

Technische Universität München

FAKULTÄT FÜR MATHEMATIK

**Neue Schätzmethode des
extremalen Index**

Diplomarbeit

von

Manuela Regler

Aufgabenstellerin: Prof. Dr. Claudia Klüppelberg

Betreuer: Dr. Gernot Müller

Abgabetermin: 15. Oktober 2004

Ich erkläre hiermit, dass ich die Diplomarbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

Taufkirchen, 15. Oktober 2004

Danksagung

Diese Arbeit entstand in Zusammenarbeit mit dem Lehrstuhl für Mathematische Statistik der Technischen Universität München. Ich möchte mich sehr herzlich bei Prof. Dr. Claudia Klüppelberg bedanken, die mir diese Arbeit ermöglicht hat und von der ich zu jeder Zeit Unterstützung erhielt. Mein besonderer Dank gilt auch Dr. Gernot Müller, meinem Diplomarbeitsbetreuer, der mich bei Problemen tatkräftig unterstützt hat.

Inhaltsverzeichnis

1	Einleitung	1
2	POT Methode	3
2.1	Standard POT Methode	3
2.2	Die POT Methode für stationäre Zeitreihen	6
3	Die Exzessfunktion	9
4	Der extreme Index	13
5	Alte Schätzmethoden des extremalen Index	19
5.1	Blockmethode	19
5.2	Der extreme Index als Kehrwert des Erwartungswertes	22
5.3	run Methode	23
6	Neue Methode 1:	
	Theoretische Momenten Methode	25
6.1	Asymptotische Verteilung	25
6.2	Schätzung des extremalen Index	27
6.2.1	Intervallschätzer	28
6.2.2	Maximum Likelihood Schätzer	33
6.3	Bootstrap Prozess	34
7	Neue Methode 2:	
	Kombination der subject-matter Methode und der run Methode	37
7.1	Zwei-Schwellen Index	38
7.2	Volatilitäts-Schwellen Index	39
7.3	Schätzung des extremalen Index θ	41
8	Simulationsstudie	43
8.1	Unabhängige Prozesse	43
8.2	ARMAX Prozess	44
8.3	AR Prozess	48
8.4	ARCH Prozess	49
8.5	GARCH Prozess	50

9	Vergleich der Methoden	53
10	Datenanalyse	57
10.1	S&P 500	57
10.2	FTSE 100	75
11	Abschliessende Bemerkungen	93
A		95
A.1	Beweis Lemma 6.5	95
A.2	Verwendete Funktionen in SPlus	98
A.2.1	Voreingestellte Funktionen von EVIS	98
A.2.2	Berechnung der Schwellenwerte	99
A.2.3	Simulationsstudie	99
A.2.4	Clusteridentifizierung mit den verschiedenen Methoden	101
A.2.5	Berechnung von θ mit den verschiedenen Methoden	106
A.2.6	Fréchet-, Weibull- und Gumbel-Verteilung	108
A.2.7	GPD	110
	Abbildungsverzeichnis	112
	Tabellenverzeichnis	119
	Literaturverzeichnis	121

Kapitel 1

Einleitung



Erdbeben, Sturmfluten und Taifune richten immer gewaltigere Schäden an, die von Versicherungen ersetzt werden müssen. Immense Kursverluste belasten die finanzielle Lage aller Unternehmen. Dies sind große Risikofaktoren für die Firmen. Krisen in der Vergangenheit haben gezeigt, dass professionelles Risikomanagement zur Überlebensfrage für Unternehmen werden kann. Dabei spielen extreme Ereignisse eine große Rolle.

Als Standardwerkzeug wird in Banken der Value at Risk (VaR) als Risikomaß verwendet, der als das 5%-Quantil der Gewinn-Verlustverteilung, auch Profit-Loss Verteilung genannt, definiert wird. Die Schätzung eines solch kleinen Quantils ist sehr schwierig, da man Aussagen über das extreme Verhalten von Datensätzen machen will. Jedoch werden hohe Risiken oft zu optimistisch bewertet.

Ein wichtiges mathematisches Mittel, um Risiken besser zu modellieren und zu messen, liefert die Extremwerttheorie. Dabei spielt der extremale Index eine spezielle Rolle. Mit Methoden wie der POT Methode versucht man den VaR besser zu schätzen. Wichtig ist es dabei, Cluster richtig zu identifizieren. Ein Cluster in den Extrema bedeutet, dass die Volatilität der Exzedenten so große Ausschläge zeigt, dass sie über einer hohen Schwelle sichtbar wird. Zur Schätzung des extremalen Index und zur Beschreibung der Cluster sind Methoden wie die Blockmethode und die run Methode entwickelt worden.

