

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN  
INSTITUT FÜR STATISTIK



Masterarbeit

# Anwendung der Extremwerttheorie zur Schätzung des Value at Risks und des Expected Shortfalls

**Autor** Diwei Lu

**Betreuer** Prof. Dr. Stefan Mittnik, Christoph Berninger

**Datum** 28.06.2018

## Zusammenfassung

Die vorliegende Masterarbeit beschäftigt sich mit der Schätzung des Value at Risks und des Expected Shortfalls und insbesondere mit Schätzungsverfahren mit der Extremwerttheorie: die *Block-Maxima*-Methode (mit und ohne extremen Index) und die *Peaks-over-Threshold*-Methode (mit und ohne GARCH-Filter). Bei einer empirischen Analyse des deutschen Aktienindex in den letzten 27 Jahren werden sie anhand von Backtests mit anderen traditionellen Methoden verglichen. Den Testergebnissen zufolge ragt die Leistung der *Peaks-over-Threshold*-Methode mit GARCH-Filter heraus. Danach folgen die GARCH-Modelle mit Studentischer t-verteilter und Skewed t-verteilter Innovation, die bei einem Signifikanzniveau von  $\alpha = 0.01$  und  $0.005$  perfekt funktionieren. Abgesehen von der Abhängigkeit der VaR-Prognosen bietet die durch den extremen Index modifizierte *Block-Maxima*-Methode auch gute Ergebnisse, obwohl diese Methode in der Finanzwelt nicht gefragt ist. Am Ende dieser Arbeit wird die zeitliche Aggregation des Value at Risks untersucht. Es kommt zur Schlussfolgerung, dass sich die primitive Square-Root-of-Time-Skalierung trotz ihrer Einfachheit als noch effektiv auf dem deutschen Markt erweist.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>7</b>
<b>2</b>	<b>Theoretischer Hintergrund zum Riskmanagement</b>	<b>10</b>
2.1	Value at Risk . . . . .	10
2.2	Expected Shortfall . . . . .	10
2.3	Schätzung des Value at Risks und des Expected Shortfalls . . . . .	11
2.3.1	Historische Simulation . . . . .	12
2.3.2	Normalverteilte Verlustverteilung . . . . .	12
2.3.3	Studentsche t-verteilte Verlustverteilung . . . . .	14
2.3.4	Skewed t-verteilte Verlustverteilung . . . . .	15
2.3.5	GARCH-Modell . . . . .	16
<b>3</b>	<b>Theoretischer Hintergrund zur Extremwerttheorie</b>	<b>18</b>
3.1	Block-Maxima-Methode . . . . .	18
3.1.1	Generalisierte Extremwertverteilung . . . . .	18
3.1.2	Schätzung der generalisierten Extremwertverteilung . . . . .	21
3.1.3	Return Level und Risikomaß . . . . .	22
3.1.4	Extremer Index . . . . .	23
3.1.5	Wahl der Größe des Blocks . . . . .	24
3.2	Peaks-over-Threshold-Methode . . . . .	26
3.2.1	Generalisierte Paretoverteilung . . . . .	26
3.2.2	Schätzung der generalisierten Pareto-Verteilung . . . . .	28
3.2.3	VaR und ES . . . . .	29
3.2.4	Wahl des Schwellenwerts . . . . .	30
3.2.5	GARCH-Filter . . . . .	34
<b>4</b>	<b>Backtesting</b>	<b>37</b>
4.1	Backtesting des Value at Risks . . . . .	37
4.1.1	Unconditional Coverage Test . . . . .	38
4.1.2	Independence Test . . . . .	39
4.1.3	Conditional Coverage Test . . . . .	40
4.2	Backtesting des Expected Shortfalls . . . . .	41
4.2.1	Acerbi und Szekelys Test 1: Testen ES nach VaR . . . . .	41
4.2.2	Acerbi und Szekelys Test 2: Testen ES unmittelbar . . . . .	43
4.2.3	Acerbi und Szekelys Test 3: Schätzen ES erneut . . . . .	43
4.2.4	P-Werts der Acerbi und Szekelys Tests . . . . .	44
4.2.5	Regression Based Expected Shortfall Backtesting . . . . .	45
4.2.6	Weitere Tests des ES . . . . .	47

<b>5</b>	<b>Anwendungsteil</b>	<b>49</b>
5.1	Block-Maxima-Methode . . . . .	51
5.2	Peaks-over-Threshold-Methode . . . . .	56
5.3	Weitere Modelle . . . . .	62
<b>6</b>	<b>Zeitliche Aggregation</b>	<b>64</b>
6.1	Square-Root-of-Time Regel . . . . .	65
6.2	Auf dem Tail-Index basierende Skalierung . . . . .	65
6.3	Tests der Anwendbarkeit der Skalierung . . . . .	66
6.4	Anwendung auf den DAX . . . . .	68
<b>7</b>	<b>Fazit</b>	<b>70</b>
<b>A</b>	<b>Anhang</b>	<b>72</b>
<b>B</b>	<b>Verwendete Software</b>	<b>79</b>
<b>C</b>	<b>Inhalt der CD</b>	<b>80</b>
<b>D</b>	<b>Urheberschaftserklärung</b>	<b>81</b>
	<b>Literaturverzeichnis</b>	<b>82</b>

# Abbildungsverzeichnis

1	Vergleich der StV und NV . . . . .	14
2	Skewed t-Verteilung . . . . .	16
3	Extremwertverteilung . . . . .	20
4	Generalisierte Pareto-Verteilung . . . . .	28
5	Sample Mean Excess Function . . . . .	32
6	Hill-Plot . . . . .	33
7	DAX . . . . .	49
8	Logarithmierter Verlust . . . . .	50
9	Norm-QQ-Plot . . . . .	51
10	Monatliche Maxima . . . . .	52
11	VaR-Prognosen der BM-Methode . . . . .	53
12	ES-Prognosen der BM-Methode . . . . .	53
13	VaR-Prognosen der BM-Methode mit dem extremen Index . . . . .	55
14	ES-Prognosen der BM-Methode mit dem extremen Index . . . . .	55
15	VaR-Prognosen der POT-Methode . . . . .	57
16	ES-Prognosen der POT-Methode . . . . .	57
17	Betrag der Log-Verluste . . . . .	59
18	Standardabweichung der Log-Verluste . . . . .	59
19	Standardisierte Residuen des GARCH(1,1)-Modells . . . . .	60
20	Autokorrelationsfunktion . . . . .	61
21	VaR-Prognosen der POT-Methode mit dem GARCH-Filter . . . . .	61
22	ES-Prognosen der POT-Methode mit dem GARCH-Filter . . . . .	62
23	VaR-Prognosen der historischen Simulation . . . . .	72
24	ES-Prognosen der historischen Simulation . . . . .	72
25	VaR-Prognosen der normalen Verteilung . . . . .	73
26	ES-Prognosen der normalen Verteilung . . . . .	73
27	VaR-Prognosen der Studentschen t-Verteilung . . . . .	74
28	ES-Prognosen der Studentschen t-Verteilung . . . . .	74
29	VaR-Prognosen der Skewed t-Verteilung . . . . .	75
30	ES-Prognosen der Skewed t-Verteilung . . . . .	75
31	VaR-Prognosen des GARCH-Modells mit der normalen Verteilung . . . . .	76
32	ES-Prognosen des GARCH-Modells mit der normalen Verteilung . . . . .	76
33	VaR-Prognosen des GARCH-Modells mit der Studentschen t-Verteilung . . . . .	77
34	ES-Prognosen des GARCH-Modells mit der Studentschen t-Verteilung . . . . .	77
35	VaR-Prognosen des GARCH-Modells mit der Skewed t-Verteilung . . . . .	78
36	ES-Prognosen des GARCH-Modells mit der Skewed t-Verteilung . . . . .	78

# Tabellenverzeichnis

1	Deskriptive Statistiken . . . . .	50
2	Shermanns Anpassungsgüte-Test . . . . .	52
3	Schätzwerte der GEV . . . . .	52
4	VaR-Backtesting der BM-Methode . . . . .	53
5	ES-Backtesting der BM-Methode . . . . .	54
6	Schätzung des extremen Index . . . . .	54
7	VaR-Backtesting der BM-Methode mit dem extremen Index . . . . .	55
8	ES-Backtesting der BM-Methode mit dem extremen Index . . . . .	56
9	Schätzwerte der GPD . . . . .	56
10	VaR-Backtesting der POT-Methode . . . . .	57
11	ES-Backtesting der POT-Methode . . . . .	58
12	Ergebnisse der Schätzung des AR(1)-GARCH(1,1)-Modells . . . . .	58
13	Ergebnisse der Schätzung des GARCH(1,1)-Modells . . . . .	58
14	VaR-Backtesting der POT-Methode mit dem GARCH-Filter . . . . .	62
15	ES-Backtesting der POT-Methode mit dem GARCH-Filter . . . . .	62
16	VaR-Backtesting der weiteren Methoden . . . . .	63
17	ES-Backtesting der weiteren Methoden . . . . .	64
18	Anwendung der SRTR . . . . .	68
19	Anwendung der modifizierten SRTR . . . . .	69
20	Anwendung der auf dem Tail-Index basierenden Skalierung . . . . .	69

# 1 Einleitung

Die Finanzwelt ist immer mit Risiko verknüpft, deswegen ist Riskmanagement immer das zentrale Thema der Finanzwelt. Aus der Sicht des Anlegers ist vor allem der zukünftige Verlust von Interesse. Die Bankaufsicht interessiert sich auch für das *Downside Risk*, um die Liquidität der Banken zu kontrollieren. Dazu wird das Risikomaß entwickelt, darunter fällt beispielsweise der Begriff *Value at Risk* (kurz: VaR). Unter dem VaR versteht man ein oberes Quantil der Verluste oder ein unteres Quantil der Renditen. Er bietet die Möglichkeit, Risiko mit nur einem Wert darzustellen. Außerdem kann der VaR in allen Risikobereichen eingesetzt werden [Ruppert, 2011, S. 553].

Während der VaR als gängiges Instrument zur Risikomessung am häufigsten angewendet wird, stößt er aber oft auf Kritik wegen seiner theoretischen Defizienz: Einerseits berücksichtigt der VaR das sogenannte *tail risk* hinter dem Quantil nicht [Yamai et al., 2002, S. 58]. Andererseits fehlt es dem VaR an einer wichtigen mathematischen Eigenschaft: Subadditivität, die von Artzner et al. [1999] zum ersten Mal mit dem Begriff „kohärentes Risikomaß“ zusammengefasst wurde. Dazu wurde der *Expected Shortfall* (kurz: ES) von Artzner et al. eingeführt, der bessere mathematische Eigenschaften aufweist. Der ES ist definiert als die Erwartung bezüglich der Verluste, die den VaR überschreiten. In diesem Sinne nannten manche Experten den VaR selbst nicht als ein Risikomaß (z.B. [Acerbi and Tasche, 2002, S. 2]). Deshalb plante der Basler Ausschuss, den VaR durch den ES zu ersetzen [Committee et al., 2013, S. 3]. Aber das Backtestingverfahren wurde noch auf Basis von VaR durchgeführt [Committee et al., 2013, S. 27], denn der ES ist schwer nachzuprüfen. Gneiting [2011] fand, dass der ES nicht *elicitable* ist, wodurch das Backtesting des ES erschwert wird. Jedoch zeigten nachfolgende Studien, dass der ES auch backtestbar ist (z.B. [Acerbi and Szekely, 2014, S. 9] haben plädiert, dass die Elizitabilität gar nichts mit dem Backtesting zu tun hat). Dazu wurden einige Tests aufgestellt. (Angelidis and Degiannakis [2006], Acerbi and Szekely [2014], Wong [2008], Emmer et al. [2015], Righi and Ceretta [2013], Moldenhauer and Pitera [2017], Acerbi and Szekely [2017], Nolde et al. [2017], Bayer and Dimitriadis [2018] usw.)

Wenn das Risiko nicht ausreichend geschätzt wird, droht ein hoher Preis. Das sieht man beim Ausbruch der Finanzkrise. Im Jahr 2008 zerstörte eine weltweite Finanzkrise zum Beispiel fast den ganzen Bankbereich. Danach wird die Bankenregulierung verschärft, wobei die Schätzung des Risikomaßes vor Herausforderungen steht. Die Finanzkrise ist ein extremer Fall, mit dem man normalerweise nicht rechnet, aber wenn der extreme Fall auftritt, hat er eine verheerende Wirkung. Jedoch sind solche extremen Verluste schwer mithilfe von konventionellen Wahrscheinlichkeitstheorien zu untersuchen, die sich hauptsächlich auf den Mittelwert richten, weil extreme Werte weit vom Mittelwert liegen. Daher kommt die Extremwerttheorie (*extreme value theory*, kurz: EVT) zur Anwendung, die im Jahr 1958 vom deutschen Mathematiker Emil Julius Gumbel zum ersten Mal systematisch dargestellt wurde. Die Extremwerttheorie wurde früher besonders in der Hydrologie und im Bauingenieurwesen

angewendet, um extremes Wetter oder Hochwasser, z.B. Jahrhundertflut, zu prognostizieren. Aber seit Jahren findet sie auch Anwendung in der Versicherungsindustrie [Coles et al., 2001, S. 1]. Dafür sind folgende Fragen von Interesse:

- Wie ist die EVT bei der Schätzung von VaR und ES einzusetzen? Gibt es eine Besonderheit, wodurch sich dieses Schätzungsverfahren von anderen konventionellen Verfahren unterscheidet?
- Kann die EVT die adäquate Schätzung von VaR und ES liefern? Wie ist die Güte der Ergebnisse einzuschätzen? Anders gefragt: wie ist der Evaluierungsprozess zu implementieren?
- Was sind die Vorteile der EVT gegenüber den anderen Theorien bei der Schätzung des Risikomaß? Oder gibt es die Möglichkeit, die EVT mit anderen Methoden zu kombinieren, um bessere Ergebnisse zu erzielen?

Im Beitrag von Longin [1996] wurde anhand einer empirischen Studie bezüglich des *New York Stock Exchange* gezeigt, dass die asymptotische Verteilung der extremen Renditen die Fréchet-Verteilung befolgt, die zur generalisierten extremen Verteilung (kurz: GEV) gehört, wobei der Tail-Index das Tail-Verhalten bestimmt. McNeil [1998] wendet die *Block-Maxima*-Methode (kurz: BM-Methode) auf die BMW-Renditen an, um das *return level* zu berechnen. Dabei berücksichtigte er auch den Cluster-Effekt der Finanzdaten, indem der extreme Index eingeführt wurde. Longin [2000] schlug einen auf der BM-Methode basierenden parametrischen Ansatz vor, um den VaR zu kalkulieren. Dabei fand er, dieser Ansatz ist konsistent mit vielen anderen statistischen Modellen für den Prozess der Renditen. Gencay and Selcuk [2004] untersuchten neun Aktienmärkte in unterschiedlichen Ländern. Dieser Studie zufolge fittet die generalisierte Pareto-Verteilung (kurz: GPD) den Schwanz der Renditeverteilung gut und die daraus entwickelte *Peaks-Over-Threshold*-Methode (kurz: POT-Methode) schneidet bei hohen Quantilen bei der VaR-Schätzung deutlich besser als die anderen Methoden ab. Allerdings setzt die Standard-POT-Methode die Unabhängigkeit der Daten voraus, die in der Praxis unrealistisch ist. Eine Gegenmaßnahme wurde von Borkovec and Klüppelberg [2000] aufgestellt, die die Standard-POT-Methode durch den extremen Index modifiziert. Im Vergleich dazu versuchten McNeil and Frey [2000] eine weitere Alternative: Sie kombinierten die POT-Methode mit dem GARCH-Prozess (*Generalized Autoregressive Conditional Heteroskedasticity*), um den GARCH-Effekt zu bereinigen. Dabei analysierten sie die durch GARCH-Filter gefilterten Residuen, statt die Verluste direkt in die GPD anzupassen. Mithilfe einer Studie auf den *Istanbul Stock Index-100* erklärten Ozun et al. [2010], dass sich die POT-Methode zur Schätzung der VaR und ES eignet, insbesondere beim Schwellenmarkt. Das Fazit wurde durch umfangreiche Backtests untermauert.

In der Literatur ist die POT-Methode beliebter als die BM-Methode zur Schätzung des VaR und des ES. Denn Marinelli et al. [2007] zufolge ist die POT-Methode signifikant besser

als die BM-Methode. Jedoch ist die EVT nicht ganz tadellos in der Finanzwelt: In Diebold [1998] wurde davor gewarnt, dass die EVT manchmal unter beschränkten Samples leidet, deren Anwendung auf den geeigneten Bereich eingegrenzt werden sollte, laut Diebold liegt der Anwendungsbereich vielleicht nicht bei 90%, 95% oder 99%, sondern höher als das 99%-Quantil.

Die vorliegende Arbeit beschäftigt sich im Wesentlichen mit der Anwendung der Extremwerttheorie im Finanzbereich, indem sie wie folgt aufgebaut wird: In Kapitel 2 werden der theoretische Hintergrund zum *Riskmanagement* und die konventionellen Modelle zur Schätzung von VaR und ES erläutert. Danach folgen in Kapitel 3 die Theoriegrundlagen der EVT und die entsprechende Vorgehensweise, wodurch VaR und ES berechnet werden. Anschließend werden die statistischen Tests in Kapitel 4 analysiert, die die Schätzungsergebnisse von VaR und ES testen und evaluieren. Kapitel 5 ist der empirische Anwendungsteil, in dem die oben erwähnten Methoden auf den DAX angewendet und anhand von angemessenen Tests ausgewertet werden. In Kapitel 6 wird der zeitlichen Aggregation nachgegangen. Zum Schluss beschließt das Fazit die Arbeit.

## 2 Theoretischer Hintergrund zum Riskmanagement

### 2.1 Value at Risk

Die Idee des Value at Risks geht auf den Anfang der 1990er Jahre zurück. Seither trat eine Serie gravierender Wirtschaftsverluste auf, zum Beispiel: der Bankrott der *Orange County*, der Zusammenbruch der Barings Bank sowie der Niedergang der Metallgesellschaft usw. Aus solchen Katastrophen haben sowohl die Regierungsaufsicht als auch das Finanzinstitut die Lektion gelernt, dass lockere Kontrolle und schlechte Verwaltung einen Verlust in Höhe von Billionen Dollars auslösen könnten. Deswegen stieg die Nachfrage nach einem leicht verständlichen Risikomaß [Jorion, 2000, S.22]. Dabei ist der Begriff VaR entstanden und seitdem ist er das bekannteste Risikomaß. Die Einfachheit des VaR liegt darin, dass der VaR auf alle Typen von Risiko anwendbar ist, z.B. Marktrisiken, Kreditrisiken, Betriebsrisiken usw [Ruppert, 2011, S. 553]. Zudem vereinfacht der VaR auch umfangreiche Risiken nur zu einem Wert, dessen Einheit identisch mit der Währung ist. Deswegen wird der VaR als die beste Lösung empfohlen, wenn ein Risikomanager seinem Chef das Risiko berichten will [Linsmeier and Pearson, 2000, S. 47].

**Definition 1** *Der VaR ist ein Quantil der Verlustfunktion gegeben Zeithorizont  $T$  und Signifikanzniveau  $\alpha$  (oder Konfidenzniveau  $1 - \alpha$ ), sodass innerhalb des gegebenen Zeithorizonts die Wahrscheinlichkeit, dass das Quantil überschritten wird, gleich  $\alpha$  ist.*

Zum Beispiel: wenn  $VaR = 100$  Euro mit  $T = 1$  und  $\alpha = 0.05$ , bleibt der Verlust am nächsten Tag mit 95% Wahrscheinlichkeit geringer als 100 Euro. Anders gesagt, die Wahrscheinlichkeit der Überschreitung beträgt gerade 5%. Im Folgenden steht die Notation  $VaR_\alpha^T$  für den VaR mit den zwei Parametern  $T$  und  $\alpha$ . Ansonsten wird die Notation  $VaR_\alpha$  benutzt, falls der Zeithorizont  $T = 1$  ist.

Laut der Definition des VaR ergibt sich:

$$F_L(VaR_\alpha) = P\{L \leq VaR_\alpha\} = 1 - \alpha, \quad (1)$$

dabei bezeichnet  $F_L$  die Verteilungsfunktion des Verlusts. Durch (1) bekommt man:

$$VaR_\alpha = \inf\{x \in \mathbb{R} : Pr(L > x) \leq \alpha\} = \inf\{x \in \mathbb{R} : F_L(x) \geq 1 - \alpha\}. \quad (2)$$

### 2.2 Expected Shortfall

Der VaR ist nicht nur in der Finanzwelt als das beliebteste Werkzeug zur Risikomessung anerkannt, sondern spielt auch bei nichtfinanziellen Organisationen eine Rolle. Seit Jahren wird jedoch Kritik am VaR geäußert.

Einerseits verweist der VaR ausschließlich auf ein Quantil, wodurch das Portfolio nur mit  $(1 - \alpha)$  Wahrscheinlichkeit garantiert wird. Daher sind keine Informationen über den Schwellenwert anzugeben. Das sogenannte *tail risk* wird dadurch nicht in Kauf genommen [Yamai et al., 2002, S. 58]. Zum Beispiel: Dem VaR gemäß tendiert ein Anleger dazu, ein Asset mit kleinerem VaR zu wählen. Hinter diesem Asset steckt vielleicht ein riesiges *tail risk*, das einen Verlust in Höhe von Millionen Euro verursachen könnte.

Andererseits wird VaR wegen des theoretischen Mangels stets kritisiert, seitdem Artzner et al. im Jahr 1999 den Begriff *coherent risk measure* eingeführt haben. In diesem Beitrag haben sie die folgenden axiomatischen Eigenschaften eines kohärenten Risikomaß aufgeführt: Monotonie, Subadditivität und positive Homogenität sowie Translationsinvarianz [Artzner et al., 1999, S. 210]. Die Subadditivität ist eine der erwünschten Eigenschaften in der Finanzwelt, die besagt: Für alle  $L_1, L_2, \dots, L_n \in \mathcal{L}$  ergibt sich:

$$\rho(L_1 + L_2 + \dots + L_n) \leq \rho(L_1) + \rho(L_2) + \dots + \rho(L_n)$$

$\mathcal{L}$  ist die Menge aller möglichen Verlustfunktionen und  $\rho$  ist ein subadditives Risikomaß. Es spiegelt die Diversifikation wider, dass das kombinierte Portfolio auf keinen Fall mehr Risiko übernimmt. Wie der wohlbekannte Aphorismus besagt: *Don't put all your eggs in one basket*.

Artzner et al. [1999] haben bewiesen, dass der VaR die Subadditivität normalerweise<sup>1</sup> nicht erfüllt. Infolge des Mangels an Subadditivität kann der VaR ein Finanzinstitut eine suboptimale Entscheidung treffen lassen [Danielsson et al., 2005, S. 2]. Aufgrund des Mangels an Subadditivität erschwert VaR außerdem die Dezentralisierung und Zentralisierung des Risikos [Artzner et al., 1999, S. 209]. Dazu wurde der *expected shortfall* als kohärentes Risikomaß aufgestellt, der die oben genannten vier Eigenschaften erfüllt.

**Definition 2** *Der ES ist der Erwartungswert der Verluste, die den VaR überschreiten:*

$$ES_\alpha = E[L|L \geq VaR_\alpha]. \quad (3)$$

Der ES mit dem Signifikanzniveau  $\alpha$  kann durch Integral analytisch kalkuliert werden [Ruppert, 2011, S. 554]:

$$ES_\alpha = \alpha^{-1} \int_0^\alpha VaR_u du. \quad (4)$$

## 2.3 Schätzung des Value at Risks und des Expected Shortfalls

In der Praxis werden VaR und ES mithilfe von vielfältigen Methoden geschätzt, die eigene Vor- und Nachteile haben. Die Auswahl der Methode ist von der Eigenschaft der Daten und Märkte sowie der Anforderung an Präzision abhängig. In diesem Abschnitt werden

---

<sup>1</sup>[Artzner et al., 1999, S. 217] haben nachgewiesen, dass VaR in spezifischen Fällen auch subadditiv ist. Diese Ausnahme wird in 2.3.2 illustriert.

konventionelle Methoden zur Schätzung von VaR und ES vorgestellt, die hauptsächlich in die zwei Kategorien „Nonparametrische Schätzung“ und „Parametrische Schätzung“ fallen.

### 2.3.1 Historische Simulation

Die Historische Simulation (kurz: HS) gehört zu der nonparametrischen Schätzung, die als die einfachste Methode gilt, denn diese Methode trifft keine statistische Verteilungsannahme für den Verlust und erfordert keine Schätzung des Parameters. Sie basiert ausschließlich auf den vergangenen Beobachtungen. Es wird angenommen, dass sich der Markt in der Zukunft wie früher entwickeln wird. Dann wird das  $(1 - \alpha)$ -Quantil von Samples als  $VaR_\alpha$  gewählt.

Gegeben  $w$  Verluste  $l_t, l_{t-1} \dots l_{t-w+1}$  ist die VaR-Prognose zum Zeitpunkt  $t$  mit dem Signifikanzniveau  $\alpha$  zu bestimmen durch:

$$\widehat{VaR}_\alpha = \hat{Q}_{1-\alpha}(\underbrace{l_t, l_{t-1} \dots l_{t-w+1}}_w), \quad (5)$$

wobei  $\hat{Q}_{1-\alpha}(l_t, l_{t-1} \dots l_{t-w+1})$  das  $(1 - \alpha)$ -Quantil der letzten  $w$  Verluste ist. Das bedeutet, bei der HS Methode beispielsweise mit  $\alpha = 0.05$  und  $w = 1000$  entspricht der VaR der 50-ten absteigenden Ordnungsstatistik der vergangenen Verluste.

Auf Basis des VaR kann ES im Rahmen der HS durch den folgenden nonparametrischen Schätzer bestimmt werden:

$$\widehat{ES}_\alpha = \frac{\sum_{i=t-w+1}^t l_i I_i}{\sum_{i=t-w+1}^t I_i}, \quad (6)$$

wobei  $I_i$  die Indikatorfunktion der Überschreitung ist:

$$I_i = \begin{cases} 1, & \text{falls } l_{t+1} > \widehat{VaR}_\alpha, \\ 0, & \text{falls } l_{t+1} \leq \widehat{VaR}_\alpha. \end{cases} \quad (7)$$

Auf der einen Seite bietet die nonparametrische Methode die Möglichkeit, das Risikomaß ohne Verteilungsnahme zu schätzen, damit eine Missspezifikation des Modells vermieden werden kann. Auf der anderen Seite versagt HS manchmal bei extremen Fällen. Zum Beispiel: Hinter dem 99%-Quantil gibt es nur 2 Verluste innerhalb eines Jahrs. Mit so wenigen Beobachtungen ist kaum eine zuverlässige statistische Inferenz zu ziehen. Außerdem führt es zu einem großen Sprung bei der VaR-Schätzung, wenn ein Ausreißer bei der Betrachtung ein- und ausgeschlossen wird [Kuester et al., 2006, S. 57]. In diesem Fall müssen die parametrischen Methoden angewendet werden.

### 2.3.2 Normalverteilte Verlustverteilung

Die Normalverteilung (kurz: NV) spielt eine wichtige Rolle in Naturwissenschaft und Technik. Selbstverständlich versucht man auch von Anfang an, die Finanzdaten durch die Nor-

malverteilung zu beschreiben. Dabei wird angenommen, dass Verlust oder Rendite mit dem Erwartungswert  $\mu$  und der Varianz  $\sigma^2$  normalverteilt ist. Danach ist der VaR durch die folgende Formel zu schätzen:

$$\widehat{VaR}_\alpha = \hat{\mu}_L + \hat{\sigma}\Phi^{-1}(1 - \alpha). \quad (8)$$

Dabei sind  $\hat{\mu}_L$  und  $\hat{\sigma}$  der Maximum-Likelihood-Schätzer von Erwartungswert und Standardabweichung der Verluste.  $\Phi^{-1}(\cdot)$  ist die inverse Funktion der Verteilungsfunktion der Standardnormalverteilung.

Laut [Ruppert, 2011, S. 557] ist dann ES durch den folgenden Schätzer zu berechnen:

$$\widehat{ES}_\alpha = \hat{\mu}_L + \hat{\sigma} \frac{\phi(\Phi^{-1}(1 - \alpha))}{\alpha}, \quad (9)$$

wobei  $\phi(\cdot)$  die Dichtefunktion der Standardnormalverteilung ist.

Aufgrund der erwünschten statistischen Eigenschaften macht die Normalverteilungsannahme das Schätzungsverfahren einfach. Außerdem ist der normalverteilte VaR unter bestimmten Voraussetzungen subadditiv: Beispielsweise gibt es zwei Portfolios,  $X$  und  $Y$ . Der Formel 8 zufolge ist der jeweilige VaR unter NV-Annahme gegeben durch:

$$VaR_\alpha(X) = \mu_X + \sigma_X\Phi^{-1}(1 - \alpha), VaR_\alpha(Y) = \mu_Y + \sigma_Y\Phi^{-1}(1 - \alpha). \quad (10)$$

Offensichtlich gilt  $\sigma_{X+Y} \leq \sigma_X + \sigma_Y$  für zwei Zufallsvariablen  $X$  und  $Y$ , weil:

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY} \leq \sigma_X^2 + \sigma_Y^2 + 2\sigma_X\sigma_Y = (\sigma_X + \sigma_Y)^2, \quad (11)$$

wobei  $\sigma_{XY} \leq \sigma_X\sigma_Y$ , das Gleichheitszeichen gilt, falls ein vollständig positiver linearer Zusammenhang zwischen  $X$  und  $Y$  besteht. Dann ergibt sich:

$$\begin{aligned} VaR_\alpha(X) + VaR_\alpha(Y) &= \mu_X + \mu_Y + (\sigma_X + \sigma_Y)\Phi^{-1}(1 - \alpha) \geq \\ &(\mu_X + \mu_Y) + \sigma_{X+Y}\Phi^{-1}(1 - \alpha) = VaR_\alpha(X + Y), \end{aligned} \quad (12)$$

unter der Voraussetzung:  $\Phi^{-1}(1 - \alpha) \geq 0$  bzw.  $\alpha \leq 0.5$  [Artzner et al., 1999, S. 217].

Aber die einfache Normalverteilungsannahme ist in der Praxis immer problematisch. Mandelbrot [1963] hat herausgestellt, dass Preis der Finanzmärkte (z.B. der Preis von Watte) nicht durch eine Normalverteilung beschrieben werden kann. Die Verteilung der Daten ist vielmehr leptokurtisch. Das bedeutet, im Vergleich zur NV konzentrieren sich die Finanzdaten mehr auf die Flanken (*fat tails*) und weniger auf das Zentrum (*sharp peak*). In diesem Sinne wird die Geeignetheit der NV-Annahme in Frage gestellt.

Angesichts der oben erwähnten Leptokurtosis ist die Studentsche t-Verteilung zur Beschreibung der Verluste denkbar, denn durch eine Einstellung des Freiheitsgrads ist die Studentsche t-Verteilung in der Lage, unterschiedliche Tail-Verhalten zu präsentieren. In

Abbildung 1 ist der Unterschied zwischen Studentischer t-Verteilung und Normalverteilung zu sehen: Insbesondere bei der rechten auf die extremen Verluste fokussierten Abbildung überholen die Beobachtungen die gefittete Normalverteilungsdichte oft. Es weist eine Unterschätzung des Risikos durch die Normalverteilungsannahme auf. Im Vergleich dazu fittet die Studentische t-Verteilung besser.

