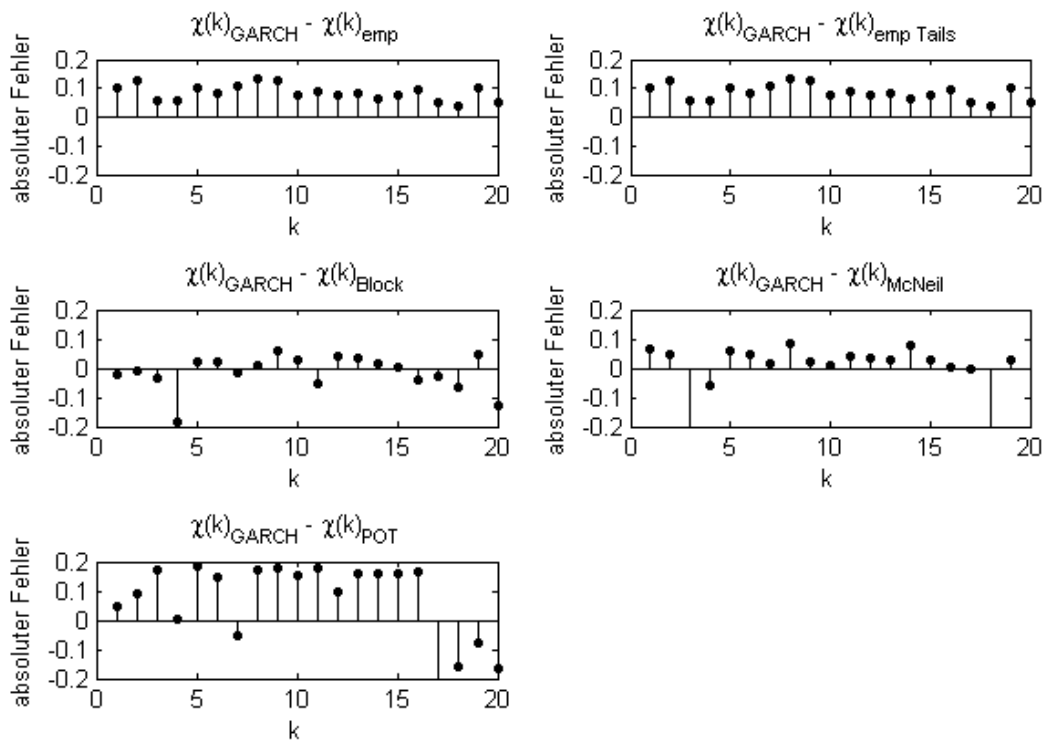


Abbildung 5.19: Absoluter Fehler der Methoden für einen simulierten quadrierten GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

Die Betrachtung des absoluten Fehlers im GARCH(1,1) Fall zeigt unter Einbezug des Kriteriums, dass ein absoluter Fehler bis  $|0.1|$  akzeptiert wird, dass die empirischen Schätzungen  $\chi(k)_{GARCH}$  am besten abbilden. Alle anderen Methoden liegen zum Teil stark über  $|0.1|$ .



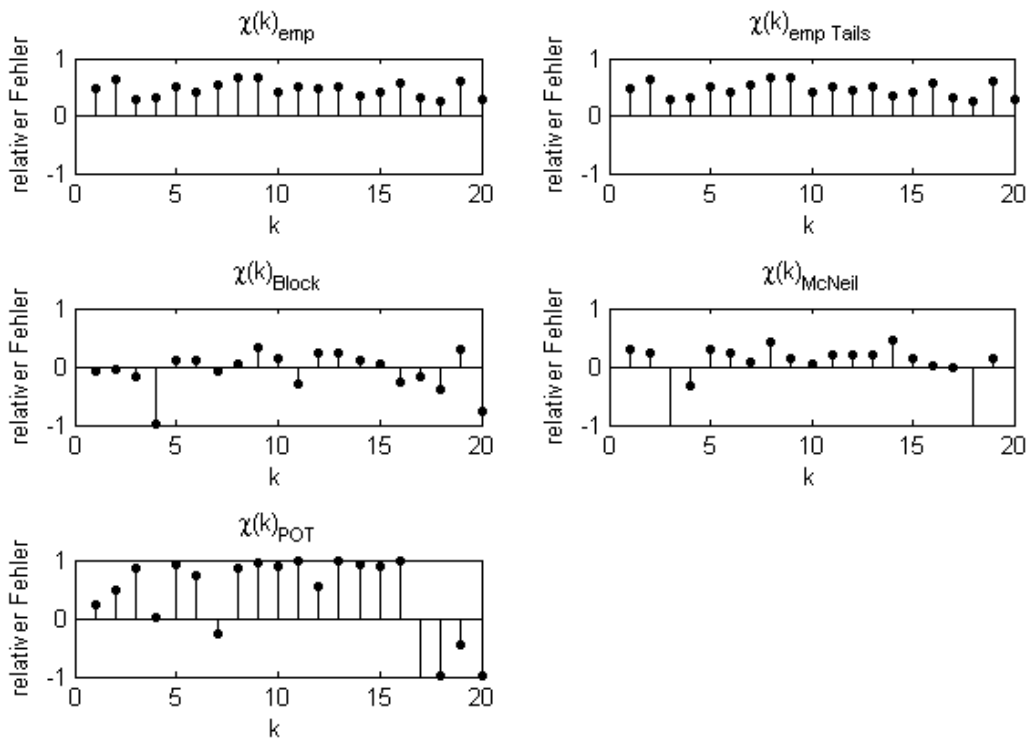


Abbildung 5.20: Relativer Fehler der Methoden für einen simulierten quadrierten GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

Die empirischen Schätzungen zeigen eine gute Näherung, unterschätzen aber  $\chi(k)_{GARCH}$ , wie sowohl in Abbildung 5.19 für den absoluten Fehler, als auch in Abbildung 5.20 für den relativen Fehler zu erkennen ist. Die Blockmaxima-Methode schwankt stärker als  $\chi(k)_{GARCH}$ , liegt aber von der Größenordnung nahe dran. Gleiches gilt für die Methode von McNeil. Der große absolute und relative Fehler der POT-Methode weist auf eine schlechte Schätzung von  $\chi(k)_{POT}$  für  $\chi(k)_{GARCH}$  hin.

### 5.1.7 AR(1) + ARCH(1) Prozess

In Abbildung 5.21 ist für einen simulierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ ,  $n = 10000$  (siehe Abbildung 4.7) die Kerndichteschätzung und die an die Momente angepasste Normalverteilung abgebildet. Es ist deutlich zu erkennen, dass der Prozess spitzgipflich ist und keiner Normalverteilung folgt.

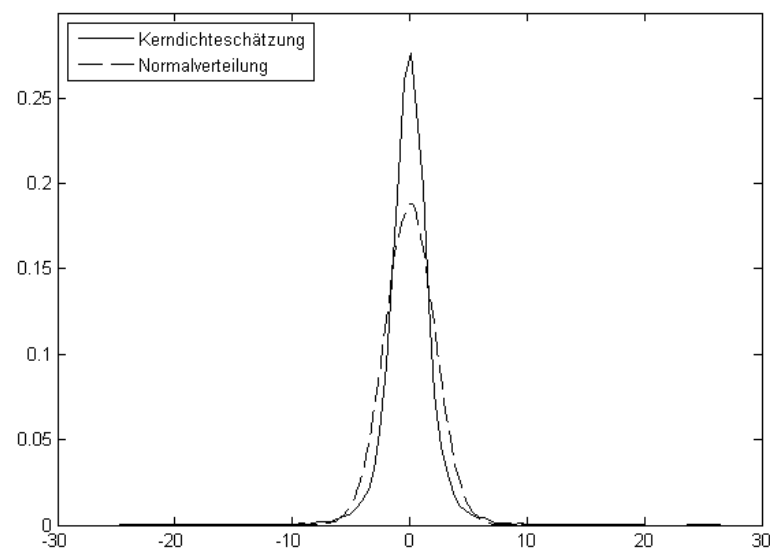


Abbildung 5.21: Kerndichteschätzung eines simulierten AR(1) + ARCH(1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable,  $n = 10000$  und Dichte der an die Momente angepasste Normalverteilung.

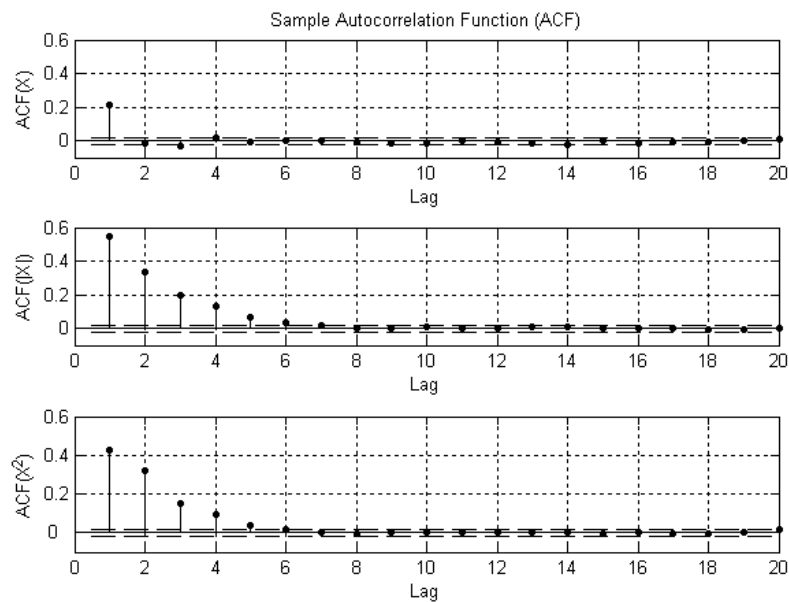


Abbildung 5.22: Oben: ACF eines simulierten AR(1) + ARCH(1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ . Mitte: ACF der absoluten Werte. Unten: ACF der quadrierten Werte.

Abbildung 5.22 zeigt die ACF eines simulierten AR(1) + ARCH(1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$  und standardnormalverteilter Noise-Variable für den einfachen, absoluten und quadrierten Prozess. Die ACF der absoluten und quadrierten Werte weicht deutlich von Null ab. Es wird ein ähnliches Verhalten für  $\chi(k)_{ARARCH}$  erwartet.

### Bestimmung mit MC Simulation

Der Tailindex in Borkovec und Klüppelberg (2001) und Kiefersbeck (1999) für den AR(1) + ARCH(1) Prozess  $(X_t)_{t \in \mathbb{N}}$  ist für den einfachen und nicht für den quadrierten Prozess berechnet worden. In den Formeln in Kapitel 4.3.2 ist dies bereits berücksichtigt, so dass das  $\kappa$  des einfachen Prozesses verwendet werden kann.

Die Wahl für den Tailindex kann mit dem Hill-Plot wie in Abbildung 5.23 erfolgen. Diese ist jedoch, wie bereits erwähnt, nicht eindeutig.