Jedoch sind diese Methoden der Schätzung des extremalen Index eher naiv. Es gibt dafür neuere Methoden, die in dieser Arbeit vorgestellt werden. Ziel dieser Methoden ist es, einen neuen Weg der Clusteridentifizierung zu finden und neue Schätzmethoden anzuwenden.

Kapitel 2

POT Methode

Die wohl bekannteste Methode um Tails oder ein Quantil zu schätzen ist die POT Methode („Peaks Over Threshold“). Die Methode setzt sich aus drei Komponenten zusammen. In diesem Kapitel soll die POT Methode nur kurz beschrieben werden, damit man den Zusammenhang zwischen dem extremalen Index und der Schätzung von Quantilen versteht. Mehr Informationen dazu findet man unter Emmer, Klüppelberg und Trüstedt [6] und Borkovec und Klüppelberg [2].

2.1 Standard POT Methode

1. Grenzprozess des Punktprozesses

Der erste Schritt ist, einen Grenzprozess für den Punktprozess der Überschreitungen von hohen Schwellen herzuleiten. Wir betrachten den Prozess der Beobachtungen $(X_i)_{i=1}^n$. Die Anzahl der Beobachtungen, die die Schwelle u überschreiten, also die Anzahl der Exzedenten, sei

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

Im Kapitel 4 (Satz 4.1) kann man sehen, dass die Verteilung von N_u gegen einen Poissonprozess mit Intensität τ konvergiert, falls $n P(X_i > u_n) \rightarrow \tau$.

2. Verallgemeinerte Paretoverteilung

Die Verteilungsfunktion der verallgemeinerten Paretoverteilung (GPD) ist gegeben durch:

$$G(x)_{\xi,\beta} = \begin{cases} 1 - (1 + \xi \frac{x}{\beta})^{-1/\xi} & \text{wenn } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{wenn } \xi = 0, \end{cases} \quad (2.1)$$

wobei

$$\begin{cases} x \geq 0 & \text{wenn } \xi \geq 0, \\ 0 \leq x \leq -\beta/\xi & \text{wenn } \xi < 0. \end{cases} \quad (2.2)$$

Die Dichte der GPD lässt sich leicht durch Differenzieren berechnen. Die GPD und die Dichte der GPD sind für verschiedenen ξ in Abbildung 2.1 und Abbildung 2.2 dargestellt.

Y sei ein Exzess und sei definiert durch $Y := X - u$. Mit Y_1, \dots, Y_{N_u} bezeichnen wir die Exzesse der Beobachtungen X_1, \dots, X_n bezüglich der Schwelle u . Man kann zeigen, dass für eine reelle Funktion $a(u)$ gilt:

$$P\left(\frac{X - u}{a(u)} > y \mid X > u\right) = (1 + \xi y)^{-1/\xi}$$

Dies ist der Tail einer GPD. Mehr Informationen zur GPD findet man unter Embrechts, Klüppelberg und Mikosch [5] (S. 165).

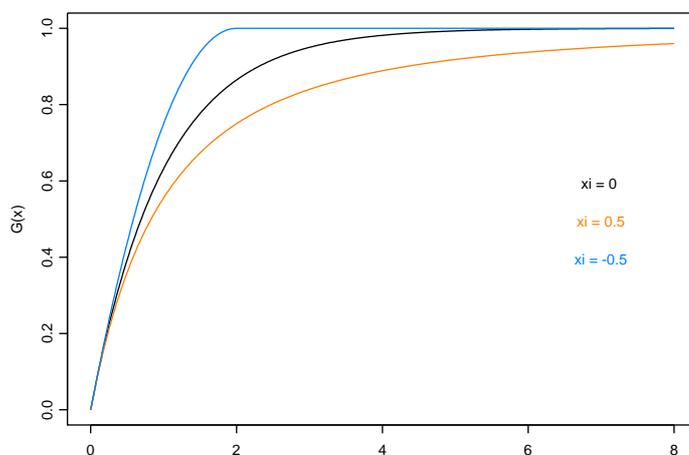


Abbildung 2.1: Verteilungsfunktion der GPD für drei verschiedene ξ und $\beta = 1$