### 2.3.3 Studentische t-verteilte Verlustverteilung

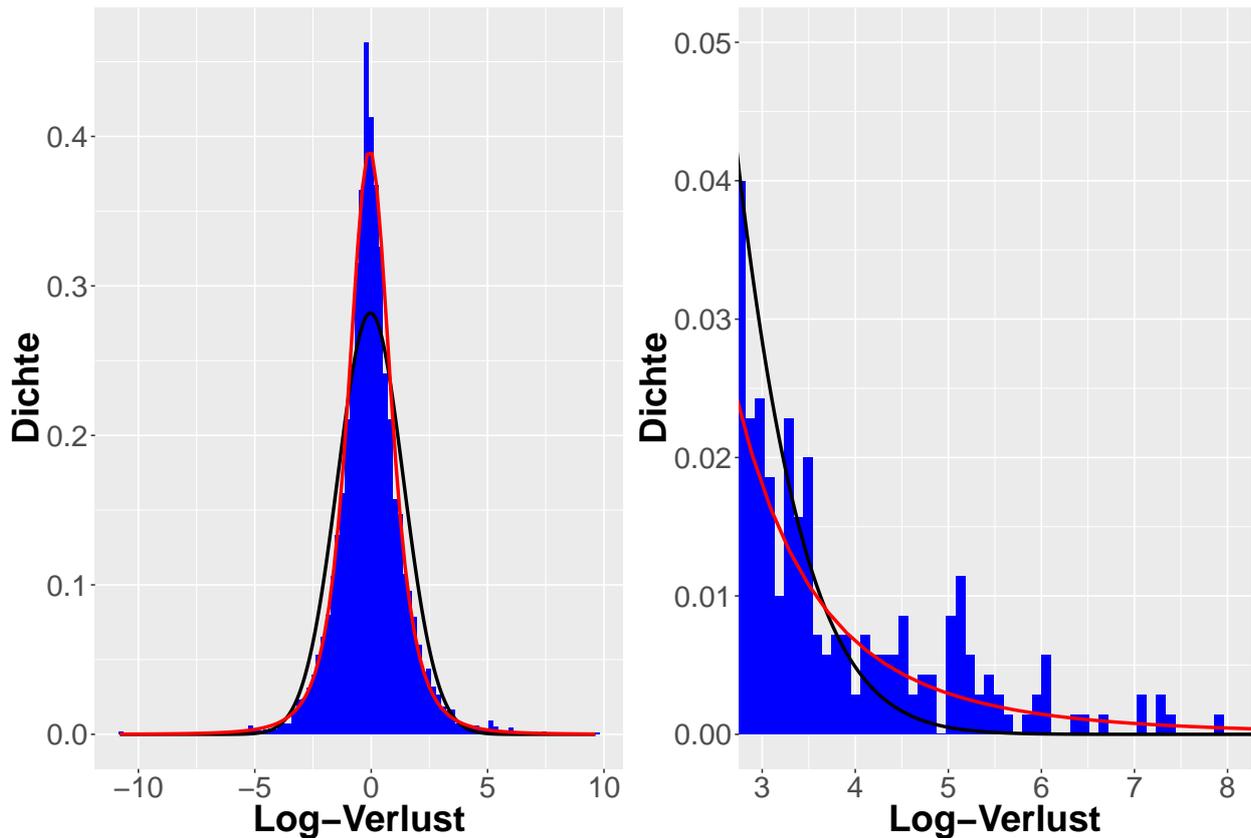


Abbildung 1: Die vorliegenden Daten werden durch Studentische t-Verteilung und Normalverteilung gefittet. Die schwarze Kurve bezeichnet die gefittete Dichtefunktion der Normalverteilung. Die rote Kurve bezeichnet die gefittete Dichtefunktion der Studentischen t-Verteilung. Die anzupassenden Daten sind die logarithmierten Verluste, die aus dem DAX vom 11.27.1990 bis zum 11.26.2017 stammen.

Unter der Studentischen t-Verteilungsannahme wird der VaR geschätzt durch:

$$\widehat{VaR}_\alpha = \hat{\mu}_L + \hat{\lambda} F_{\hat{\nu}}^{-1}(1 - \alpha), \quad (13)$$

wobei  $\hat{\mu}_L$ ,  $\hat{\lambda}$  und  $\hat{\nu}$  geschätzter Erwartungswert, geschätzter Skaleparameter und geschätzter Freiheitsgrad sind.  $F_{\hat{\nu}}^{-1}(\cdot)$  entspricht der inversen Studentischen t-Verteilungsfunktion mit  $\hat{\nu}$  Freiheitsgrad [Ruppert, 2011, S. 557].

Der ES mit dem Signifikanzniveau  $\alpha$  ist [Ruppert, 2011, S. 557]:

$$\widehat{ES}_\alpha = \hat{\mu}_L + \hat{\lambda} \left( \frac{f_{\hat{\nu}}\{F_{\hat{\nu}}^{-1}(1-\alpha)\}}{\alpha} \times \frac{\hat{\nu} + (F_{\hat{\nu}}^{-1}(1-\alpha))^2}{\hat{\nu} - 1} \right), \quad (14)$$

wobei  $f_{\hat{\nu}}(\cdot)$  die Dichtefunktion der Studentischen t-Verteilung mit Freiheitsgrad  $\hat{\nu}$  ist.

### 2.3.4 Skewed t-verteilte Verlustverteilung

Die Studentische t-Verteilung bietet eine gute Alternative, um die leptokurtischen Finanzdaten zu beschreiben. Die Dicke des Tails kann auch durch unterschiedliche Freiheitsgrade  $\nu$  konfiguriert werden. Aber die Finanzdaten sind nicht immer symmetrisch, daher ist die symmetrische t-Verteilung nicht ausreichend. Um dieses Problem zu lösen, wurde die Schiefe in die symmetrische Verteilung von Fernández and Steel [1998] eingeführt:

$$f(x|\gamma) = \frac{2}{\gamma + \gamma^{-1}} \left\{ f\left(\frac{x}{\gamma}\right)I_{[0,\infty)}(x) + f(\gamma x)I_{(-\infty,0)}(x) \right\}, \quad (15)$$

mit dem Skewing-Parameter  $\gamma > 0$ . Dabei ist  $f(x)$  eine symmetrische und unimodale Dichtefunktion bzw.  $f(x) = f(|x|)$ . Die Dichtefunktion  $f(|x|)$  ist sinkend in  $|x|$ . Der Skewing-Parameter  $\gamma$  bestimmt die Schiefe der Verteilung,  $\gamma > 1$  bezeichnet die Rechtsschiefe, dagegen  $\gamma < 1$  die Linksschiefe. Offensichtlich ist die symmetrische Verteilung der Sonderfall der Asymmetrisierung, denn:

$$f(x|\gamma = 1) = f(x). \quad (16)$$

Mit der Einführung des Skewing-Parameters in die Studentische t-Verteilung erhält man die sogenannte Skewed Studentische t-Verteilung (kurz: sktV). Fernández and Steel [1998] zufolge ermöglicht die neue Verteilung eine flexible Modellierung von *skewness* und *fat tails*. In Abbildung 2 ist die Überlegenheit der Skew-t-Verteilung gegenüber den anderen zwei Verteilungen zu sehen, insbesondere beim Schwanz.

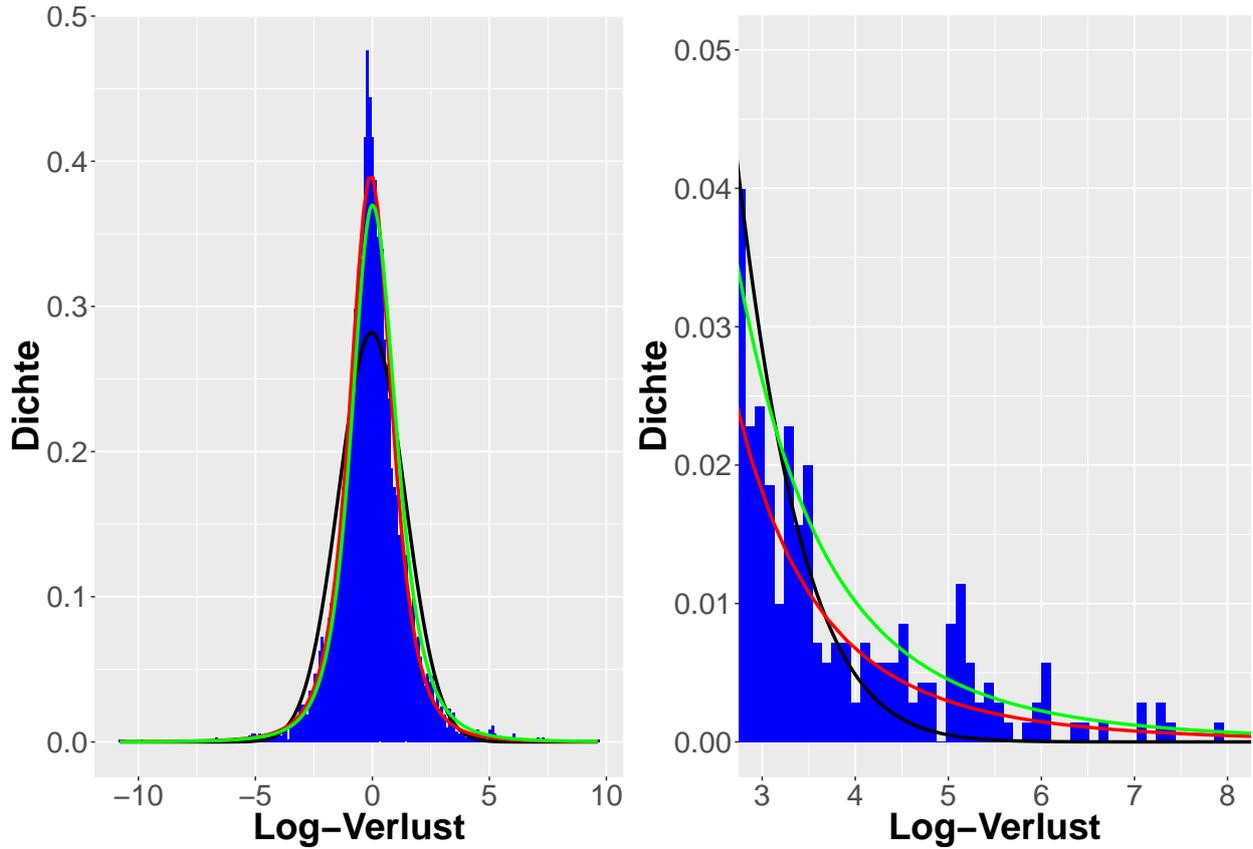


Abbildung 2: Auf der Basis von Abbildung 1 wird die gefittete Kurve der Skewed  $t$ -Verteilung hierbei grün gekennzeichnet.

Laut [Giot and Laurent, 2003, S. 7] ist der VaR zu schätzen durch:

$$\widehat{VaR}_\alpha = \begin{cases} -\hat{\gamma} F_{\hat{\nu}}^{-1}\left(\frac{\alpha}{2}(1 + \hat{\gamma}^{-2})\right), & \text{falls } \alpha \leq \frac{1}{1 + \hat{\gamma}^2}, \\ \frac{1}{\hat{\gamma}} F_{\hat{\nu}}^{-1}\left(\frac{1 - \alpha}{2}(1 + \hat{\gamma}^2)\right), & \text{falls } \alpha > \frac{1}{1 + \hat{\gamma}^2}, \end{cases} \quad (17)$$

wobei  $\hat{\nu}, \hat{\gamma}$  der geschätzte Freiheitsgrad und der geschätzte Skewing-Parameter sind,  $F_{\hat{\nu}}^{-1}(\cdot)$  entspricht der inversen Studentischen  $t$ -Verteilungsfunktion mit  $\hat{\nu}$  Freiheitsgrad.

Die explizite Darstellung des ES ist kompliziert, aber die Kalkulation des ES kann durch die Lösung des Integrals 4 erfolgen.

### 2.3.5 GARCH-Modell

In der Praxis sind die sogenannten stilisierten Fakten (*stylized facts*) bei Finanzdaten zu beobachten. Unter den stilisierten Fakten versteht man neben der Non-Normalität, die oben erläutert worden ist, auch das Volatilitätscluster (*volatility clustering*). Das bedeutet, die Volatilitäten haben die Tendenz, immer zusammen aufzutreten [Peitz, 2016, S. 5]. Wie [Mandelbrot, 1963, S. 418] sagte: „[...] large changes tend to be followed by large changes – of either sign – and small changes tend to be followed by small changes [...]“

Das von Bollerslev [1986] eingeführte GARCH (*generalized autoregressive conditional heteroskedasticity*)-Modell ist der am häufigsten verwendete Ansatz, um die variierende Volatilität zu beschreiben. Während der Erwartungswert der Finanzzeitreihe durch das ARMA (*auto-regressive and moving average*)-Modell analysiert wird, wird die Varianz mithilfe des GARCH-Modells modelliert. Deswegen wird das ARMA(p,q)-GARCH(r,s)-Modell verwendet, um die Verlustzeitreihe genauer zu beschreiben:

$$l_t = \alpha_0 + \sum_{i=1}^p \alpha_i l_{t-i} + \sum_{j=1}^q \beta_j \epsilon_{t-j} + \sigma_t z_t, \quad (18)$$

$$\sigma_t^2 = c_0 + \sum_{i=1}^r c_i \epsilon_{t-i}^2 + \sum_{j=1}^s d_j \sigma_{t-j}^2, \quad (19)$$

dabei ist die Innovation  $z_t$  unabhängig und identisch verteilt (i.i.d), mit einer marginalen Dichtefunktion  $f_z(z) \sim (0, 1)$  [McNeil and Frey, 2000, S. 275]. Die Dichtefunktion  $f_z(z)$  kann verschiedene Verteilungen aufweisen, z.B. Normalverteilung, Studentsche t-Verteilung, Skewed t-Verteilung usw. Die dazugehörigen Parameter werden mithilfe der Maximum-Likelihood-Methode unter spezifischer Verteilungsannahme geschätzt.

In der Studie von Aboura [2014] war das ARMA(1,1)-GARCH(1,1)-Modell nach der Abwägung zwischen Maximierung der Kriterien und Einfachheit des Modells am besten. Zudem sahen McNeil and Frey [2000], Berkowitz and O'Brien [2002] das sparsame GARCH(1,1)-Modell auch als effektiv an.

Die derzeitigen Prognosen von VaR und ES mit dem Signifikanzniveau  $\alpha$  sind zu schätzen durch:

$$\widehat{VaR}_\alpha = \mu_{t+1|t} + \sigma_{t+1|t} F^{-1}(1 - \alpha), \quad (20)$$

$$\widehat{ES}_\alpha = \mu_{t+1|t} + \sigma_{t+1|t} E[Z|Z > F^{-1}(1 - \alpha)], \quad (21)$$

dabei bezeichnet  $F^{-1}(\cdot)$  die inverse Verteilungsfunktion der Innovation  $Z$ . Die Berechnung von  $\mu_{t+1|t}$  und  $\sigma_{t+1|t}$  hat durch die Ein-Schritt-Prognose zu erfolgen, wodurch sich diese bedingte Methode von den vorherigen unbedingten Methoden unterscheidet.

## 3 Theoretischer Hintergrund zur Extremwerttheorie

Die Extremwerttheorie ist ein spezieller Bereich der Statistik, die sich mit den Ausreißern beschäftigt. Obwohl solche extremen Fälle im wirklichen Leben fast unmöglich sind, könnten sie aber weltweit Schlagzeilen machen, falls sie auftreten. In der Finanzwelt nennt man solche negativen Extrema Finanzkrisen. Eine Finanzschwierigkeit könnte ein Unternehmen ruinieren, aber eine große Finanzkrise könnte die ganze Finanzindustrie sowie einen Staat an den Rand drängen. Das Hauptprinzip von Riskmanagement ist, solchen Katastrophen vorzubeugen, deshalb ist die Rolle der EVT beim Riskmanagement nicht zu ignorieren.

Unter der EVT gibt es zwei Methoden: die *Block-Maxima*-Methode und die *Peaks-over-Threshold*-Methode. Im Folgenden werden die zwei Methoden erläutert. Der theoretische Inhalt des Kapitels orientiert sich an Coles et al. [2001], Embrechts et al. [1997], McNeil et al. [2015].

### 3.1 Block-Maxima-Methode

Die *Block-Maxima*-Methode nennt man auch die klassische Extremwerttheorie, die auf einer fundierten theoretischen Basis fusst und eine lange Geschichte hat. Denn die Methode geht auf 1920er Jahre zurück, als Fréchet, Fisher und Tippett ihre frühen Versionen nacheinander aufstellten. Aber erst im Jahr 1948 wurde die Theorie von Gnedenko bekannt gemacht.

#### 3.1.1 Generalisierte Extremwertverteilung

Unter *Block-Maxima* versteht man das Maximum in einem Zeitblock. Das Maximum wird definiert als:

$$M_n = \max\{X_1, X_2, \dots, X_n\}, \quad (22)$$

wobei  $X_1, X_2, \dots, X_n$  eine Folge von unabhängig und identisch mit einer Verteilungsfunktion  $F(\cdot)$  verteilten Zufallsvariablen sind, z.B. eine Reihe von Verlusten<sup>2</sup>. Wenn  $n$  die Anzahl der Geschäftstage eines Jahres ist, dann ist  $M_n$  das jährliche Maximum.

Unter der i.i.d. Annahme gilt:

$$Pr(M_n \leq z) = Pr(X_1 \leq z, X_2 \leq z, \dots, X_n \leq z) = F^n(z). \quad (23)$$

Wenn die zugrunde liegende Verteilung  $F(\cdot)$  bekannt ist, ist die Verteilung von  $M_n$  laut 23 ohne Schwierigkeit herzuleiten, aber  $F(\cdot)$  ist meistens unbekannt. Eine Alternative ist, die asymptotische Eigenschaft von  $F^n(\cdot)$  zu analysieren. Analog zum Zentralen Grenzwertsatz, der besagt, dass der Mittelwert von i.i.d. Zufallsvariablen annähernd normalverteilt ist, konvergiert das Maximum von i.i.d. Zufallsvariablen auch zu einer bestimmten Verteilung.

---

<sup>2</sup>In der Praxis ist die i.i.d. Annahme normalerweise nicht gerechtfertigt, aber in 3.1.4 ist die Annahme auf den nicht i.i.d. Fall zu erweitern.

Um das asymptotische Verhalten von  $F^n(\cdot)$  zu untersuchen, wird der Fall von  $n \rightarrow \infty$  betrachtet, dabei tritt aber ein Problem auf: Sei  $z_+$  der rechte Endpunkt von  $F(\cdot)$ <sup>3</sup>, dann gilt die folgende Gleichung für jede  $z < z_+$  :

$$Pr(M_n \leq z) = F^n(z) \rightarrow 0, n \rightarrow \infty. \quad (24)$$

Für jede  $z > z_+$ , gilt:

$$Pr(M_n \leq z) = F^n(z) \rightarrow 1, n \rightarrow \infty. \quad (25)$$

Das bedeutet, dass die Verteilung von  $M_n$  auf jeden Fall zur 0 oder 1 für  $n \rightarrow \infty$  degeneriert. Diese Schwierigkeit ist zu beheben, indem die Maxima  $M_n$  durch zwei Konstanten  $a_n > 0$  und  $b_n \in \mathbb{R}$  standardisiert werden:

$$M_n^* = \frac{M_n - b_n}{a_n}, \quad (26)$$

und die asymptotische Verteilung von  $M_n^*$  lässt sich durch das Fisher–Tippett–Gnedenko-Theorem definieren, das das Pendant zum zentralen Grenzwertsatz in Extremwertbereich ist:

**Theorem 1 (Fisher–Tippett–Gnedenko-Theorem)** *Wenn es  $a_n > 0$ ,  $b_n \in \mathbb{R}$  und eine nicht degenerierte Verteilung  $H(\cdot)$  gibt, dann gilt:*

$$\lim_{n \rightarrow \infty} Pr \left( \frac{M_n - b_n}{a_n} \leq z \right) = H(z), z \in \mathbb{R}$$

und die Verteilung  $H(\cdot)$  gehört zu einer der folgenden Verteilungsfamilien:

$$\begin{aligned} \text{Gumbel-Verteilung: } H_1(z) &= \exp(-e^{-z}), z \in \mathbb{R}; \\ \text{Fréchet-Verteilung: } H_2(z) &= \begin{cases} \exp(-z^{-\alpha}), & z > 0, \\ 0, & z \leq 0, \end{cases} \alpha > 0; \\ \text{Weibull-Verteilung: } H_3(z) &= \begin{cases} 1, & z > 0, \\ \exp(-(-z)^\alpha), & z \leq 0, \end{cases} \alpha > 0. \end{aligned}$$

Der Beweis des Theorems findet sich in [Embrechts et al., 1997, S. 122].

Das Fisher–Tippett–Gnedenko-Theorem gibt Aufschluss darüber, dass die Verteilung der Extremalwerten anders als die Verteilung der originalen Daten modelliert werden soll. Die Dichtekurven der drei Verteilungen sind in Abbildung 3 zu erkennen.

---

<sup>3</sup>Endpunkt  $z_+$  ist der kleinste Wert von  $z$ , sodass  $F(z) = 1$ .

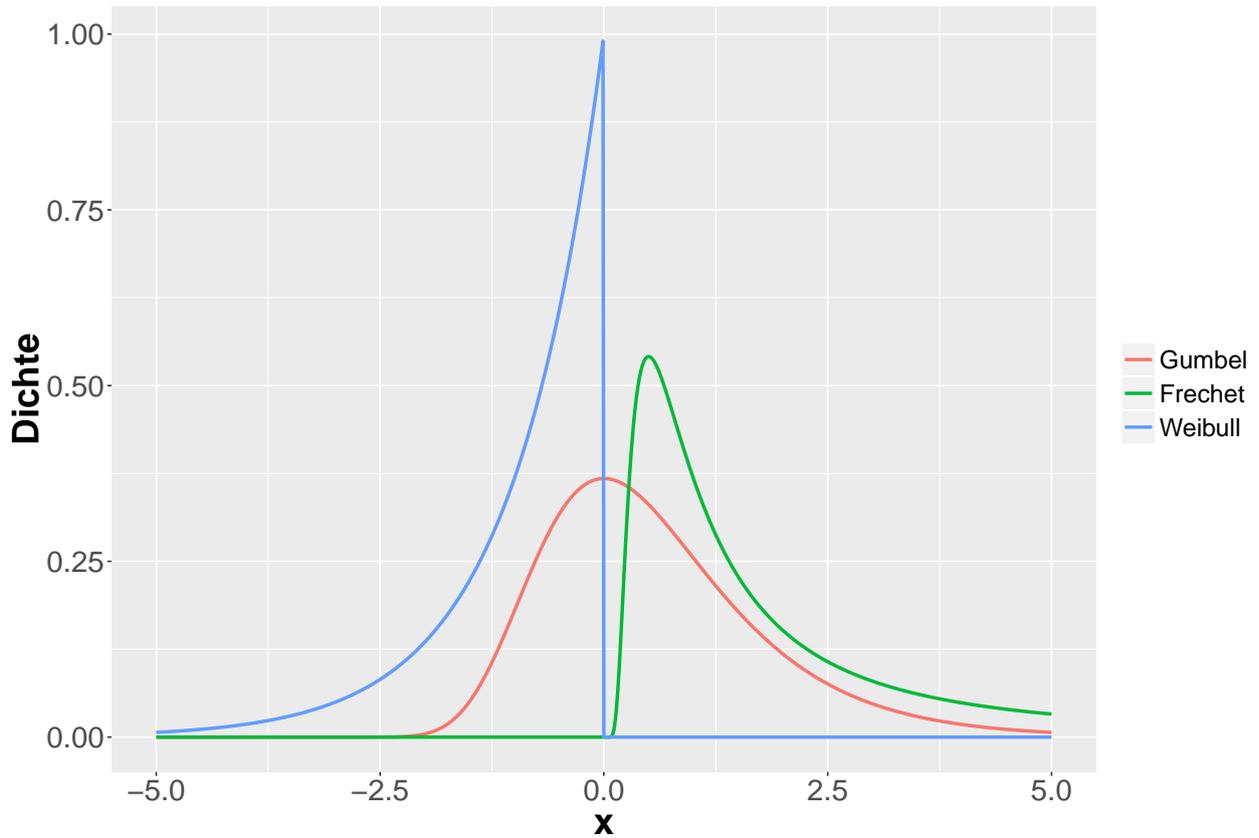


Abbildung 3: Die Dichtekurven der Extremwertverteilung.  $\alpha = 1$  für die Fréchet-Verteilung und die Weibull-Verteilung.

Mit der Einführung des Lageparameters (*location parameter*)  $\mu$  und des Skalenparameters (*scale parameter*)  $\sigma$  kann man die Verteilungsfunktion umschreiben als:

$$\text{Gumbel-Verteilung : } H_1(z) = \exp(-e^{-\frac{z-\mu}{\sigma}}), z \in \mathbb{R};$$

$$\text{Fréchet-Verteilung : } H_2(z) = \begin{cases} \exp(-(\frac{z-\mu}{\sigma})^{-\alpha}), & z > 0, \\ 0, & z \leq 0, \end{cases} \alpha > 0;$$

$$\text{Weibull-Verteilung : } H_3(z) = \begin{cases} 1, & z > 0, \\ \exp(-(-(\frac{z-\mu}{\sigma}))^\alpha), & z \leq 0, \end{cases} \alpha > 0.$$

Die drei Extremwertverteilungen haben ganz unterschiedliche Eigenschaften: die Weibull-Verteilung hat einen finiten Endpunkt  $z_+$ , dagegen haben die Fréchet-Verteilung und die Gumbel-Verteilung einen infiniten Endpunkt  $z_+ = \infty$ . Beim rechten Tail fällt die Gumbel-Verteilung außerdem exponentiell, während die Fréchet-Verteilung polynomial sinkt. Aber in der Praxis ist es kompliziert, drei unterschiedliche Verteilungen gleichzeitig an die Daten anzupassen, weil man vergleichen muss, welche Verteilung die beste Anpassungsgüte liefert. Eine gute Lösung ist, die drei Verteilungen zu einer gemeinsamer Form zusammenzufassen:

**Theorem 2 (Generalisierte Extremwertverteilung)** Wenn es  $a_n > 0$ ,  $b_n \in \mathbb{R}$  und eine

nicht degenerierte Verteilung  $G(\cdot)$  gibt, dann gilt:

$$\lim_{n \rightarrow \infty} \Pr \left( \frac{M_n - b_n}{a_n} \leq z \right) = G(z), z \in \mathbb{R}$$

und die Verteilung  $H(\cdot)$  gehört zu der folgenden Verteilungsfamilie:

$$G(z) = \exp \left\{ - \left( 1 + \xi \frac{z - \mu}{\sigma} \right)^{-1/\xi} \right\}, 1 + \frac{\xi(z - \mu)}{\sigma} > 0, \quad (27)$$

mit  $\mu, \xi \in \mathbb{R}$  und  $\sigma > 0$ .  $G(\cdot)$  nennt man generalisierte Extremwertverteilung (*generalized extreme value distribution*, kurz: GEV).  $\xi$  ist der Formparameter (*shape parameter*), der den Typ der Extremwertverteilung bestimmt. Wenn  $\xi = 0$ , ist  $G(\cdot)$  eine Gumbel-Verteilung. Wenn  $\xi > 0$ , ist  $G(\cdot)$  eine Fréchet-Verteilung. Wenn  $\xi < 0$ , ist  $G(\cdot)$  eine Weibull-Verteilung.

### 3.1.2 Schätzung der generalisierten Extremwertverteilung

Unter der i.i.d. Annahme können die unbekannt Parameter der GEV durch die MLE-Methode geschätzt werden. Sei  $g(\cdot)$  die Dichtefunktion der GEV mit den Maxima  $\mathbf{Z} = (z_1, z_2, \dots, z_m)$ , dann ist die Log-Likelihood-Funktion gegeben durch:

$$\ell(\mu, \sigma, \xi \mid \mathbf{Z}) = \log \left( \prod_{i=1}^m g(z_i) I_{[1 + \xi(z_i - \mu)\sigma > 0]} \right). \quad (28)$$

Im Falle von  $\xi \neq 0$  ist die Log-Likelihood-Funktion:

$$\ell(\mu, \sigma, \xi \mid \mathbf{Z}) = -m \log \sigma - (1 + 1/\xi) \sum_{i=1}^m \log \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^m \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right]^{-1/\xi}. \quad (29)$$

Im Falle von  $\xi = 0$  ist die Log-Likelihood-Funktion:

$$\ell(\mu, \sigma \mid \mathbf{Z}) = -m \log \sigma - \sum_{i=1}^m \left( \frac{z_i - \mu}{\sigma} \right) - \sum_{i=1}^m \exp \left( - \frac{z_i - \mu}{\sigma} \right). \quad (30)$$

Die geschätzten Werte von  $\mu, \sigma, \xi$  bekommt man durch die Maximierung der Gleichung 29 oder 30:

$$(\hat{\mu}, \hat{\sigma}, \hat{\xi}) = \operatorname{argmax}_{\mu, \sigma, \xi} \ell(\mu, \sigma, \xi \mid \mathbf{Z}), \quad (31)$$

$$(\hat{\mu}, \hat{\sigma}) = \operatorname{argmax}_{\mu, \sigma} \ell(\mu, \sigma \mid \mathbf{Z}). \quad (32)$$

Abbildung 3 zeigt, dass im Falle des  $\xi \neq 0$  die Extremwertverteilungen einen rechten oder linken Endpunkt haben, wobei die Regularitätsbedingungen der MLE verletzt sein könnten.

Deshalb ist die Anwendung der MLE an manchen Stellen problematisch. [Smith, 1985, S. 67] hat das Problem studiert und die Schlussfolgerung wie folgt zusammengefasst:

- Wenn  $\xi > -0.5$ , ist der ML-Schätzer regulär. Der Schätzer hat asymptotische Normalität.
- Wenn  $-1 < \xi < -0.5$ , existiert der ML-Schätzer. Aber die normale asymptotische Eigenschaft ist verloren.
- Wenn  $\xi < -1$ , existiert der ML-Schätzer nicht.

Aber der Fall  $\xi \leq -0.5$ , der einem sehr kurzen Schwanz der Verteilung entspricht, tritt in der Praxis selten auf. Longin [1996] hat anhand einer empirischen Studie gezeigt, dass die Renditen asymptotisch fréchetverteilt sind, wobei  $\xi > 0$  ist. Somit ist die Anwendung der GEV auf Finanzdaten normalerweise unproblematisch.

### 3.1.3 Return Level und Risikomaß

Sei  $z_{1-\alpha}$  das  $(1 - \alpha)$ -Quantil von  $G(\cdot)$ , dann ergibt sich aus 27:

$$G(z_{1-\alpha}) = 1 - \alpha = \exp \left\{ - \left( 1 + \xi \frac{z_{1-\alpha} - \mu}{\sigma} \right)^{-1/\xi} \right\}, \quad (33)$$

$$z_{1-\alpha} = \mu - \frac{\sigma}{\xi} (1 - (-\log(1 - \alpha))^{-\xi}). \quad (34)$$

Das Quantil  $z_{1-\alpha}$  nennt man auch *return level* mit *return period*  $1/\alpha$ . Unter Einbeziehung der Größe des Blocks  $n$  wird es auch als  $z_{n,1-\alpha}$  geschrieben. Zum Beispiel:  $M_n$  ist das jährliche Maximum, dann ist  $z_{n,1-\alpha}$  das *return level*, das im Durchschnitt alle  $1/\alpha$  Jahre überschritten wird:

$$1 - G(z_{n,1-\alpha}) = Pr(M_n > z_{n,1-\alpha}) = \alpha. \quad (35)$$

In mancher Literatur schreibt man auch  $z_{n,1-1/k}$ , sodass  $z_{n,1-1/k}$  das  $k$ -jährige *return level* ist. Diese Definition wird häufig in der Hydrologie und für Wettervorhersagen verwendet, um beispielsweise eine Jahrhundertflut vorherzusagen. Der Zeitraum, in dem das *return level* überschritten wird, wird *stress period* genannt.

In diesem Fall beschreibt die Extremwertverteilung  $G(\cdot)$  nur das Verhalten der Maxima  $M_n$ , liefert aber keine Information zum Quantil. Unter der i.i.d. Annahme in 23 von Zufallsvariable  $X$  gilt jedoch [McNeil, 1998, Longin, 2000]:

$$(1 - \alpha)^{1/n} = Pr^{1/n}(M_n \leq z_{n,1-\alpha}) = F(z_{n,1-\alpha}), \quad (36)$$

das heißt,  $z_{n,1-\alpha}$  ist das  $(1 - \alpha)^{1/n}$ -Quantil von  $X$ . Zum Beispiel: Ein 10-monatiges *return*

level ist das 99.50%-Quantil von  $X^4$ .

Nach der Umrechnung aus 36 ist der  $VaR_\alpha$  zu kalkulieren durch:

$$\widehat{VaR}_\alpha = G^{-1}((1 - \alpha)^n), \quad (37)$$

wobei  $G^{-1}(\cdot)$  die inverse GEV ist. Zum Beispiel: Beim monatlichen Block entspricht  $VaR_{0.01}$  dem 80.97%-Quantil von  $G(\cdot)$  oder dem 5.26-monatigen *return level*<sup>5</sup>.

Auf Basis von 37 hat die Schätzung des entsprechenden  $ES_\alpha$  durch 4 zu erfolgen.

### 3.1.4 Extremer Index

Im vorherigen Text 3.1.1 und 3.1.3 wird vorausgesetzt, dass die Zufallsvariable  $X$  i.i.d. ist. Diese Annahme ist bei einer Versicherung oder bei Betriebsrisiko unproblematisch, aber sie gilt nicht für finanzielle Renditen [McNeil et al., 2015, S. 150]. Es ist vorstellbar, dass in einer *stress period* mehr Extremverluste auftreten. Dies ist auf den Cluster-Effekt der Finanzdaten zurückzuführen. In diesem Teil wird die i.i.d. Annahme gelockert und der sogenannte extreme Index (*extremal index*) eingeführt, dadurch ist das Modell auf die stationäre Zeitreihe zu erweitern, wo der Cluster-Effekt besteht.