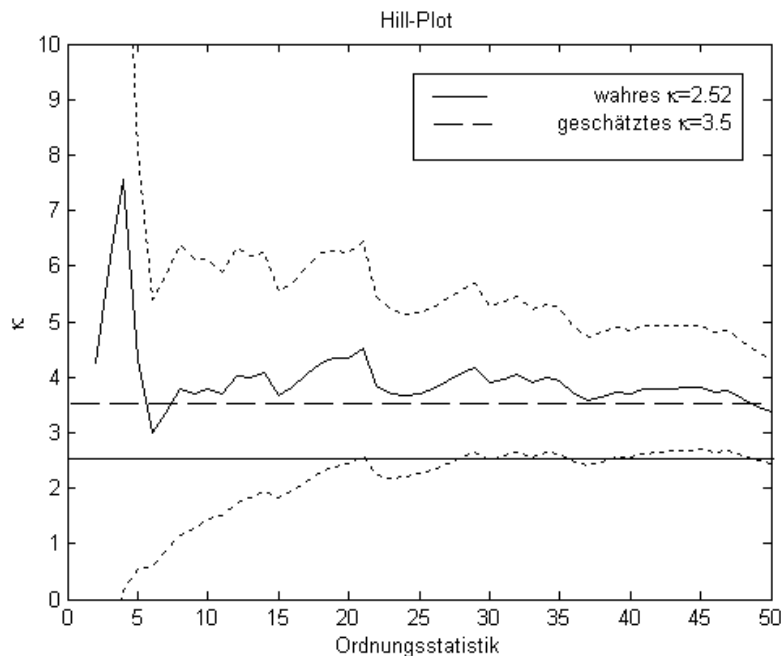


Abbildung 5.23: Hill-Plot für einen simulierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

$\chi(k)_{ARARCH}$  wird, wie auch beim ARCH(1) und GARCH(1,1) Prozess, mittels MC Simulation mit 10000 Wiederholungen unter Anwendung von Formel (4.17) bestimmt. Diese ist direkt in Formel (4.15) überführbar. Dadurch wird die Beziehung zwischen ARCH(1) und AR(1) + ARCH(1) Prozess verdeutlicht. Für ein AR(1) + ARCH(1) Modell mit  $\alpha_0 = 1$ ,  $\lambda = 0.2$ ,  $\alpha_1 = 0.8$  beträgt der Tailindex  $\kappa$  für standardnormalverteilten Noise 2.52. Vergleicht man Abbildung 5.24 mit Abbildung 5.3, dann erkennt man sehr gut den Zusammenhang zwischen den beiden Prozessen. In Tabelle A.4 sind die Werte zu Abbildung 5.24 angegeben, die die Berechnung von  $\chi(k)_{ARARCH}$  mittels MC Simulation zeigt.

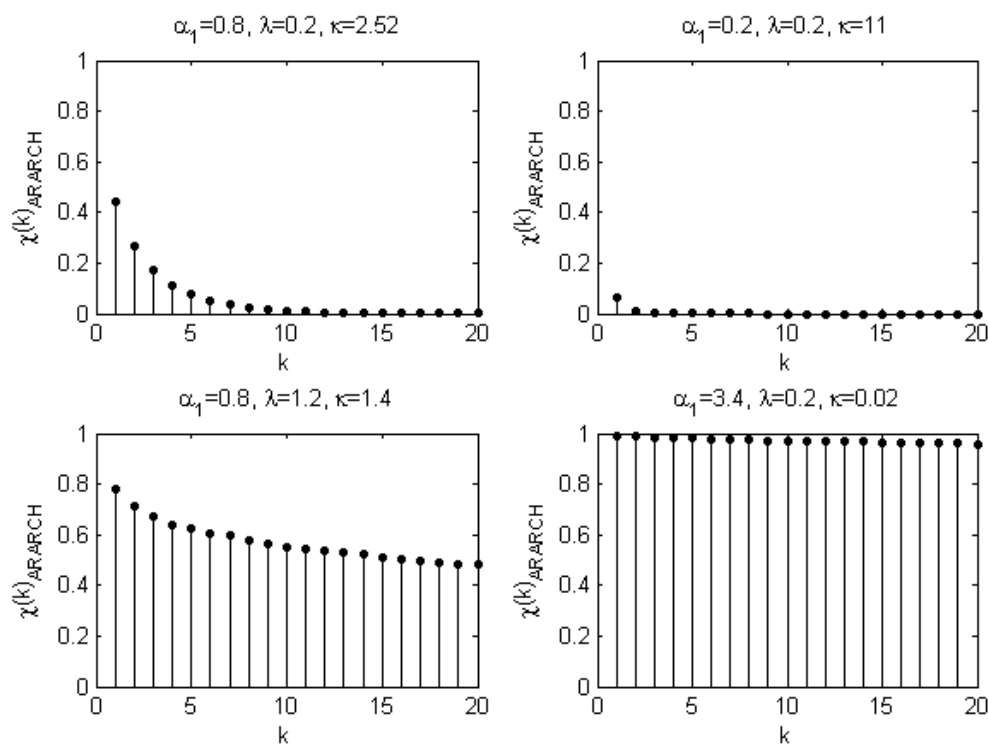


Abbildung 5.24:  $\chi(k)_{ARARCH}$  für einen simulierten quadrierten AR(1) + ARCH(1) Prozesse mit mit verschiedenen  $\alpha_0$ ,  $\alpha_1$ ,  $\lambda$  und  $\kappa$  für eine standardnormalverteilte Noise-Variable.

In Abbildung 5.25 sind die Ergebnisse für die verschiedenen Methoden im Vergleich zu  $\chi(k)_{ARARCH}$  dargestellt. Im Vergleich mit der ACF in Abbildung 5.22 fällt auf, dass die ACF der quadrierten Daten schneller abfällt wie  $\chi(k)_{ARARCH}$ .

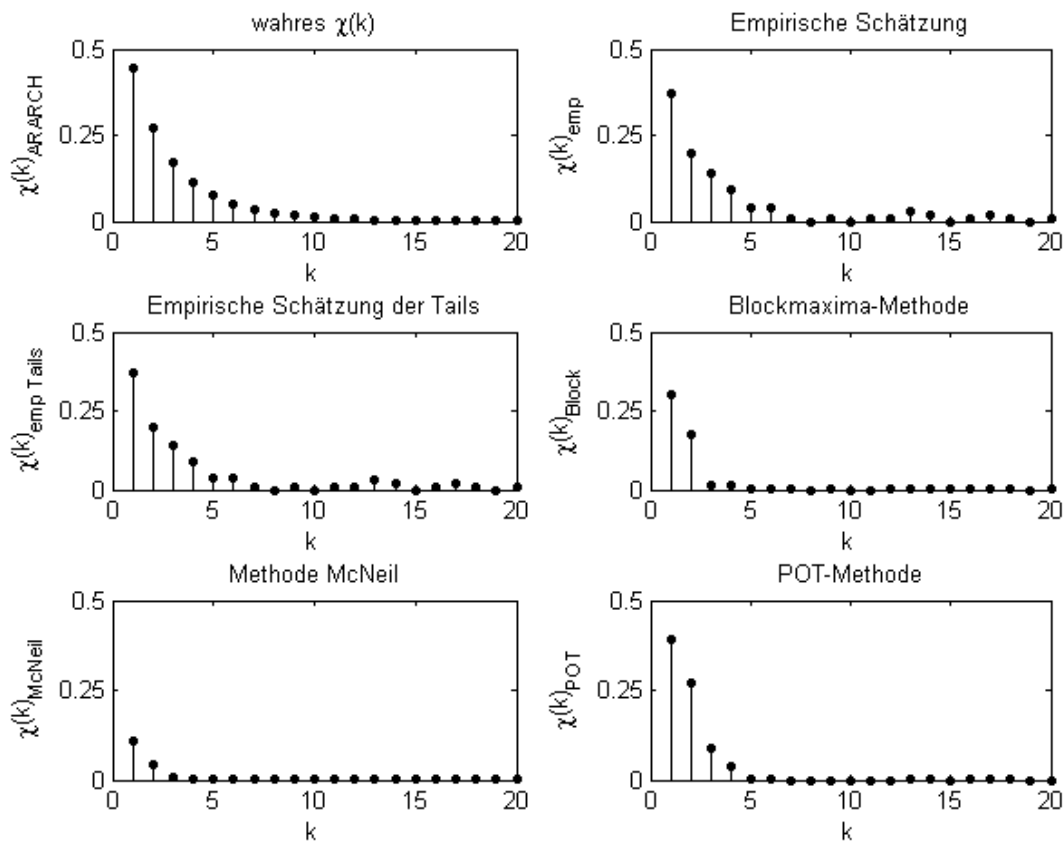


Abbildung 5.25: Plots der Methoden für einen simulierten quadrierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$  und  $n = 10000$ .

Das Verhalten aller Methoden ist sehr ähnlich. Dies ist in Abbildung 5.25 ersichtlich. Jedoch unterschätzen alle Methoden  $\chi(k)_{ARARCH}$ .

### Empirische Schätzung und Empirische Schätzung der Tails

Die empirische Schätzung erfolgt analog zur empirischen Schätzung im ARCH(1) und GARCH(1,1) Fall.  $\chi(k)_{emp}$  unterscheidet sich beim AR(1) + ARCH(1) Prozess wiederum nicht von dem Ergebnis für  $\chi(k)_{emp\ Tails}$ . Die Ergebnisse für diese beiden Methoden sind in Abbildung 5.25 rechts oben und in der Mitte links sowie in Tabelle A.4 zusammengefasst. Wählt man die Schwelle niedriger z.B.  $u = 10\%$  sind die Ergebnisse für  $\chi(k)_{emp\ Tails}$  für  $k = 1, 2, 3$  sehr nahe an den Werten von  $\chi(k)_{ARARCH}$ , allerdings fällt die Schätzung für  $\chi(k)_{emp\ Tails}$  dann auch langsamer gegen Null. Dies zeigt wiederum die Problematik der Schwellenwahl.

### Blockmaxima-Methode

Für das AR(1) + ARCH(1) Modell zeigt die Blockmaxima-Methode aus Kapitel 3.3.2 ein gutes Ergebnis. Es fällt auf, dass hier eine Blockgröße von  $r = 50$  wiederum sehr gut geeignet ist, da die Nullhypothesen des  $\mathcal{LM}$ - und  $\mathcal{LB}$ -Tests nicht abgelehnt werden. Die Ergebnisse in Abbildung 5.25 in der Mitte rechts sind der Darstellung der ACF der quadrierten Werte in Abbildung 5.2 ähnlich. Allerdings sind die Werte von  $\chi^{(k)}_{Block}$ , wie man in Tabelle A.4 erkennt, niedriger als die der ACF des quadrierten Prozesses. Für den Formparameter ergab die Schätzung im Mittel für  $X_n$  und  $Y_{k,j}$   $\hat{\xi} = 0.4$ . Dies entspricht ungefähr dem Kehrwert von  $\kappa$ . Allerdings weicht der Schätzer für  $X_n$  mit  $\hat{\xi} = 0.7576$  stark ab. Dies gilt auch für die  $\hat{\xi}$  für  $Y_{k,j}$ .

### Methode von McNeil

Wie bei den bereits betrachteten Modellen ist auch beim AR(1) + ARCH(1) Modell das Ergebnis für  $\chi^{(k)}_{Block}$ ,  $\chi^{(k)}_{McNeil}$  und  $\chi^{(k)}_{POT}$  von der Wahl für  $\alpha_0$  abhängig. Wählt man  $\alpha_0$  sehr klein, ist das Ergebnis sehr klein und weist imaginäre Anteile auf. Zur Überprüfung der Modellwahl wird auch hier der QQ-Plot in Abbildung 5.26 herangezogen. Die Abweichungen der Residuen sind hier deutlich größer als im ARCH(1) Fall, aber dennoch akzeptabel. Als Blockgröße wurde  $r = 50$  und als Schwelle  $u = 1\%$  gewählt. Es ergibt sich für den Extremalen Index ungefähr 0.7. Für die Schätzung des Quantils wurde  $m = 3$  gewählt. Es wird das 98.5%-Quantil geschätzt.

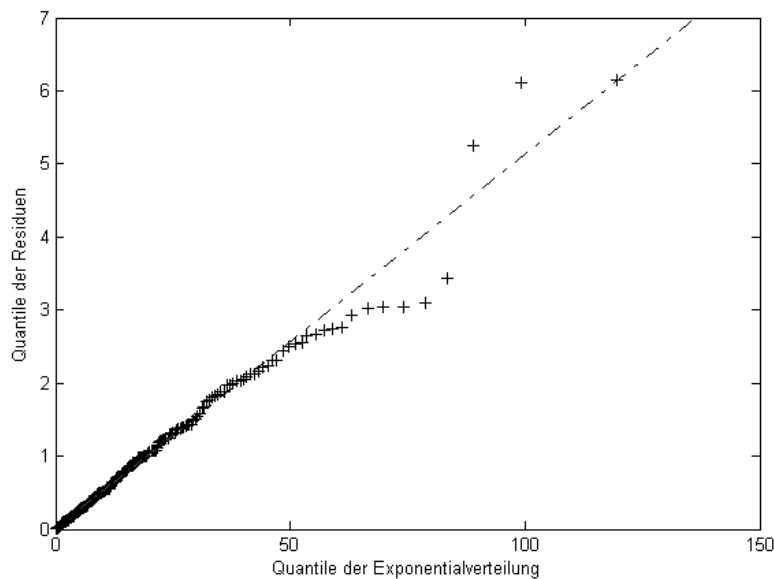


Abbildung 5.26: QQ-Plot der Residuen eines simulierten quadrierten AR(1) + ARCH(1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$  und  $n = 10000$ .