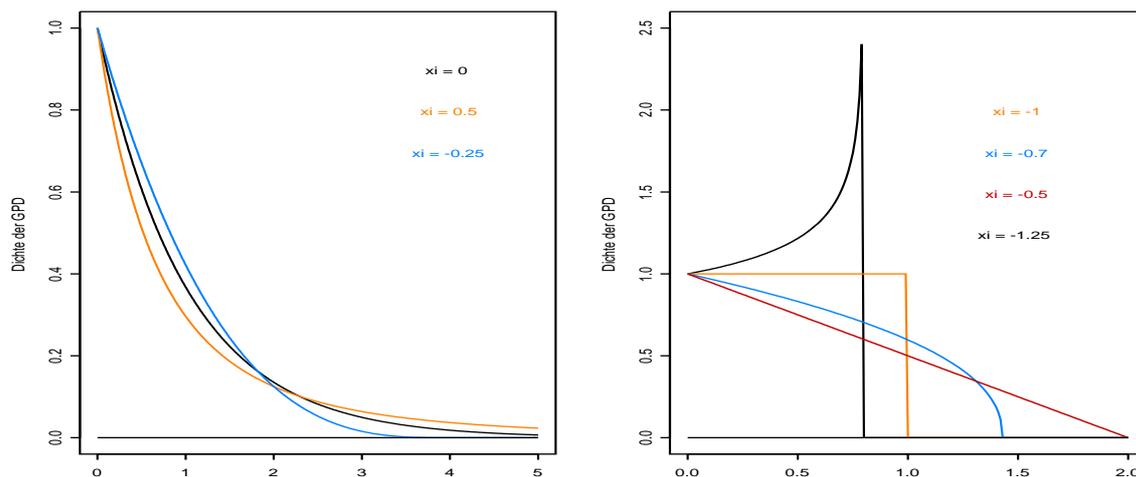


Abbildung 2.2: Dichte der GPD für verschiedene ξ und $\beta = 1$

3. Unabhängigkeit

Sei $\bar{F} = 1 - F$ und $\bar{F}_u(y) = P(Y_1 > y | X > u)$. Dann ist klar, dass

$$\bar{F}_u(y) = P(X - u > y | X > u) = \frac{\bar{F}(u + y)}{\bar{F}(u)}.$$

Folglich gilt, dass

$$\bar{F}(u + y) = \bar{F}(u) \bar{F}_u(y). \quad (2.3)$$

Schätzt man nun $\bar{F}(u)$ mit

$$\hat{\bar{F}} = \frac{N_u}{n}$$

und $\bar{F}_u(y)$ mit

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

so gilt wegen (2.3)

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

Somit ist

$$\hat{x}_q = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1 - q) \right)^{-\hat{\xi}} - 1 \right) \quad (2.4)$$

das q -Quantil der Verteilung F .

$\hat{\beta}$ und $\hat{\xi}$ werden durch die Maximum Likelihood Methode (ML) oder den Hillschätzer (vgl. Embrechts, Klüppelberg und Mikosch [5] (S. 330ff, 336)) geschätzt. Da in die Schätzung nur die Exzesse Y_1, \dots, Y_{N_u} einfließen, hat man zur Schätzung oft nur wenig Daten zur Verfügung. Aus diesem Grund kann man sich nicht auf die asymptotischen Optimalitätseigenschaften des MLE verlassen.

Die Verteilung der Exzesse $(Y_i)_{i=1}^{N_u}$ sei eine GPD mit der Dichte:

$$f(y) = \frac{1}{\beta} \left(1 + \xi \frac{y}{\beta}\right)^{-\frac{1}{\xi}-1}, \quad (2.5)$$

mit $\xi \neq 0$ und y wie in (2.2).

Die Maximum Likelihood Funktion von β und ξ ergibt sich demnach aus:

$$l((\xi, \beta); Y) = -N_u \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{N_u} \ln \left(1 + \frac{\xi}{\beta} (X_i - u)\right)$$

Setzt man nun für $q = 5\%$ so bekommt man das 5%-Quantil der Verteilung der Daten und somit das Downside Risk Capital (DRC). Der VaR ist das DRC korrigiert durch den Erwartungswert. Damit hat man $DRC = \hat{x}_{0.05}$ und $VaR = E(X) - DRC$, dessen Schätzung bei der Risikomessung das Ziel ist.

2.2 Die POT Methode für stationäre Zeitreihen

Das ganze ändert sich nun ein wenig bei stationären Zeitreihen. Im stationären Fall haben wir meist keine iid verteilten Zufallsvariablen, also bildet man eine assoziierte Folge von iid verteilten Zufallsvariablen. Jedoch die Vorgehensweise der POT Methode in diesem Fall ist dieselbe.

Sei $(X_i)_{i=1}^n$ eine Folge stationärer Beobachtungen, $(\tilde{X}_i)_{i=1}^n$ die assoziierte unabhängige Folge und F sei die dazugehörige Verteilungsfunktion.