Zuerst nehmen wir an,  $X_n$  ist eine stationäre, aber nicht i.i.d. Zeitreihe mit einer marginalen Verteilungsfunktion  $F(\cdot)$ . Die Definition 22 gilt wie vorher. Gleichzeitig wird eine andere Zeitreihe  $\tilde{X}_n$  definiert, die dieselbe marginale Verteilungsfunktion  $F(\cdot)$  hat, aber i.i.d. ist. Ähnlich wie  $X_n$  wird das Maximum von  $\tilde{X}_n$  angegeben durch:

$$\tilde{M}_n = \max\{\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n\}. \quad (38)$$

Dann ist die formale Definition des extremen Index  $\theta$  anzugeben:

**Definition 3** Sei  $X_n$  eine stationäre Zeitreihe und  $0 < \theta < 1$ . Für jede  $\tau > 0$  gibt es eine Folge  $u_n$  in  $\mathbb{R}$ , sodass:

$$\lim_{n \rightarrow \infty} n(1 - F(u_n)) = \tau, \quad (39)$$

$$\lim_{n \rightarrow \infty} Pr(\tilde{M}_n \leq u_n) = \exp(-\tau), \quad (40)$$

$$\lim_{n \rightarrow \infty} Pr(M_n \leq u_n) = \exp(-\theta\tau). \quad (41)$$

$\theta$  ist der extreme Index von  $X_n$  und nicht von  $u_n$  abhängig.

<sup>4</sup>Hierbei wird angenommen, dass es 21 Geschäftstage in einem Monat gibt. Mit  $k = 10$  und  $n = 21$  ist  $(1 - 1/10)^{1/21} \approx 0.9950$ .

<sup>5</sup>Für  $n = 21$  und  $\alpha = 0.01$  gilt:  $(1 - 0.01)^{21} \approx 0.8097$  und  $(1 - 0.8179)^{-1} \approx 5.26$ .

Aus der Definition des extremen Index  $\theta$  ergibt sich:

$$Pr(M_n \leq u_n) \approx Pr^\theta(\tilde{M}_n \leq u_n) = F^{n\theta}(u_n). \quad (42)$$

Bei der Gleichung 42 kann man sehen, die Maxima von  $n$  Beobachtungen, die aus  $X_n$  mit dem extremen Index  $\theta$  stammen, sind äquivalent zu den Maxima von  $n\theta$  Beobachtungen, die aus der i.i.d. Reihe  $\tilde{X}_n$  stammen. Dabei funktioniert  $\theta^{-1}$  wie die durchschnittliche Größe des Clusters [McNeil, 1998, S. 7]. Beispielsweise für  $\theta = 1$  beträgt die durchschnittliche Größe des Clusters 1, sodass es kein Cluster gibt. Für  $\theta = 0.5$  gibt es  $n/2$  Clusters, wobei durchschnittlich jeweils 2 Beobachtungen ein Cluster bilden. In einem Wort, je kleiner der extreme Index  $\theta$  ist, desto stärker ist der Cluster-Effekt.

Der extreme Index  $\theta$  kann mithilfe der Block-Methode geschätzt werden, dabei werden die Daten  $X$  in  $m$  äquidistanten Blöcke unterteilt, deren Größe  $n$  ist. Danach wird ein Schwellenwert  $u$  gewählt, sodass die Überschreitungen gezählt werden können. Dann wird  $\theta$  durch den folgenden asymptotischen Schätzer bestimmt [McNeil, 1998, S. 12]:

$$\hat{\theta} = n^{-1} \frac{\log(1 - K_u/m)}{\log(1 - N_u/(mn))}, \quad (43)$$

wobei  $N_u$  die Anzahl der Überschreitungen des Schwellenwerts  $u$  ist.  $K_u$  ist die Anzahl des Blocks, in dem der Schwellenwert  $u$  überschritten wird. Die Herleitung des Schätzers findet sich in [Embrechts et al., 1997, S. 419]. Zu diesem Schätzer hat Mcneil gewarnt, dass sowohl  $m$  als auch  $n$  groß genug sein sollen, somit wird der vierteljährliche oder der halbjährliche Block in der Praxis in Betracht gezogen [McNeil, 1998, S. 12].

Der Schätzer des VaR wird auf Basis von 37 berechnet, aber durch den extremem Index  $\theta$  modifiziert:

$$\widehat{VaR}_\alpha = G^{-1}((1 - \alpha)^{n\theta}). \quad (44)$$

Beispielsweise für  $\theta = 0.5$  ist  $VaR_{0.01}$  das 90.34%-Quantil von  $G(\cdot)$  und das 10.45-monatige *return level*. Zu beachten ist, dass der geschätzte modifizierte VaR im Vergleich zum vorherigen Beispiel deutlich zunimmt. Das impliziert, dass die Ignorierung des Cluster-Effekts zur Unterschätzung des Risikos führen könnte.

### 3.1.5 Wahl der Größe des Blocks

Angenommen, vorhanden sind die i.i.d. täglichen Daten  $X$  im Zeitraum  $[1, T]: x_1, x_2, \dots, x_T$ . Bei der BM-Methode werden die Daten in  $m$  äquidistante Blöcke unterteilt, damit die Größe des Blocks gleich  $n$  ist. Dazu gilt:  $T = mn$ . Das Maximum  $M_i$  definiert sich als der größte

Wert von  $x_{(i-1)n+1}, x_{(i-1)n+2}, \dots, x_{in}$  für  $i = 1, \dots, m$ :

$$\underbrace{x_1, \dots, x_n}_{M_1}, \underbrace{x_{n+1}, \dots, x_{2n}}_{M_2}, \underbrace{x_{2n+1}, \dots, x_{3n}}_{M_3}, \dots, \underbrace{x_{(m-1)n+1}, \dots, x_{mn}}_{M_m}. \quad (45)$$

Das Fisher-Tippett-Gnedenko-Theorem besagt, dass die Verteilung des Maximums  $M_i$  gegen die GEV konvergiert. Allerdings stößt die Anwendung des Theorems in der Praxis auf Probleme: Das Theorem setzt  $n \rightarrow \infty$  voraus, deshalb muss man die Größe des Blocks  $n$  groß genug halten, damit das Maximum des Blocks wirklich ein Extremwert ist. Zum anderen führt eine größere Blockgröße bei einer gegebenen Anzahl der Beobachtungen  $T$  definitiv zu weniger Maxima  $M_i$ , die durch die GEV modelliert werden können. Dies hat eine übergroße Varianz und eine unzuverlässige Inferenz zur Folge. Dazu soll man einen Ausgleich zur Wahl der Größe des Blocks finden: Einerseits muss das Fisher-Tippett-Gnedenko-Theorem bei dieser Blockgröße gelten, andererseits soll es so viele Maxima wie möglich geben.

Zunächst werden die GEV an die Maxima, die unterschiedliche Frequenzen haben, angepasst, z.B. wöchentliche, monatliche, vierteljährliche und halbjährliche sowie jährliche Maxima. Dann wird die Anpassungsgüte der GEV mithilfe von Tests geprüft, sodass die Frage beantwortet werden kann: Sind die Maxima durch die GEV richtig beschrieben? Aber wie ist der Test aufzubauen, der für die GEV machbar ist?

Im Jahr 1957 analysierte Sherman die asymptotische Eigenschaft der sogenannten  $\omega_n$  Statistik und entwickelte einen Test, damit man die  $\omega_n$  testen kann. Die Statistik  $\omega_n$  ist:

$$\omega_n = \frac{1}{2} \sum_{k=1}^{n+1} \left| L_k - \frac{1}{n+1} \right|. \quad (46)$$

$n$  ist die Anzahl der im Einheitsintervall  $(0, 1)$  uniformverteilten Punkte, dadurch wird das Einheitsintervall in  $n + 1$  Subintervalle unterteilt.  $L_k$  ist die Länge des  $k$ -ten Subintervalls. Gemäß [Sherman, 1957, S. 261] ist die  $\omega_n$  normalverteilt mit dem Erwartungswert  $E_n$  und der Varianz  $D_n$ :

$$E_n = \left( \frac{n}{n+1} \right)^{n+1} \xrightarrow{n \rightarrow \infty} \frac{1}{e}, \quad (47)$$

$$D_n \approx \frac{2e - 5}{e^2} \cdot \frac{1}{n}, \quad (48)$$

wobei  $e$  die Eulersche Zahl ist [Sherman, 1957, S. 261].

Da das Quantil einer beliebigen Verteilung im Einheitsintervall  $(0, 1)$  uniformverteilt ist, ist dieser Test auch auf die GEV zu erweitern. Sei  $(Z'_{n,i})_{i=1, \dots, m}$  die Ordnungsstatistik des

Maximums  $M_i$ , dann ergibt sich aus 46 die Statistik [Gumbel, 1958, S. 38]:

$$\Omega_m = \frac{1}{2} \sum_{i=0}^m \left| F_{Z_n}^{asympt}(Z'_{n,i+1}) - F_{Z_n}^{asympt}(Z'_{n,i}) - \frac{1}{m+1} \right|, \quad (49)$$

dabei wird definiert, dass  $F_{Z_n}^{asympt}(Z'_{n,0}) = 0$  und  $F_{Z_n}^{asympt}(Z'_{n,m+1}) = 1$  sind. Bei diesem Test behauptet die Nullhypothese, dass die Maxima die Verteilung  $F_{Z_n}^{asympt}$  befolgen bzw. die GEV. Laut [Longin, 2000, S. 1106] kann die Statistik  $\Omega_m$  als die Diskrepanz zwischen der geschätzten Verteilung und der beobachteten Verteilung interpretiert werden, deswegen deutet ein niedriger Wert von  $\Omega_m$  darauf hin, dass die Maxima durch die GEV gut gefittet werden. Im Gegensatz dazu bedeutet ein hoher Wert von  $\Omega_m$  eine schlechte Anpassung. Wegen der asymptotischen normalen Eigenschaft der Statistik ist der kritische Wert anhand 48 zu kalkulieren. Wenn  $\Omega_m$  den kritischen Wert überschreitet, wird die  $H_0$  abgelehnt. Dies ist hauptsächlich auf die zu kleine Blockgröße  $n$  zurückzuführen. In diesem Fall soll man die Blockgröße vergrößern und den Test wieder vornehmen [Longin, 2000, S. 1106].

## 3.2 Peaks-over-Threshold-Methode

In Kapitel 3.1 wird über die BM-Methode diskutiert, die eine gute Alternative bietet, um den Extremwert zu analysieren. Dennoch hat die klassische Methode einen großen Nachteil: die Verschwendung der Daten. Die Daten werden im Voraus in Blöcke unterteilt und in jedem Block wird nur ein Maximum betrachtet. Unter der Voraussetzung  $n \rightarrow \infty$  des Fisher–Tippett–Gnedenko-Theorems steht nur eine beschränkte Anzahl von Maxima zur Verfügung. Innerhalb eines Blocks wird insbesondere in der *stress period* auf zahlreiche Extremwerte verzichtet, selbst wenn sie viel größer als manche Maxima in anderen Blöcken sein könnten. Deswegen wird die sogenannte *Peaks-over-Threshold-Methode* (kurz: POT-Methode) in der Praxis oft gegenüber der BM-Methode bevorzugt.

### 3.2.1 Generalisierte Paretoverteilung

Bei der POT-Methode kommt der Exzess über einen Schwellenwert in Betracht, statt die Maxima zu analysieren. Sei  $X_1, X_2, \dots, X_n$  die i.i.d. Zufallsvariablen mit einer marginalen Verteilungsfunktion  $F(\cdot)$ . Dabei werden solche  $X_i$ , die einen Schwellenwert  $u$  überschreiten, als Extremwerte verstanden. Wir definieren  $Y_1, Y_2, \dots, Y_{N_u}$  als die Exzesse, wobei  $N_u$  die Anzahl der Überschreitungen ist. Offensichtlich gilt  $X = Y + u$ .

**Definition 4 (Bedingte Exzessverteilung)** Die Verteilungsfunktion des Exzesses über

den Schwellenwert  $u$  ist durch die bedingte Wahrscheinlichkeit gegeben:

$$F_u(y) = Pr(Y \leq y \mid X > u) = Pr(X - u \leq y \mid X > u), \quad (50)$$

$$= \frac{Pr(X \leq u + y, X > u)}{Pr(X > u)} = \frac{F(y + u) - F(u)}{1 - F(u)}, \quad (51)$$

wobei der Exzess  $y > 0$  ist.

Wenn die zugrunde liegende Verteilung  $F(\cdot)$  bekannt ist, ist die Verteilung des Exzesses durch 51 zu bestimmen. Aber im wirklichen Leben ist dies nicht der Fall, dann muss man sich an die asymptotische Eigenschaft der Exzessverteilung halten.

**Theorem 3 (Pickands–Balkema–de Haan-Theorem)** Seien  $X_1, X_2, \dots, X_n$  i.i.d. Zufallsvariablen mit einer marginalen Verteilungsfunktion  $F(\cdot)$ . Sei  $F_u(\cdot)$  die bedingte Exzessverteilung bezüglich eines hohen Schwellenwerts  $u$ , dann ist die  $F_u(\cdot)$  durch die generalisierte Pareto-Verteilung (kurz: GPD)  $G_{\xi, \beta}(\cdot)$  approximiert:

$$F_u(y) \xrightarrow{u \rightarrow \infty} G_{\xi, \beta}(y). \quad (52)$$

Der Beweis des Pickands–Balkema–de Haan-Theorems findet sich in [Embrechts et al., 1997, S. 165].

**Definition 5 (Generalisierte Pareto-Verteilung)** Die Verteilungsfunktion der generalisierten Pareto-Verteilung  $G_{\xi, \beta}(\cdot)$  mit dem Formparameter  $\xi \in \mathbb{R}$  und dem Skalenparameter  $\beta > 0$  ist gegeben durch:

$$G_{\xi, \beta}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta}\right)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp\left(-\frac{y}{\beta}\right), & \xi = 0. \end{cases} \quad (53)$$

wobei  $y \geq 0$  wenn  $\xi \geq 0$ , oder  $y \in [0, -\beta/\xi]$  wenn  $\xi < 0$ .

Für  $\xi = 0$  bekommt man die Exponentialverteilung. Für  $\xi > 0$  bekommt man die *heavy-tailed* gewöhnliche Pareto-Verteilung. Für  $\xi < 0$  bekommt man die *short-tailed* Pareto-II-Verteilung. Die Dichtefunktion sowie die Verteilungsfunktion der GPD sind in Abbildung 4 veranschaulicht.

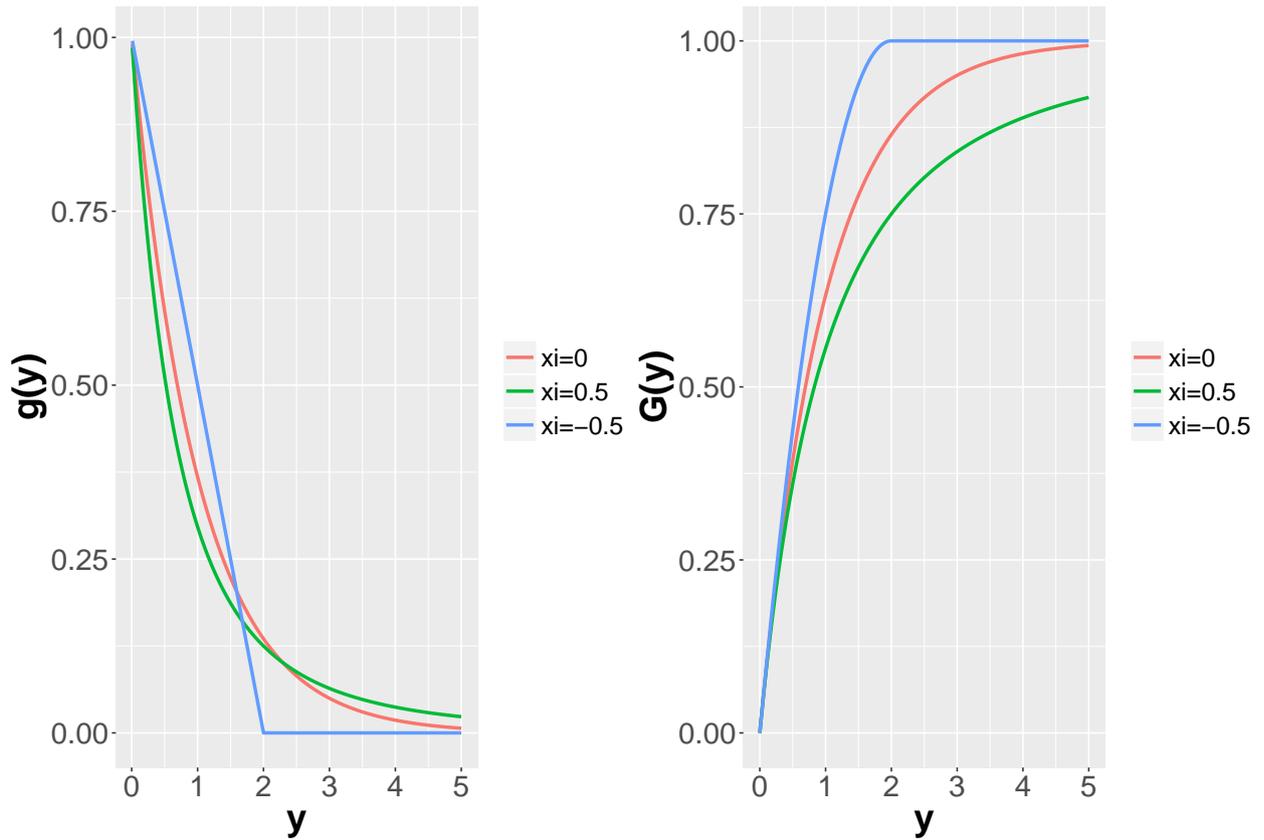


Abbildung 4: Die generalisierte Pareto-Verteilung. Die linke Abbildung zeichnet die Dichtefunktion der GPD mit unterschiedlichen  $\xi$ . Die rechte Abbildung zeichnet die Verteilungsfunktion der GPD mit unterschiedlichen  $\xi$ .

### 3.2.2 Schätzung der generalisierten Pareto-Verteilung

Unter der i.i.d. Annahme können die unbekannt Parameter der GPD durch die MLE-Methode geschätzt werden. Sei  $g_{\beta,\xi}(\cdot)$  die Dichtefunktion der GPD mit den Exzessen  $\mathbf{Y} = (y_1, y_2, \dots, y_{N_u})$ , dann ist die Log-Likelihood-Funktion gegeben durch:

$$\ell(\beta, \xi | \mathbf{Y}) = \log \left( \prod_{i=1}^{N_u} g_{\beta,\xi}(y_i) \right). \quad (54)$$

Im Falle von  $\xi \neq 0$  ist die Log-Likelihood-Funktion:

$$\ell(\beta, \xi | \mathbf{Y}) = -n \log \beta - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{N_u} \log \left(1 + \xi \frac{y_i}{\beta}\right), \quad (55)$$

wobei  $1 + \xi \frac{y_i}{\beta} > 0$  für  $i = 1, 2, \dots, N_u$  sind.

Im Falle von  $\xi = 0$  ist die Log-Likelihood-Funktion:

$$\ell(\beta \mid \mathbf{Y}) = -n \log \beta - \frac{1}{\beta} \sum_{i=1}^{N_u} y_i. \quad (56)$$

Dann sind die Parameter  $\beta, \xi$  geschätzt durch die Maximierung der 55 oder 56:

$$(\hat{\beta}, \hat{\xi}) = \operatorname{argmax}_{\beta, \xi} \ell(\beta, \xi \mid \mathbf{Y}), \quad (57)$$

$$\hat{\beta} = \operatorname{argmax}_{\beta} \ell(\beta \mid \mathbf{Y}). \quad (58)$$

### 3.2.3 VaR und ES

Die Schätzung von VaR und ES ist aus der bedingten Wahrscheinlichkeit 51 herzuleiten:

$$F(y + u) = F_u(y) - F(u)F_u(y) + F(u), \quad (59)$$

$$1 - F(y + u) = 1 - F_u(y) + F(u)F_u(y) - F(u), \quad (60)$$

$$1 - F(x) = (1 - F(u))(1 - F_u(y)), \quad (61)$$

wobei  $x = y + u$  und  $y \geq 0$  ist.

Laut des Pickands-Balkema-de Haan-Theorems befolgt der Exzess asymptotisch die GPD, dann ist  $F_u(y)$  in 61 durch  $G_{\beta, \xi}(y)$  zu ersetzen:

$$1 - F(x) = (1 - F(u))(1 - G_{\beta, \xi}(y)). \quad (62)$$

Somit ist die Verteilungsfunktion von  $X$  geschätzt durch:

$$\hat{F}(x) = 1 - (1 - \hat{F}(u))(1 - \hat{G}_{\beta, \xi}(y)). \quad (63)$$

Bei der Gleichung 63 wird der Term  $(1 - \hat{F}(u))$  als  $Pr(X > u)$  verstanden. In der Praxis kann die Wahrscheinlichkeit der Überschreitung durch  $N_u/n$  approximiert werden. Aus der Definition der GPD in 53 ergibt sich der Schätzer:

$$\hat{F}(x) = \begin{cases} 1 - \frac{N_u}{n} \left( 1 + \hat{\xi} \frac{x - u}{\hat{\beta}} \right)^{-1/\hat{\xi}}, & \xi \neq 0; \\ 1 - \frac{N_u}{n} \exp \left( -\frac{x - u}{\hat{\beta}} \right), & \xi = 0. \end{cases} \quad (64)$$

Sei  $x_{1-\alpha}$  das  $(1 - \alpha)$ -Quantil von Interesse, dann ist der Schätzer umzuschreiben als:

$$\hat{F}(x_{1-\alpha}) = 1 - \alpha = \begin{cases} 1 - \frac{N_u}{n} \left( 1 + \hat{\xi} \frac{x_{1-\alpha} - u}{\hat{\beta}} \right)^{-1/\hat{\xi}}, & \xi \neq 0; \\ 1 - \frac{N_u}{n} \exp \left( -\frac{x_{1-\alpha} - u}{\hat{\beta}} \right), & \xi = 0. \end{cases} \quad (65)$$

Da der  $VaR_\alpha$  tatsächlich das  $(1 - \alpha)$ -Quantil von  $X$  ist, bekommt man den Schätzer vom  $VaR_\alpha$  durch die Umschreibung von 65:

$$\widehat{VaR}_\alpha = \hat{x}_{1-\alpha} = \begin{cases} u + \frac{\hat{\beta}}{\hat{\xi}} \left[ \left( \frac{n\alpha}{N_u} \right)^{-\hat{\xi}} - 1 \right], & \xi \neq 0; \\ u - \hat{\beta} \log \left( \frac{n\alpha}{N_u} \right), & \xi = 0. \end{cases} \quad (66)$$

Auf Basis von 66 ist der entsprechende  $ES_\alpha$  geschätzt mithilfe von 4:

$$\widehat{ES}_\alpha = \alpha^{-1} \int_0^\alpha \widehat{VaR}_u du = \frac{\widehat{VaR}_\alpha}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1 - \hat{\xi}}, \quad (67)$$

wobei  $\xi < 1$  ist.

### 3.2.4 Wahl des Schwellenwerts

In 3.1.5 haben wir über die Wahl der Blockgröße diskutiert, die vor einem widersprüchlichen Problem steht. Während die BM-Methode das Maxima als Extremwert definiert, betrachtet die POT-Methode die über den Schwellenwert stehenden Exzesse als Extremwerte. Trotzdem ist die POT-Methode mit einer ähnlichen Schwierigkeit konfrontiert: Zunächst erfordert das Pickands-Balkema-de Haan-Theorem einen ausreichend hohen Schwellenwert  $u$ , damit die Voraussetzung  $u \rightarrow \infty$  nicht verletzt ist, ansonsten sind die auf der GPD beruhende statistische Schätzung und die Inferenz verzerrt. Zum anderen sind über einen überhohen Schwellenwert nur wenige oder keine Werte zu beobachten, die zu einer übergroßen Varianz führen könnten. Deswegen ist die Abwägung zwischen Verzerrung und Varianz notwendig. Grundsätzlich gibt es zwei Methoden, um den geeigneten Wert von  $u$  zu wählen: durch die *Mean Excess* Funktion oder durch die Hill-Methode.

**Mean Excess Funktion** Bei dieser Methode wird der Erwartungswert des Exzesses in Betracht gezogen. Dabei wird die sogenannte *Mean Excess* Funktion aufgebaut, die eine Funktion ausschließlich von  $u$  ist:

**Definition 6 (Mean Excess Funktion)** Die Mean Excess Funktion  $e(u)$  definiert sich als:

$$e(u) = E(X - u \mid X > u), \quad (68)$$

sie wird bei der Survival-Analyse auch *Mean Residual Life* genannt. Im Falle der GPD ist die *mean excess* Funktion durch die folgende Formel zu kalkulieren:

$$e(u) = \frac{\beta}{1 - \xi} + \frac{\xi u}{1 - \xi}. \quad (69)$$

Der Beweis dazu findet sich in [Embrechts et al., 1997, S. 162], dabei ist es zu beachten, dass die  $e(u)$  linear in  $u$  ist. Das bedeutet, wenn die Exzesse durch die GPD gut gefittet werden, ist die  $e(u)$  linear abhängig von  $u$ . Dieser Zusammenhang bietet eine Möglichkeit, um den geeigneten Wert von  $u$  zu bestimmen, dazu wird der *Mean Residual Life Plot* betrachtet:  $\{(u, e(u)), u < x_{max}\}$ , der in der Praxis durch den *Sample Mean Residual Life Plot* geschätzt wird:

$$\left\{ \left( u, \frac{1}{N_u} \sum_{i=1}^{N_u} (x_i - u) \right), u < x_{max} \right\}. \quad (70)$$

wobei  $x_1, x_2, \dots, x_{N_u}$  die über dem Schwellenwert  $u$  stehenden Beobachtungen sind. Anschließend wird der Schwellenwert als  $u^*$  im Plot gewählt, sodass für  $u > u^*$  die Kurve von  $e(u)$  (annähernd) linear ist. Allerdings ist die optimale Wahl von  $u$  in der praktischen Durchführung nicht immer eindeutig. Als Beispiel wird die folgende Abbildung 5 angeführt: Diese knifflige Kurve hat viele unterschiedlichen Tendenzen. Ein relativ klarer linearer Zusammenhang ist aber im Intervall  $[3.5, 5]$  zu beobachten. Für  $u > 5$  gibt es zu wenige Beobachtungen, deshalb wird  $u = 3.5$  als der Schwellenwert gewählt.

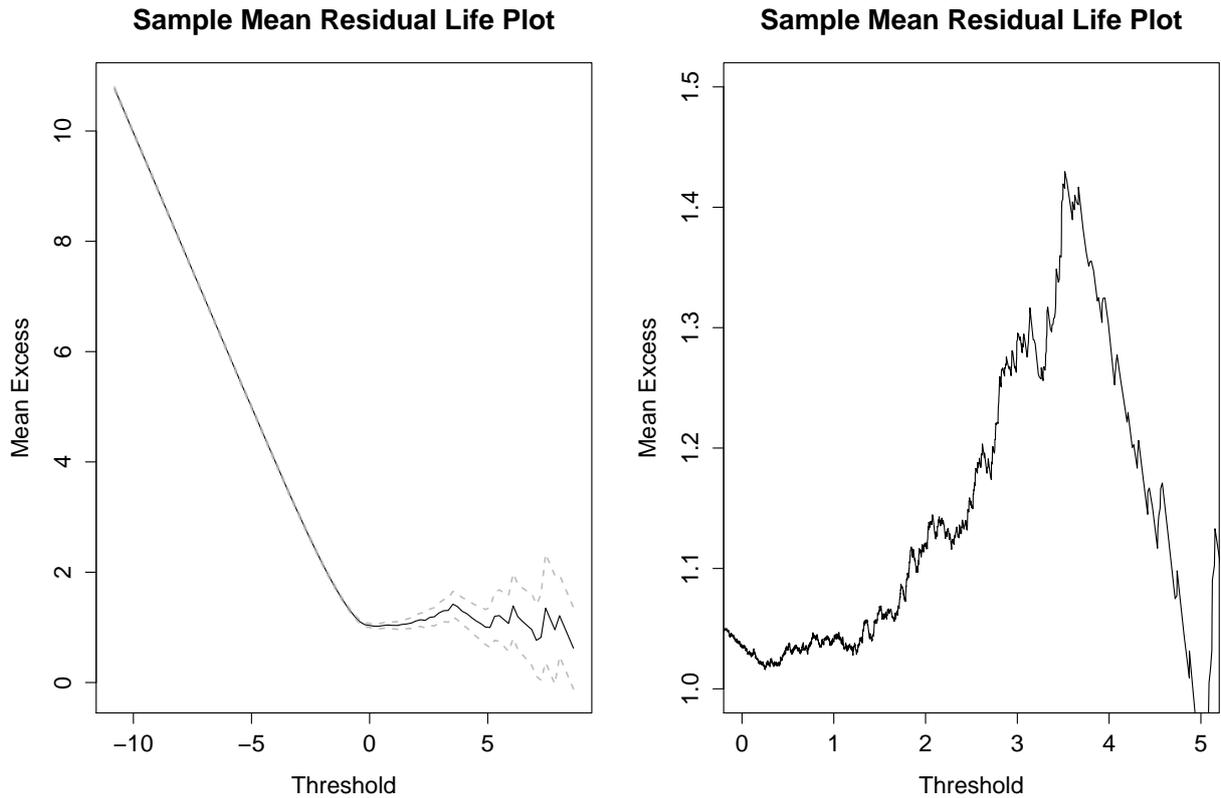


Abbildung 5: *Sample Mean Excess Function*. Bei der rechten Grafik wird die Kurve im Intervall  $[0, 5]$  gezeigt. Die Daten sind die logarithmierten Verluste aus dem DAX vom 11.27.1990 bis zum 11.26.2017.

**Hill-Methode** Neben der *Mean Excess* Funktion ist die berühmte Hill-Methode eine weitere Alternative, um den Schwellenwert zu bestimmen. Diese Methode geht auf Hill et al. [1975] zurück und das eigentliche Ziel dieser Methode ist, den sogenannten Tail-Index  $\iota$  der *fat-tailed* Verteilung herauszufinden.

Für  $\xi > 0$  ist der Hill-Schätzer gegeben durch:

$$\hat{\iota} = \frac{1}{\hat{\xi}} = \frac{1}{N_u} \sum_{i=1}^{N_u} (\log x_{(i)} - \log x_{(N_u+1)}), \quad (71)$$

wobei  $x_{(1)}, x_{(2)}, \dots, x_{(N_u)}$  die abnehmenden Ordnungsstatistiken sind<sup>6</sup>. Hierbei spielt die Ordnungsstatistik  $x_{(N_u+1)}$  die Rolle des Schwellenwerts. Um den optimalen Schwellenwert zu finden, wird der geschätzte Parameter  $\hat{\iota}$  auf der y-Achse gegen unterschiedliche Werte von  $N_u$  auf der x-Achse gezeichnet:  $\{(N_u, \hat{\xi}), N_u = 1, \dots, n - 1\}$ . Dabei wird der Schwellenwert  $x_{(N_u+1)}$  so gewählt, dass in der Nähe von  $x_{(N_u+1)}$  der Wert von y-Achse stabil ist. Als Beispiel wird die Abbildung 6 angeführt.

<sup>6</sup>In mancher Literatur wird  $N_u$  durch  $N_u - 1$  ersetzt.