Die Ergebnisse für  $\chi^{(k)}_{McNeil}$  sind in Abbildung 5.25 links unten und in Tabelle A.4 zusammengefasst.

### POT-Methode

Bei der POT-Methode hat sich  $u = 1\%$  als gute Wahl gezeigt. Die Exzesse sind hier scheinbar unabhängig laut  $\mathcal{LB}$ - und  $\mathcal{LM}$ -Test. Wie man in Abbildung 5.25 und Tabelle A.4 erkennt, ist das Verhalten von  $\chi(k)_{POT}$  dem Verhalten der ACF der quadrierten Werte in Abbildung 5.22 und  $\chi(k)_{ARARCH}$  in Abbildung 5.24 sehr ähnlich. Bei der POT-Methode ergibt sich für  $\hat{\xi}$  im Mittel wieder der Kehrwert von  $\kappa$ . Der Wert für  $X_n$  weicht mit  $\hat{\xi} = 0.7383$  stark ab. Dies ist auch bei der Blockmaxima-Methode bereits aufgefallen. Die Werte für  $\hat{\xi}$  der GPD korrespondieren wie erwartet mit den  $\hat{\xi}$  der GEV. In Abbildung 5.27 sind verschiedene Plots eines simulierten quadrierten AR(1) + ARCH(1) Prozesses im Bezug auf die GPD dargestellt. Im Plot der Daten und der Quantile der GPD sind Abweichungen erkennbar. Diese liegen allerdings im Rahmen. Daher kann die Verteilungsannahme akzeptiert werden.

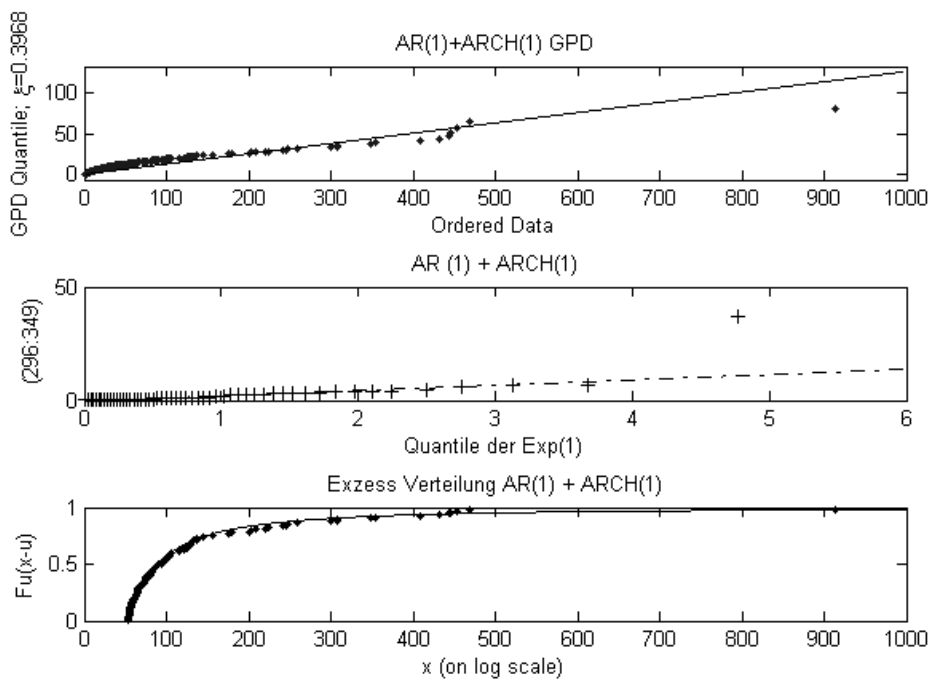


Abbildung 5.27: Oben: GPD Quantile für einen simulierten quadrierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ ,  $n = 10000$  und standardnormalverteilter Noise-Variable. Mitte: Quantile der Exponentialverteilung und Zwischenankunftszeiten ( $X_{296} : X_{349}$ ). Unten: Exzess-Verteilung.

## 5.1.8 Beurteilung für den AR(1) + ARCH(1) Prozess

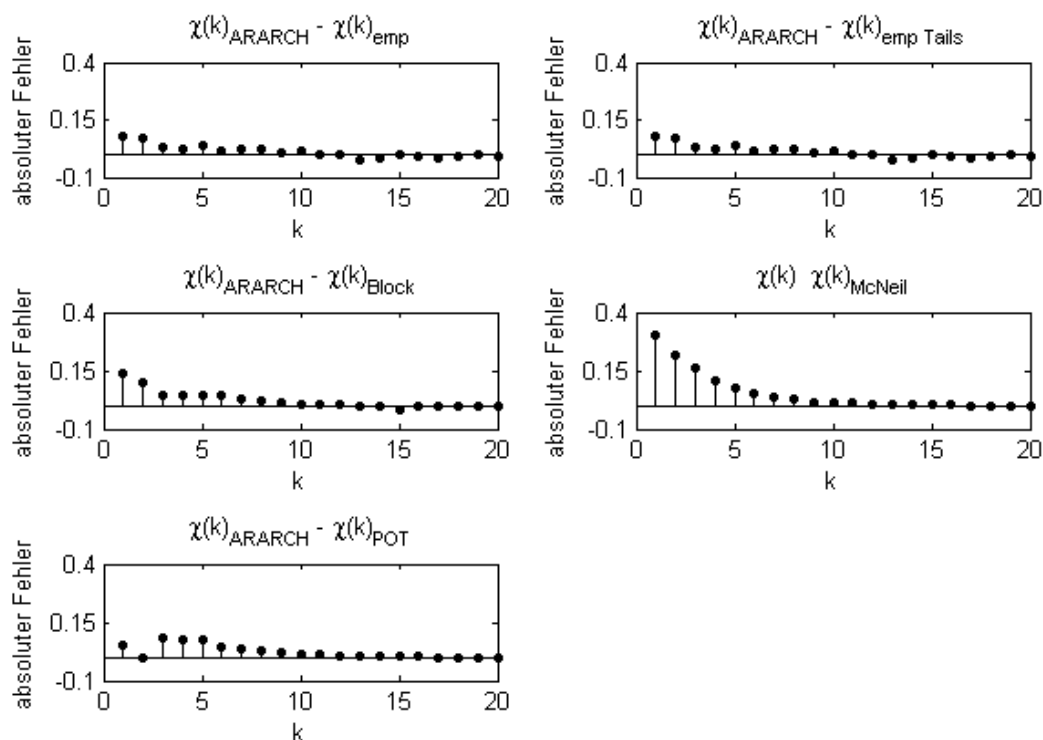


Abbildung 5.28: Absoluter Fehler der Methoden für einen simulierten quadrierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

In Abbildung 5.28 erkennt man, dass der absolute Fehler für alle Methoden signifikant vorhanden ist. Vor allem in den ersten Lags sind die Fehler hoch (besonders bei der Vorgehensweise von McNeil). Da eine Abweichung von  $|0.1|$  als akzeptabel angesehen wird, befinden sich alle Methoden bis auf die Methode von McNeil im Rahmen der Fehlertoleranz.



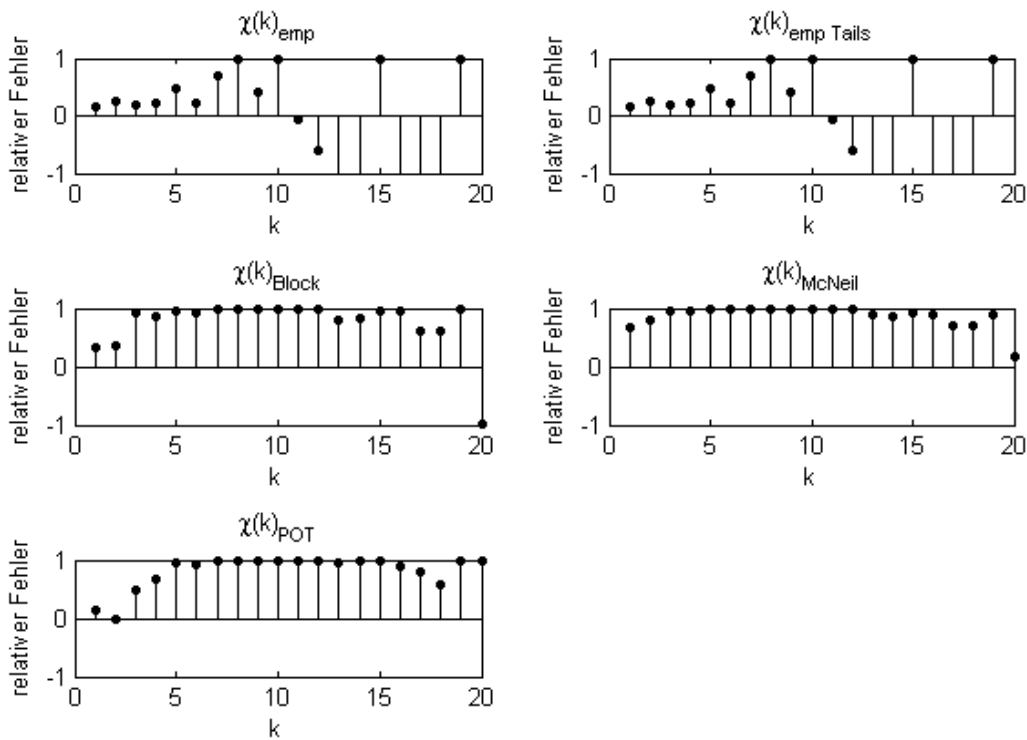


Abbildung 5.29: Relativer Fehler der Methoden für einen simulierten quadrierten AR(1)+ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

In Abbildung 5.29 zeigt sich bei den empirischen Schätzungen ein besonders hoher negativer Fehler in späteren Lags. Die relativen Fehler weisen für alle Methoden zum Teil einen Wert von  $|1|$  auf. Dies entspricht einer Abweichung von 100%. Die Werte der absoluten und relativen Fehler sind in Tabelle A.5 und A.6 dargestellt.

Im Folgenden werden die Methoden auf einen realen Datensatz angewendet.

## 5.2 Anwendung auf einen Datensatz

Für diese Arbeit wurden neben den Daten der Allianz Aktie auch Hochfrequenzdaten und Versicherungsdaten betrachtet. Dabei handelt es sich um Log>Returns der Intelaktie im Intervall von fünf Minuten. Ein Tag umfasst 77 Log>Returns. Insgesamt umfasst der Datensatz  $n = 14476$  Daten. Allerdings bestand dabei das Problem, dass diese nicht eindeutig einer Verteilung mit schweren Tails zuzuordnen waren und damit die Schätzung teilweise nicht die gewünschten Ergebnisse lieferte.

Da Versicherungsdaten als unabhängig gelten, wurden diese verwendet, um die Methoden

zu überprüfen. Es wurden die Daten der dänischen Feuerversicherung vom 03.01.1980 - 31.12.1990 betrachtet. Die Daten haben eine Länge von  $n = 2167$ . Wie erwartet lieferte die Anpassung dieser Daten an ein Modell aus Kapitel 4 Parameter der Größenordnung Null und damit wurde auch  $\chi(k)_{dan}$  Null. Auch für die Methoden aus Kapitel 3.3 ergaben sich stets sehr kleine von Null kaum abweichende Ergebnisse. Dies wurde auch erwartet. Es zeigt sich also das Problem einen geeigneten Datensatz zu finden.