Sei $M_n = \max(X_1, \dots, X_n)$, $\tilde{M}_n = \max(\tilde{X}_1, \dots, \tilde{X}_n)$ und u_n sei der Schwellenwert der Daten X_1, \dots, X_n .

Es gilt wieder

$$nP(X_i > u_n) \rightarrow \tau,$$

da hierzu die Daten nicht unabhängig sein müssen. Somit geht die Verteilung von N_u wieder gegen einen Poissonprozess mit Intensität τ .

Für die assoziierten unabhängigen Folgen gilt weiter (vgl. Embrechts, Klüppelberg und Mikosch [5], Prop. 3.1.1, S. 116), dass

$$P(\tilde{M}_n \leq u_n) \rightarrow e^{-\tau}.$$

Für die nicht iid verteilten Zufallsvariablen gilt demnach, wegen der Definition des extremalen Index (Definition 4.3)

$$P(M_n \leq u_n) \approx P(\tilde{M}_n \leq u_n)^\theta \rightarrow (e^{-\tau})^\theta = e^{-\theta\tau}. \quad (2.6)$$

Für eine hohe Schwelle u sei nun

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

und der Punktprozess ist wegen des Satzes von Poisson (vgl. Satz 4.1) ein Poissonprozess mit Intensität $\theta\tau$. Dies zeigt, dass die Cluster unabhängig voneinander sind.

Zur Annäherung der GPD werden nun die Clustermaxima herangezogen, d.h.

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

wobei $M^{(i)}$ das Maximum im i -ten Cluster ist.

Mit den Methoden, die später in dieser Arbeit näher beschrieben werden, können die Cluster identifiziert werden bzw. die Maxima berechnet und θ geschätzt werden.

Nun schätzt man $\overline{F}(u)$ mit

$$\widehat{\overline{F}} = \frac{N_u^C}{n \widehat{\theta}_u},$$

d.h. es wird mit $1/\widehat{\theta}_u$, der mittleren Clustergröße ¹, multipliziert.

Weiterhin gilt wie bei nichtstationären Zeitreihen $\widehat{\overline{F}}_u(y) = \left(1 + \widehat{\xi} \frac{y}{\widehat{\beta}}\right)^{-1/\widehat{\xi}}$ und damit

$$\widehat{\overline{F}}(u + y) = \frac{N_u^C}{n \widehat{\theta}_u} \left(1 + \widehat{\xi} \frac{y}{\widehat{\beta}}\right)^{-1/\widehat{\xi}}.$$

Somit kann man das q -Quantil x_q wieder mit

$$\hat{x}_q = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left(\left(\frac{n \widehat{\theta}_u}{N_u^C} (1 - q) \right)^{-\widehat{\xi}} - 1 \right) \quad (2.7)$$

bestimmen, wobei $\widehat{\beta}$ und $\widehat{\xi}$ wieder durch Maximum Likelihood oder den Hill-Schätzer geschätzt werden können. Die Schätzmethoden für θ werden in den folgenden Kapiteln besprochen.

Man zieht zur Schätzung nur die Clustermaxima $M^{(i)}$ heran, die laut Borkovec und Klüppelberg [2] (S. 17) pseudo-iid verteilt sind. Die Verteilung der Maxima sei eine GPD und die Dichte hat die Gestalt (2.5).

Die Maximum Likelihood Funktion von β und ξ ergibt sich demnach aus:

$$l((\xi, \beta); X) = -N_u^C \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{N_u^C} \ln\left(1 + \frac{\xi}{\beta}(M^{(i)} - u)\right)$$

Für $q = 5\%$ bekommt man nach der Schätzung der drei Parameter und Berechnung vom DRC wieder den Schätzer für den VaR.

Besondere Beachtung verdient folgende Tatsache

Je nach Datensatz ergibt sich ein unterschiedlicher Risikobereich, der entweder im 5%-Quantil oder im 95%-Quantil liegt. Untersucht man z.B. Schadensdaten durch Naturkatastrophen, so sind hohe Ausschläge nachteilig. Also liegt das Risiko oberhalb des 95%-Quantils. Zur VaR Berechnung muß man dann $q = 0.95$ setzen. Untersucht man jedoch Indexkurse, so ist es ein großes Risiko, falls der Kurs fällt, negativ ist. Demnach liegt das Risiko im 5%-Quantil und $q = 0.05$ zu wählen. In Kapitel 10 werden die negativen Log Return von Indexkursen analysiert, darum ist wieder das 95%-Quantil für die Schätzung des VaR zu wählen.

Eine Beispielrechnung der POT Methode angewendet an die Datensätze S&P 500 und FTSE 100 findet man unter Kapitel 10.