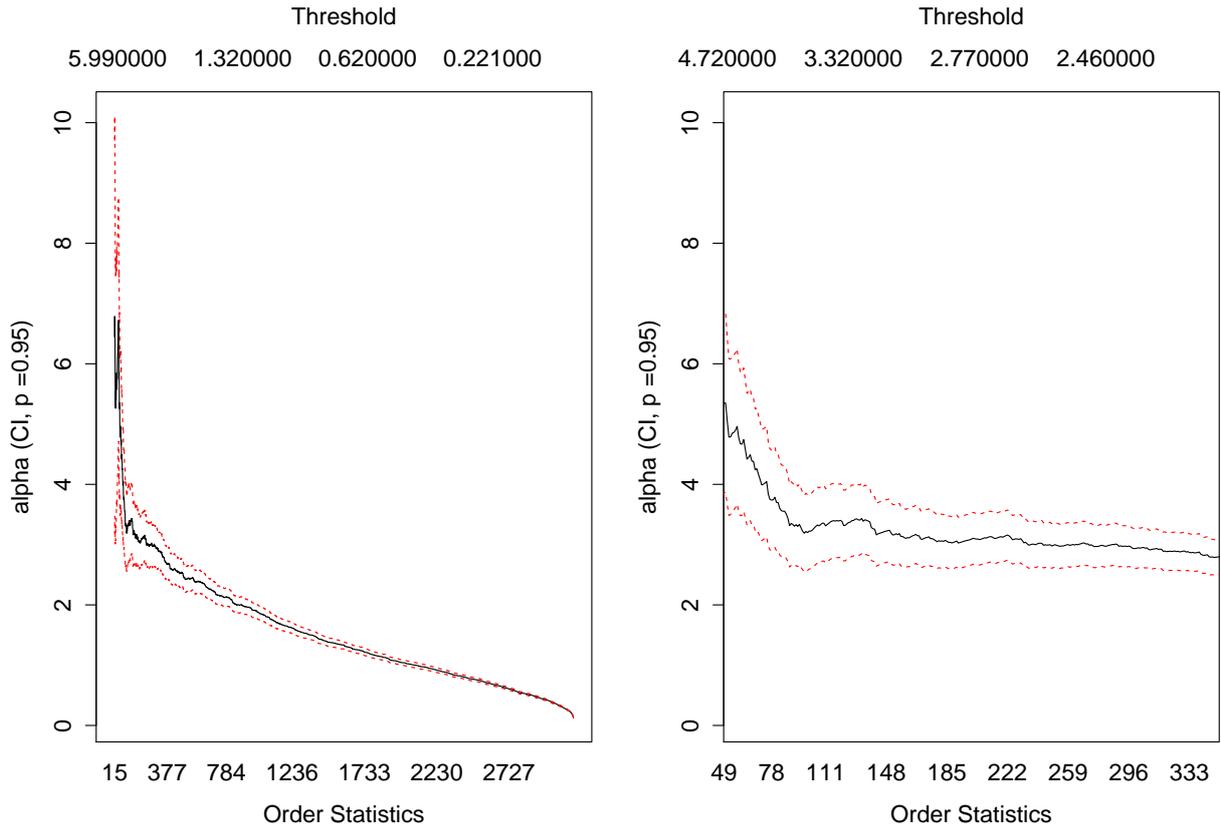


Abbildung 6: Hill-Plot mit einem 95%-Konfidenzintervall. Der rechte Plot fokussiert auf das Intervall  $[60, 340]$ . Die Daten sind die logarithmierten Verluste aus dem DAX vom 11.27.1990 bis zum 11.26.2017. Es gibt insgesamt 6826 Beobachtungen.

Wie der linke Plot in 6 darstellt, sind heftige Schwankungen bei einem kleinen Wert von  $N_u$  (entspricht hohem Schwellenwert) zu erkennen und für große  $N_u$  nimmt der geschätzte Parameter  $\hat{\alpha}$  eindeutig ab. Theoretisch befindet sich der ideale Schwellenwert zwischen den beiden Phasen, aber anhand des originalen Hill-Plots ist keine klare Entscheidung zu treffen.

[McNeil et al., 2015, S. 159] schlugen vor, dass der Wert von  $N_u$  klein sein soll, z.B. 10-50 aus 1000 Beobachtungen. Deswegen ist der Hill-Plot auf das Intervall  $[60, 340]$  vergrößert. Gemäß des rechten Plots wird  $N_u = 100$  bestimmt und die Ordnungsstatistik  $x_{100}$  entspricht ungefähr 3.50. Durch 71 kann der Formparameter  $\hat{\xi}$  auch direkt erhalten werden, aber der Hill-Schätzer gilt nur für den *heavy-tailed* Fall bzw.  $\xi > 0$ . [McNeil and Frey, 2000, S. 287] haben davor gewarnt, dass es doch eine Periode gibt, in der die Verteilung der Finanzdaten auch *light-tailed* ist. Anhand einer Simulationsstudie empfahlen Mcneil und Frey, die Formparameter  $\xi$  durch die GPD zu berechnen, weil die GPD-Methode viel stabiler und in den meisten Fällen effizienter als die Hill-Methode ist. Außerdem zeigten Embrechts, Klüppelberg und Mikosch ein paar *horror plots* der Hill-Methode, z.B. auf Seite 270 von Embrechts et al. [1997], dass die Güte der Hill-Methode durch die Abhängigkeit der Daten stark gestört wird. Auf Seite 344 von Embrechts et al. [1997] wurde die Untauglichkeit des Hill-Schätzers bei

DAX-Daten untermauert.

Die Schätzung durch sowohl einen *Sample Mean Residual Life* Plot als auch einen Hill-Plot hat jedoch durch die sogenannte *eyeball method* zu erfolgen [Danielsson, 2011, S. 174], die manchmal zu subjektiv ist. Dazu gibt es doch die auf der Maximierung der mittleren quadratischen Abweichung (*mean squared error*) basierende *Subsample-Bootstrap*-Methode, deren Implementierung aber kompliziert und zeitaufwändig ist<sup>7</sup>.

### 3.2.5 GARCH-Filter

Bisher ist die Anwendung der POT-Methode nur auf die i.i.d. Daten eingegrenzt, aber wie Kapitel 2.3.5 erläutert, eignen sich die Finanzzeitreihen nicht dazu, durch so ein einfaches Modell modelliert zu werden. Um die stilisierten Fakten zu vermeiden, wird die BM-Methode bereits in 3.1.4 durch den extremen Index modifiziert. In diesem Abschnitt wird der von McNeil and Frey [2000] aufgestellte GARCH-Filter vorgestellt, womit die gefilterten Daten statt der originalen Daten zu betrachten sind, dann ist die POT-Methode auch auf die Finanzzeitreihe anwendbar.

Sei  $X_1, X_2, \dots, X_t$  eine starke stationäre Zeitreihe, z.B. die logarithmierten Renditen oder Verluste. Dabei wird angenommen, dass diese Zeitreihe durch die folgende Gleichung beschrieben werden kann [McNeil and Frey, 2000, S. 275]:

$$X_t = \mu_t + \sigma_t Z_t, \quad (72)$$

wobei die Innovation  $Z_t$  ein *strict white noise process* ist, mit einer marginalen Verteilungsfunktion  $F_Z(z) \sim (0, 1)$ . Die  $F_Z(\cdot)$  erlaubt unterschiedliche Verteilungsannahmen, z.B. Normalverteilung, Studentsche t-Verteilung usw. In diesem Fall ist die Verteilung  $F_X(x)$  von Interesse, damit die Risikomaße geschätzt werden können. Mithilfe von 2 ist die Prognose des VaR zum Zeitpunkt  $t$  mit dem Signifikanzniveau  $\alpha$  gegeben durch:

$$VaR_\alpha = \inf\{x \in \mathbb{R} : F_{X_{t+1}|\mathcal{G}_t}(x) \geq 1 - \alpha\}, \quad (73)$$

wobei  $\mathcal{G}_t$  die verfügbare Information zum Zeitpunkt  $t$  bezeichnet und  $F_{X_{t+1}|\mathcal{G}_t}$  die prognostische Verteilung von  $X_{t+1}$  ist. Mithilfe von 3 ist die Prognose des ES zum Zeitpunkt  $t$  mit dem Signifikanzniveau  $\alpha$  gegeben durch:

$$ES_\alpha = E[X_{t+1} | X_{t+1} \geq VaR_\alpha]. \quad (74)$$

Gemäß Gleichung 72 ergibt sich die prognostische Verteilungsfunktion von  $X_{t+1}$  [McNeil

---

<sup>7</sup>Siehe [Danielsson et al., 2001]. Das dazugehörige Verfahren kann mithilfe vom *tea* Paket (Version: 1.0) durchgeführt werden. Aber die Laufzeit ist sehr lang und von der Rechenleistung stark abhängig.

and Frey, 2000, S. 276]:

$$F_{X_{t+1}|\mathcal{G}_t} = Pr(\mu_{t+1} + \sigma_{t+1}Z_{t+1} \leq x | \mathcal{G}_t) \quad (75)$$

$$= P\left(Z_{t+1} \leq \frac{x - \mu_{t+1}}{\sigma_{t+1}} | \mathcal{G}_t\right), \quad (76)$$

$$= F_Z\left(\frac{x - \mu_{t+1}}{\sigma_{t+1}}\right). \quad (77)$$

Vor der Schätzung der Risikomaße muss die Zeitreihe durch das ARMA(p,q)-GARCH(r,s)-Modell spezifiziert werden. Als Beispiel führten McNeil and Frey [2000] das einfache, aber effektive AR(1)-GARCH(1,1)-Modell an. Mithilfe von 19 ist der bedingte Erwartungswert von  $X_t$  bestimmt durch:

$$\mu_t = \alpha_0 + \alpha X_{t-1}. \quad (78)$$

Die bedingte Varianz von  $X_t$  ist:

$$\sigma_t^2 = c_0 + c_1 \epsilon_{t-1}^2 + d \sigma_{t-1}^2, \quad (79)$$

wobei  $\epsilon_t = X_t - \mu_t$ ,  $c_0 > 0$ ,  $c_1 > 0$  und  $d > 0$  sind. Bei den traditionellen Vorgehensweisen werden die Parameter des GARCH-Modells ( $\hat{\alpha}_0, \hat{\alpha}, \hat{c}_0, \hat{c}_1, \hat{d}$ ) durch die MLE-Methode geschätzt, wobei die Innovation  $Z_t$  beispielsweise als standardnormalverteilt angenommen wird. Aber diese primitive Annahme ist in der Praxis immer nicht gerechtfertigt und unterschätzt das Quantil insbesondere für kleine  $\alpha$ . Feinere Verteilungsannahmen riskieren sich auch, die Dynamik fälschlicherweise zu spezifizieren, deswegen schlugen [McNeil and Frey, 2000, S. 278] vor, die Parameter durch die sogenannte *Pseudo-Maximum-Likelihood*-Methode (Kurz: PML-Methode)<sup>8</sup> zu schätzen. Die PML-Methode macht keine Verteilungsannahme für die Innovation  $Z_t$ , aber liefert konsistente und asymptotische normale Schätzer für die Parameter [Gourieroux et al., 1984, S. 681].

Nachdem man die Parameter des GARCH-Modells geschätzt hat, können die bedingten Erwartungswerte ( $\hat{u}_1, \hat{u}_2, \dots, \hat{u}_t$ ) und die bedingten Varianzen ( $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_t^2$ ) zu jedem Zeitpunkt durch 78 und 79 rekursiv kalkuliert werden.

Im Anschluss daran bekommt man die standardisierten Residuen mithilfe von 72:

$$(z_1, z_2, \dots, z_t) = \left( \frac{x_1 - \hat{u}_1}{\hat{\sigma}_1}, \frac{x_2 - \hat{u}_2}{\hat{\sigma}_2}, \dots, \frac{x_t - \hat{u}_t}{\hat{\sigma}_t} \right), \quad (80)$$

wenn das angepasste GARCH-Modell geeignet ist, soll die gefilterte Folge  $(z_1, z_2, \dots, z_t)$  i.i.d. sein [McNeil and Frey, 2000, S. 278]. Danach gehen die  $(z_1, z_2, \dots, z_t)$  direkt in die GPD-Methode ein, wie Kapitel 3.2.1 bis 3.2.4 schon erläutern. Die daher geschätzten VaR und

<sup>8</sup>In mancher Literatur wird diese Methode auch *Quasi-Maximum-Likelihood* genannt.

ES der gefilterten Daten mit dem Signifikanzniveau  $\alpha$  werden als  $\hat{z}_\alpha$  und  $\hat{S}_\alpha$  gekennzeichnet [McNeil and Frey, 2000, S. 293].

Schließlich werden mithilfe von 77 die prognostischen Risikomaße der originalen Daten zum Zeitpunkt  $t$  geschätzt durch [McNeil and Frey, 2000, S. 276]:

$$\widehat{VaR}_\alpha = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} \hat{z}_\alpha, \quad (81)$$

$$\widehat{ES}_\alpha = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} \hat{S}_\alpha. \quad (82)$$

## 4 Backtesting

Im vorherigen Kapitel wurden zahlreiche Schätzungsmethoden für VaR und ES vorgestellt. Ein geeignetes Modell zu selektieren, ist sowohl für die Bank als auch für die Bankenübersicht am wichtigsten, denn ein schlechtes Modell könnte das Risiko unterschätzen oder überschätzen. Auf der einen Seite droht der Bank, falls das Risiko unterschätzt wird, die Insolvenz. Darüber hinaus wird die Bank, wenn sie ein zur Unterschätzung führendes Risikomodell benutzt, durch die Bankenübersicht mit einem hohen Multiplikationsfaktor bestraft [Kerkhof and Melnberg, 2004, S. 2]. Auf der anderen Seite verliert die Bank, wenn sie das Risiko überschätzt, Potential, mehr Rendite zu erbringen.

Dazu kommen die Fragen: Welche Methode bietet angemessene Prognosen? Wie definiert man die Güte der Prognose? Gibt es ein bestimmtes Kriterium dazu? Um solche Fragen zu beantworten, wird das Backtesting zur Anwendung gebracht. In diesem Kapitel sind unterschiedliche Möglichkeiten des Backtestings zur Evaluierung der Güte der VaR- und ES-Prognose zu beschreiben.

Allerdings erfordert das Backtestingverfahren viele Beobachtungen, da Verletzungen von Prognosen nicht häufig auftreten [Daniëlsson, 2011, S. 143], somit wird am Anfang dieses Kapitels die Definition von Zeitfenster erklärt:

**Definition 7 (Zeitfenster)** *Das Zeitfenster  $w$  ist die Anzahl der sukzessiven Beobachtungen, die für die Schätzung zur Verfügung stehen.*

Das heißt, um das Risikomaß zum Zeitpunkt  $t + 1$  zu prognostizieren, werden die Verluste  $l_t, l_{t-1}, \dots, l_{t-w+1}$  verwendet. Gegeben Verluste  $\{l\}_{t=1}^T$  soll das Zeitfenster  $w \ll T$  sein [McNeil and Frey, 2000, S. 287].

### 4.1 Backtesting des Value at Risks

Dieser Abschnitt orientiert sich im Wesentlichen an Christoffersen [1998], die einen umfassenden Evaluierungsprozess zur Intervall-Prognose entwickelte. Es handelt sich um eine Folge von den Verlusten  $\{l\}_{t=1}^T$  und eine *out-of-sample* Intervall-Prognose  $(L_{t|t-1}(p), U_{t|t-1}(p))$ , die zum Zeitpunkt  $t - 1$  gemacht wird. Dabei ist  $p$  die Überdeckungswahrscheinlichkeit (*coverage probability*) des Intervalls,  $L_{t|t-1}(p)$  und  $U_{t|t-1}(p)$  sind die Unter- bzw. die Obergrenze.

Im Rahmen der Arbeit von Christoffersen wurde der prognostische VaR als ein einseitiges Intervall definiert:  $(L_{t|t-1}(\alpha), +\infty)$ , wobei  $p = \alpha$  ist. Für die zum Zeitpunkt  $t - 1$  gemachte VaR-Prognose ist die folgende Indikatorvariable  $I_t$  definiert als:

$$I_t = \begin{cases} 1, & \text{falls } l_t \in (L_{t|t-1}(\alpha), +\infty), \\ 0, & \text{falls } l_t \notin (L_{t|t-1}(\alpha), +\infty). \end{cases} \quad (83)$$

Dabei ist zu sehen, dass  $I_t = 1$  auf eine Verletzung bzw. eine Überschreitung der Prognose hindeutet. Mit dieser Indikatorvariable kann ein Kriterium in [Christoffersen, 1998, S. 843] etabliert werden:

**Definition 8** Die VaR-Prognosen  $\{(L_{t|t-1}(\alpha), +\infty)\}_{t=1}^T$  werden als effizient definiert, wenn  $E[I_t | \mathcal{G}_{t-1}] = \alpha$  für alle  $t$  gilt, wobei  $\mathcal{G}_{t-1}$  die verfügbare Information zum Zeitpunkt  $t - 1$  bezeichnet.

Das heißt, dass die Verletzungswahrscheinlichkeit eines effizienten VaR-Modells gleich dem Signifikanzniveau  $\alpha$  sein soll. Unter dieser Bedingung ergibt sich die Verteilung der Indikatorvariable  $I_t$ :

$$\{I_t\} \stackrel{\text{i.i.d.}}{\sim} \text{Bern}(\alpha), \forall t. \quad (84)$$

Das bedeutet, falls  $E[I_t | \mathcal{G}_{t-1}] = E[I_t | I_1, I_2, \dots, I_{t-1}] = \alpha$ , soll die Indikatorvariable  $I_t$  identisch unabhängig bernoulliverteilt mit dem Parametern  $\alpha$  sein. Der Beweis findet sich in [Christoffersen, 1998, S. 844].

Im folgenden Text werden drei auf Basis von 84 entwickelte Likelihood-Quotienten-Tests vorgestellt, die einander ergänzen.

#### 4.1.1 Unconditional Coverage Test

Beim unbedingten Überdeckungstest (*Unconditional Coverage Test*) wird getestet, ob die VaR-Prognose das Intervall richtig überdeckt. Unter der Bernoulli-Verteilungsannahme ergeben sich die Nullhypothese und die Alternativhypothese:

$$H_0 : E[I_t] = \alpha \text{ vs. } H_1 : E[I_t] \neq \alpha,$$

wobei  $t = 1, 2, \dots, T$ . Unter der Nullhypothese ist die Likelihood-Funktion gegeben durch:

$$L(\alpha | I_1, I_2, \dots, I_T) = (1 - \alpha)^{n_0} \alpha^{n_1}, \quad (85)$$

wobei  $n_0, n_1$  die Anzahlen von  $I_t = 0$  und  $I_t = 1$  sind. Unter der Alternativhypothese ist die Indikatorvariable  $I_t$  bernoulliverteilt, aber mit dem Parameter  $\pi$ , deswegen sieht die Likelihood-Funktion so aus:

$$L(\pi | I_1, I_2, \dots, I_T) = (1 - \pi)^{n_0} \pi^{n_1}. \quad (86)$$

Nach [Christoffersen, 1998, S. 845] wird die Likelihood-Quotienten-Statistik für den unbedingten Überdeckungstest wie folgt formuliert:

$$\text{LQ}_{\text{UC}} = 2[\log(L(\hat{\pi} | I_1, I_2, \dots, I_T) - \log(L(\alpha | I_1, I_2, \dots, I_T))] \stackrel{\text{asy}}{\sim} \chi^2(1). \quad (87)$$

wobei  $\hat{\pi} = n_1/(n_0 + n_1)$  der ML-Schätzer von  $\pi$  ist. Da die LQ-Statistik asymptotisch  $\chi^2$ -verteilt mit dem Freiheitsgrad 1 ist, lehnt man bei der Entscheidung für den Test die Nullhypothese mit dem Signifikanzniveau  $\alpha_1$  ab, falls die LQ-Statistik das Quantil  $\chi_{1-\alpha_1}^2(1)$  überschreitet<sup>9</sup>.

Zu diesem Test wies Christoffersen aber darauf hin, dass die zeitabhängige Struktur von  $I_t$  nicht berücksichtigt ist, nur die Anzahlen von  $n_0$  und  $n_1$  haben einen Einfluss auf das Ergebnis des Tests. Dies könnte dazu führen, dass die Nullhypothese, selbst wenn das Auftreten von Verletzungen zu unterschiedlichen Zeitpunkten voneinander abhängig ist, trotzdem beibehalten wird. Deswegen ist der Test der Unabhängigkeit als Ergänzung notwendig.

#### 4.1.2 Independence Test

Beim Unabhängigkeitstest (*Independence Test*) wird die zeitliche Unabhängigkeit der Indikatorvariablen  $\{I_t\}_{t=1}^T$  getestet, dazu wird die binäre Markovkette mit der Übergangsmatrix  $\Pi_1$  in die Untersuchung einbezogen:

$$\Pi_1 = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}, \quad (88)$$

wobei die Übergangswahrscheinlichkeit  $\pi_{ij} = P(I_t = j | I_{t-1} = i)$  mit  $i = 0, 1$  und  $j = 0, 1$  ist. Die Übergangsmatrix  $\Pi_1$  impliziert, dass die Übergangswahrscheinlichkeit vom jetzigen Zustand abhängig ist. Daraus folgt die Likelihood-Funktion unter der Abhängigkeit:

$$L(\Pi_1 | I_1, I_2, \dots, I_T) = (1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}}, \quad (89)$$

wobei  $n_{ij}$  die Anzahl der Übergänge von  $I = i$  auf  $I = j$  ist. Mithilfe der MLE-Methode bekommt man den Schätzer von  $\Pi_1$ :

$$\hat{\Pi}_1 = \begin{bmatrix} \frac{n_{00}}{n_{00} + n_{01}} & \frac{n_{01}}{n_{00} + n_{01}} \\ \frac{n_{10}}{n_{10} + n_{11}} & \frac{n_{11}}{n_{10} + n_{11}} \end{bmatrix}, \quad (90)$$

Unter der Nullhypothese der Unabhängigkeit gibt es die folgende Übergangsmatrix  $\Pi_0$ :

$$\Pi_0 = \begin{bmatrix} 1 - \pi_0 & \pi_0 \\ 1 - \pi_0 & \pi_0 \end{bmatrix}, \quad (91)$$

wobei die Übergangswahrscheinlichkeit nicht vom jetzigen Zustand abhängig ist, deswegen

---

<sup>9</sup>Hierbei wird das Signifikanzniveau vom Test als  $\alpha_1$  bezeichnet, damit es nicht mit dem Signifikanzniveau vom VaR verwechselt wird.

werden  $\pi_{01}$  und  $\pi_{11}$  hierbei als  $\pi_0$  zusammengefasst und lassen sich durch den ML-Schätzer  $\hat{\pi}_0 = (n_{01} + n_{11}) / (n_{01} + n_{11} + n_{00} + n_{10})$  schätzen. Danach ergibt sich die auf der Nullhypothese basierende Likelihood-Funktion:

$$L(\Pi_0 \mid I_1, I_2, \dots, I_T) = (1 - \pi_0)^{n_{00} + n_{10}} \pi_0^{n_{01} + n_{11}} \quad (92)$$

Nach [Christoffersen, 1998, S. 846] wird die LQ-Statistik für den Unabhängigkeitstest wie folgt formuliert:

$$LQ_{\text{Ind}} = 2[\log(L(\hat{\Pi}_1 \mid I_1, I_2, \dots, I_T) - \log(L(\hat{\Pi}_0 \mid I_1, I_2, \dots, I_T))] \stackrel{\text{asy}}{\approx} \chi^2(1). \quad (93)$$

Diese Teststatistik befolgt wiederum die  $\chi^2$ -Verteilung mit dem Freiheitsgrad 1. Bei einem Signifikanzniveau von  $\alpha_1$  wird die Folge  $I_t$  als abhängig beurteilt, falls die LQ-Statistik das Quantil  $\chi_{1-\alpha_1}^2(1)$  überschreitet.

Beim Unabhängigkeitstest spielt die Überdeckungswahrscheinlichkeit  $\alpha$  aber keine Rolle, nur die Abhängigkeit von  $I_t$  hat Einfluss. Deshalb hat Christoffersen einen weiteren Test aufgestellt, der es ermöglicht, die Überdeckung und die Abhängigkeit gleichzeitig zu testen.

### 4.1.3 Conditional Coverage Test

Beim bedingten Überdeckungstest (*Conditional Coverage Test*) werden die oben genannten beiden Tests kombiniert. Dabei wird die Nullhypothese des unbedingten Überdeckungstests gegen die Alternativhypothese des Unabhängigkeitstests getestet. Aus 85 und 89 wird die Likelihood-Funktion vom bedingten Überdeckungstest wie folgt formuliert:

$$LQ_{\text{CC}} = 2[\log(L(\hat{\Pi}_1 \mid I_1, I_2, \dots, I_T) - \log(L(\alpha \mid I_1, I_2, \dots, I_T))] \stackrel{\text{asy}}{\approx} \chi^2(2). \quad (94)$$

Der Beweis findet sich in [Christoffersen, 1998, S. 861]. Wenn die Teststatistik  $LQ_{\text{CC}}$  das Quantil  $\chi_{1-\alpha_1}^2(2)$  nicht überschreitet, wird die Schlussfolgerung gezogen, dass die VaR-Prognosen sowohl voneinander unabhängig als auch richtig überdeckt sind. Zu beachten ist, dass in den Abschnitten 4.1.1 und 4.1.2 die Übergangswahrscheinlichkeiten durch  $\hat{\pi} = n_1 / (n_0 + n_1)$  und  $\hat{\pi}_0 = (n_{01} + n_{11}) / (n_{01} + n_{11} + n_{00} + n_{10})$  geschätzt werden. Christoffersen hat herausgefunden, dass, wenn man die erste Indikatorvariable  $I_1$  ignoriert, gilt  $\hat{\pi} = \hat{\pi}_0$ . Aus diesem Grund gilt der folgende Zusammenhang zwischen den drei Teststatistiken [Christoffersen, 1998, S. 847]:

$$LQ_{\text{CC}} = LQ_{\text{Ind}} + LQ_{\text{UC}}. \quad (95)$$

Trotz der Gleichung 95 ist es doch empfehlenswert, die zwei untergeordneten Tests durchzuführen. Denn der kombinierte Test hat eine geringere Power, wenn die VaR-Prognosen voneinander abhängig oder nicht richtig überdeckt sind [Daniélsson, 2011, S. 159]. Wenn die kombinierte Teststatistik  $LQ_{\text{CC}}$  so groß ist, dass man die Nullhypothese ablehnen muss, weiß

man außerdem nicht, welche untergeordnete Teststatistik dazu beigetragen hat.

## 4.2 Backtesting des Expected Shortfalls

In Kapitel 2.2 wird erläutert, dass der ES im Vergleich zum VaR bessere mathematische Eigenschaften hat. Im Jahr 2013 hat der Basler Ausschuss (*Basel Committee on Banking Supervision*, kurz: BCBS) deshalb eine Reform geplant, in der der  $VaR_{0,99}$  durch den  $ES_{0,975}$  ersetzt werden sollte [Committee et al., 2013, S. 3]. Allerdings ist der VaR heutzutage noch das beliebteste Risikomaß im Finanzbereich und das Backtestingverfahren wird meist anhand des VaR vorgenommen [Committee et al., 2013, S. 27], da der ES im Vergleich zum VaR, für dessen Test nur die Verletzungsindikatorfunktion benötigt wird, schwieriger backzutesten ist. Während der VaR tatsächlich ein Quantil ist, ist der ES ein Erwartungswert, ein Backtesting des Erwartungswerts ist wenig intuitiv und erfordert mehr Beobachtungen. Die theoretische Rechtfertigung gegen das Backtesting des ES wurde von Gneiting [2011] gegeben: Zunächst hat Gneiting die Definition der Elizitabilität (*Elicitability*) aufgestellt, einfach ausgedrückt [Acerbi and Szekely, 2014, S. 1]:

**Definition 9** *Eine Statistik  $Y$  der Zufallsvariable  $X$  wird als elicitable definiert, wenn sie den Erwartungswert einer Score Funktion  $S(\cdot)$  minimiert:*

$$Y = \operatorname{argmin}_y E[S(y, X)]. \quad (96)$$

Es gibt exemplarische *elicitable* Statistiken, z.B. der Mittelwert  $\bar{y} = \operatorname{argmin}_y E[(y - X)^2]$  und der Medianwert  $\hat{y} = \operatorname{argmin}_y E[|y - X|]$  [Acerbi and Szekely, 2014, S. 2]. Gneiting [2011] hat auch bewiesen, dass der  $VaR_\alpha$  als  $(1 - \alpha)$ -Quantil *elicitable* ist, mit der Score Funktion  $S(y, X) = (\mathbf{1}(X \geq y) - \alpha)(X - y)$ . Der ES hat aber keine Elizitabilität. (Der Beweis kann in [Gneiting, 2011, S. 23] nachgelesen werden.) Der Mangel an Elizitabilität erschwert das Backtestingverfahren [Carver, 2013].

Jedoch zeigten nachfolgende Studien, dass das Backtesting des ES doch möglich ist. [Kerkhof and Melenberg, 2004, S. 24] haben sogar behauptet, dass das Backtesting des ES im Vergleich zum VaR eine höhere Power hat. Acerbi und Szekely haben dafür plädiert, dass die Elizitabilität gar nichts mit dem Backtesting zu tun hat. Als Gegenargument führen sie an, dass der VaR niemals mithilfe der Score Funktion backgetestet worden ist [Acerbi and Szekely, 2014, S. 2]. Dazu haben sie drei nonparametrische Tests des ES entwickelt.

### 4.2.1 Acerbi und Szekelys Test 1: Testen ES nach VaR

Der erste Test von Acerbi und Szekely basiert auf der bedingten Verlusterwartung. Aus 3 folgt:

$$ES_{\alpha,t} = E[L_t | L_t - VaR_{\alpha,t} > 0]. \quad (97)$$

Danach ergibt sich die folgende Gleichung [Acerbi and Szekely, 2014, S. 3]:

$$E \left[ \frac{L_t}{ES_{\alpha,t}} - 1 \mid L_t - VaR_{\alpha,t} > 0 \right] = 0. \quad (98)$$

Angenommen, dass der VaR bereits getestet worden ist, dann ist mit der gegebenen Folge von Verlusten  $\{L\}_{t=1}^T$  die Teststatistik definiert durch [Acerbi and Szekely, 2014, S. 3]:

$$Z_1 = \frac{\sum_{t=1}^T \frac{L_t I_t}{ES_{\alpha,t}}}{N_T} - 1, \quad (99)$$

wobei  $N_T = \sum_{t=1}^T I_t > 0$  ist.  $I_t$  ist die Indikatorfunktion aus 83. Die entsprechenden Hypothesen wurden von [Acerbi and Szekely, 2014, S. 3] wie folgt formuliert:

$$\begin{aligned} H_0 : P_t^{[\alpha]} &= F_t^{[\alpha]}, \forall t; \\ H_1 : ES_{\alpha,t}^F &\geq ES_{\alpha,t}, \forall t \text{ und } > \text{ f\"ur einige } t, \\ VaR_{\alpha,t}^F &= VaR_{\alpha,t}, \forall t, \end{aligned} \quad (100)$$

wobei  $F_t$  die zugrunde liegende Verteilung des Verlusts ist.  $P_t$  ist die prognostische Verteilung des Verlusts.  $P_t^{[\alpha]}$  bezeichnet den Schwanz der Verteilung  $P_t$ :  $P_t^{[\alpha]}(L) = \min(1, (1 - P_t(L))/\alpha)$  für  $L > VaR_{\alpha,t}$ . Die Nullhypothese besagt, dass die Verteilung beim Schwanz richtig modelliert ist. Die Alternativhypothese berücksichtigt nur den Fall der Unterschätzung, somit ist dieser Test einseitig. Zu beachten ist, dass im Falle der Ablehnung der  $H_0$  der VaR noch als richtig gesehen wird, weil dieser Test voraussetzt, dass die Modellierung das VaR-Backtesting im Voraus bestanden hat [Acerbi and Szekely, 2014, S. 4].

Unter diesen Bedingungen haben wir die folgende Gleichungen, die zur Testentscheidung dienen:

$$E_{H_0}[Z_1 \mid N_T > 0] = 0 \text{ und } E_{H_1}[Z_1 \mid N_T > 0] > 0. \quad (101)$$

Der Beweis findet sich in [Roccioletti, 2015, S. 61]. Gemäß 101 soll der Erwartungswert der Teststatistik  $Z_1$  gleich Null sein, falls das Modell den ES richtig prognostiziert. Ein positiver Wert weist auf Probleme hin.

Es gibt auch andere Variationen dieses Tests, z.B. McNeil and Frey [2000] entwickelten einen auf dem Bootstrap basierten Test vor dem Hintergrund des GARCH-Modells. Dabei wurden die Residuen definiert als:

$$r_t = \frac{l_t - \widehat{ES}_{\alpha,t}}{\hat{\sigma}_t}, \quad (102)$$

wobei  $\hat{\sigma}_t$  die Standardabweichung zum Zeitpunkt  $t$  ist, deren Schätzung durch die PML-Methode erfolgt (Siehe Kapitel 3.2.5). Das Testprinzip ist, dass, wenn der ES richtig mo-

delliert wird, der Erwartungswert der Residuen gleich Null sein soll:  $E_{H_0}[r_t] = 0$ . Außerdem haben Mcneil und Frey durch die Detektion der Grafik der Residuen festgestellt, dass die Annahme der bedingten Normalverteilung des GARCH-Modells für die Prognose des ES „nutzlos“ ist [McNeil and Frey, 2000, S. 294]. Zu diesem Test haben Bayer und Dimitriadis darauf hingewiesen, dass bei diesem Test sowohl der ES und der VaR als auch die Volatilität gleichzeitig getestet werden. Das bedeutet, dieser Test ist ein gemeinsamer Test (*joint test*) und eine Ablehnung der  $H_0$  des gemeinsamen Tests muss nicht zur Schlussfolgerung kommen, dass ausschließlich das Modell des ES misspezifiziert ist. Die Ablehnung könnte auch auf andere Faktoren zurückzuführen sein [Bayer and Dimitriadis, 2018].