Die Ergebnisse der vorangehenden Kapitel werden nun auf die Log>Returns der Allianz Aktie angewendet. Dabei wird die Modellwahl überprüft sowie die Aussagekraft der Ergebnisse. Es wird erwartet, dass Anzeichen auftreten, dass diese mit einem GARCH Modell dargestellt werden können, da Finanzmarktdaten mit einem GARCH Modell häufig sehr gut modelliert werden können. Die Verteilung der Log>Returns in diesem Kapitel ist nicht bekannt. Es wird jedoch angenommen, dass diese schwere Tails besitzen. Die Stylized Facts legen nahe, dass die Log>Returns von Finanzzeitreihen schwere Tails haben. Als Hilfsmittel kann man den ME-Plot aus Abbildung 3.2 heranziehen. Dieser bestätigt die Annahme der schweren Tails. Die Aktienkursdaten, die im Rahmen dieser Arbeit verwendet werden, basieren auf Tagesschlusskursen von Bloomberg im Zeitraum vom 02.01.1990 bis zum 30.12.2004. Betrachtet werden die Kurse der Allianz (ALL) mit einer Stichprobenanzahl  $n$  von 3774. Die Daten wurden um Wochenenden und Feiertage sowie um sonstige Tage ohne Kurse bereinigt. Ein beispielhafter Plot für die Aktienkursdaten im Vergleich zu den Log>Returns ist in Abbildung 5.30 zu finden. Zum Vergleich ist unten der Plot eines GARCH(1,1) Prozesses mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.01$  und  $\beta = 0.98$  dargestellt.

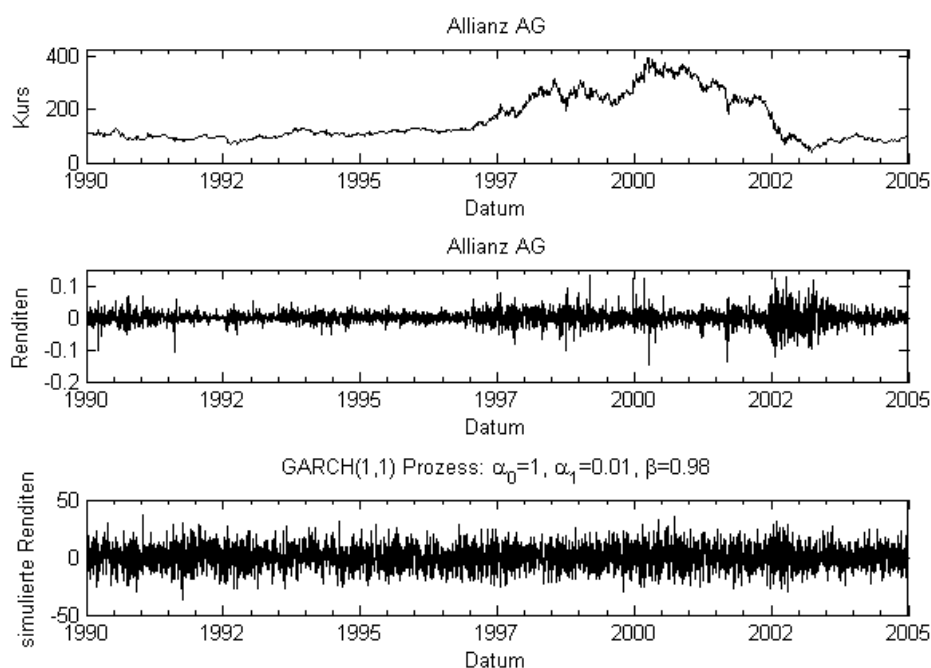


Abbildung 5.30: Kurse und Log>Returns der Aktienkursdaten der Allianz AG, sowie simulierte Renditen.

Die vorgestellten Daten wurden als Log>Returns auch mit deskriptiven und induktiven Statistiken analysiert. Tabelle 5.9 zeigt mit der Stichprobengröße  $n$ , dem kleinsten ( $Min$ ) und dem größten Wert ( $Max$ ) sowie den standardisierten Momenten, dem empirischen Mittelwert ( $\mu$ ), dem Median ( $x_{0.5}$ ), der Standardabweichung ( $s$ ), der Varianz ( $s^2$ ), der Schiefe ( $\mathcal{S}$ ) und der Kurtosis ( $\mathcal{K}$ ) die deskriptiven Statistiken des verwendeten Datensatzes.

	$n$	Min	Max	$\mu$	$x_{0.5}$	$s$	$s^2$	$\mathcal{S}$	$\mathcal{K}$
ALL	3773	-0.1493	0.1319	0.0000	0.0000	0.0220	0.0005	-0.0703	8.4691

Tabelle 5.9: Deskriptive Statistiken der Log>Returns der Allianz Aktie.

Der Mittelwert ist Null. Die Varianz ist über den gesamten Zeitraum berechnet ebenfalls nicht sehr groß. Bei der Schiefe ist festzustellen, dass die Daten eine Linksschiefe im Vergleich zur Normalverteilung aufweisen. Die Daten sind mehr oder weniger deutlich im positiven Bereich und somit leptokurtisch im Vergleich zur Normalverteilung. Die Angabe des Minimalwertes bzw. des Maximalwertes dient dazu, die maximalen Ausschläge der Zeitreihe zu erkennen.

In Abbildung 4.1 erkennt man, dass die Allianz Log>Returns schwerere Tails als die Normalverteilung haben. Dies bestätigt die Ergebnisse der deskriptiven Statistik in Tabelle 5.9 und legt die Vermutung nahe, dass die Daten mit einem nicht-linearen Zeitreihenmodell modelliert werden können.

Das Verhalten der ACF der quadrierten Werte in Abbildung 5.31 ist dem Verhalten der ACF des quadrierten GARCH(1,1) Prozesses in Abbildung 5.13 sehr ähnlich.

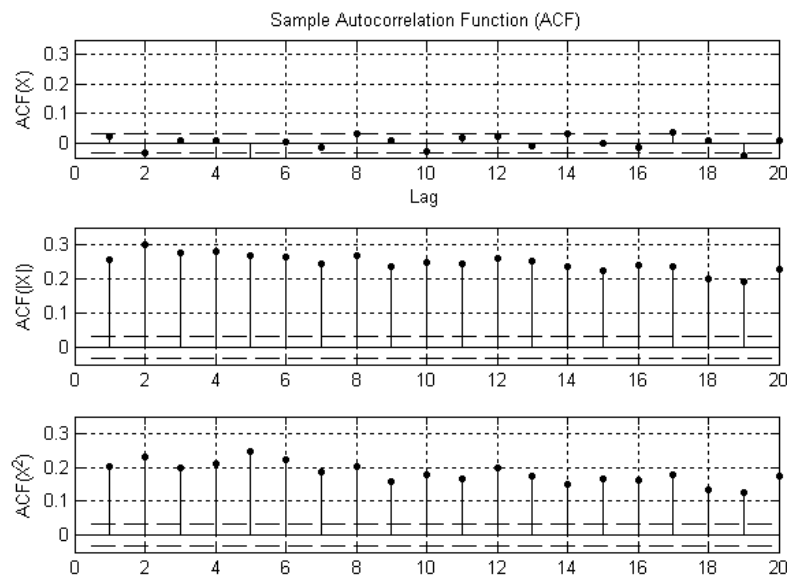


Abbildung 5.31: Oben: ACF der Allianz Log>Returns. Mitte: ACF der absoluten Werte der Allianz Log>Returns. Unten: ACF der quadrierten Werte.

Für die induktiven Statistiken gibt es verschiedene Tests, mit denen man überprüfen kann, ob ein Datensatz der Normalverteilung folgt bzw. welches Korrelationsverhalten vorliegt. Diese sind in Tabelle 5.10 angegeben. Enthalten ist der Jarque-Bera Test ( $\mathcal{JB}$ ) auf Normalität nach Jarque und Bera (1987). Mit dem Ljung-Box Test ( $\mathcal{LB}$ ) wird auf die zeitliche Korrelation der Returns und mit dem Ljung-Box Test mit quadrierten Returns ( $\mathcal{LB}^2$ ) auf Autokorrelation nach Ljung und Box (1978) getestet. In der letzten Spalte der Tabelle 5.10 befindet sich der von Engle (1982) vorgeschlagene  $\mathcal{LM}$ -Test auf (G)ARCH-Effekte. Folgende Notation wird verwendet:  $\mathcal{LB}(\sqrt{n}) =$  Teststatistik des Ljung-Box Tests mit Lag  $(\sqrt{n})$ .

	$\mathcal{JB}$	$\mathcal{LB}(\sqrt{n})$	$\mathcal{LB}^2(\sqrt{n})$	$\mathcal{LM}(\sqrt{n})$
<i>ALL</i> ( <i>p</i> -Wert)	4705.33 (0,000)	119.8 (0,000)	5144.9 (0,000)	715.29 (0,000)

Tabelle 5.10: Induktive Statistiken der Log>Returns der Allianz Aktie.

Bei den ausgewählten Daten zeigt der Jarque-Bera Test, dass die Nullhypothese auf Normalverteilung mit einem  $p$ -Wert  $< 0.05$  deutlich abgelehnt wird. Die Teststatistik liegt stark über den kritischen Werten. Wendet man den Ljung-Box Test auf die quadrierten Log>Returns an, so erhält man eine Ablehnung der Nullhypothese. Dieses Ergebnis weist auf Autokorrelation in den quadrierten Log>Returns hin. Erwartungsgemäß lehnt auch der  $\mathcal{LM}$ -Test die Nullhypothese, also das Vorhandensein keiner ARCH-Effekte, ab. In den Abbildungen 5.32 erkennt man, dass die Log>Returns der Allianz keiner Normalverteilung folgen. Dies wird aufgrund der deskriptiven und induktiven Statistiken auch erwartet.

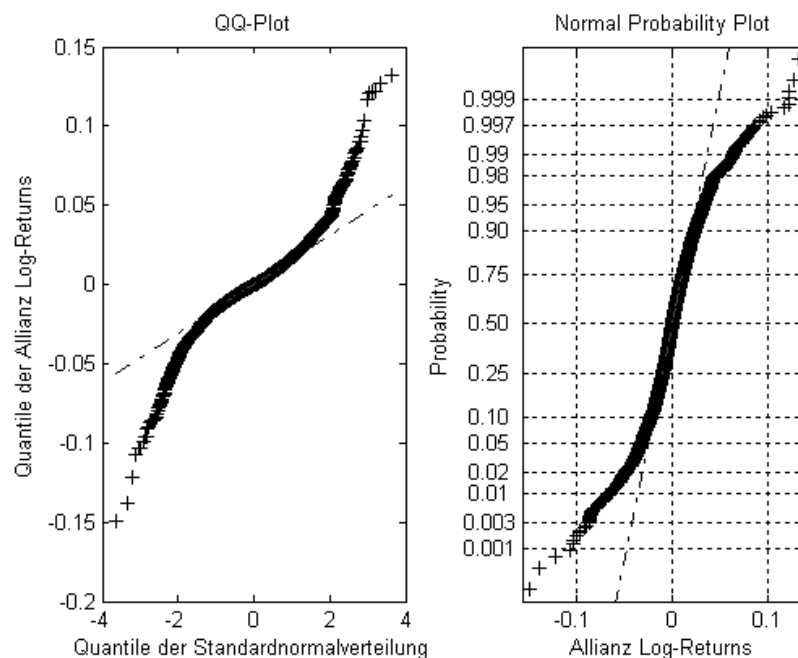


Abbildung 5.32: QQ- und Normal Probability Plot der Allianz Log>Returns.

Zusammenfassend kann man für die Analyse der Allianz Log>Returns festhalten, dass sie, wie erwartet wurde, nach deskriptiver, induktiver und graphischer Analyse ein ähnliches Verhalten wie ein simulierter ARCH(1) oder GARCH(1,1) Prozess aufweisen.