#### 4.2.2 Acerbi und Szekelys Test 2: Testen ES unmittelbar

Der zweite Test von Acerbi und Szekely basiert auf der unbedingten Erwartung [Acerbi and Szekely, 2014, S. 4]:

$$ES_{\alpha,t} = E \left[ \frac{L_t I_t}{\alpha} \right]. \quad (103)$$

Daraus folgt die Teststatistik des zweiten Tests [Acerbi and Szekely, 2014, S. 4]:

$$Z_2 = \sum_{t=1}^T \frac{L_t I_t}{\alpha T ES_{\alpha,t}} - 1. \quad (104)$$

Die entsprechenden Hypothesen sind ähnlich wie 100 [Acerbi and Szekely, 2014, S. 4]:

$$\begin{aligned} H_0 : P_t^{[\alpha]} &= F_t^{[\alpha]}, \forall t; \\ H_1 : ES_{\alpha,t}^F &\geq ES_{\alpha,t}, \forall t \text{ und } > \text{ für einige } t, \\ VaR_{\alpha,t}^F &\geq VaR_{\alpha,t}, \forall t, \end{aligned}$$

aber bei der  $H_1$  muss der VaR nicht richtig modelliert werden. Wie bei 101 gibt es die folgenden Gleichungen:

$$E_{H_0}[Z_2] = 0 \text{ und } E_{H_1}[Z_2] > 0. \quad (105)$$

Der Beweis findet sich in [Roccioletti, 2015, S. 63]. Aus 99 und 104 gibt es den folgenden Zusammenhang zwischen  $Z_1$  und  $Z_2$  [Acerbi and Szekely, 2014, S. 4]:

$$Z_2 = (1 + Z_1) \frac{N_T}{\alpha T} - 1. \quad (106)$$

#### 4.2.3 Acerbi und Szekelys Test 3: Schätzen ES erneut

Der dritte Test von Acerbi und Szekely wurde durch die Idee von Berkowitz [2001] inspiriert, dass, wenn die prognostische Verteilung die Richtige ist, der Funktionswert der kumulierten

Verteilungsfunktion  $P_t$  im Intervall  $(0, 1)$  identisch unabhängig uniform verteilt sein soll. Dabei werden die Funktionswerte  $U_t = P_t(L_t)$  als *realized ranks* beschrieben [Acerbi and Szekely, 2014, S. 5]. Es gilt:

$$U_t \stackrel{\text{i.i.d.}}{\sim} U(0, 1), \forall t. \quad (107)$$

Um den Test zu ermöglichen, haben Acerbi und Szekely einen auf den i.i.d. Zufallsstichproben  $\vec{Y} = \{Y_t\}$  basierenden Schätzer für den  $ES_\alpha$  konstruiert:

$$\widehat{ES}_\alpha^{(N)}(\vec{Y}) = \frac{1}{N\alpha} \sum_i^{[N\alpha]} Y_{i:N}, \quad (108)$$

wobei  $Y_{i:N}$  die Ordnungsstatistiken bezeichnet.  $[x]$  ist der Integerteil von  $x$ .

Vor diesem Hintergrund ist die dritte Teststatistik gegeben durch [Acerbi and Szekely, 2014, S. 5]:

$$Z_3 = \frac{1}{T} \sum_{t=1}^T \frac{\widehat{ES}_\alpha^{(T)}(P_t^{-1}(\vec{U}))}{E_V \left[ \widehat{ES}_\alpha^{(T)}(P_t^{-1}(\vec{V})) \right]} - 1, \quad (109)$$

wobei der Vektor  $\vec{V}$  i.i.d.  $U(0, 1)$  ist.

Aber für diesen Test werden die ganzen Verteilungen des Verlusts in die Hypothesen einbezogen [Acerbi and Szekely, 2014, S. 5]:

$$\begin{aligned} H_0 : P_t &= F_t, \forall t; \\ H_1 : P_t &\succeq F_t, \forall t \text{ und } \succ \text{ für einige } t, \end{aligned}$$

wobei  $\succeq$  die stochastische Dominanz erster Ordnung bezeichnet<sup>10</sup>.

In diesem Fall gibt es auch die folgenden Gleichungen der Teststatistik:

$$E_{H_0}[Z_3] = 0 \text{ und } E_{H_1}[Z_3] > 0. \quad (110)$$

Der Beweis findet sich in [Roccioletti, 2015, S. 67].

Zu diesem Test haben Acerbi und Szekely folgendes behauptet: Im Vergleich zu Test 1 und Test 2 ist der dritte Test wenig intuitiv, aber der Test 3 ist sehr allgemein [Acerbi and Szekely, 2014, S. 5].

#### 4.2.4 P-Werts der Acerbi und Szekelys Tests

Die Verteilungen der Teststatistiken  $Z_1, Z_2$  und  $Z_3$  sind nicht bekannt, deswegen ist der p-Wert durch eine Monte-Carlo-Simulation zu schätzen [Acerbi and Szekely, 2014]. Dabei

<sup>10</sup>Die Verteilung A hat die stochastische Dominanz erster Ordnung (*first-order stochastic dominance*) gegenüber der Verteilung B, wenn  $F_A(x) \leq F_B(x), \forall x$ .

wird angenommen, dass  $H_0$  richtig ist. Dann ist die Simulation durch die folgenden Schritte vorzunehmen [Acerbi and Szekely, 2014, S. 6]:

- Ziehe  $L_t^i$  aus der prognostischen Verteilung  $P_t$ , wobei  $L_t^i \sim P_t, \forall t, \forall i = 1, 2, \dots, M$ .
- Berechne die Teststatistik für jede Wiederholung:  $Z^i = Z(\vec{L}^i)$ .
- Schätze den p-Wert durch:  $p = \sum_{i=1}^M (\mathbf{1}(Z^i > Z(\vec{l}))) / M$ .

Dabei ist  $M$  die Anzahl der Wiederholungen, die groß genug sein soll.  $\vec{l}$  ist der Vektor der realisierten Verluste. Bei einem gegebenen Signifikanzniveau von  $\alpha_1$  wird die Nullhypothese abgelehnt, falls  $p < \alpha_1$ . Bei diesem Test ist zu sehen, dass der ES-Test im Vergleich zum Backtesting des VaR, bei dem nur die Verletzungsfunktion der Prognosen genügt, zusätzlich die Speicherung der prognostischen Verteilungen  $P_t$  erfordert. Allerdings haben Acerbi und Szekely durch eine Simulationsstudie<sup>11</sup> herausgefunden, dass die Statistik  $Z_2$  einen stabilen kritischen Wert für einen breiten Bereich des Tail-Index aufweist, wobei  $Z_{2*} = 0.70$  und  $Z_{2*} = 1.80$  für die Signifikanzniveaus 0.05 und 0.0001<sup>12</sup>, deshalb hielten sie die Speicherung der Verteilung beim zweiten Test für nicht nötig [Acerbi and Szekely, 2014, S. 9]. Des Weiteren hat [Roccioletti, 2015, S. 68] behauptet, dass der zweite Test die größte Power im Vergleich zu den anderen Tests hat. Während der erste Test die richtige VaR-Prognose voraussetzt und der dritte Test sich als ein ergänzender Test eignet, beurteilten Acerbi und Szekely den zweiten Test als selbstständig funktionierend [Acerbi and Szekely, 2014, S. 12]. Aus solchen Gründen wird der zweite Test gegenüber den anderen zwei Tests bevorzugt.

#### 4.2.5 Regression Based Expected Shortfall Backtesting

Der sogenannte *Regression Based Expected Shortfall* (kurz: ESR) Backtest wurden von Bayer und Dimitriadis im Jahr 2018 vorgeschlagen. Um diesen neuartigen Test durchzuführen, benötigt man ausschließlich den Prognosewert des ES. Dabei wird getestet, ob die Folge der Prognosewerte des ES  $\{\hat{e}_t, t = 1, \dots, T\}$  bezüglich der realisierten Renditen  $\{R_t, t = 1, \dots, T\}$  richtig spezifiziert ist [Bayer and Dimitriadis, 2018]. Von diesem Test gibt es zwei Variationen: den bivariate ESR-Test und den Intercept-ESR-Test.

**Bivariater ESR-Test** Bei diesem Test werden die realisierten Renditen  $\{R_t, t = 1, \dots, T\}$  auf die Prognosewerte des ES  $\{\hat{e}_t, t = 1, \dots, T\}$  regressiert [Bayer and Dimitriadis, 2018,

<sup>11</sup>Diese Simulationsstudie wurden auf Basis von der Studentischen t-Verteilung mit umfangreichen Freiheitsgraden  $\nu$  durchgeführt, weil durch Einteilung des Freiheitsgrads  $\nu$  unterschiedliche Tail-Verhalten simuliert werden können.

<sup>12</sup>In der Arbeit von Acerbi und Szekely bezog sich der Test auf der Rendite statt des Verlusts, deswegen waren die kritischen Werte  $Z_{2*} = -0.70$  und  $Z_{2*} = -1.80$  für die Signifikanzniveaus 0.05 und 0.0001 in ihrer Arbeit.

S. 4]:

$$R_t = a + b\hat{e}_t + \hat{u}_t^e. \quad (111)$$

Zu beachten ist, dass diese Gleichung für das ES-Funktional gilt, deswegen ist sie anders als die gewöhnliche Regression, bei der die Erwartungen von der erklärenden Variable und der zu erklärenden Variable auf beiden Seiten stehen. Daraus entstehen die folgenden auf dem ES basierenden Gleichungen:

$$ES_\alpha[\hat{u}_t^e | \mathcal{G}_{t-1}] = 0, \quad (112)$$

$$ES_\alpha[Y_t | \mathcal{G}_{t-1}] = a + b\hat{e}_t, \quad (113)$$

wobei  $\mathcal{G}_{t-1}$  die verfügbare Information zum Zeitpunkt  $t - 1$  bezeichnet. Wenn die Prognose des ES richtig ist, soll der Intercept  $a = 0$  und die Steigung  $b = 1$  sein [Bayer and Dimitriadis, 2018, S. 2]. Aus 113 gibt es:  $\hat{e}_t = ES_\alpha[Y_t | \mathcal{G}_{t-1}]$ . Im Anschluss daran werden die folgenden Hypothesen bei diesem Test getestet:

$$H_0 : (a, b) = (0, 1) \text{ vs. } H_1 : (a, b) \neq (0, 1).$$

Da dieser Test die Parameter  $(a, b)$  gleichzeitig testet, wird er als bivariater ESR-Test bezeichnet [Bayer and Dimitriadis, 2018, S. 5]. Wegen der Nicht-Elizitabilität des ES können die Parameter  $(a, b)$  nicht direkt erhalten werden [Dimitriadis and Bayer, 2017]. Durch eine neue Regression-Technik, in der die Regressionsgleichung des Quantils mit der Regressionsgleichung des ES kombiniert wird, sind die Parameter semiparametrisch zu schätzen (vgl. [Bayer and Dimitriadis, 2018, S. 18] und [Dimitriadis and Bayer, 2017]). Danach ist eine Wald-Statistik zu formulieren [Bayer and Dimitriadis, 2018, S. 5]:

$$T_{ESR} = \left( (\hat{a}, \hat{b})^T - (0, 1)^T \right)^T \hat{\Sigma}_{ES}^{-1} \left( (\hat{a}, \hat{b})^T - (0, 1)^T \right)^T, \quad (114)$$

wobei  $\hat{\Sigma}_{ES}$  die geschätzte asymptotische Kovarianzmatrix der Parameter  $(a, b)$  ist. [Bayer and Dimitriadis, 2018, S. 5] haben bewiesen, dass diese Teststatistik die  $\chi^2$ -Verteilung mit dem Freiheitsgrad 2 befolgt:

$$T_{ESR} \stackrel{\text{asy}}{\sim} \chi^2(2). \quad (115)$$

Dieser Test kann mithilfe von Rs Funktion `esr_backtest()` aus dem `esback` Paket (Version: 0.1.1) durchgeführt werden.

**Intercept-ESR-Test** Der im vorherigen Text vorgestellte bivariate ESR-Test lässt nur zweiseitige Hypothesen zu, weil der Einfluss der Fehlprognose des ES auf die Parameter  $(a, b)$  unklar ist [Bayer and Dimitriadis, 2018, S. 6]. Jedoch tendieren die Finanzinstitute oft dazu, das Risiko zu unterschätzen, um wenig Kapitalbedarf zu veranschlagen. Deshalb

hielten Bayer und Dimitriadis es für ausreichend, nur die Unterschätzung des Risikos zu bestrafen. Dazu wurde der sogenannte Intercept ESR-Test, der die einseitigen Hypothesen auch erlaubt, aufgestellt.

Dieser Test basiert wieder auf der Regressionsgleichung 111, aber hierbei wird die Steigung  $b$  auf 1 festgesetzt und die Regression auf diese Weise vorgenommen [Bayer and Dimitriadis, 2018, S. 6]:

$$R_t - \hat{e}_t = a + \hat{u}_t^e. \quad (116)$$

Dabei wird getestet, ob der Intercept  $a = 0$  ist. Dann sind die entsprechenden einseitigen Hypothesen möglich:

$$H_0 : a = 0 \text{ vs. } H_1 : a < 0,$$

und die zweiseitigen Hypothesen sind:

$$H_0 : a = 0 \text{ vs. } H_1 : a \neq 0,$$

Im Anschluss daran wird ein auf der asymptotischen Kovarianz basierender t-Test entwickelt (Details dazu finden sich in [Bayer and Dimitriadis, 2018, S. 6]).

Dieser Test kann mithilfe von Rs Funktion `esr_backtest_intercept()` aus dem `esback` Paket (Version: 0.1.1) durchgeführt werden.

#### 4.2.6 Weitere Tests des ES

Eine empirische Vorgehensweise zum ES-Backtesting, die auf der Approximation einiger Quantile basiert, wurde von [Emmer et al., 2015] vorgeschlagen. Aus 4 ist der  $ES_\alpha$  durch die VaR-Prognose mit vier unterschiedlichen Signifikanzniveaus zu approximieren:

$$\begin{aligned} ES_\alpha &= \alpha^{-1} \int_0^\alpha VaR_u \, du \\ &\approx \frac{1}{4} [VaR_\alpha + VaR_{0.75\alpha} + VaR_{0.5\alpha} + VaR_{0.25\alpha}]. \end{aligned} \quad (117)$$

Beispielsweise wird  $ES_{0.05}$  durch den Mittelwert von  $VaR_{0.05}$ ,  $VaR_{0.0375}$  und  $VaR_{0.025}$  sowie  $VaR_{0.0125}$  approximiert. In diesem Fall wird die Prognose von  $ES_{0.05}$  nur akzeptiert, wenn der VaR in vier bestimmten Signifikanzniveaus erfolgreich prognostiziert wird. Im Vergleich zu den Acerbi und Szekelys nonparametrischen Tests erfordert diese Methode keine Monte-Carlo-Simulation und wegen der Einfachheit ist sie leicht zu implementieren [Emmer et al., 2015, S. 20]. Offensichtlich setzt das Verfahren richtiges VaR-Modell voraus.

Es gibt auch weitere Testmöglichkeiten, z.B. die *Conditional Calibration Backtests* von Nolde et al. [2017]. Allerdings ist dieser Test ähnlich wie der Mcneil und Freys Test (siehe 4.2.1) auch ein gemeinsamer Test. Als ein weiterer Nachteil des gemeinsamen Test führten Bayer und Dimitriadis an, dass nicht alle dazu benötigten Informationen bei der Aufsichts-

stelle eingereicht werden, wodurch der gemeinsame Test in der Praxis nicht einfach implementiert werden kann [Bayer and Dimitriadis, 2018, S.2].

## 5 Anwendungsteil

Im diesem Kapitel werden die vorgestellten Methoden auf die empirischen Daten angewendet. Die Daten sind die Schlusskurse des DAX<sup>13</sup> vom 11.27.1990 bis zum 11.26.2017, wobei es insgesamt 6827 Geschäftstage gibt:

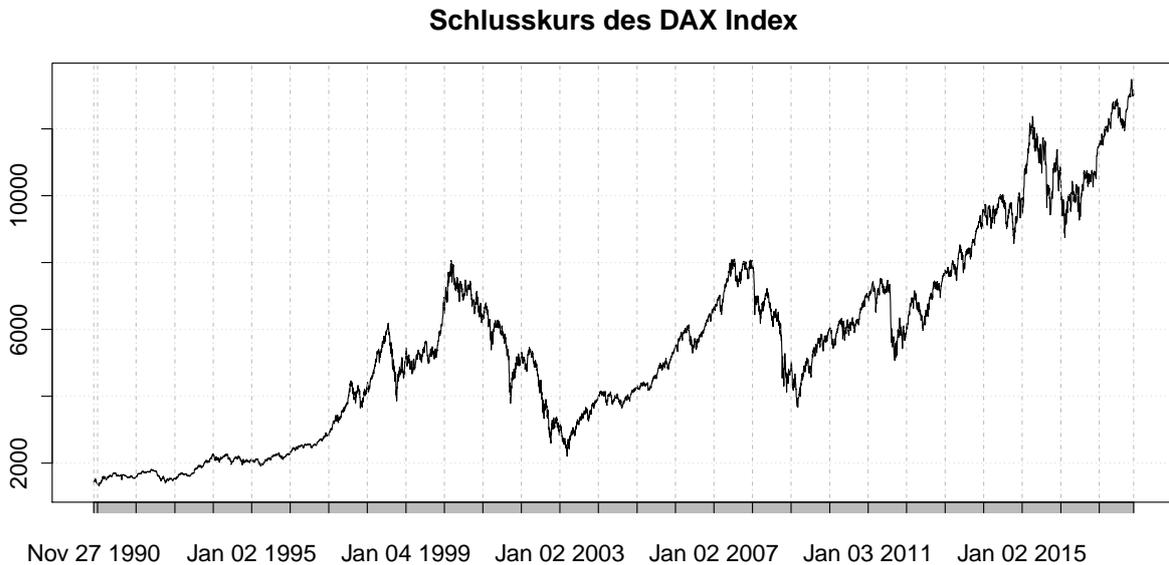


Abbildung 7: *Der Schlusskurs des DAX.*

In Abbildung 7 ist zu erkennen, dass sich die Preise immer nach oben schieben. Diese Erscheinung erschwert die Zeitreihenanalyse, deswegen werden die Renditen statt der Preise betrachtet:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (118)$$

Somit werden insgesamt 6826 Renditen gebildet. Im Anschluss daran wird der Logarithmus der Rendite definiert:

$$r_t = \log \left( \frac{P_t}{P_{t-1}} \right). \quad (119)$$

Der Wertebereich der logarithmierten Rendite ist symmetrisch und beträgt  $(-\infty, \infty)$ , wohingegen die Rohdaten der Rendite im Bereich  $[-100\%, \infty)$  liegen. Es gilt:

$$r_t = \log(1 + R_t) \approx R_t \quad \text{für} \quad R_t \ll 1. \quad (120)$$

Eines der wichtigen Ziele der Risikomaße ist, dem großen Aktiensturz vorzubeugen, deswegen wird der logarithmierte Verlust statt der logarithmierten Rendite im Folgenden Text

<sup>13</sup>Der DAX steht für den Deutschen Aktienindex, der die Wertentwicklung der 30 größten deutschen Unternehmen misst. Die Daten stammen aus der Webseite [finance.yahoo.com](http://finance.yahoo.com).

in Betracht bezogen:  $l_t = -r_t$ . Die logarithmierten Verluste sind in Abbildung 8 zu sehen: Sie weisen keine monoton steigende oder fallende Tendenz auf. Die Stationarität der neuen Zeitreihe wird auch durch den *Augmented Dickey-Fuller* (kurz: ADF)-Test überprüft, wobei der p-Wert  $< 0.01$  ist.

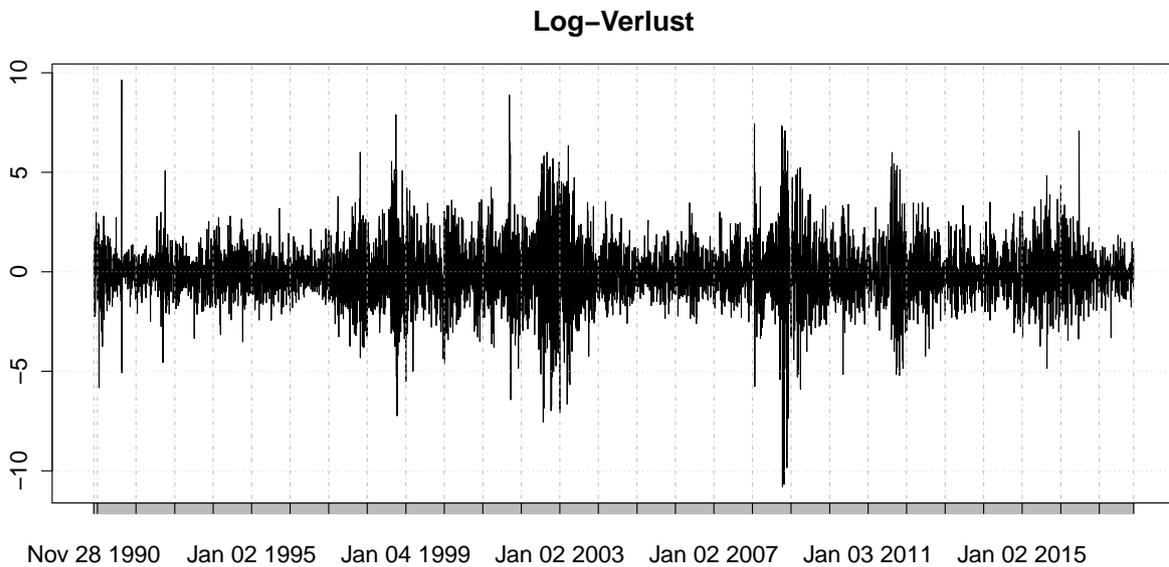


Abbildung 8: *Logarithmierter Verlust des DAX.*

Anzahl	Mittel	Median	Min.	Max.	Std.Abw.	Schiefe	Wölbung
6826	-0.0324	-0.0859	-10.7975	9.6277	1.4156	0.1354	7.7844

Tabelle 1: Deskriptive Statistiken des logarithmierten Verlusts des DAX.

Tabelle 1 gibt deskriptive Statistiken des logarithmierten Verlusts an. Die Schiefe = 0.1354 und die Wölbung = 7.7844 deuten darauf hin, dass die Daten rechtsschief und steilgipflig gegenüber der Normalverteilung sind (Siehe auch Abbildungen 1 und 2), während die Normalverteilung die Schiefe = 0 und die Wölbung = 3 hat. Die Abweichung von der Normalverteilung spiegelt sich auch beim Norm-QQ-Plot 9 wider: Ist die Normalverteilungsannahme gerechtfertigt, sollten die Datenpunkte möglichst auf der Winkelhalbierenden des Normal-QQ-Plots liegen. Am Anfang und am Ende sind jedoch große Nichtübereinstimmungen auszumachen.

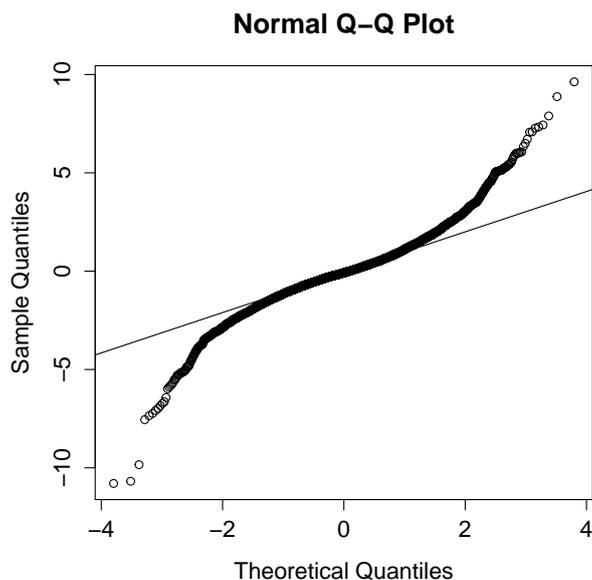


Abbildung 9: Norm-QQ-Plot des logarithmierten Verlusts des DAX.

Im Folgenden wird die Größe des Zeitfensters auf 1050 gesetzt, was ungefähr 4 jährigen Daten entspricht<sup>14</sup>. Somit werden 5776 Prognosen gemacht, die zur Untersuchung der Güte der unterschiedlichen Modelle dienen.

## 5.1 Block-Maxima-Methode

In diesem Abschnitt wird die BM-Methode benutzt, dabei ist die GEV an die Daten anzupassen. In jedem Zeitfenster gibt es 1050 Beobachtungen  $l_t, l_{t-1}, \dots, l_{t-1050}$ .

Zunächst ist die Frage zu klären, wie groß der Block sein soll. Normalerweise sind die folgenden Möglichkeiten anzubieten: ein Jahr, ein Halbjahr und ein Quartal sowie ein Monat. Theoretisch sollen die Blöcke äquidistant sein. In der Tat ist dies aber nicht der Fall, jedes Jahr und sogar jeder Monat haben unterschiedliche Anzahlen von Geschäftstagen, deswegen werden in der folgenden empirischen Analyse die Anzahlen der Geschäftstage von einem Jahr, einem Halbjahr und einem Quartal sowie einem Monat als jeweils 252, 126, 63 und 21 definiert. Danach ist der Test der Anpassungsgüte von Shermann (Siehe 3.1.5) einzusetzen, um zu bestimmen, ob die Maxima durch die GEV modelliert werden können. In Tabelle 2<sup>15</sup> sind die p-Werte der Teststatistiken  $\Omega_m$  für unterschiedliche Größen des Blocks anzugeben.<sup>16</sup>

<sup>14</sup>Die Größe des Zeitfensters ist 1050 statt 1000, weil die BM-Methode einen Anspruch an die Teilbarkeit der Größe stellt (siehe 5.1).

<sup>15</sup>In folgenden Tabellen 2-20 werden alle p-Werte  $> 0.05$  fett geschrieben.

<sup>16</sup>Bei diesem Test werden die ersten 6804 Beobachtungen betrachtet, weil 6804 durch 252 teilbar ist.  $m = 324, 108, 54, 27$  sind die Anzahlen der Blöcke.

	$\Omega_{324}$	$\Omega_{108}$	$\Omega_{54}$	$\Omega_{27}$
p-Wert	<b>0.88</b>	<b>0.78</b>	<b>0.57</b>	<b>0.66</b>

Tabelle 2: Shermanns Anpassungsgüte-Test

In Tabelle 2 sind alle p-Werte größer als 0.05. Die Ergebnisse weisen darauf hin, dass es bei einem gegebenen Signifikanzniveau von 0.05 in allen Fällen keinen ausreichenden Beweis dagegen gibt, dass die GEV an die Daten gut gefittet ist. Im Weiteren wird somit  $n = 21$  als die Größe des Blocks gewählt, damit mehr Maxima zur Verfügung stehen können. Die monatlichen Maxima sind in Abbildung 10 dazustellen.

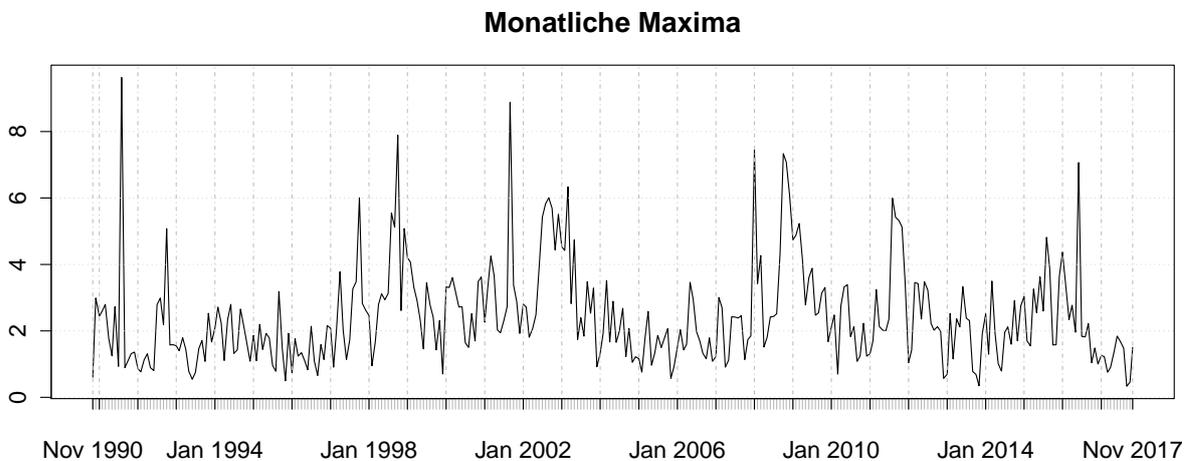


Abbildung 10: *Monatliche Maxima.*

Nachdem die monatlichen Maxima aus den Daten herausgefiltert werden, ist die GEV an die Maxima anzupassen. Die durch die MLE-Methode geschätzten Schätzwerte für die Parameter der GEV sind in Tabelle 3 aufgelistet. Dabei ist es auffällig, dass der Formparameter  $\hat{\xi} > 0$  ist. Gemäß 27 weist der positive Formparameter darauf hin, dass die monatlichen Maxima fréchetverteilt sind. Das Ergebnis deckt sich mit der Entdeckung von Longin [1996], in der sich die Extremwerte des US-Aktienmarkts auch als fréchetverteilt beweisen ließen. Die 5776 VaR- und ES-Prognosen sind in den Abbildungen 11 und 12 dargestellt. Die entsprechenden Testergebnisse sieht man in den Tabellen 4<sup>17</sup> und 5<sup>18</sup>.

$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$
1.7408	0.9703	0.0977

Tabelle 3: Schätzwerte der GEV

<sup>17</sup>Der R-Code zum Christoffersens Test beruht auf [Danielsson, 2011, S. 155,157].

<sup>18</sup>Wegen der Einheitlichkeit ist der einseitige Intercept-ESR-Test statt dem bivariaten ESR-Test anzuwenden, da die Tests von Acerbi und Szekely einseitig sind.

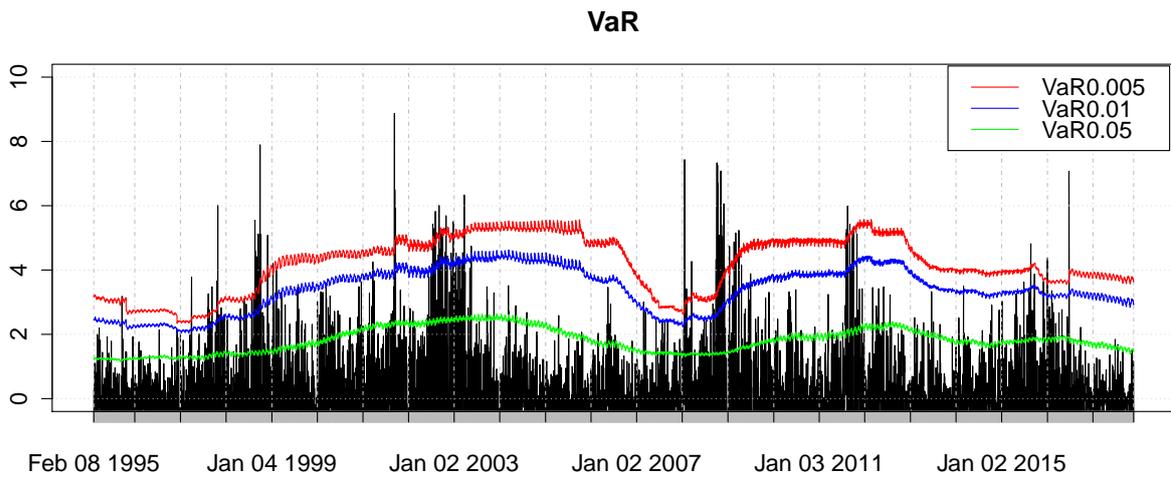


Abbildung 11: *VaR-Prognosen der BM-Methode.*

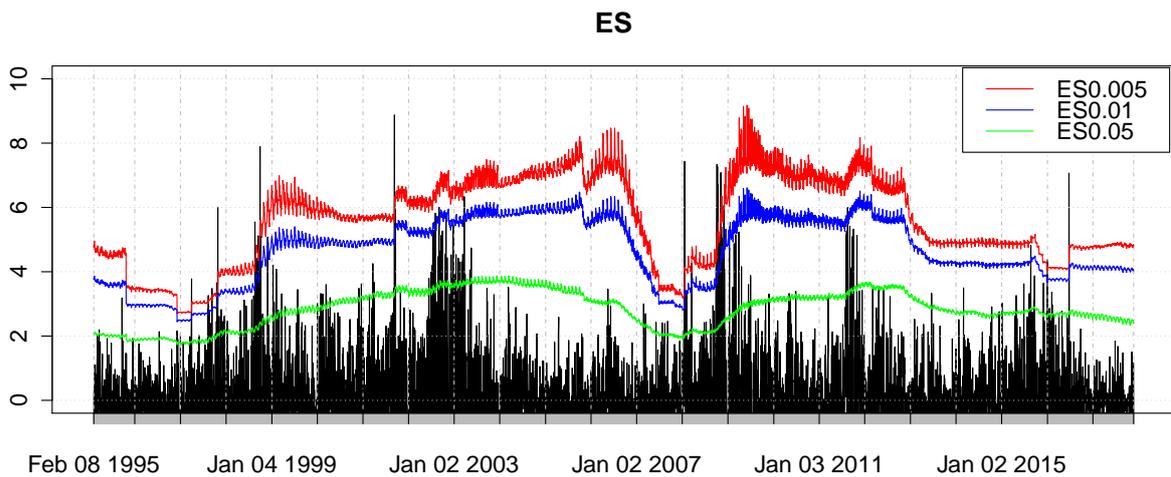


Abbildung 12: *ES-Prognosen der BM-Methode.*

$VaR_\alpha$	#Verletzung	%Verletzung	U.C.	Ind.	C.C.
$VaR_{0.05}$	505	8.74	0.00	0.00	0.00
$VaR_{0.01}$	121	2.09	0.00	0.00	0.00
$VaR_{0.005}$	63	1.09	0.00	0.04	0.00

Tabelle 4: VaR-Backtesting der BM-Methode.

$ES_\alpha$	$Z_2^{19}$	I_ESR
$ES_{0.05}$	< 0.05	0.00
$ES_{0.01}$	< 0.05	0.00
$ES_{0.005}$	< 0.05	0.00

Tabelle 5: ES-Backtesting der BM-Methode

Die Ergebnisse in den Tabellen 4 und 5 sind nicht zufriedenstellend, da keine p-Werte größer als 0.05 sind. Dies ist wohl auf das Ignorieren des Cluster-Effekts zurückzuführen. Um diesen Effekt auszugleichen, ist der extreme Index  $\theta$  einzuführen. Gemäß 3.1.4 ist der extreme Index  $\theta$  durch die BM-Methode zu schätzen, dazu wird Tabelle 6 zur Schätzung des extremen Index erstellt:

$(m, n)$	$N_u$	15	20	25	30	40	50	100	200
Quartal	$K_u$	11	14	14	15	16	17	28	58
(112,60)	$\hat{\theta}$	0.77	0.75	0.60	0.54	0.43	0.37	0.32	0.41
Halbjahr	$K_u$	11	11	11	11	12	13	21	40
(56,120)	$\hat{\theta}$	0.82	0.61	0.49	0.41	0.34	0.30	0.26	0.36

Tabelle 6: Schätzung des extremen Index mit der BM-Methode.

Ähnlich wie [McNeil, 1998, S. 12] stellen die Quartalsdaten stabilere Ergebnisse (zwischen 0.77 und 0.32 mit dem Erwartungswert 0.52) im Vergleich zu den Halbjahresdaten dar (zwischen 0.82 und 0.26 mit dem Erwartungswert 0.45). Deswegen entscheidet sich der extreme Index für den Erwartungswert der Quartalsergebnisse 0.52. Die Veranschaulichung der 5776 VaR- und ES-Prognosen findet sich in den Abbildungen 13 und 14. Die entsprechenden Testergebnisse sieht man in den Tabellen 7 und 8.

<sup>19</sup>Man bekommt die p-Werte vom Acerbi und Szekelys zweiten Test, indem man die Teststatistik  $Z_2$  mit dem festgehaltenen kritischen Wert  $Z_{2*} = 0.70$  vergleicht (Siehe 4.2.4).

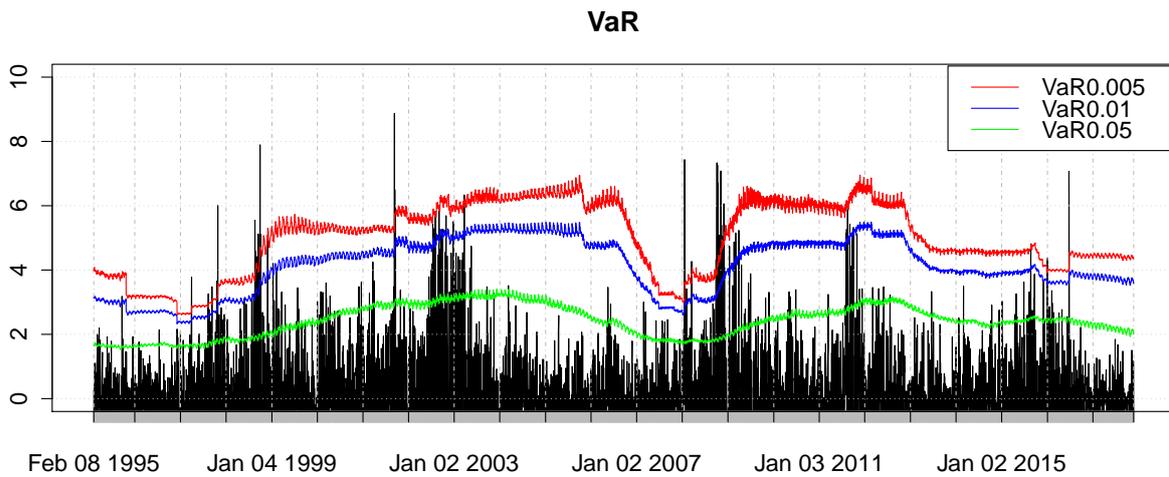


Abbildung 13: VaR-Prognosen der BM-Methode mit dem extremen Index.

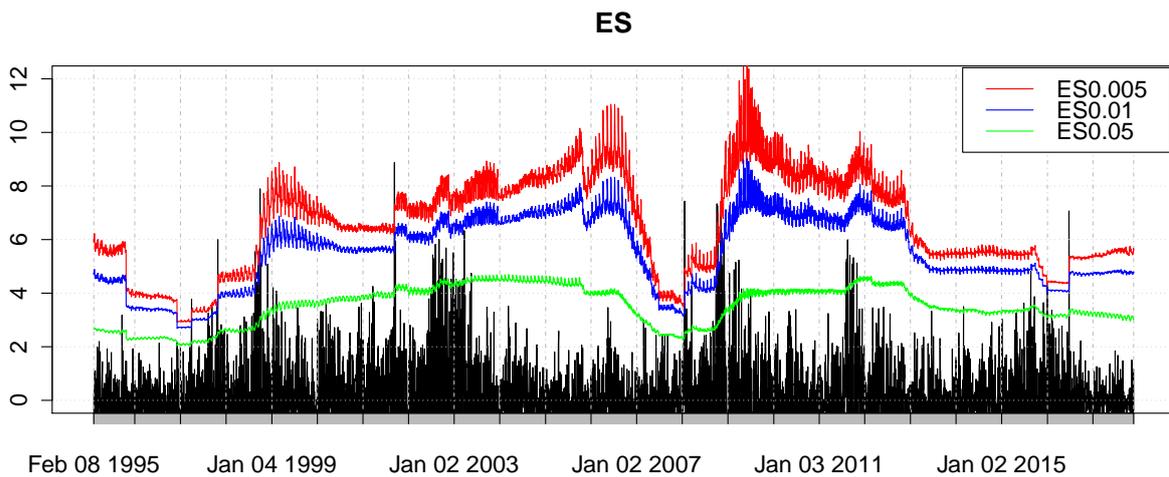


Abbildung 14: ES-Prognosen der BM-Methode mit dem extremen Index.

$VaR_\alpha$	#Verletzung	%Verletzung	U.C.	Ind.	C.C.
$VaR_{0.05}$	289	5.00	<b>0.99</b>	0.00	0.00
$VaR_{0.01}$	63	1.09	<b>0.49</b>	0.04	<b>0.09</b>
$VaR_{0.005}$	38	0.66	<b>0.10</b>	<b>0.25</b>	<b>0.14</b>

Tabelle 7: VaR-Backtesting der BM-Methode mit dem extremen Index.

$ES_\alpha$	$Z_2$	IESR
$ES_{0.05}$	> <b>0.05</b>	<b>0.16</b>
$ES_{0.01}$	> <b>0.05</b>	<b>0.14</b>
$ES_{0.005}$	> <b>0.05</b>	<b>0.12</b>

Tabelle 8: ES-Backtesting der BM-Methode mit dem extremen Index.

Aus den Tabellen 7 und 8 ist zu erkennen, dass die Ergebnisse nach der Modifikation durch den extremen Index deutlich verbessert sind: Die modifizierte BM-Methode hat den unbedingten Überdeckungstest und den ES-Test auf allen Signifikanzniveaus bestanden. Aber die Unabhängigkeit gilt nur, wenn  $\alpha = 0.005$ .

## 5.2 Peaks-over-Threshold-Methode

In diesem Abschnitt wird die POT-Methode angewendet. In Kapitel 3.2.4 ist erläutert worden, wie der Schwellenwert  $u$  bestimmt wird und beim *Mean Excess Plot* (Abbildung 5) beträgt der Schwellenwert der Daten ungefähr 3.5. Dann sind die Daten auf die GPD anzupassen. Die durch die MLE-Methode geschätzten Schätzwerte für die Parameter  $\hat{\beta}, \hat{\xi}$  sind in Tabelle 9 aufgelistet. Der Formparameter  $\hat{\xi} < 0$ , das bedeutet, dass die Hill-Methode in diesem Fall nicht anwendbar ist [McNeil and Frey, 2000, S. 287].

$\hat{\beta}$	$\hat{\xi}$
1.6701	-0.1813

Tabelle 9: Schätzwerte der GPD

Allerdings ist es mit 5776 Prognosen fast unmöglich, für jedes Zeitfenster einen neuen Plot zu detektieren und einen neuen Schwellenwert zu schätzen. Deswegen wird das 90%-Sample-Quantil der 1050 vergangenen Verluste im weiteren Verlauf als der Schwellenwert festgesetzt, wie [McNeil and Frey, 2000, S. 288] vorgingen, wobei 10% der Daten immer als Exzesse gewählt wurden. Die VaR- und ES-Prognosen finden sich in den Abbildungen 15 und 16. Die entsprechenden Testergebnisse sieht man in den Tabellen 10 und 11.

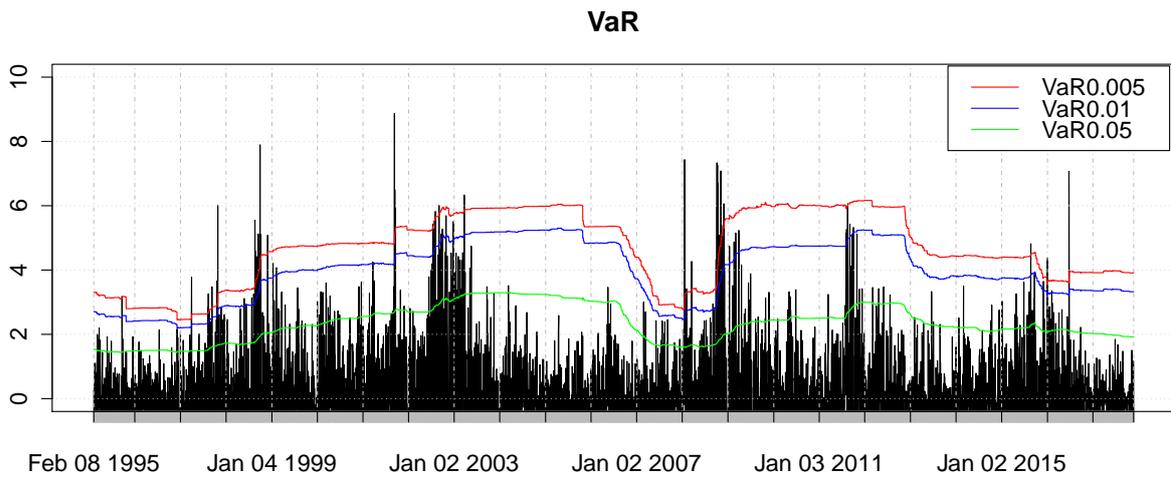


Abbildung 15: *VaR-Prognosen der POT-Methode.*

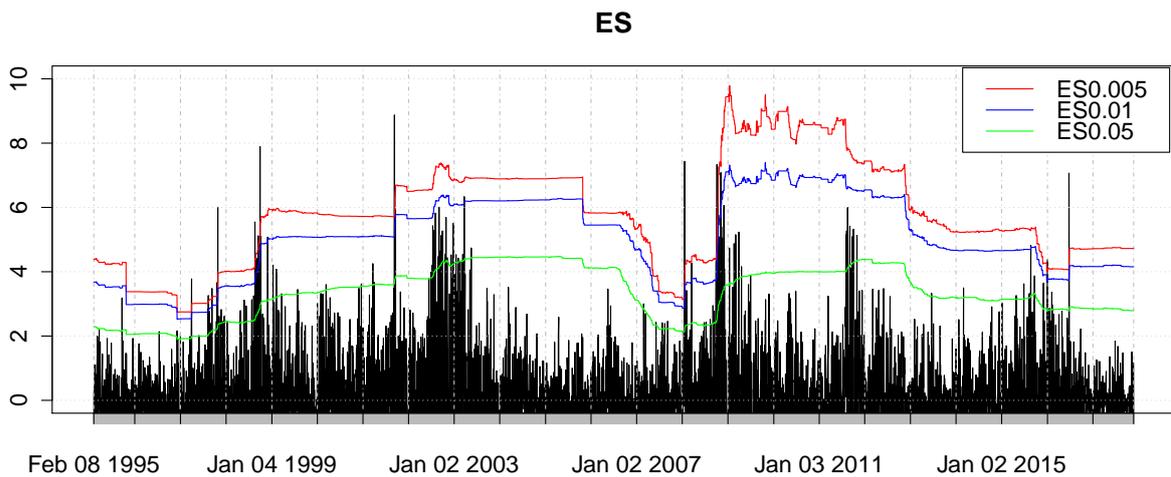


Abbildung 16: *ES-Prognosen der POT-Methode.*

$VaR_\alpha$	#Verletzung	%Verletzung	U.C.	Ind.	C.C.
$VaR_{0.05}$	325	5.63	0.03	0.00	0.00
$VaR_{0.01}$	73	1.26	<b>0.05</b>	0.00	0.00
$VaR_{0.005}$	42	0.73	0.02	<b>0.32</b>	0.04

Tabelle 10: VaR-Backtesting der POT-Methode.

$ES_\alpha$	$Z_2$	I.ESR
$ES_{0.05}$	> <b>0.05</b>	0.00
$ES_{0.01}$	> <b>0.05</b>	0.01
$ES_{0.005}$	> <b>0.05</b>	0.01

Tabelle 11: ES-Backtesting der POT-Methode.

Die Ergebnisse der Tests zeigen, dass das POT-Modell nicht gut abschneidet. Der Intercept-ESR-Test und der Acerbi und Szekelys Test bieten widersprüchliche Folgerungen. Um diese Methode zu verbessern, ist der in Kapitel 3.2.5 vorgestellte GARCH-Filter einzusetzen. Zuerst wird ein AR(1)-GARCH(1,1)-Modell für die 6826 Log-Verluste angepasst. Wie Kapitel 3.2.5 zeigt, wird die PML-Methode für die Schätzung der Parameter des GARCH-Prozesses verwendet. Die dazugehörigen Schätzwerte sind in Tabelle 12 angegeben:

Parameter	$\hat{\alpha}_0$	$\hat{\alpha}$	$\hat{c}_0$	$\hat{c}_1$	$\hat{d}$
Schätzer	-0.071181	-0.003151	0.027693	0.082845	0.902120
Std.Abw.	0.013689	0.012785	0.008591	0.009249	0.010042
p-Wert	0.00	<b>0.81</b>	0.00	0.00	0.00

Tabelle 12: Ergebnisse der Schätzung des AR(1)-GARCH(1,1)-Modells

Da der AR-Term nicht signifikant ist, ist das Modell auf GARCH(1,1) ohne AR-Term zu reduzieren (Siehe Tabelle 13). In den Abbildungen 17 und 18 sind die absoluten Beträge und die Standardabweichungen der Log-Verluste zu sehen.

Parameter	$\hat{\alpha}_0$	$\hat{c}_0$	$\hat{c}_1$	$\hat{d}$
Schätzer	-0.070930	0.027690	0.082876	0.902093
Std.Abw.	0.013486	0.008593	0.009251	0.010041
p-Wert	0.00	0.00	0.00	0.00

Tabelle 13: Ergebnisse der Schätzung des GARCH(1,1)-Modells.

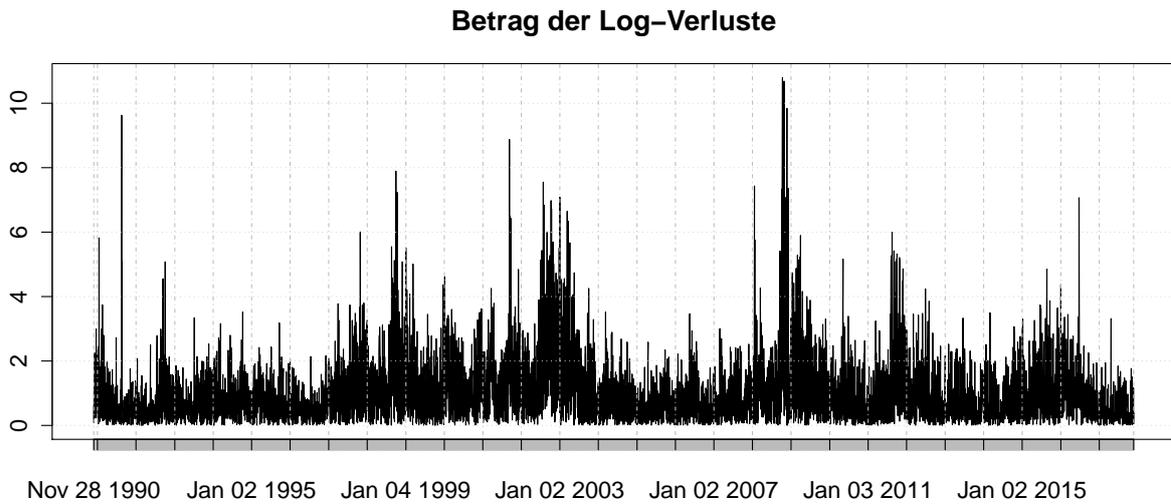


Abbildung 17: *Betrag der Log-Verluste.*

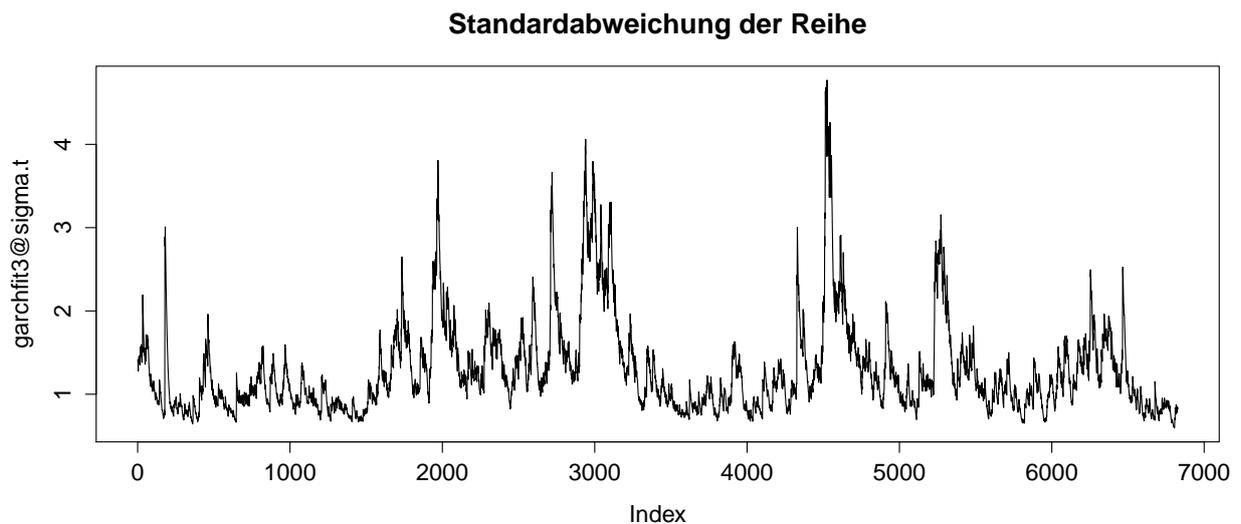


Abbildung 18: *Standardabweichung der Log-Verluste.*

Im Anschluss daran werden die standardisierten Residuen mithilfe von 80 berechnet, die sich in Abbildung 19 zeigen, wo fast kein Cluster-Effekt zu erkennen ist. In Abbildung 20 sind die Autokorrelationsfunktionen (kurz: ACF) der originalen Log-Verluste und der standardisierten Residuen zu finden, die darauf hinweisen, dass die originalen Log-Verluste offensichtlich nicht i.i.d. sind, während die standardisierten Residuen vielleicht die i.i.d. Voraussetzung erfüllen. Diese Schlussfolgerung wird auch durch den Ljung-Box-Test untermauert, wobei mit dem p-Wert 0.53 die Unabhängigkeit der standardisierten Residuen nicht abgelehnt werden kann.

Danach wird die GPD auf die standardisierten Residuen  $z_t$  angepasst. Die entsprechenden Risikomaße sind durch 82 zu berechnen. Die VaR- und ES-Prognosen sind in den Abbildungen 21 und 22 zu finden, wegen der Zickzackartigkeit der Kurve werden dabei nur die ersten 1000 Prognosen als Beispiel dargestellt. In den Tabellen 14 und 15 ist eine deutliche Verbesserung gegenüber dem primitiven POT-Modell zu erkennen. Gemäß der Testergebnisse hat das modifizierte POT-Modell alle Tests bestanden. Selbst wenn dieses Modell das Risiko bei  $\alpha = 0.01$  und  $\alpha = 0.005$  überschätzen könnte, sind die Abweichungen aber nicht signifikant.

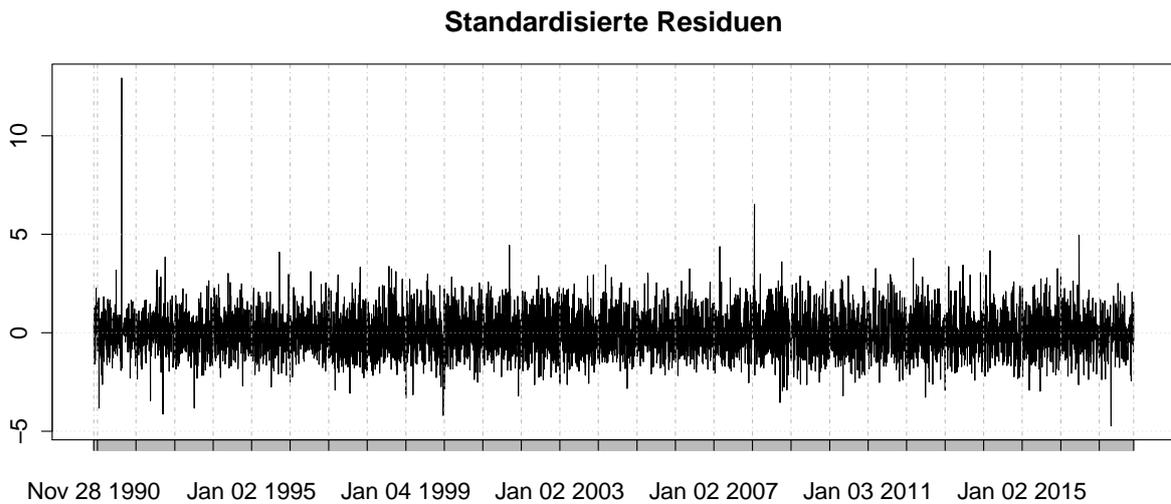


Abbildung 19: *Standardisierte Residuen des GARCH(1,1)-Modells.*

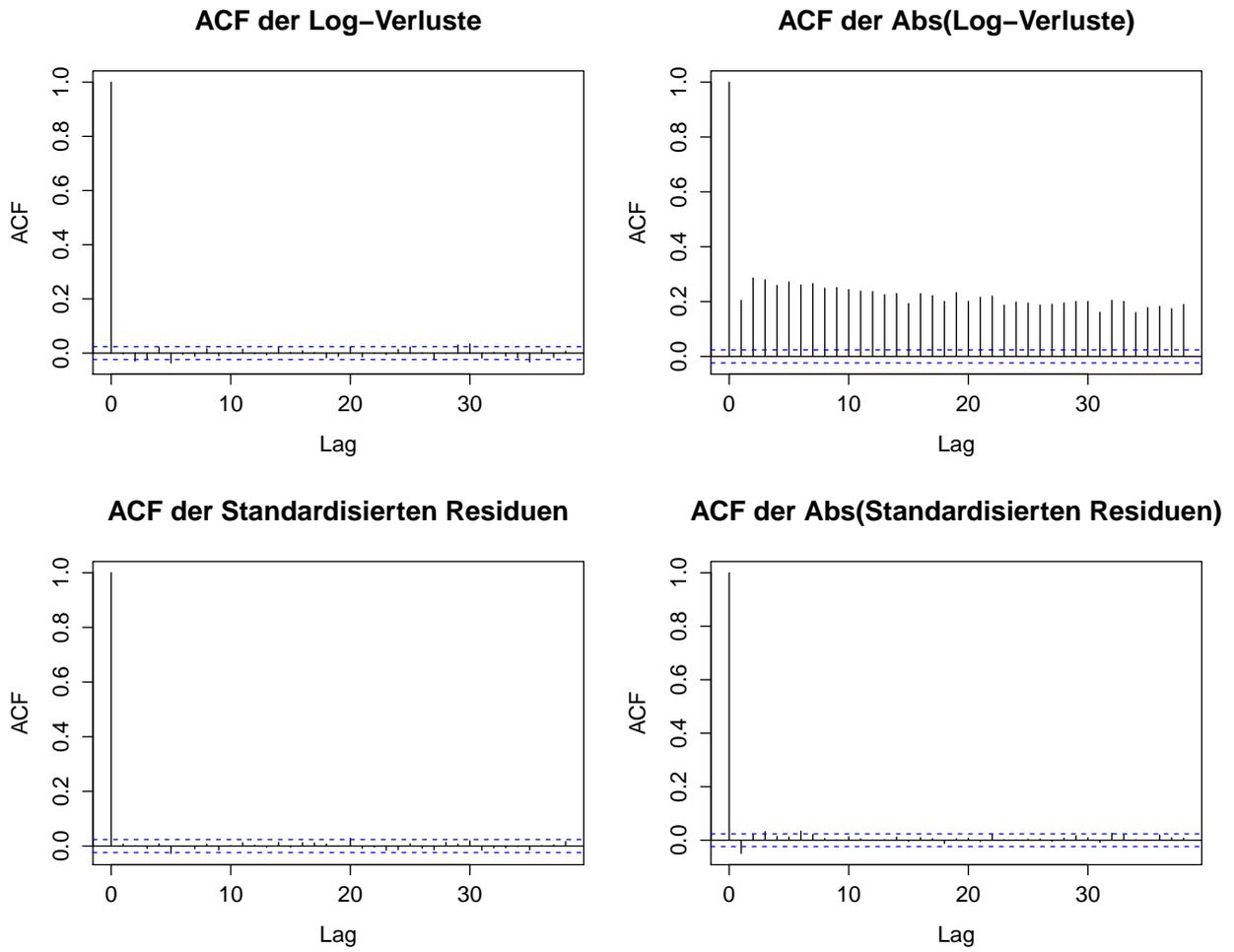


Abbildung 20: Die Autokorrelationsfunktion der originalen Log-Verluste ist oben zu finden, die Autokorrelationsfunktion der standardisierten Residuen unten.

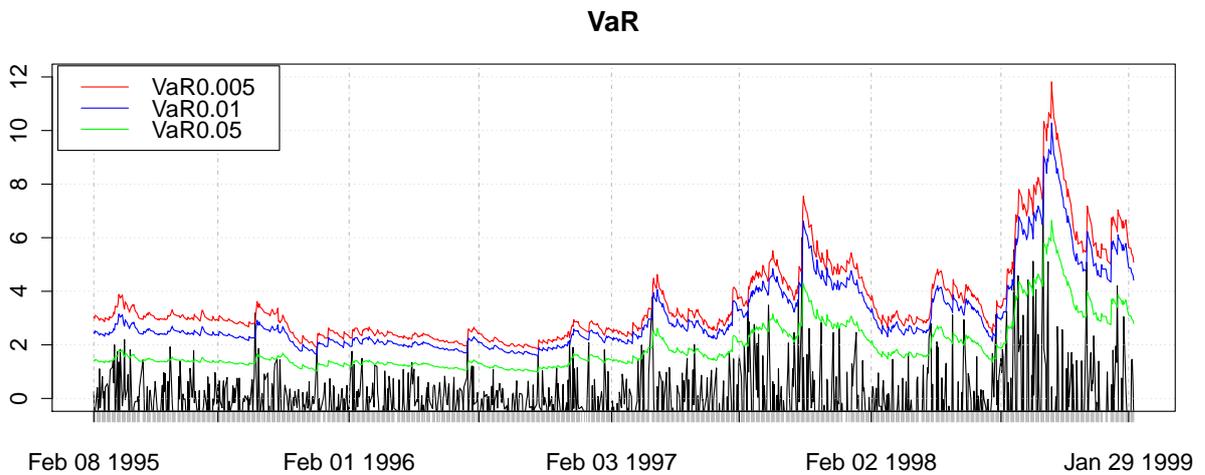


Abbildung 21: Die ersten 1000 VaR-Prognosen der POT-Methode mit dem GARCH-Filter.

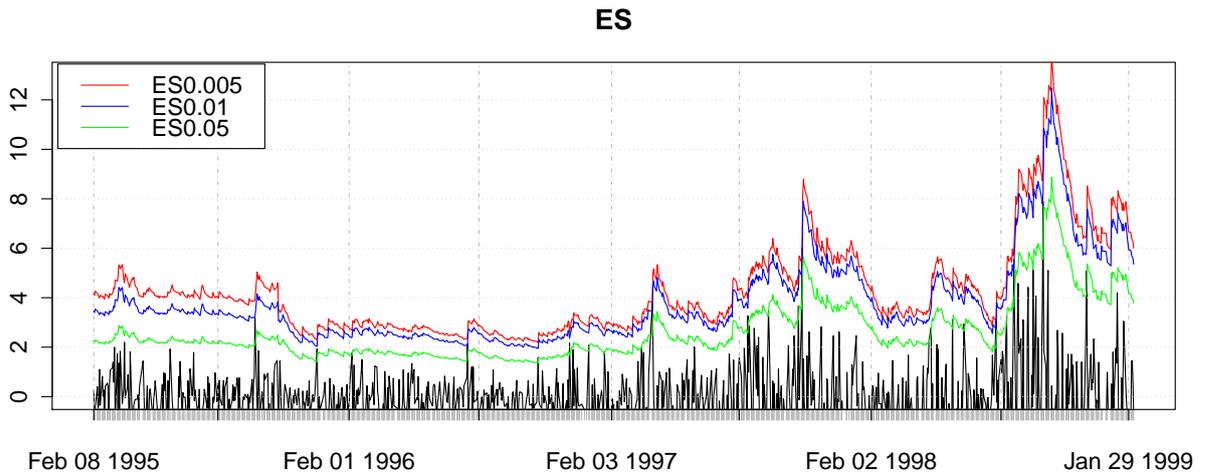


Abbildung 22: Die ersten 1000 ES-Prognosen der POT-Methode mit dem GARCH-Filter.

$VaR_\alpha$	#Verletzung	%Verletzung	U.C.	Ind.	C.C.
$VaR_{0.05}$	311	5.38	<b>0.19</b>	<b>0.47</b>	<b>0.32</b>
$VaR_{0.01}$	51	0.88	<b>0.36</b>	<b>0.48</b>	<b>0.51</b>
$VaR_{0.005}$	26	0.45	<b>0.58</b>	<b>0.11</b>	<b>0.24</b>

Tabelle 14: VaR-Backtesting der POT-Methode mit dem GARCH-Filter.

$ES_\alpha$	$Z_2$	I-ESR
$ES_{0.05}$	> <b>0.05</b>	<b>0.16</b>
$ES_{0.01}$	> <b>0.05</b>	<b>0.28</b>
$ES_{0.005}$	> <b>0.05</b>	<b>0.12</b>

Tabelle 15: ES-Backtesting der POT-Methode mit dem GARCH-Filter.

### 5.3 Weitere Modelle

Die Testergebnisse der weiteren Modelle sind in den Tabellen 16 und 17 zusammengefasst. Die entsprechenden Grafiken der Prognosen befinden sich in Anhang A.