### Bestimmung mit MC Simulation

Zunächst werden die Allianz Log>Returns an ein ARCH(1) Modell angepasst. Es ergibt sich  $\alpha_0 = 0.00033098$  und  $\alpha_1 = 0.30974$ . Als Tailindex wird der Wert aus Tabelle 4.1 für  $\alpha_1 = 0.3$ , also  $\kappa = 4.180$  verwendet. Da für Finanzmarktdaten die Modellierung mit einem GARCH(1,1) Modell meist passender ist, werden die Allianz Log>Returns auch an ein GARCH(1,1) Modell angepasst. Es ergeben sich dafür folgende Parameter:  $\alpha_0 = 3.696 \cdot 10^{-6}$ ,  $\alpha_1 = 0.073084$  und  $\beta = 0.92042$ . Für den Tailindex  $\kappa$  wird der Wert 5 verwendet, der mit dem Hill-Plot geschätzt wurde (siehe dazu Abbildung 5.33). Wenn man die Daten an ein AR(1) + ARCH(1) Modell anpasst, erhält man folgende Parameter:  $\alpha_0 = 3.2303 \cdot 10^{-4}$ ,  $\alpha_1 = 0.3357$ ,  $\lambda = 0.0966$ . Mit dem Hill-Plot ergibt sich für  $\kappa = 4.5$ . Die Parameter wurden sowohl mit der *garchfit* Routine aus Matlab als auch mit einem selbst geschriebenen ML-Schätzer geschätzt. Beide Methoden ergaben dieselben Parameter. Die Noise-Variable ist standardnormalverteilt.

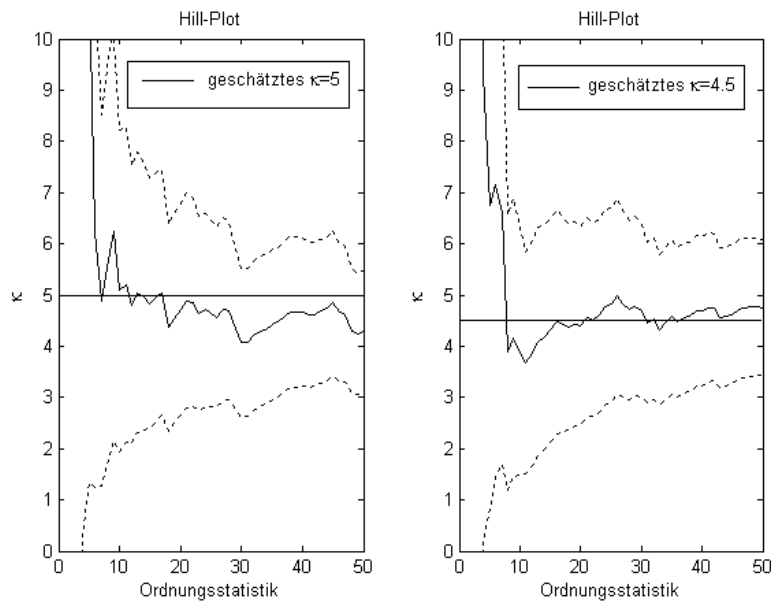


Abbildung 5.33: Hill-Plot der angepassten Prozesse der Allianz Log>Returns an ein GARCH(1,1) Modell mit  $\alpha_0 = 3.696 \cdot 10^{-6}$ ,  $\alpha_1 = 0.073084$ ,  $\beta = 0.92042$  (links) und ein AR(1) + ARCH(1) Modell mit  $\alpha_0 = 3.2303 \cdot 10^{-4}$ ,  $\alpha_1 = 0.3357$ ,  $\lambda = 0.0966$  (rechts).

Nun wird  $\chi(k)_{Prozess}$  für die angepassten quadrierten Log>Returns der Allianz bestimmt. Die wahren  $\chi(k)_{Prozess}$  sind in Abbildung 5.34 dargestellt. Erwartungsgemäß ist  $\chi(k)_{GARCH}$  beinahe konstant, während die anderen beiden bis zum Lag 3 abfallen.

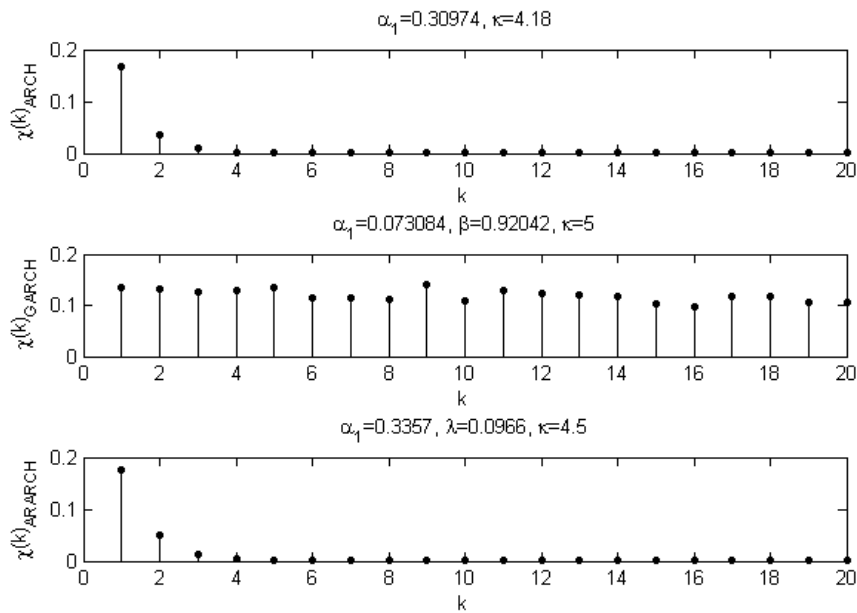


Abbildung 5.34:  $\chi(k)_{Prozess}$  mit Formel (4.15) für ARCH(1), Formel (4.16) für GARCH(1,1) und Formel (4.17) für AR(1) + ARCH(1) mit den Parametern für die angepassten Log>Returns.

Die Schätzungen für  $\chi(k)$  mit den verschiedenen Methoden aus Kapitel 3 sind in Abbildung 5.36 dargestellt. Dazu werden die quadrierten Log>Returns der Allianz Aktie verwendet. Diese sind in Abbildung 5.35 dargestellt.

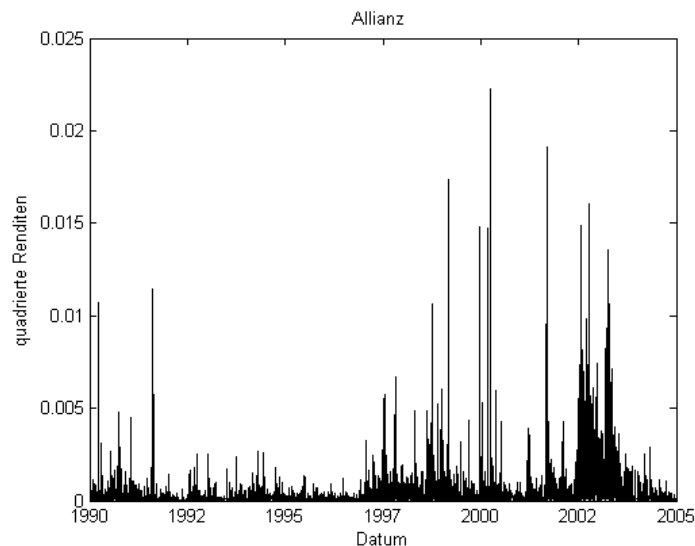


Abbildung 5.35: Quadrierte Allianz Log>Returns.

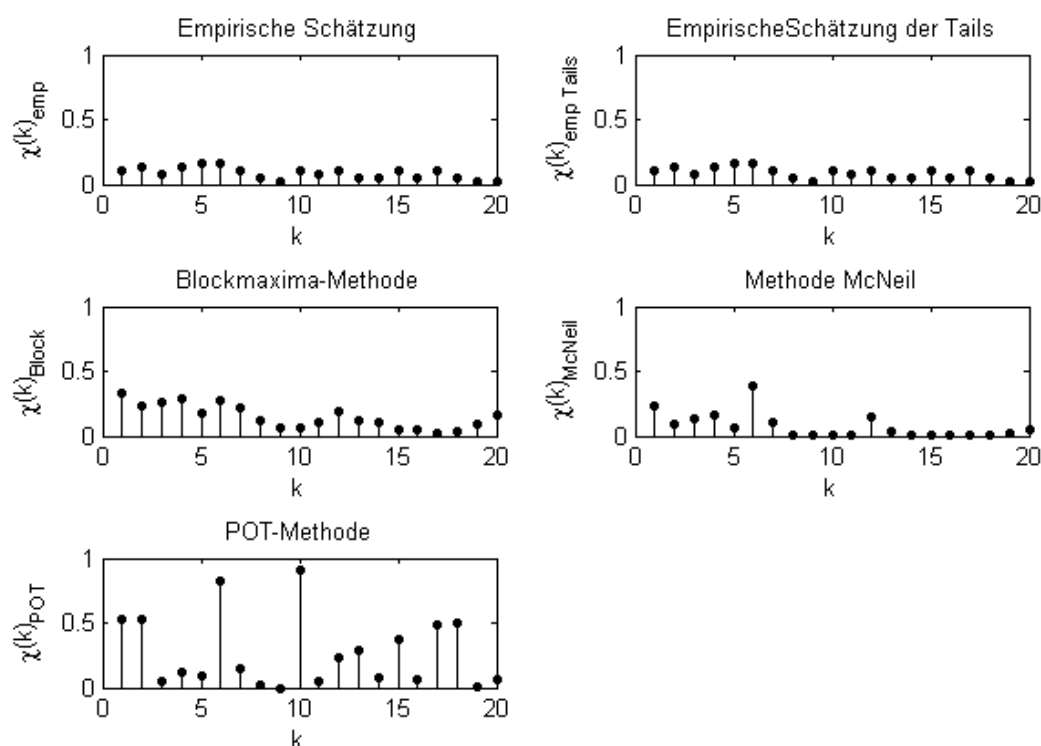


Abbildung 5.36: Ergebnisse für  $\chi(k)_{Methode}$  mit den verschiedenen Schätzmethoden für die quadrierten Allianz Log>Returns.

Die Ergebnisse für die einzelnen Methoden werden im Folgenden beschrieben.

### Empirische Schätzung und empirische Schätzung der Tails

Die empirische Schätzung und die empirische Schätzung der Tails liefern erwartungsgemäß das gleiche Ergebnis. In Abbildung 5.36 und Abbildung 5.15 kann man eine Ähnlichkeit zwischen der empirischen Schätzung für die Allianz Daten und  $\chi(k)_{GARCH}$  erkennen. Daher liegt die Vermutung nahe, dass sich ein GARCH(1,1) Modell zur Modellierung der Allianz Log>Returns eignet. Dieses Ergebnis ist plausibel, da Finanzmarktdaten wie bereits erwähnt, häufig über ein GARCH(1,1) Modell dargestellt werden (siehe Kapitel 4.1).

### Blockmaxima-Methode

Bei der Blockmaxima-Methode werden für eine Blockgröße  $r = 100$  die Nullhypothesen des  $\mathcal{LM}$ - und  $\mathcal{LB}$ -Tests in fast allen Fällen nicht abgelehnt. Daher wird diese Blockgröße verwendet, da dann noch 37 Blöcke bzw. Blockmaxima zur Verfügung stehen. Wie bei der

Betrachtung des simulierten quadrierten GARCH(1,1) Prozesses, weicht das  $\hat{\xi}$  der GEV stark vom  $\xi$  der POT-Methode ab.

### Methode von McNeil

Es wurde  $r = 100$  gewählt. Bei einer Blockgröße von  $r = 100$  sind fast alle Blockmaxima unabhängig. Es wurde  $m = 1$  gewählt. Dies entspricht der Schätzung des 97%-Quantils. Für die Schätzung des Extremalen Index wurde  $u = 1\%$  gewählt. Es liegen dann aufgrund der Clusterbildung noch 4-9 Blöcke über der Schwelle. Für den Extremalen Index ergibt sich  $\hat{\theta} = 0.2$ . In Abbildung 5.37 ist der QQ-Plot der Quantile der Exponentialverteilung und der Residuen der quadrierten Allianz Log>Returns dargestellt. Diese liegen zum Großteil auf einer Geraden.