Modell	$VaR_\alpha$	#Verletzung	%Verletzung	U.C.	Ind.	C.C.
Historische Simulation	$VaR_{0.05}$	323	5.59	0.04	0.00	0.00
	$VaR_{0.01}$	80	1.39	0.01	0.00	0.00
	$VaR_{0.005}$	41	0.71	0.03	<b>0.30</b>	<b>0.06</b>
Normale Verteilung	$VaR_{0.05}$	331	5.73	0.01	0.00	0.00
	$VaR_{0.01}$	144	2.49	0.00	0.00	0.00
	$VaR_{0.005}$	88	1.52	0.00	0.00	0.00
Studentsche t-Verteilung	$VaR_{0.05}$	377	6.53	0.00	0.00	0.00
	$VaR_{0.01}$	86	1.49	0.00	0.00	0.00
	$VaR_{0.005}$	44	0.76	0.01	0.05	0.00
Skewed t-Verteilung	$VaR_{0.05}$	342	5.92	0.00	0.00	0.00
	$VaR_{0.01}$	73	1.26	<b>0.05</b>	0.00	0.00
	$VaR_{0.005}$	41	0.71	0.03	0.04	0.01
GARCH mit NV	$VaR_{0.05}$	344	5.96	0.00	<b>0.40</b>	0.00
	$VaR_{0.01}$	93	1.61	0.00	<b>0.69</b>	0.00
	$VaR_{0.005}$	59	1.02	0.00	<b>0.64</b>	0.00
GARCH mit tV	$VaR_{0.05}$	366	6.34	0.00	<b>0.47</b>	0.00
	$VaR_{0.01}$	72	1.25	<b>0.07</b>	<b>0.91</b>	<b>0.19</b>
	$VaR_{0.005}$	36	0.62	<b>0.20</b>	<b>0.23</b>	<b>0.21</b>
GARCH mit sktV	$VaR_{0.05}$	353	6.11	0.00	<b>0.71</b>	0.00
	$VaR_{0.01}$	61	1.06	<b>0.67</b>	<b>0.68</b>	<b>0.84</b>
	$VaR_{0.005}$	26	0.45	<b>0.58</b>	<b>0.11</b>	<b>0.24</b>

Tabelle 16: VaR-Backtesting der weiteren Methoden.

Modell	$ES_\alpha$	$Z_2$	I.ESR
Historische Simulation	$ES_{0.05}$	> <b>0.05</b>	0.00
	$ES_{0.01}$	> <b>0.05</b>	0.01
	$ES_{0.005}$	> <b>0.05</b>	0.01
Normale Verteilung	$ES_{0.05}$	> <b>0.05</b>	0.00
	$ES_{0.01}$	< 0.05	0.00
	$ES_{0.005}$	< 0.05	0.00
Studentsche t-Verteilung	$ES_{0.05}$	> <b>0.05</b>	0.00
	$ES_{0.01}$	> <b>0.05</b>	0.00
	$ES_{0.005}$	> <b>0.05</b>	0.01
Skewed t-Verteilung	$ES_{0.05}$	> <b>0.05</b>	0.01
	$ES_{0.01}$	> <b>0.05</b>	<b>0.08</b>
	$ES_{0.005}$	> <b>0.05</b>	<b>0.09</b>
GARCH mit NV	$ES_{0.05}$	> <b>0.05</b>	0.00
	$ES_{0.01}$	< 0.05	0.00
	$ES_{0.005}$	< 0.05	0.00
GARCH mit tV	$ES_{0.05}$	> <b>0.05</b>	0.00
	$ES_{0.01}$	> <b>0.05</b>	<b>0.10</b>
	$ES_{0.005}$	> <b>0.05</b>	<b>0.08</b>
GARCH mit sktV	$ES_{0.05}$	> <b>0.05</b>	<b>0.08</b>
	$ES_{0.01}$	> <b>0.05</b>	<b>0.56</b>
	$ES_{0.005}$	> <b>0.05</b>	<b>0.46</b>

Tabelle 17: ES-Backtesting der weiteren Methoden.

## 6 Zeitliche Aggregation

Bisher wird nur über den täglichen Zeithorizont diskutiert, denn in der Praxis ist die Halte-dauer meistens ein Tag [Daniélsson, 2011, S. 89]. Daneben stellt man auch den Anspruch an die mehrtägigen Risikomaße, insbesondere an die zehntägigen, die durch den BCBS im Jahr 1996 vorgesehen wurde [on Banking Supervision, 1996, S. 4], weil die Anleger dadurch ihre Position liquidieren können [Degiannakis and Potamia, 2017, S. 176]. Aber die Berechnung der Risikomaße für einen langen Zeithorizont ist anspruchsvoller als die für einen Tag, weil beispielsweise für den zehntägigen VaR mindestens 10-jährige Daten verwendet werden sollen [Daniélsson, 2011, S. 89]. Zudem ist es, wenn man den Zeithorizont auf ein Jahr verlängert, unmöglich, den jährlichen VaR direkt zu kalkulieren, denn es gibt fast keinen Markt, der eine mehr als 200-jährige Geschichte hat. Eine Gegenmaßnahme dazu ist die *scaling law*, womit man das mehrtägige Risikomaß auf Tagesbasis bekommt.

## 6.1 Square-Root-of-Time Regel

Eine der bekanntesten *scaling laws* ist die durch den BCBS vorgeschlagene *square-root-of-time* Regel (kurz: SRTR), die es ermöglicht, den täglichen VaR auf den zehntägigen VaR umzurechnen. Bei dieser Regel wird das Risiko durch die Quadratwurzel der Zeitlänge skaliert. Das heißt, mit gegebenem  $VaR_\alpha^1$  wird der n-tägige VaR berechnet durch:

$$VaR_\alpha^n = \sqrt{n}VaR_\alpha^1. \quad (121)$$

Wegen der Einfachheit und der Instruktion durch den BCBS ist die SRTR sehr populär in der Finanzwelt. Aber die theoretische Geeignetheit der SRTR wird oft infrage gestellt, weil die SRTR voraussetzt, dass die Daten identisch unabhängig normalverteilt sind. Allerdings ist das nicht der Fall. Zum Versagen der SRTR und darüber, ob sie das Risiko unterschätzt oder überschätzt, äußern Wissenschaftler widersprüchliche Meinungen: Bei [McNeil and Frey, 2000, S. 298] und [Danielsson and Zigrand, 2006, S. 1] wurde die Unterschätzung durch diese Regel festgestellt, dagegen sprachen [Danielsson et al., 1998, S. 2] und [Danielsson, 2011, S. 178] aber für die Überschätzung des Risikos, die durch die SRTR eingeführt wurde.

Dazu haben Wang et al. [2011] eine umfassende Studie durchgeführt, in der die unterschiedlichen stilisierten Eigenschaften der Finanzdaten überprüft worden sind. Sie haben herausgefunden, dass die Leistung der SRTR durch solche Faktoren uneinheitlich beeinflusst wird: Die Abhängigkeit der Zeitreihe und der dicke Schwanz der Daten führen zu einer gravierenden Verzerrung beim Ergebnis der SRTR, während der Sprung, das Volatilitätscluster und der Non-Null-Mittelwert eine vergleichsweise nur unauffällige Abweichung erzeugen [Wang et al., 2011, S. 1158]. Aufgrund der Komplexität der vielfältigen Confoundings ist die Auswertung der SRTR nicht einfach. Obwohl die SRTR bei der Untersuchung der globalen Märkte das Risiko im Durchschnitt unterschätzte, waren die Ergebnisse aber nach Land dramatisch unterschiedlich. Gemäß der empirischen Analyse scheint die Anwendung der SRTR auf die entwickelten Länder unproblematisch zu sein [Wang et al., 2011, S. 1168]. Darüber hinaus fanden Brummelhuis and Kaufmann [2004] die SRTR akzeptabel für den AR(1)-GARCH(1,1)-Prozess, falls der autoregressive Term nicht zu groß ist.

## 6.2 Auf dem Tail-Index basierende Skalierung

Danielsson et al. [1998] haben eine weitere Art von Skalierung vorgestellt, dabei erhält man den mehrtägigen VaR dadurch, dass der tägliche VaR mithilfe des Tail-Index skaliert wird. Anders als die SRTR wird der n-tägige VaR hierbei geschätzt durch:

$$TVaR_\alpha^n = n^{1/\hat{\iota}}VaR_\alpha^1, \quad (122)$$

wobei  $\hat{\iota}$  den geschätzte Tail-Index der Daten bezeichnet, der durch den Hill-Plot herausge-

funden werden kann (siehe Abbildung 6). Mit  $\hat{\iota} = 3.50$  bekommt man den Skalierungsterm  $n^{1/3.5}$ , der deutlich kleiner als  $\sqrt{n}$  der SRTR ist.

### 6.3 Tests der Anwendbarkeit der Skalierung

In diesem Abschnitt wird die Anwendbarkeit der Skalierung auf den DAX vom 28.11.1990 bis zum 21.08.2014 ( $T = 6000$ ) untersucht. Als Beispiele sind die Zeithorizonte  $n = 10$  und  $n = 30$  mit den Signifikanzniveaus  $\alpha = 0.05, 0.01$  und  $0.005$  zu betrachten. Um möglichst viele Daten zur Verfügung zu stellen, ist die von Wang et al. [2011] vorgeschlagene *Subsample*-Methode einzusetzen. Die folgende Vorgehensweise folgt [Wang et al., 2011, S. 1160].

Zunächst sind  $n$  *Subsamples*  $\{L_{s,k}^n\}_{s=1}^{T/h-1}$  zu bilden. Die Log-Verluste lassen sich direkt summieren:

$$\underbrace{L_k + \dots + L_{n+k-1}}_{L_{1,k}^n}, \underbrace{L_{n+k} + \dots + L_{2n+k-1}}_{L_{2,k}^n}, \dots, \underbrace{L_{(s-1)n+k} + \dots + L_{sn+k-1}}_{L_{s,k}^n}, \dots, \quad (123)$$

wobei  $k = 1, 2, \dots, n$ . Für jede  $k$  gibt es einen entsprechenden  $VaR_\alpha^n(k)$ . Der sogenannte *Benchmark*-VaR für die Frequenz  $n$  wird berechnet durch:

$$\overline{VaR}_\alpha^n = \frac{1}{n} \sum_{k=1}^n VaR_\alpha^n(k). \quad (124)$$

Falls die SRTR gut funktioniert<sup>20</sup>, soll  $\overline{VaR}_\alpha^n \approx \sqrt{n} VaR_\alpha^1$  sein. Um die Abweichung der SRTR zu vermessen, ist die folgende Bias-Funktion  $\zeta(n)$  zu definieren:

$$\zeta(n) = \frac{\sqrt{n} VaR_\alpha^1}{\overline{VaR}_\alpha^n} - 1. \quad (125)$$

Ein positiver (negativer) Wert der  $\zeta(n)$  weist auf eine Überschätzung (Unterschätzung) der SRTR hin. Um die Anwendbarkeit der SRTR zu überprüfen, werden zwei Tests durchgeführt: der Varianz-Ratio-Test und der *Subsample-Based*-Test.

Der Varianz-Ratio-Test wurde von Lo and MacKinlay [1988] entwickelt. Dabei wird die Varianz-Ratio-Statistik mit der Frequenz  $n$  wie folgt definiert:

$$\overline{VR}^n = 1 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \hat{\rho}_k, \quad (126)$$

wobei  $\hat{\rho}_k$  den Schätzer des Autokorrelationskoeffizienten der  $k$ -ten Ordnung bezeichnet. Wenn die Varianz-Ratio-Statistik  $\overline{VR}^n > 1 (< 1)$  ist, weist sie auf eine positive (negati-

<sup>20</sup>Aus Gründen der Einfachheit ist die originale SRTR in folgenden Verfahren der Tests als Beispiel anzuführen. Die Formeln gelten auch für andere Skalierungen.

ve) Abhängigkeit der Zeitreihe hin, wodurch die SRTR zur Unterschätzung (Überschätzung) führt [Wang et al., 2011, S. 1164]. Aufgrund der  $\overline{VR}^n$  ist die standardisierte Teststatistik  $\psi(n)$  gegeben durch:

$$\psi(n) = \frac{\sqrt{T}(\overline{VR}^n - 1)}{\sqrt{\hat{\vartheta}_n}} \stackrel{\text{asy}}{\sim} N(0, 1), \quad (127)$$

wobei  $\psi(n)$  die asymptotische Varianz der  $\overline{VR}^n$  ist (Details dazu finden sich in [Wang et al., 2011, S. 1164] und [Lo and MacKinlay, 1988, S. 50]). Obwohl der Varianz-Ratio-Test nur die Abhängigkeit der Zeitreihe detektiert, ist er aber normalerweise ausreichend, denn die Abhängigkeit der Daten spielt eine wesentliche Rolle bei der Fehlschätzung der SRTR [Wang et al., 2011, S. 1164].

Darüber hinaus wurde der *Subsample-Based-Test* (kurz: SB-Test) als ein Gesamttest von Wang et al. [2011] aufgestellt. Dabei haben sie herausgefunden, dass gemäß des zentralen Grenzwertsatzes der standardisierte Bias-Term standardnormalverteilt ist:

$$SB(n) = \frac{\sqrt{n}VaR_\alpha^1 - \overline{VaR}_\alpha^n}{\Sigma(n)} \stackrel{\text{asy}}{\sim} N(0, 1), \quad (128)$$

wobei  $\Sigma(n)$  die *bootstrapped* Standardabweichung von  $(\sqrt{n}VaR_\alpha^1 - \overline{VaR}_\alpha^n)$  ist, die durch das folgende Verfahren berechnet wird:

- Die 6000 Beobachtungen sind in 5991 äquidistante Blöcke zu unterteilen:  
 $\{l_1, l_2, \dots, l_{10}\}, \{l_2, l_3, \dots, l_{11}\}, \dots, \{l_{5991}, l_{5992}, \dots, l_{6000}\}$ .
- Resampling wird vorgenommen, dabei werden 600 Blöcke aus den 5991 Blöcken zufällig gezogen, um ein *bootstrapped* Sample zu erzeugen.
- Für jedes *bootstrapped* Sample werden  $VaR_\alpha^1$ ,  $\overline{VaR}_\alpha^n$  und  $(\sqrt{n}VaR_\alpha^1 - \overline{VaR}_\alpha^n)$  gespeichert.
- Die Schritte 2 und 3 werden 5000 Mal wiederholt, dadurch bekommt man die *bootstrapped* Verteilung des Bias-Terms. Im Anschluss daran ist die *bootstrapped* Standardabweichung  $\Sigma(n)$  zu berechnen.

**Modifizierte SRTR** Da der hauptsächliche Fehler der SRTR auf die Abhängigkeit der Daten zurückzuführen ist, empfohlen [Wang et al., 2011, S. 1164] eine modifizierte SRTR-Methode (kurz: MSRTR), die auf der Varianz-Ratio-Statistik 126 basiert:

$$MVaR_\alpha^n = \sqrt{n \times \overline{VR}^n} \times VaR_\alpha^1. \quad (129)$$

Bei der modifizierten SRTR lässt sich die Unterschätzung oder die Überschätzung durch die Varianz-Ratio-Statistik kompensieren. Im Falle von  $\overline{VR}^n = 1$  ist die MSRTR identisch mit

der SRTR.

## 6.4 Anwendung auf den DAX

In diesem Abschnitt werden die oben genannten Skalierungsmethoden auf den deutschen Markt angewendet. Die geschätzten VaRs und die Testergebnisse<sup>21</sup> sind in den Tabellen 18,19 und 20 aufzulisten. In Tabelle 18 ist eine systematische Überschätzung der SRTR bei allen  $\alpha$ -Niveaus zu sehen, da alle Werte von  $\zeta(n)\%$  positiv sind. Im Vergleich dazu nähern sich die Werte von  $\zeta(n)\%$  in Tabelle 19 an Null an, insbesondere wenn  $n = 30$ . Diesem Unterschied ist zu entnehmen, dass die Ergebnisse durch die Modifizierung einigermaßen verbessert werden. Bei den letzten zwei Spalten der Tabellen 18 und 19 kann man sehen, dass alle p-Werte größer als 0.05 sind, das heißt, sowohl die SRTR als auch die MSRTR sind geeignet im Falle des DAX. Dieses Fazit stimmt mit der Feststellung von Wang et al. [2011] überein.

Im Hinblick auf Tabelle 20 ist eine eindeutige und gravierende Unterschätzung festzustellen. Die durch die Tail-Index-Skalierung eingeführte Verzerrung ist sogar viel größer als die anderen Skalierungen. Die Ursache liegt darin, dass die Tail-Index-Skalierung nur den dicken Schwanz der Finanzdaten berücksichtigt. Die durch den dicken Schwanz verursachte Aufwärtsverzerrung wird dadurch kompensiert, aber die anderen Faktoren bleiben unbehandelt. Die Anwendbarkeit der Tail-Index-Skalierung wurde auch von [McNeil and Frey, 2000, S. 298] bestritten.

$\alpha$	$VaR_\alpha^1$	$\overline{VaR}_\alpha^n$	$\sqrt{n}VaR_\alpha^1$	$\zeta(n)\%$	$\psi(n)$	$SB(n)$
n=10						
0.05	2.29	6.75	7.23	7.11	-1.56( <b>0.12</b> )	1.53( <b>0.13</b> )
0.01	4.31	13.34	13.64	2.25	-1.56( <b>0.12</b> )	0.33( <b>0.74</b> )
0.005	5.15	16.16	16.28	0.74	-1.56( <b>0.12</b> )	0.10( <b>0.92</b> )
n=30						
0.05	2.29	12.13	12.53	3.30	-0.85( <b>0.40</b> )	0.40( <b>0.69</b> )
0.01	4.31	22.87	23.62	3.28	-0.85( <b>0.40</b> )	0.37( <b>0.71</b> )
0.005	5.15	25.55	28.19	10.33	-0.85( <b>0.40</b> )	1.09( <b>0.28</b> )

Tabelle 18: Anwendung der SRTR auf den DAX vom 28.11.1990 bis zum 21.08.2014.

<sup>21</sup>Der Varianz-Ratio-Test dient nur zur Überprüfung der Anwendbarkeit der SRTR, deswegen ist die Varianz-Ratio-Statistik in den Tabellen 18,19 und 20 unverändert.

$\alpha$	$VaR_\alpha^1$	$\overline{VaR}_\alpha^n$	$MVaR_\alpha^n$	$\zeta(n)\%$	$\psi(n)$	$SB(n)$
n=10						
0.05	2.29	6.75	6.83	1.19	-1.56( <b>0.12</b> )	0.26( <b>0.79</b> )
0.01	4.31	13.34	12.88	-3.45	-1.56( <b>0.12</b> )	-0.53( <b>0.60</b> )
0.005	5.15	16.16	15.38	-4.83	-1.56( <b>0.12</b> )	-0.67( <b>0.50</b> )
n=30						
0.05	2.29	12.13	11.84	-2.39	-0.85( <b>0.40</b> )	-0.29( <b>0.77</b> )
0.01	4.31	22.87	22.32	-2.40	-0.85( <b>0.40</b> )	-0.27( <b>0.79</b> )
0.005	5.15	25.55	26.64	4.27	-0.85( <b>0.40</b> )	0.45( <b>0.65</b> )

Tabelle 19: Anwendung der modifizierten SRTR auf den DAX vom 28.11.1990 bis zum 21.08.2014.

$\alpha$	$VaR_\alpha^1$	$\overline{VaR}_\alpha^n$	$TVaR_\alpha^n$	$\zeta(n)\%$	$\psi(n)$	$SB(n)$
n=10						
0.05	2.29	6.75	4.42	-34.52	-1.56( <b>0.12</b> )	-6.95(0.00)
0.01	4.31	13.34	8.32	-37.63	-1.56( <b>0.12</b> )	-5.93(0.00)
0.005	5.15	16.16	9.94	-38.49	-1.56( <b>0.12</b> )	-5.17(0.00)
n=30						
0.05	2.29	12.13	6.04	-50.21	-0.85( <b>0.40</b> )	-5.63(0.00)
0.01	4.31	22.87	11.39	-50.20	-0.85( <b>0.40</b> )	-5.84(0.00)
0.005	5.15	25.55	13.60	-46.77	-0.85( <b>0.40</b> )	-4.82(0.00)

Tabelle 20: Anwendung der auf dem Tail-Index basierenden Skalierung auf den DAX vom 28.11.1990 bis zum 21.08.2014.

## 7 Fazit

Die Hauptergebnisse in der vorliegenden Arbeit lassen sich in folgende Punkte zusammenfassen:

- Wie die Ergebnisse in Tabelle 16 zeigen, sind alle traditionellen unbedingten Modelle in allen Fällen gescheitert. Da die unbedingten Modelle den Effekt des Volatilitätsclusters der Finanzdaten ignorieren, sind sie alle beim VaR Unabhängigkeitstest durchgefallen. Obwohl die Studentsche t-Verteilung oder die Skewed t-Verteilung im Vergleich zur Normalverteilung bezüglich der unbedingten Überdeckung eine kleine Verbesserung bietet, ist diese aber noch nicht ausreichend. Alle unbedingten Modelle unterschätzen VaR und die Unterschätzung ist manchmal erheblich (siehe Tabelle 16).
- Die Einführung des GARCH-Modells, sogar des einfachsten GARCH(1,1)-Modells, bringt deutliche Verbesserungen für die Schätzung des VaR im Vergleich zum unbedingten Modell (siehe Tabelle 16). Da das GARCH-Modell die variierende Volatilität beschreiben kann, ist es in der Lage, i.i.d. VaR-Prognosen zu produzieren (siehe die Ergebnisse des Unabhängigkeitstests in den Tabellen 14 und 16) und gilt als wichtiges Hilfsmittel bei der Risiko-Schätzung. Das GARCH(1,1)-Modell mit Studentsche t-verteilter oder Skewed t-verteilter Innovation macht taugliche Prognosen bei  $\alpha = 0.01$  und  $\alpha = 0.005$  für VaR und ES (siehe Tabelle 16 und 17).
- Die rohe BM-Methode schneidet bei der Schätzung von VaR und ES schlecht ab (siehe Tabellen 4 und 5). Wenn man den extremen Index einrechnet, bietet die neue BM-Methode zufriedenstellende Ergebnisse bei der unbedingten Überdeckung (siehe Tabelle 7 und 8). Allerdings spiegelt der extreme Index ausschließlich das durchschnittliche Niveau des Volatilitätsclusters wider. Deswegen ist er unfähig, die individuelle Volatilität zu beachten. Dies führt dazu, dass die modifizierte BM-Methode beim Unabhängigkeitstest nicht gut abschneidet (siehe Tabelle 7). Beim Auswahlverfahren des Extremwerts werden außerdem viele Daten weggelassen. Solche Faktoren haben die Unpopularität der BM-Methode bei der Risiko-Schätzung zur Folge.
- Die originale POT-Methode zeigt auch eine mangelhafte Leistung bei der Schätzung von VaR und ES (siehe Tabellen 10 und 11), da die grundlegende i.i.d. Verteilungsannahme in der Praxis verletzt ist. Dieser Defekt ist dadurch zu beheben, dass sich die Daten durch den GARCH-Filter vorbehandeln lassen. Die mit dem GARCH-Filter kombinierte POT-Methode ragt aus allen verglichenen Methoden heraus und gilt als der Spitzenreiter in dieser Arbeit (siehe Tabellen 14 und 15).
- Eine Schwierigkeit bei der POT-Methode ist die Bestimmung des Schwellenwerts, die normalerweise mithilfe von *Mean Residual Life Plot* oder *Hill-Plot* erfolgt. Dieses *eyeball* Verfahren ist aber subjektiv und nicht automatisch, somit ist es nicht machbar

bei der Schätzung mit dem gleitenden Zeitfenster. Wie die empirischen Ergebnisse in dieser Arbeit zeigen (siehe Tabellen 14 und 15), ist es in der Praxis jedoch ausreichend, das 90%-Sample-Quantil als Schwellenwert zu wählen.

- Die auf der Extremwerttheorie basierenden modifizierten Methoden bieten eine gute Leistung bei  $\alpha = 0.05, 0.01$  und  $0.005$  bezüglich der unbedingten Überdeckungsfähigkeit, zudem hat die modifizierte POT-Methode sogar alle Tests bestanden (Siehe Tabellen 7, 8, 14 und 15). Daraus lässt sich folgern, dass die Extremwerttheorie eine breite Anwendung bei der Risiko-Schätzung hat und nicht auf ein minimales Quantil eingegrenzt werden muss. Dieses Ergebnis spricht gegen [Diebold, 1998, S. 11], die uns vor Missbrauch der EVT warnte, weil die EVT-Methode nur zu sehr extremen Fällen, z.B.  $\alpha = 0.001$ , gehöre.
- Im Anwendungsteil werden zwei ES-Tests durchgeführt. Während der Intercept-ESR-Test die Ergebnisse bietet, die fast mit dem VaR unbedingten Überdeckungstest übereinstimmen, zeigt der Acerbi und Szekelys zweite Test aber unzureichende Power, um unzuverlässige Modelle abzulehnen (siehe Tabellen 16 und 17).
- Obwohl die primitive *square-root-of-time* Regel wegen ihrer unrealistischen Annahme oft für Kritik sorgt, erweist sie sich überraschenderweise effektiv für den DAX (siehe Tabelle 18). Die gute Leistung der SRTR lässt sich aber nicht bei allen Märkten beobachten, deswegen ist die SRTR erst auf die Daten anzuwenden, nachdem ihre Geeignetheit durch den Vortest geprüft wurde. Nach der Kompensation der seriellen Abhängigkeit durch die Varianz-Ratio-Statistik geht die SRTR noch einen kleinen Schritt weiter (siehe Tabelle 19), wohingegen die Tail-Index-Skalierung das Risiko drastisch unterschätzt (siehe Tabelle 20).

Bei weiteren Studien könnte ferner der folgende Punkt von Interesse sein:

- Die aus der empirischen Analyse dieser Arbeit hervorgehenden Schlussfolgerungen basieren auf dem gleitenden Zeitfenster mit 1050 Tagen. Es ist vorstellbar, dass die Leistungen der verschiedenen Ansätze je nach Größe des Zeitfensters unterschiedlich sein könnten. Zu diesem Punkt haben Kuester et al. [2006] den Einfluss der unterschiedlichen Größen des Zeitfensters dadurch untersucht, dass sie die Fälle der Größen  $w = 250, 500$  und  $1000$  hinsichtlich der Überdeckungsfähigkeit verglichen haben. Dabei sind sie zur Schlussfolgerung gekommen, dass sich die meisten Modelle konstant verhalten, während andere wechselnde Leistungen zeigen (siehe [Kuester et al., 2006, S. 82]). Aber die größeren Zeitfenster, insbesondere für die BM-Methode, die mehrere Beobachtungen erfordert, bleiben noch unerforscht.