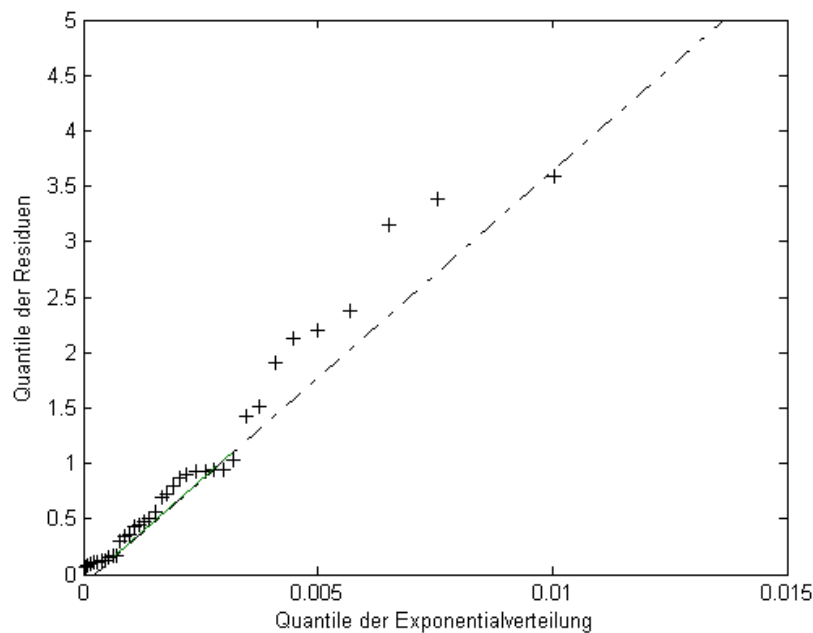


Abbildung 5.37: QQ-Plot der Residuen der quadrierten Allianz Log>Returns.

### POT-Methode

Für  $u = 1\%$  ergeben sich bei der Schätzung der Parameter der GPD teilweise negative  $\hat{\xi}$  und daher keine brauchbaren Werte für  $\chi(k)_{POT}$ , da für die ML-Schätzung zu wenig Daten vorhanden sind. Für die Schwelle  $u = 2\%$  ergeben sich verwendbare Werte, die jedoch stark schwanken. Das Ergebnis der POT-Methode ähnelt dem Ergebnis für einen simulierten quadrierten GARCH(1,1) Prozess in Abbildung 5.16 rechts unten. Die Zeiten zwischen den Exzessen werden mit einem QQ-Plot auf die Exponentialverteilung überprüft. In Abbildung 5.38 unten erkennt man, dass diese Annahme im Großen und Ganzen eingehalten wird. Die Allianz Log>Returns über der Schwelle  $u = 2\%$  folgen der kumulierten Dichtefunktion der geschätzten GPD. Das geschätzte Modell ist als Kurve geplottet.



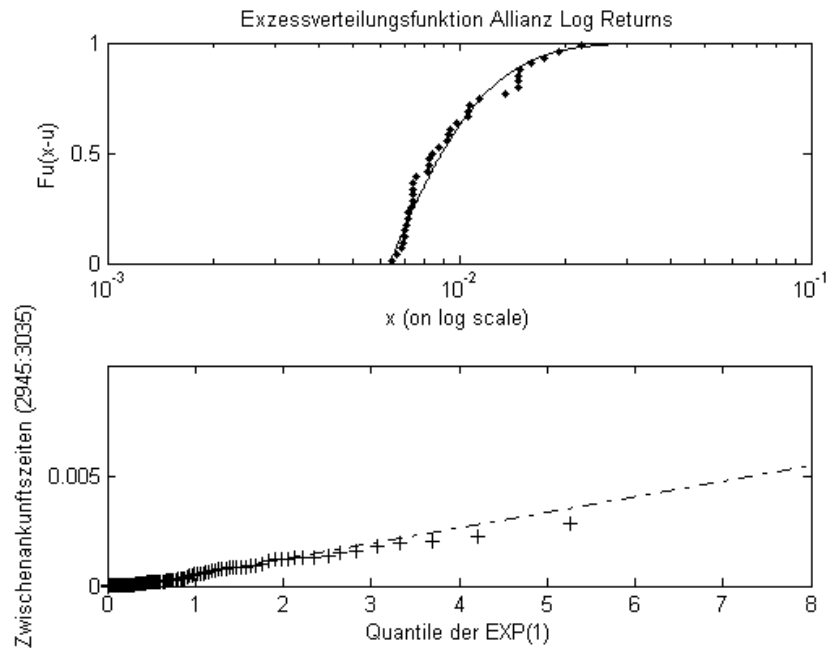


Abbildung 5.38: Oben: Exzess Verteilung für die quadrierten Allianz Log>Returns. Unten: Quantile der Exponentialverteilung.

In Abbildung 5.39 sind die Allianz Log>Returns gegen die Quantile der GPD geplottet. Als Formparameter wurden dabei die Kehrwerte der Tailindizes der angepassten Modelle verwendet. Die Abweichungen sind zum Teil erheblich.

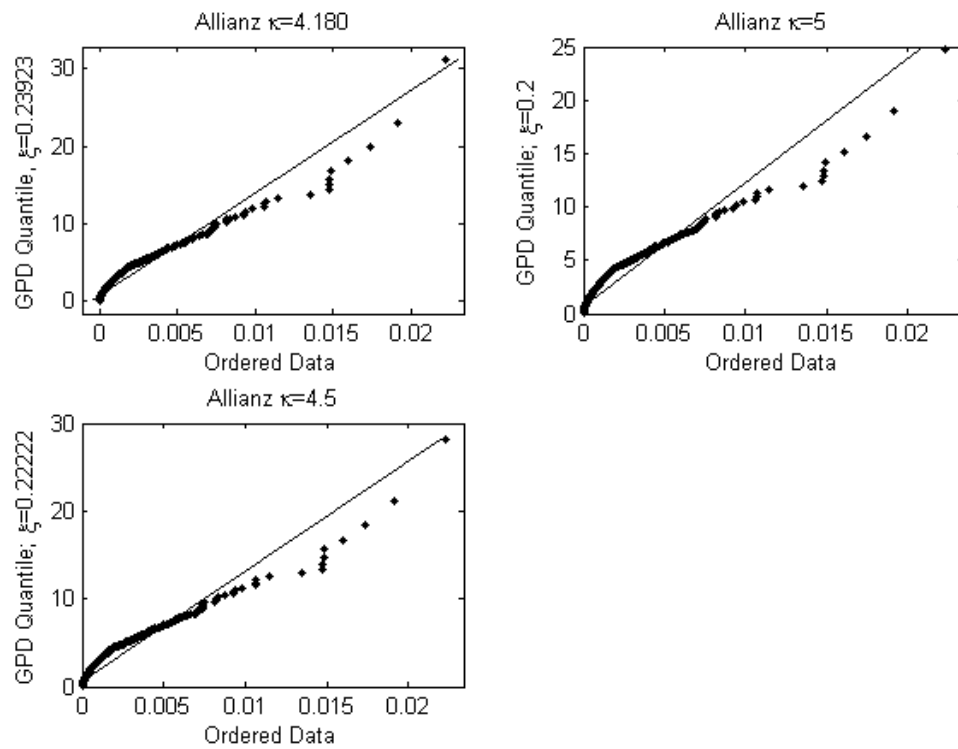


Abbildung 5.39: Oben: GPD Plots für die quadrierten Allianz Log>Returns mit dem  $\kappa$  des angepassten ARCH(1) Prozesses. Mitte: GARCH(1,1) Prozesses. Unten: AR(1) + ARCH(1) Prozess.

In Tabelle 5.11 sind die Ergebnisse für die verschiedenen Methoden in numerischer Form zusammengefasst.

Lag $k$	$\chi^{(k)}_{emp}$	$\chi^{(k)}_{emp\ Tails}$	$\chi^{(k)}_{Block}$	$\chi^{(k)}_{McNeil}$	$\chi^{(k)}_{POT}$
1	0.1081	0.1081	0.3280	0.2258	0.5289
2	0.1351	0.1352	0.2273	0.0939	0.5226
3	0.0811	0.0811	0.2551	0.1352	0.0435
4	0.1351	0.1353	0.2928	0.1667	0.1144
5	0.1622	0.1624	0.1818	0.0613	0.0913
6	0.1622	0.1624	0.2763	0.3884	0.8268
7	0.1081	0.1083	0.2156	0.1050	0.1427
8	0.0541	0.0542	0.1237	0.0111	0.0269
9	0.0270	0.0271	0.0702	0.0011	0.0000
10	0.1081	0.1084	0.0612	0.0011	0.9153
11	0.0811	0.0813	0.1008	0.0058	0.0449
12	0.1081	0.1085	0.1874	0.1439	0.2340
13	0.0541	0.0542	0.1226	0.0321	0.2823
14	0.0541	0.0543	0.1070	0.0024	0.0824
15	0.1081	0.1085	0.0433	0.0002	0.3730
16	0.0541	0.0543	0.0489	0.0004	0.0570
17	0.1081	0.1086	0.0271	0.0001	0.4907
18	0.0541	0.0543	0.0366	0.0002	0.4962
19	0.0270	0.0272	0.0901	0.0151	0.0135
20	0.0270	0.0272	0.1690	0.0525	0.0698

Tabelle 5.11: Ergebnisse der Methoden für die quadrierten Allianz Log&gt;Returns.

Die Ergebnisse für die quadrierten Allianz Log>Returns sind weder dem ARCH(1) Fall noch dem GARCH(1,1) oder AR(1) + ARCH(1) Fall eindeutig zuzuordnen. Eine Entscheidung zur Modellwahl ist daher nicht eindeutig möglich.

# Kapitel 6

## Zusammenfassung und Ausblick

Ein Problem der Extremwerttheorie ist die Verfügbarkeit einer ausreichend großen Anzahl an geeigneten Daten. Da Extrema die Folge von seltenen Ereignissen sind, ist es von großer Bedeutung so viele Daten wie möglich zu verwenden, da nicht die gesamte Stichprobe verwendet werden kann. Dieses Problem war auch in dieser Arbeit gegeben. Obwohl eine Stichprobe von 10000 Daten verwendet wird, treten teilweise Probleme bei der Schätzung der Parameter der GEV und GPD auf. Je nachdem wie  $Y_{k,j}$  bestimmt wird, sind nur noch sehr wenig Daten für die ML-Schätzung vorhanden. Damit ergeben sich dann für die Schätzung von  $\hat{\xi}$  und damit für  $\chi(k)$  teilweise unbrauchbare Werte.

Ein Problem sowohl bei der POT- als auch bei der Blockmaxima-Methode ist durch die Wahl von  $x$  bzw.  $y = x - u$  in den Formeln für den Tail (2.2), (2.6) und (3.2) gegeben. Für beide Werte gilt, dass sie gegen  $\infty$  streben. In dieser Arbeit wurde für beide Werte jeweils das Maximum von  $X_n$  gewählt. Ein weiteres Problem stellt die Erfüllung der Annahmen dar. Bei der Blockmaxima-Methode wird angenommen, dass die Blockmaxima unabhängig sind. Dies wurde überprüft, indem die Blockmaxima auf Korrelation und ARCH-Effekte mit dem  $\mathcal{LB}$ - und  $\mathcal{LM}$ -Test getestet wurden. Ebenso sollen die Exzesse der POT-Methode unabhängig sein. Auch dies wurde mit der Korrelation überprüft. Es ergab sich jedoch das Problem, dass selbst bei einer kleinen Schwelle von  $u = 1\%$  die Exzesse teilweise korreliert, also abhängig sind.

Eine Regel für die Blockgrößen- und Schwellenwahl fehlt und ist schwer aufzustellen. Es wurden daher stets verschiedene Tests und Plots angewandt, um eine geeignete Schwelle bzw. Blockgröße herauszufinden. Die Vergleichbarkeit der Methoden ist daher schwierig. Aufgrund der in dieser Arbeit vorgestellten Methoden ist es schwer möglich eine Modellauswahl basierend auf einer Schätzung von  $\chi(\cdot)$  zu treffen. Wenn man ein solches Vorgehen trotzdem wählen möchte, so sollte man die empirische Schätzung anwenden, denn obwohl bei den empirischen Schätzmethode die Abhängigkeitsstruktur der verwendeten Daten nicht berücksichtigt wird, sind die Ergebnisse hier oft sehr nahe an den Werten von  $\chi(\cdot)_{\text{Prozess}}$ .

Bei der Betrachtung der Allianz Daten wird deutlich, dass eine Stichprobengröße von  $n = 3774$  für die Extremwerttheorie schon sehr klein ist, da bei der Unabhängigkeit der Exzesse und der Blockmaxima eine geeignete Wahl der Schwelle bzw. der Blockgröße

schwierig ist. Ein weiteres Problem ist die Anpassung eines geeigneten Modells an die Daten. Die Ungenauigkeiten, die sich dabei ergeben können, fließen auch in die Berechnung von  $\chi(\cdot)_{Prozess}$  ein.

Zusammenfassend lässt sich festhalten, dass nur mit der Betrachtung der absoluten und relativen Fehler, der hier verwendeten Methoden, keine Entscheidung über ein passendes Modell gefällt werden kann. Daher sollte man weitere Kennzahlen bzw. Möglichkeiten der Schätzung in Betracht ziehen. Eine weitere Vorgehensweise wäre die Schätzung von  $\chi(\cdot)$  durch Schätzungen der Näherungen von oben ( $\bar{\chi}(\cdot)$ ) und von unten ( $\underline{\chi}(\cdot)$ ).

Ein nächster Schritt kann die Entwicklung einer Teststatistik sein, indem man die Hypothese, dass ein bestimmtes Modell geeignet ist, ablehnt, wenn der Fehler eine bestimmte Größe erreicht hat.

# Anhang A

## Abbildungen und Tabellen

### A.1 GARCH(1,1) Prozess

Lag $k$	$\chi^{(k)}_{GARCH}$	$\chi^{(k)}_{emp}$	$\chi^{(k)}_{emp\ Tails}$	$\chi^{(k)}_{Block}$	$\chi^{(k)}_{McNeil}$	$\chi^{(k)}_{POT}$
1	0.2097	0.1100	0.1100	0.2290	0.1466	0.1608
2	0.1948	0.0700	0.0700	0.2028	0.1500	0.1013
3	0.1980	0.1400	0.1400	0.2305	0.4093	0.0276
4	0.1886	0.1300	0.1301	0.3757	0.2482	0.1857
5	0.1989	0.1000	0.1001	0.1801	0.1375	0.0143
6	0.2013	0.1200	0.1201	0.1763	0.1538	0.0559
7	0.1963	0.0900	0.0901	0.2139	0.1830	0.2470
8	0.2007	0.0700	0.0701	0.1925	0.1165	0.0310
9	0.1840	0.0600	0.0601	0.1262	0.1598	0.0064
10	0.1732	0.1000	0.1001	0.1466	0.1668	0.0183
11	0.1820	0.0900	0.0901	0.2354	0.1436	0.0011
12	0.1668	0.0900	0.0901	0.1267	0.1332	0.0730
13	0.1602	0.0800	0.0801	0.1251	0.1296	0.0006
14	0.1730	0.1100	0.1102	0.1543	0.0963	0.0149
15	0.1741	0.1000	0.1002	0.1686	0.1473	0.0165
16	0.1650	0.0700	0.0701	0.2080	0.1622	0.0000
17	0.1612	0.1100	0.1102	0.1908	0.1638	0.7450
18	0.1603	0.1200	0.1202	0.2240	0.3906	0.3184
19	0.1707	0.0700	0.0701	0.1218	0.1453	0.2496
20	0.1685	0.1200	0.1202	0.2996	0.4512	0.3349

Tabelle A.1: Ergebnisse der Methoden für einen simulierten quadrierten GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

## Beurteilung für den GARCH(1,1) Prozess

Methode Lag $k$	Emp.	Emp. Tails	Block	McNeil	POT
1	0.0997	0.0996	-0.0194	0.0631	0.0489
2	0.1248	0.1248	-0.0080	0.0448	0.0936
3	0.0580	0.0580	-0.0325	-0.2113	0.1704
4	0.0586	0.0585	-0.1871	-0.0596	0.0029
5	0.0989	0.0988	0.0188	0.0614	0.1846
6	0.0813	0.0813	0.0250	0.0475	0.1455
7	0.1063	0.1062	-0.0176	0.0133	-0.0507
8	0.1307	0.1306	0.0082	0.0842	0.1697
9	0.1240	0.1239	0.0578	0.0242	0.1776
10	0.0732	0.0731	0.0266	0.0064	0.1549
11	0.0920	0.0919	-0.0534	0.0384	0.1809
12	0.0768	0.0767	0.0401	0.0336	0.0938
13	0.0802	0.0801	0.0352	0.0306	0.1596
14	0.0630	0.0629	0.0187	0.0767	0.1581
15	0.0741	0.0740	0.0056	0.0268	0.1576
16	0.0950	0.0949	-0.0429	0.0028	0.1650
17	0.0512	0.0510	-0.0296	-0.0026	-0.5838
18	0.0403	0.0401	-0.0638	-0.2303	-0.1581
19	0.1007	0.1005	0.0488	0.0254	-0.0790
20	0.0485	0.0483	-0.1310	-0.2827	-0.1664

Tabelle A.2: Absoluter Fehler für die Berechnung von  $\chi(k)_{GARCH}$  für einen simulierten quadrierten GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

Methode Lag $k$	Emp.	Emp. Tails	Block	McNeil	POT
1	0.4753	0.4753	-0.0924	0.3008	0.2333
2	0.6407	0.6406	-0.0409	0.2300	0.4802
3	0.2929	0.2927	-0.1642	-1.0672	0.8608
4	0.3106	0.3103	-0.9924	-0.3163	0.0152
5	0.4972	0.4969	0.0944	0.3087	0.9282
6	0.4039	0.4036	0.1243	0.2360	0.7225
7	0.5415	0.5411	-0.0899	0.0676	-0.2583
8	0.6512	0.6509	0.0410	0.4195	0.8457
9	0.6738	0.6735	0.3139	0.1313	0.9653
10	0.4226	0.4221	0.1537	0.0370	0.8946
11	0.5054	0.5048	-0.2937	0.2108	0.9941
12	0.4605	0.4599	0.2404	0.2016	0.5624
13	0.5007	0.5000	0.2195	0.1911	0.9961
14	0.3642	0.3633	0.1081	0.4434	0.9140
15	0.4257	0.4248	0.0319	0.1540	0.9050
16	0.5758	0.5751	-0.2602	0.0170	0.9999
17	0.3177	0.3165	-0.1835	-0.0160	-3.6212
18	0.2513	0.2500	-0.3977	-1.4369	-0.9864
19	0.5898	0.5890	0.2862	0.1486	-0.4626
20	0.2880	0.2865	-0.7776	-1.6773	-0.9873

Tabelle A.3: Relativer Fehler für die Berechnung von  $\chi(k)_{GARCH}$  für einen simulierten quadrierten GARCH(1,1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.1$ ,  $\beta = 0.89$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .



**A.2 AR(1) + ARCH(1) Prozess**

Lag $k$	$\chi^{(k)}_{ARARCH}$	$\chi^{(k)}_{emp}$	$\chi^{(k)}_{emp\ Tails}$	$\chi^{(k)}_{Block}$	$\chi^{(k)}_{McNeil}$	$\chi^{(k)}_{POT}$
1	0.4450	0.3700	0.3700	0.3032	0.1451	0.3893
2	0.2701	0.2000	0.2000	0.1689	0.0513	0.2719
3	0.1727	0.1400	0.1400	0.1262	0.0085	0.0886
4	0.1140	0.0900	0.0900	0.0716	0.0052	0.0363
5	0.0767	0.0400	0.0400	0.0314	0.0020	0.0029
6	0.0513	0.0400	0.0400	0.0069	0.0014	0.0047
7	0.0348	0.0100	0.0100	0.0043	0.0003	0.0000
8	0.0253	0.0000	0.0000	0.0053	0.0001	0.0000
9	0.0172	0.0100	0.0100	0.0006	0.0001	0.0000
10	0.0121	0.0000	0.0000	0.0047	0.0001	0.0000
11	0.0095	0.0100	0.0100	0.0007	0.0001	0.0000
12	0.0062	0.0100	0.0100	0.0008	0.0001	0.0000
13	0.0041	0.0300	0.0300	0.0038	0.0005	0.0002
14	0.0030	0.0200	0.0200	0.0039	0.0004	0.0001
15	0.0024	0.0000	0.0000	0.0185	0.0002	0.0000
16	0.0020	0.0100	0.0100	0.0059	0.0002	0.0002
17	0.0015	0.0200	0.0200	0.0054	0.0005	0.0003
18	0.0008	0.0100	0.0100	0.0052	0.0002	0.0003
19	0.0007	0.0000	0.0000	0.0022	0.0001	0.0000
20	0.0006	0.0100	0.0100	0.0004	0.0005	0.0000

Tabelle A.4: Ergebnisse der Methoden für einen simulierten quadrierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

**Beurteilung für den AR(1) + ARCH(1) Prozess**

Methode Lag $k$	Emp.	Emp. Tails	Block	McNeil	POT
1	0.0750	0.0750	0.1418	0.2999	0.0557
2	0.0701	0.0701	0.1012	0.2188	-0.0018
3	0.0327	0.0326	0.0464	0.1642	0.0841
4	0.0240	0.0240	0.0424	0.1088	0.0777
5	0.0367	0.0367	0.0454	0.0747	0.0738
6	0.0113	0.0113	0.0444	0.0498	0.0465
7	0.0248	0.0248	0.0305	0.0345	0.0347
8	0.0253	0.0253	0.0200	0.0252	0.0253
9	0.0072	0.0072	0.0166	0.0171	0.0172
10	0.0121	0.0121	0.0074	0.0120	0.0121
11	-0.0005	-0.0005	0.0087	0.0094	0.0095
12	-0.0038	-0.0038	0.0054	0.0061	0.0062
13	-0.0259	-0.0259	0.0003	0.0036	0.0039
14	-0.0170	-0.0170	-0.0009	0.0026	0.0029
15	0.0024	0.0024	-0.0161	0.0022	0.0024
16	-0.0080	-0.0080	-0.0039	0.0018	0.0018
17	-0.0185	-0.0185	-0.0039	0.0011	0.0012
18	-0.0092	-0.0093	-0.0045	0.0005	0.0005
19	0.0007	0.0007	-0.0014	0.0007	0.0007
20	-0.0094	-0.0095	0.0001	0.0001	0.0005

Tabelle A.5: Absoluter Fehler für die Berechnung von  $\chi(k)_{ARARCH}$  für einen simulierten quadrierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

Methode Lag $k$	Emp.	Emp. Tails	Block	McNeil	POT
1	0.1685	0.1684	0.3187	0.6738	0.1252
2	0.2595	0.2594	0.3747	0.8102	-0.0066
3	0.1892	0.1889	0.2689	0.9507	0.4871
4	0.2107	0.2104	0.3720	0.9544	0.6817
5	0.4788	0.4785	0.5910	0.9737	0.9620
6	0.2199	0.2194	0.8658	0.9719	0.9077
7	0.7125	0.7123	0.8756	0.9910	0.9988
8	1.0000	1.0000	0.7904	0.9952	0.9991
9	0.4197	0.4191	0.9648	0.9924	0.9999
10	1.0000	1.0000	0.6148	0.9918	1.0000
11	-0.0544	-0.0555	0.9224	0.9923	1.0000
12	-0.6007	-0.6027	0.8709	0.9778	0.9968
13	-6.2623	-6.2718	0.0708	0.8816	0.9472
14	-5.6280	-5.6373	-0.2995	0.8675	0.9751
15	1.0000	1.0000	-6.6823	0.9176	1.0000
16	-3.9490	-3.9569	-1.9429	0.9028	0.8884
17	-12.0287	-12.0509	-2.5401	0.7000	0.8123
18	-12.0231	-12.0465	-5.8191	0.6988	0.5916
19	1.0000	1.0000	-1.9380	0.8937	0.9967
20	-17.1171	-17.1534	0.2480	0.1593	0.9927

Tabelle A.6: Relativer Fehler für die Berechnung von  $\chi(k)_{ARARCH}$  für einen simulierten quadrierten AR(1) + ARCH(1) Prozess mit  $\alpha_0 = 1$ ,  $\alpha_1 = 0.8$ ,  $\lambda = 0.2$ , standardnormalverteilter Noise-Variable und  $n = 10000$ .