# A Anhang

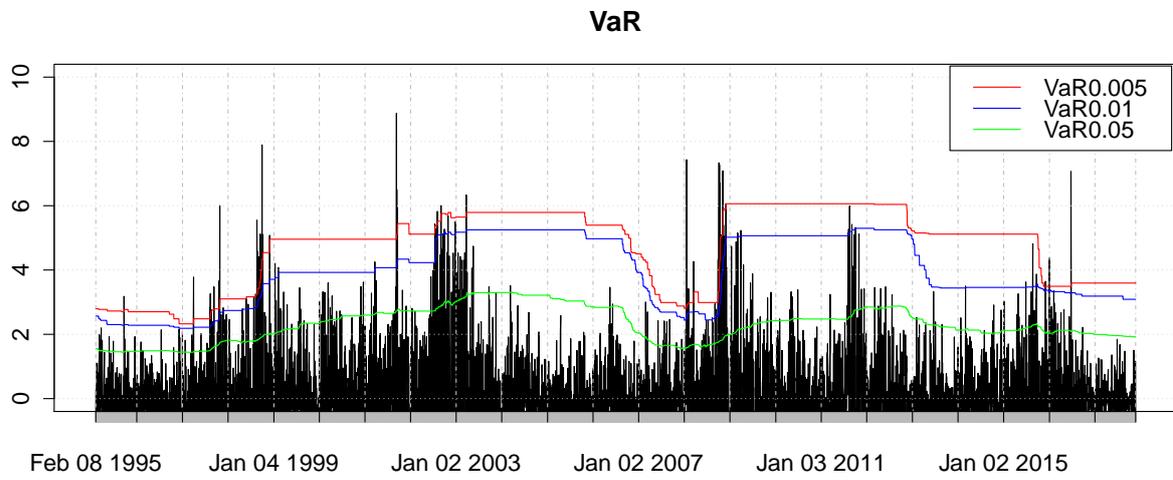


Abbildung 23: *VaR-Prognosen der historischen Simulation.*

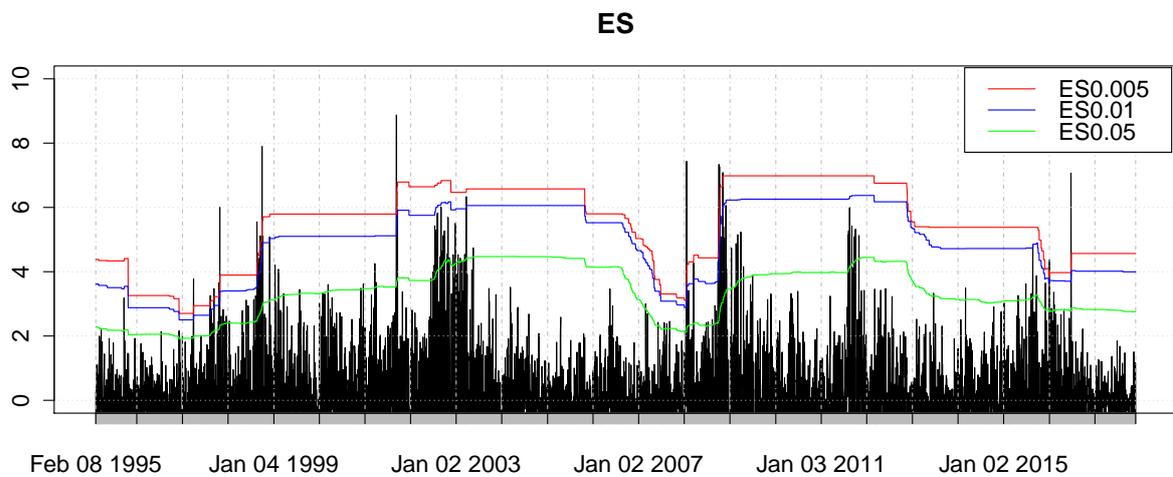


Abbildung 24: *ES-Prognosen der historischen Simulation.*

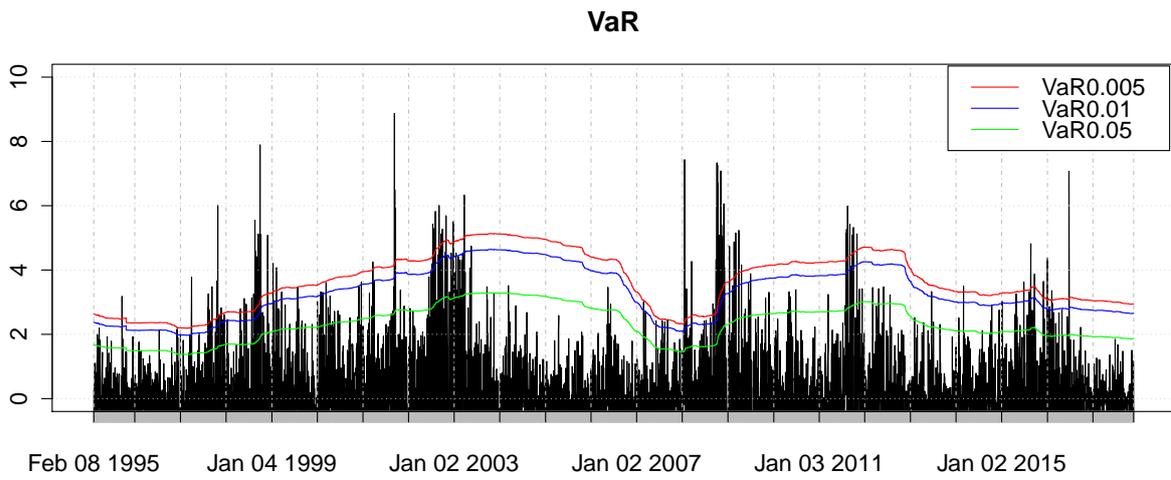


Abbildung 25: *VaR-Prognosen der normalen Verteilung.*

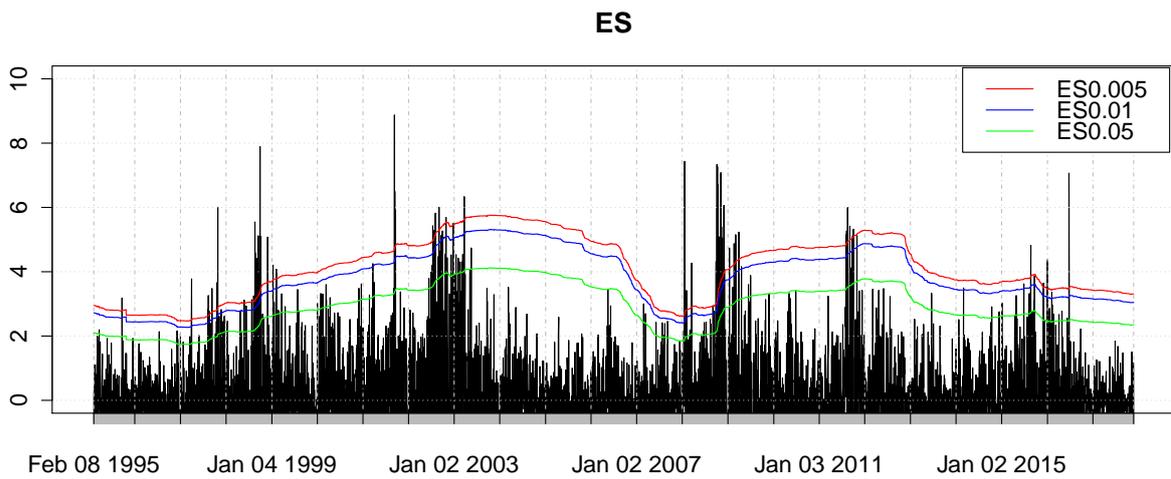


Abbildung 26: *ES-Prognosen der normalen Verteilung.*

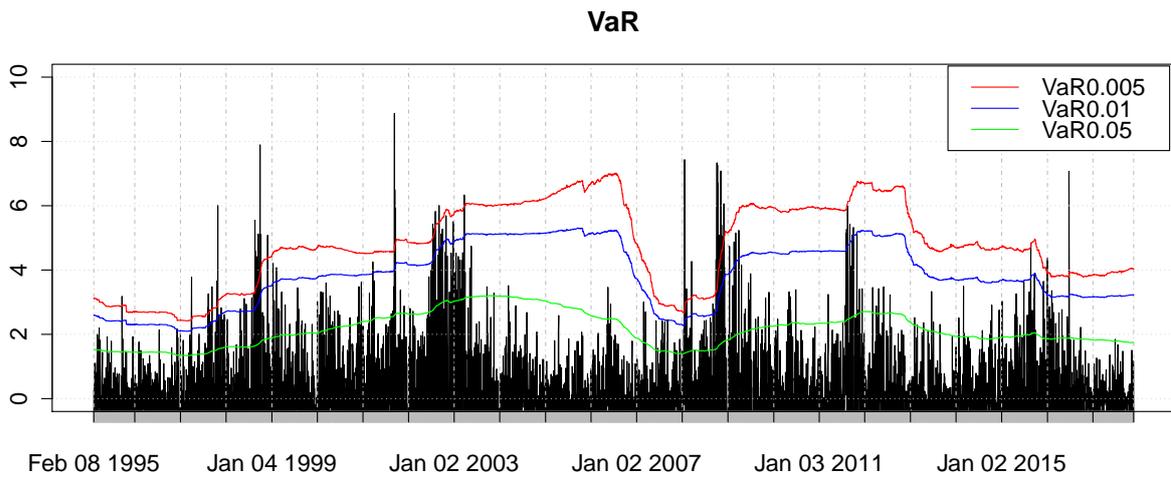


Abbildung 27: *VaR-Prognosen der Studentischen  $t$ -Verteilung.*

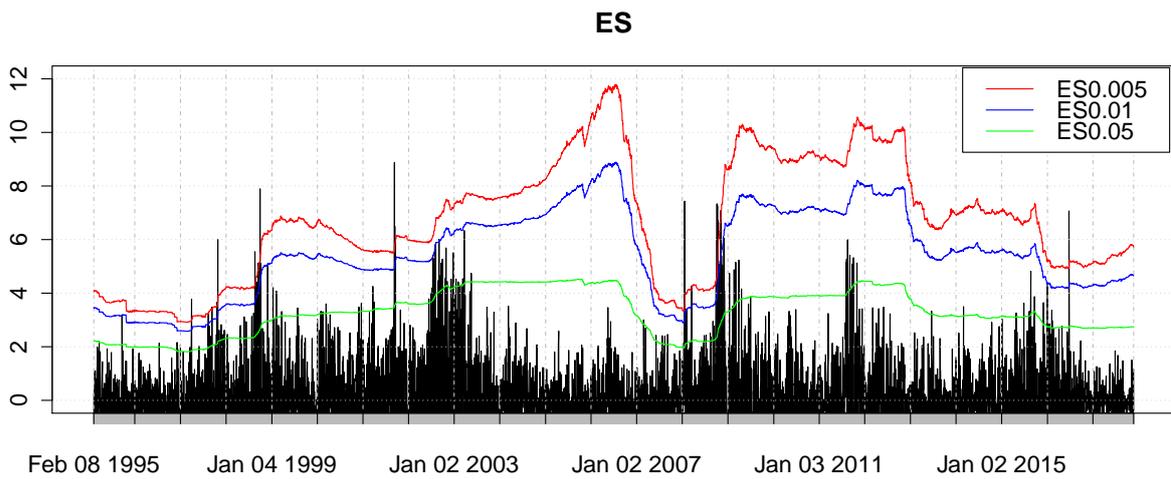


Abbildung 28: *ES-Prognosen der Studentischen  $t$ -Verteilung.*

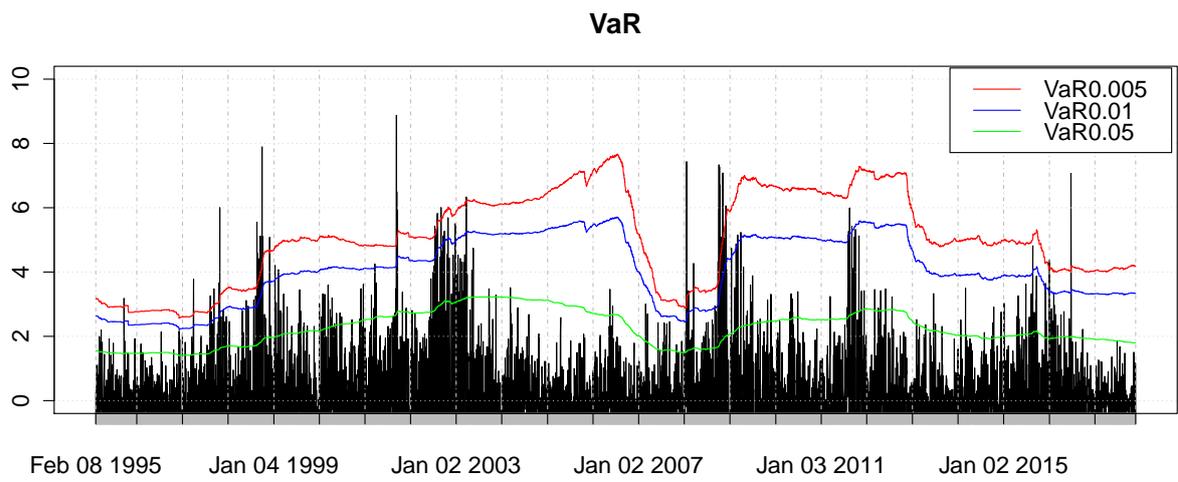


Abbildung 29: *VaR-Prognosen der Skewed  $t$ -Verteilung.*

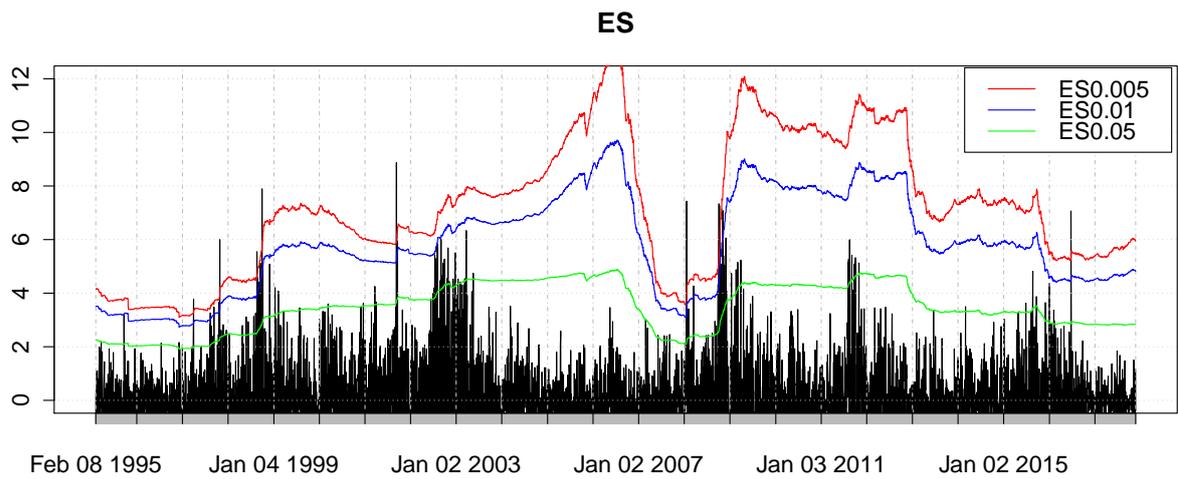


Abbildung 30: *ES-Prognosen der Skewed  $t$ -Verteilung.*

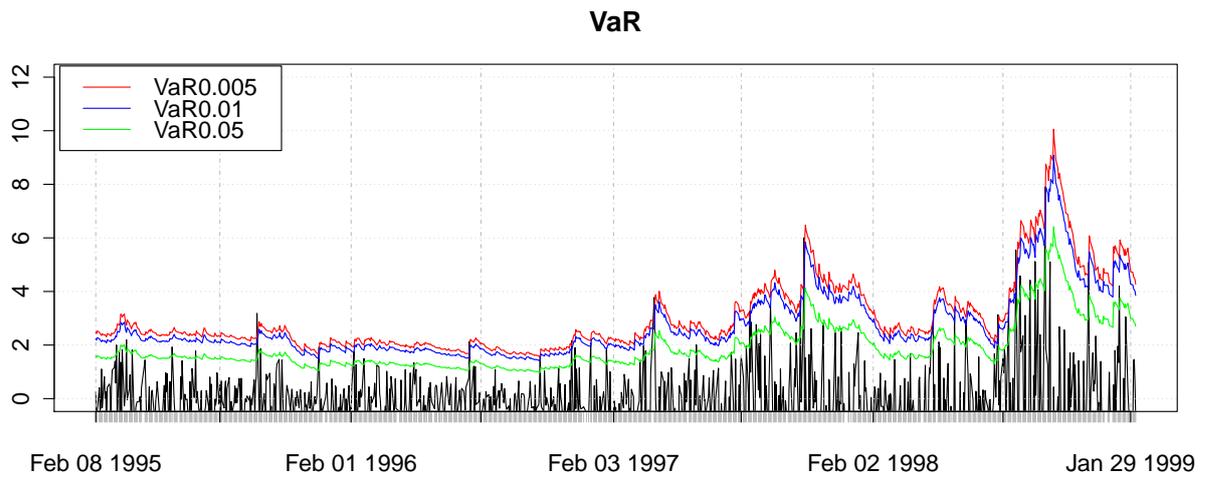


Abbildung 31: Die ersten 1000 VaR-Prognosen des GARCH-Modells mit der normalen Verteilung.

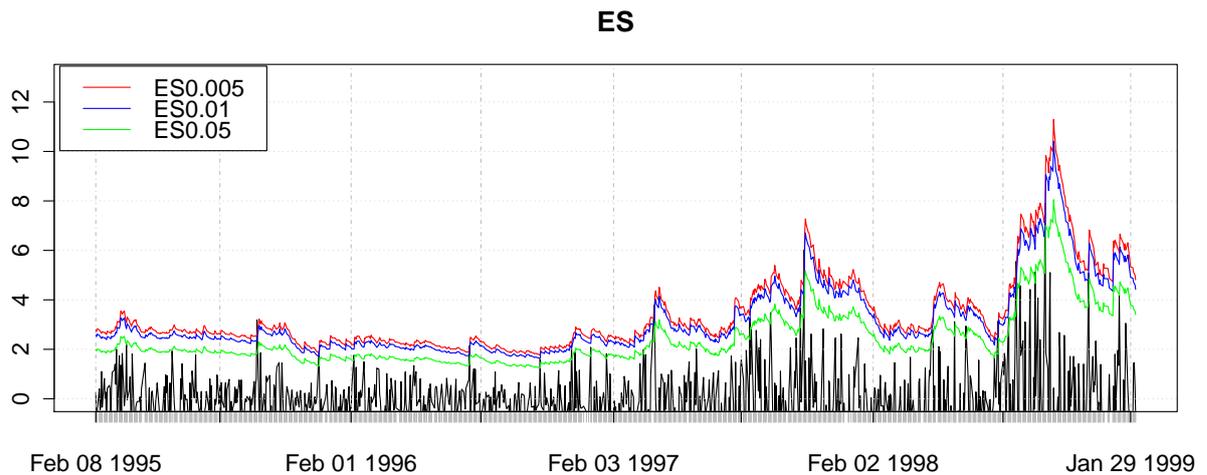


Abbildung 32: Die ersten 1000 ES-Prognosen des GARCH-Modells mit der normalen Verteilung.

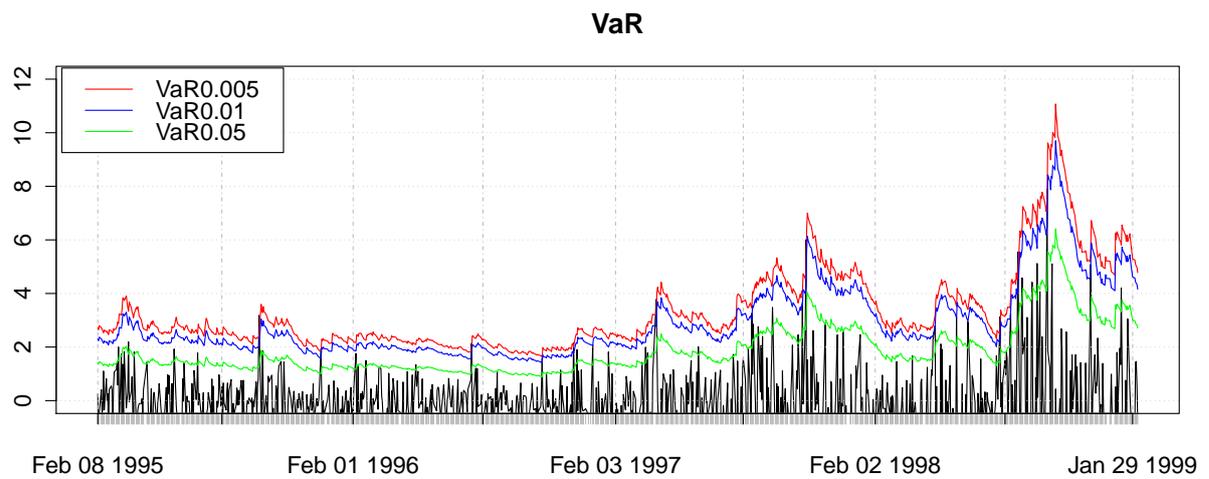


Abbildung 33: Die ersten 1000 VaR-Prognosen des GARCH-Modells mit der Studentschen  $t$ -Verteilung.

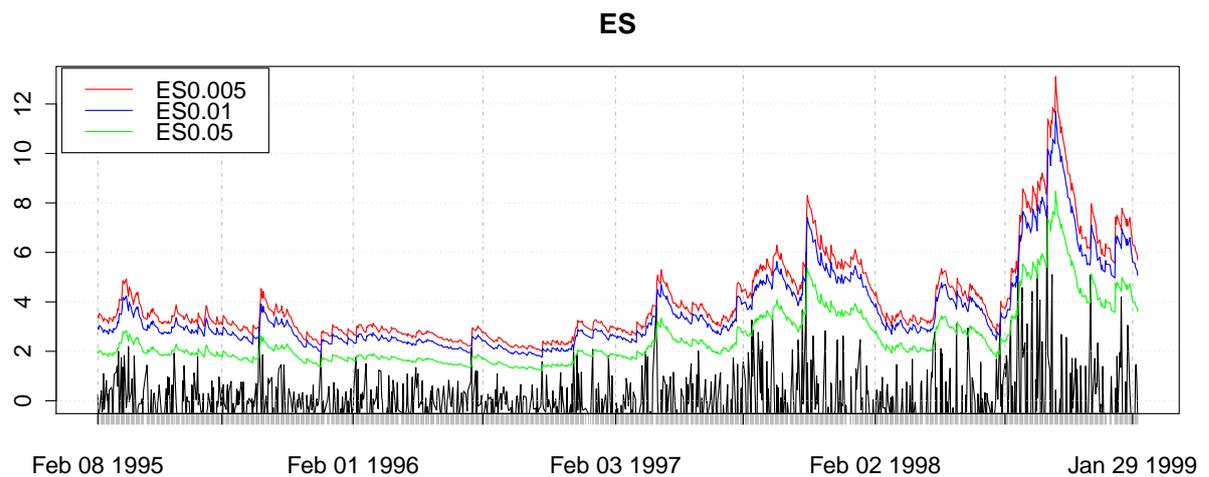


Abbildung 34: Die ersten 1000 ES-Prognosen des GARCH-Modells mit der Studentschen  $t$ -Verteilung.

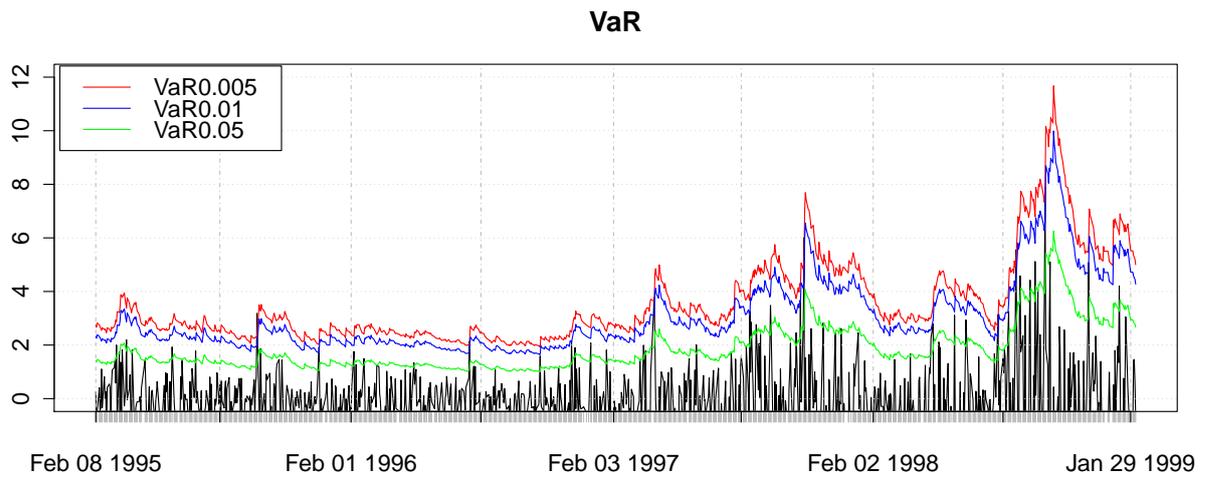


Abbildung 35: Die ersten 1000 VaR-Prognosen des GARCH-Modells mit der Skewed  $t$ -Verteilung.

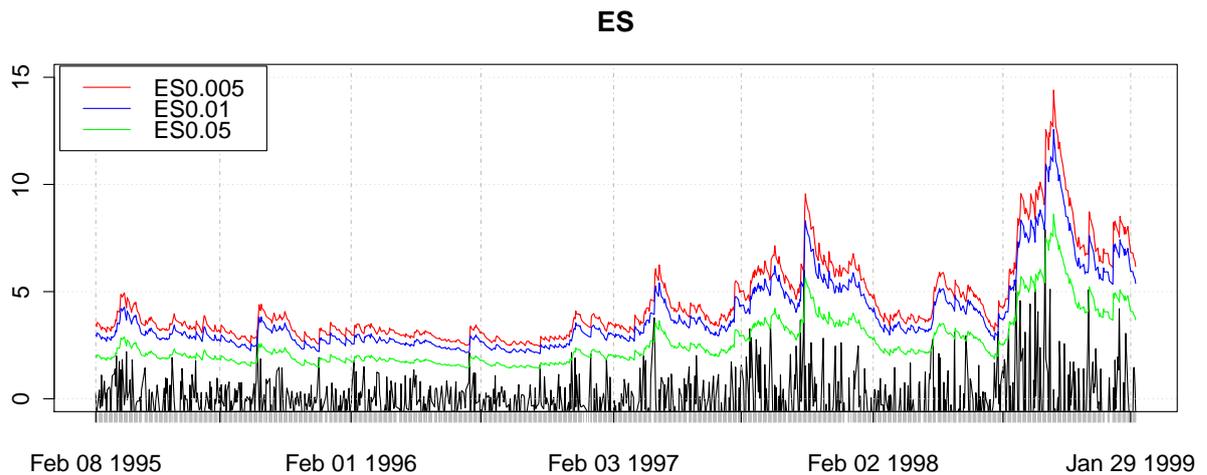


Abbildung 36: Die ersten 1000 ES-Prognosen des GARCH-Modells mit der Skewed  $t$ -Verteilung.

## B Verwendete Software

Alle Berechnungen und Grafiken in dieser Arbeit sind mit der freizugänglichen statistischen Software RStudio durchgeführt worden.

Abbildung 1 wurde mithilfe der Funktion *fitdistr()* aus dem *MASS* Paket (Version: 7.3-45) und der Funktion *dnorm()* aus dem *stats* Paket (Version: 3.2.5) erstellt. Zum Fitten der Daten durch die Skewed t-Verteilung wurde die Funktion *sstdFit()* aus dem *fGarch* Paket (Version: 3010.82.1) verwendet. Die gefittete Kurve in Abbildung 2 wurde durch die Funktion *dskt()* aus dem *skewt* Paket (Version: 0.1) erzeugt. Die Wahrscheinlichkeitsdichte der Extremwertverteilung 3 wurde mithilfe der Funktion *devd()* aus dem *extRemes* Paket (Version: 2.0-8) gezeichnet. Abbildung 4 wurde mittels der Funktion *devd()* aus dem *extRemes* Paket (Version: 2.0-8) erstellt. Die linke Grafik in Abbildung 5 wurde mithilfe der Funktion *mrlplot()* aus dem *extRemes* Paket (Version: 2.0-8) erstellt. Die rechte Grafik in Abbildung 5 wurde mithilfe der Funktion *meplot()* aus dem *evir* Paket (Version: 1.7-3) erstellt. Abbildung 6 wurde mittels der Funktion *hill()* aus dem *evir* Paket (Version: 1.7-3) gezeichnet. Zum Regression Based Expected Shortfall Backtesting wurden die Funktionen *esr\_backtest()* und *esr\_backtest\_intercept()* aus dem *esback* Paket (Version: 0.1.1) implementiert. Das Erstellen der monatlichen Maxima in Abbildung 10 wurde mithilfe der Funktion *apply.monthly()* aus dem *xts* Paket (Version: 0.9-7) gemacht. Die Schätzung der Parameter der GEV in Tabelle 3 war mithilfe der Funktion *fevd()* aus dem *extRemes* Paket (Version: 2.0-8) erfolgt. Die Schätzung der Parameter der GPD in Tabelle 9 wurde mithilfe der Funktion *fevd()* aus dem *extRemes* Paket (Version: 2.0-8) gemacht. Die Schätzung der Parameter des GARCH-Modells in Tabelle 12 wurde mithilfe der Funktion *garchFit()* aus dem *fGarch* Paket (Version: 3010.82.1) gemacht. Die Abbildung der Autokorrelationsfunktion 20 wurde mittels der Funktion *acf()* aus dem *stats* Paket (Version: 3.2.5) erstellt. Zur Implementierung des Varianz-Ratio-Tests in Kapitel 6.4 war die Funktion *Lo.Mac()* aus dem *vrtest* Paket (Version: 0.97) nötig. Zum Erstellen der Abbildungen 1, 2 und 3 sowie 4 wurde das *ggplot2* Paket (Version: 2.2.1) auch verwendet.

## C Inhalt der CD

Auf der beigelegten CD befinden sich folgende Ordner und Dateien:

- Ordner Masterarbeit:  
Enthält die angefertigte Masterarbeit im PDF-Format.
- Ordner Daten und R-Codes:  
Enthält die DAX-Daten vom 11.27.1990 bis zum 11.26.2017 im CSV-Format und die R-Programmcodes:
  - DAX.csv
  - Abschlussarbeit.R
  - AbbildungErstellen.R
  - workspace.RData
- Ordner Grafiken:  
Enthält die Grafiken im PDF-Format, die in der vorliegenden Masterarbeit dargestellt werden.
- Die Datei ReadMe.txt beschreibt den Inhalt der CD.

## D Urheberschaftserklärung

Hiermit erkläre ich, dass die vorliegende Masterarbeit von mir selbstständig verfasst wurde und dass keine anderen als die angegebenen Hilfsmittel benutzt wurden. Die Stellen der Arbeit, die anderen Werken dem Wortlaut oder Sinn nach entnommen sind, sind in jedem einzelnen Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht. Diese Erklärung erstreckt sich auch auf etwa in der Arbeit enthaltene Zeichnungen, Kartenskizzen und bildliche Darstellungen.

---

Ort, Datum

---

Unterschrift

## Literatur

- S. Aboura. When the us stock market becomes extreme? *Risks*, 2(2):211–225, 2014.
- C. Acerbi and B. Szekely. Back-testing expected shortfall. *Risk*, page 76, 2014.
- C. Acerbi and B. Szekely. General properties of backtestable statistics. 2017.
- C. Acerbi and D. Tasche. Expected shortfall: a natural coherent alternative to value at risk. *Economic notes*, 31(2):379–388, 2002.
- T. Angelidis and S. A. Degiannakis. Backtesting var models: An expected shortfall approach. 2006.
- P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.
- S. Bayer and T. Dimitriadis. Regression based expected shortfall backtesting. *arXiv preprint arXiv:1801.04112*, 2018.
- J. Berkowitz. Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics*, 19(4):465–474, 2001.
- J. Berkowitz and J. O’Brien. How accurate are value-at-risk models at commercial banks? *The journal of finance*, 57(3):1093–1111, 2002.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- M. Borkovec and C. Klüppelberg. Extremwerttheorie für finanzzeitreihen—ein unverzichtbares werkzeug im risikomanagement. *Handbuch Risikomanagement*, 1:219–244, 2000.
- R. Brummelhuis and R. Kaufmann. Time scaling for garch (1, 1) and ar (1)-garch (1, 1) processes. *Preprint, ETH Zurich*, 2004.
- L. Carver. Mooted var substitute cannot be back-tested, says top quant. *Risk*, March, 8, 2013.
- P. F. Christoffersen. Evaluating interval forecasts. *International economic review*, pages 841–862, 1998.
- S. Coles, J. Bawa, L. Trenner, and P. Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.
- B. Committee et al. Fundamental review of the trading book: A revised market risk framework. *Consultative Document, October*, 2013.

- J. Danielsson. *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab*, volume 588. John Wiley & Sons, 2011.
- J. Danielsson and J.-P. Zigrand. On time-scaling of risk and the square-root-of-time rule. *Journal of Banking & Finance*, 30(10):2701–2713, 2006.
- J. Danielsson, P. Hartmann, and C. de Vries. The cost of conservatism. *Risk*, 11(1):101–103, 1998.
- J. Danielsson, L. de Haan, L. Peng, and C. G. de Vries. Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate analysis*, 76(2):226–248, 2001.
- J. Danielsson, B. N. Jorgensen, S. Mandira, G. Samorodnitsky, and C. G. De Vries. Sub-additivity re-examined: the case for value-at-risk. Technical report, Cornell University Operations Research and Industrial Engineering, 2005.
- S. Degiannakis and A. Potamia. Multiple-days-ahead value-at-risk and expected shortfall forecasting for stock indices, commodities and exchange rates: Inter-day versus intra-day data. *International Review of Financial Analysis*, 49:176–190, 2017.
- F. Diebold. Pitfalls and opportunities in the use of extreme value theory in risk management [working paper]. *Wharton School, University of Pennsylvania*, 1998.
- T. Dimitriadis and S. Bayer. A joint quantile and expected shortfall regression framework. *arXiv preprint arXiv:1704.02213*, 2017.
- P. Embrechts, C. Kluppelberg, and T. Mikosch. Modelling extremal events: With a view towards insurance and finance, 1997.
- S. Emmer, M. Kratz, and D. Tasche. What is the best risk measure in practice? a comparison of standard measures. 2015.
- C. Fernández and M. F. Steel. On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371, 1998.
- R. Gencay and F. Selcuk. Extreme value theory and value-at-risk: Relative performance in emerging markets. *International Journal of Forecasting*, 20(2):287–303, 2004.
- P. Giot and S. Laurent. Value-at-risk for long and short trading positions. *Journal of Applied Econometrics*, 18(6):641–663, 2003.
- T. Gneiting. Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494):746–762, 2011.

- C. Gouriéroux, A. Monfort, and A. Trognon. Pseudo maximum likelihood methods: Theory. *Econometrica: Journal of the Econometric Society*, pages 681–700, 1984.
- E. Gumbel. Statistics of extremes. 1958. *Columbia Univ. press, New York*, page 247, 1958.
- B. M. Hill et al. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 3(5):1163–1174, 1975.
- P. Jorion. Value at risk. 2000.
- J. Kerkhof and B. Melenberg. Backtesting for risk-based regulatory capital. *Journal of Banking & Finance*, 28(8):1845–1865, 2004.
- K. Kuester, S. Mittnik, and M. S. Paolella. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89, 2006.
- T. J. Linsmeier and N. D. Pearson. Value at risk. *Financial Analysts Journal*, 56(2):47–67, 2000.
- A. W. Lo and A. C. MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, 1(1):41–66, 1988.
- F. M. Longin. The asymptotic distribution of extreme stock market returns. *Journal of business*, pages 383–408, 1996.
- F. M. Longin. From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, 24(7):1097–1130, 2000.
- B. Mandelbrot. Bb mandelbrot, j. business 36, 394 (1963). *J. Business*, 36:394, 1963.
- C. Marinelli, S. d’Addona, and S. T. Rachev. A comparison of some univariate models for value-at-risk and expected shortfall. *International Journal of Theoretical and Applied Finance*, 10(06):1043–1075, 2007.
- A. J. McNeil. Calculating quantile risk measures for financial return series using extreme value theory. Technical report, ETH Zurich, 1998.
- A. J. McNeil and R. Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, 7(3):271–300, 2000.
- A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: Concepts, techniques and tools*. Princeton university press, 2015.
- F. Moldenhauer and M. Pitera. Backtesting expected shortfall: is it really that hard? 2017.

- N. Nolde, J. F. Ziegel, et al. Elicitability and backtesting: Perspectives for banking regulation. *The Annals of Applied Statistics*, 11(4):1833–1874, 2017.
- B. C. on Banking Supervision. *Overview of the amendment to the capital accord to incorporate market risks*. Bank for International Settlements, 1996.
- A. Ozun, A. Cifter, and S. Yilmazer. Filtered extreme-value theory for value-at-risk estimation: evidence from turkey. *The Journal of Risk Finance*, 11(2):164–179, 2010.
- C. Peitz. *Die parametrische und semiparametrische Analyse von Finanzzeitreihen: neue Methoden, Modelle und Anwendungsmöglichkeiten*. Springer-Verlag, 2016.
- M. B. Righi and P. S. Ceretta. Individual and flexible expected shortfall backtesting. 2013.
- S. Roccioletti. *Backtesting Value at Risk and Expected Shortfall*. Springer, 2015.
- D. Ruppert. *Statistics and data analysis for financial engineering*, volume 13. Springer, 2011.
- B. Sherman. Percentiles of the  $\omega_n$  statistic. *The Annals of Mathematical Statistics*, 28(1):259–261, 1957.
- R. L. Smith. Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90, 1985.
- J.-N. Wang, J.-H. Yeh, and N. Y.-P. Cheng. How accurate is the square-root-of-time rule in scaling tail risk: A global study. *Journal of Banking & Finance*, 35(5):1158–1169, 2011.
- W. K. Wong. Backtesting trading risk of commercial banks using expected shortfall. *Journal of Banking & Finance*, 32(7):1404–1415, 2008.
- Y. Yamai, T. Yoshihara, et al. On the validity of value-at-risk: comparative analyses with expected shortfall. *Monetary and economic studies*, 20(1):57–85, 2002.