# Anhang B

## Punkt-Prozesse

Punkt-Prozesse sind ein mathematisches Modell zur Beschreibung der Zeitpunkte bestimmter Ereignisse. Die bekanntesten Arten von Punkt-Prozessen sind Poisson-Prozesse. Statt dem Begriff Punkt-Prozess wird gelegentlich auch Zählprozess verwendet. Denn ein Punkt-Prozess ist ein bei stetiger Zeit beobachteter Prozess, der zu jedem Zeitpunkt die Anzahl, der bis zu einschließlich diesem Zeitpunkt eingetretenen Ereignisse (Punkte), zählt. Ereignisse sind z.B. der Eingang eines Telefonanrufs in einer Telefonzentrale, die Ankunft eines Kunden an einem Schalter u.ä. Punkt-Prozess-Techniken sind hilfreich bei Extremwerten von Zeitreihen. Dieser Abschnitt basiert auf Applebaum (2004), Resnick (1987) und Sato (1999).

### Definition B.1 (Stoppzeit)

Eine ZV  $T : \Omega \rightarrow \mathbb{N}$  heißt Stoppzeit, falls  $\{T \leq n\} \in \mathcal{F}_n$ ,  $\forall n \in \mathbb{N}$  gilt.  $(\mathcal{F}_n)_{n \geq 0}$  bezeichnet dabei eine Filtrierung.

### Definition B.2 (Punkt-Prozess)

Ein Punkt-Prozess auf einem Wahrscheinlichkeitsraum ist eine Folge von Stoppzeiten  $T_n$ ,  $n \in \mathbb{N}$ , mit Werten in  $(0, \infty)$ , so dass

- (1)  $T_0 = 0$  f.s.,
- (2)  $T_n < T_{n+1}$  f.s., auf  $\{T_n < \infty\}$ ,
- (3)  $T_n = T_{n+1}$  f.s. auf  $\{T_n = \infty\}$ .

Ein Punkt-Prozess ist eindeutig beschrieben durch seinen Zählprozess  $N = (N_t, t \in \mathbb{R}_+)$ :

$$N_t = \begin{cases} n & \text{falls } t \in [T_n, T_{n+1}), \quad n \geq 0 \\ \infty & \text{falls } t \geq T_\infty \end{cases}$$

Mit  $T_\infty$  wird der Grenzwert der Folge  $T_n$  beschrieben.  $N_t$  zählt also die Punkte des Punkt-Prozesses im Intervall  $(0, t]$ .

**Bemerkung B.3**

$N$  hat folgende Eigenschaften:

- $N$  ist ein monoton wachsender Prozess mit Werten in  $\{0, 1, \dots, \infty\}$ .
- $N_0 = 0$ .
- $N$  ist rechtsstetig mit linksseitig existierenden Grenzwerten (càdlàg) und Sprüngen der Größe 1.

**Definition B.4 (Poisson-Prozess)**

Ein  $\mathbb{N}_0$ -wertiger stochastischer Prozess  $(N_t)_{t \in \mathbb{R}_+}$  heißt Poisson-Prozess zum Parameter  $\mu > 0$ , falls

- (1)  $N_0 = 0$ ,
- (2)  $N_t$  rechtsseitig stetige Pfade mit Sprüngen der Höhe 1 hat,
- (3) er unabhängige Zuwächse besitzt, d.h.  $N_{t_1} - N_{t_0}, \dots, N_{t_n} - N_{t_{n-1}}$  sind unabhängig für  $0 \leq t_1 < t_2 < \dots < t_n$ ,
- (4)  $N_t - N_s$  Poisson-verteilt zum Parameter  $\mu(t - s)$  für  $s \leq t$  ist. Der Parameter  $\mu$  wird auch als Rate oder Intensität des Poisson-Prozesses bezeichnet.

Ein Zählprozess heißt homogener Poisson-Prozess, falls die Zuwächse des Prozesses stationär, unabhängig und Poisson-verteilt sind. Ändert sich die Rate  $\mu$  mit der Zeit, d.h.  $\mu = \mu(t)$ , spricht man von einem inhomogenen Poisson-Prozess. In diesem Fall liegt auch keine Stationarität mehr vor, d.h. die statistischen Eigenschaften ändern sich mit der Wahl des Startzeitpunktes.

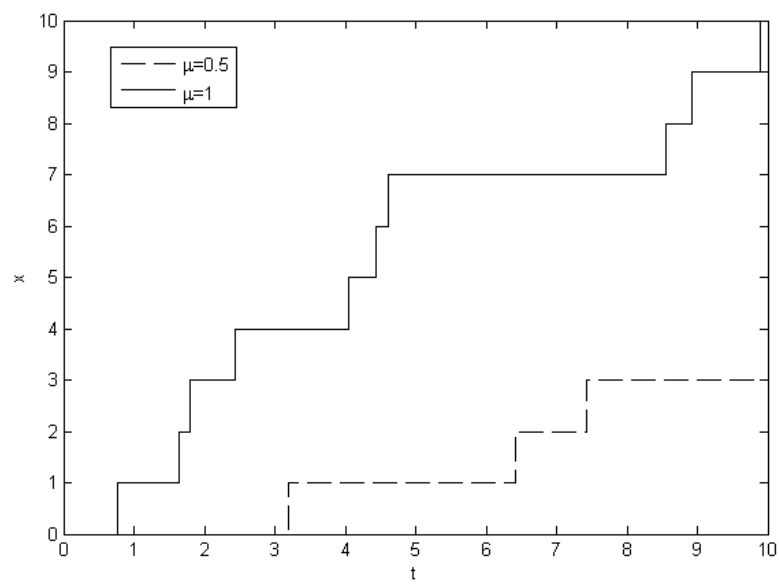


Abbildung B.1: Realisierungen eines Poisson-Prozesses mit verschiedenen Intensitäten  $\mu$ .

# Literaturverzeichnis

- Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge.
- Beirlant, J., Goegebeur, Y., Segers, J. und Teugels, J. L. (2004). *Statistics of Extremes*. Wiley, Chichester, 1. Aufl.
- Beirlant, J., Teugels, J. L. und Vynckier, P. (1996). *Practical Analysis of Extreme Values*. Leuven University Press, Leuven, 1. Aufl.
- Bingham, N., Goldie, C. und Teugels, J. (1987). *Regular Variation*, Bd. 27 von *Encyclopedia of Mathematics and its applications*. Cambridge University Press, Cambridge.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, Bd. 31 Nr. 3, S. 307–327.
- Bollerslev, T., Engle, R. F. und Nelson, D. B. (1994). ARCH Modells. In: R. F. Engle und D. L. McFadden (Hg.), *Handbook of Econometrics, Vol. IV*. Elsevier Science, S. 2959–3038.
- Borkovec, M. (2000). Extremal behaviour of the autoregressive process with ARCH(1) Errors. *Stochastic Processes and their applications*, Bd. 85, S. 189–207.
- Borkovec, M. und Klüppelberg, C. (2000). Extremwerttheorie für Finanzzeitreihen – ein unverzichtbares Werkzeug im Risikomanagement. In: B. Rudolph und L. Johanning (Hg.), *Handbuch Risikomanagement*. Uhlenbruch Verlag, Bad Soden, S. 219–241.
- Borkovec, M. und Klüppelberg, C. (2001). The Tail of the Stationary Distribution of an Autoregressive Process with ARCH(1) Errors. *The Annals of Applied Probability*, Bd. 11 Nr. 4, S. 1220–1241.
- Breiman, L. (1965). On some limit theorems similar to the arc-sin law. *Theory of Probability and its Applications*, Bd. 10 Nr. 2, S. 323–331.
- Brockwell, P. J., Chadraa, E. und Lindner, A. (2005). Continuous time GARCH processes. *The Annals of Applied Probability*, Bd. 16 Nr. 2, S. 790–826.
- Brockwell, P. J. und Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer Series in Statistics. Springer, New York, 2. Aufl.

- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, Bd. 1 Nr. 2, S. 223–236.
- de Haan, L., Resnick, S. I., Rootzén, H. und Vries, C. G. d. (1989). Extremal Behaviour of Solutions to a Stochastic Difference Equation with Applications to ARCH Processes. *Stochastic Processes and their Applications*, Bd. 32, S. 213–224.
- Embrechts, P., Klüppelberg, C. und Mikosch, T. (2003). *Modelling Extremal Events for Insurance and Finance*. Springer.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, Bd. 50 Nr. 4, S. 987–1008.
- Fasen, V. und Klüppelberg, C. (2006). High-level dependencies in time series models. *In Preparation*.
- Gençay, R., Selçuk, F. und Ulugülyağci, A. (2001). EVIM: A Software Package for Extreme Value Analysis in MATLAB (erhältlich unter: <http://www-m4.ma.tum.de/courses/index.de.html>).
- Goldie, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *The Annals of Applied Probability*, Bd. 1 Nr. 1, S. 126–166.
- Gomes, I., de Haan, L. und Pestana, D. (2004). Joint exceedances of the ARCH process. *Journal of Applied Probability*, Bd. 41, S. 919–926.
- Hill, B. M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *The Annals of Statistics*, Bd. 3 Nr. 5, S. 1163–1174.
- Jarque, C. M. und Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, Bd. 55 Nr. 5, S. 163–172.
- Kiefersbeck, N. (1999). Stationarität und Tailindex bei zeitdiskreten Volatilitätsmodellen. Diplomarbeit, Centre of Mathematical Sciences, Munich University of Technology, Garching bei München.
- Klüppelberg, C. (2003). Risk Management with Extreme Value Theory. In: B. Finkenstädt und H. Rootzén (Hg.), *Extreme Values in Finance, Telecommunications, and the Environment*. Chapman and Hall/CRC, S. 101–168.
- Ljung, G. M. und Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika*, Bd. 65 Nr. 2, S. 297–303.
- Mari, D. und Kotz, S. (2001). *Correlation and Dependence*. Imperial College Press, London.
- McNeil, A. J. (1998). Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory. <http://e-collection.ethbib.ethz.ch/show?type=bericht&nr=85>. ETH E-Collection.



- Mikosch, T. (2004). Modeling Dependence and Tails of Financial Time Series. In: B. Finkenstädt und H. Rootzén (Hg.), *Extreme Values in Finance, Telecommunications, and the Environment*. Chapman & Hall/CRC, S. 185–286.
- Mikosch, T. und Stărică, C. (2000). Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process. *The Annals of Statistics*, Bd. 28 Nr. 5, S. 1427–1451.
- Regler, M. (2004). Neue Schätzmethode des extremalen Index. Diplomarbeit, Centre of Mathematical Sciences, Munich University of Technology, Garching bei München.
- Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
- Smith, R. L. (1984). Threshold Methods for Sample Extremes. In: J. Tiago de Oliveira (Hg.), *Statistical Extremes and Applications*. D. Reidel Publishing Company, Dordrecht, Boston, Lancaster, S. 621–638.
- Vervaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Advances in Applied Probability*, Bd. 11 Nr. 4, S. 750–783.