¹Die mittlere Clustergröße wird im Kapitel 4 näher erläutert

Kapitel 3

Die Exzessfunktion

Eine exzellente explorative Methode, die Tails zu unterscheiden, ist die Exzessfunktion, die für positive Daten entwickelt wurde. Um sie daher auf den Datensatz S&P 500 und FTSE 100 anwenden zu können, nehmen wir zur Untersuchung nur die positiven Daten der Zeitreihe. Nähere Beschreibungen findet man im Kapitel 10 in der Datenanalyse.

Sei $z^+ = \max(z, 0)$ der Positivteil von z und $\#A$ die Anzahl der Elemente in einer endlichen Menge A . Die empirische Funktion:

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

schätzt die Exzessfunktion:

$$e(u) = E(X - u | X > u), \quad u \geq 0$$

Die Tabelle 3.1 beinhaltet einen kleinen Überblick über einige Standardverteilungen. Davon sind vier Exzessfunktionen in Abbildung 3.1 zusammengefasst.

Exponential	λ^{-1}
Gamma	$\beta^{-1} \left(1 + \frac{\alpha-1}{\beta u} + o\left(\frac{1}{u}\right) \right)$
Lognormal	$\frac{\sigma^2 u}{\ln u - \mu} (1 + o(1))$
Pareto	$\frac{\kappa+u}{\alpha-1}, \alpha > 1$
Weibull	$\frac{u^{1-\tau}}{c\tau} (1 + o(1))$

Tabelle 3.1: empirische Exzessfunktion von verschiedenen Standardverteilungen

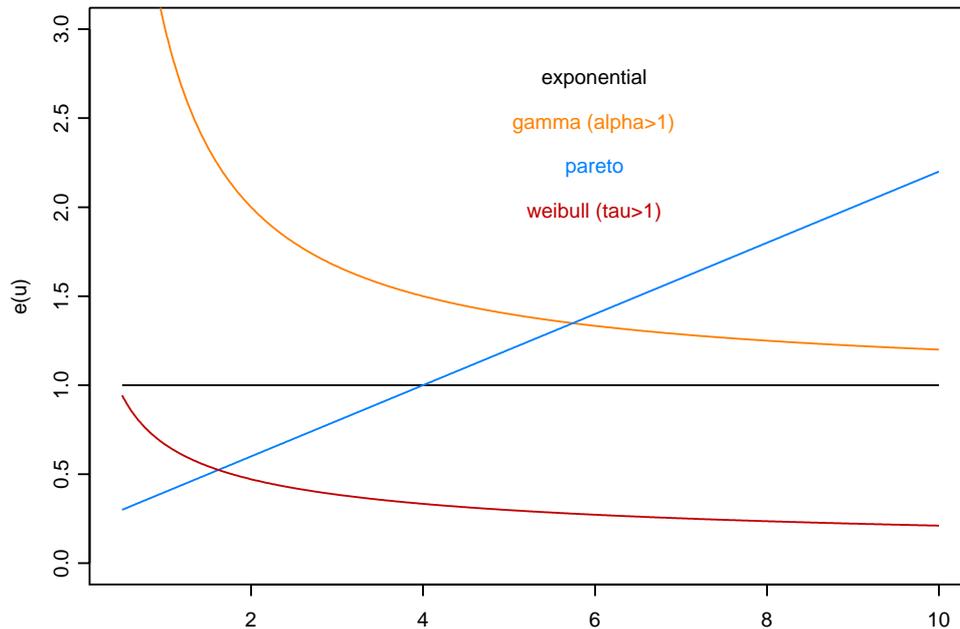


Abbildung 3.1:

Die empirische Exzessfunktion von verschiedenen Standardverteilungen unter Vernachlässigung der Restdaten.

Die Exzessfunktion der Exponentialverteilung ist eine Konstante, nämlich gerade ihr Parameter. Die Exzessfunktion von Verteilungen, deren Tail leichter ist als der Tail der Exponentialverteilung, fällt gegen Null. Bei Verteilungen mit fetterem Tail strebt die Exzessfunktion gegen unendlich.

Um eine gute Schätzung für den VaR zu bekommen, muß der Schwellenwert u geeignet gewählt werden. Hierfür ist die Exzessfunktion nützlich. Im Falle einer verallgemeinerten Paretoverteilung steigt die Exzessfunktion monoton linear. Eine mögliche Wahl von u ist somit der Wert, ab dem die empirische Exzessfunktion linear nach oben strebt. Es ist natürlich schwierig eine eindeutige Wahl zu treffen.

So kann man zum Beispiel für den Datensatz S&P 500 $u = 0.009$ (vgl. Abbildung 10.5) und für FTSE 100 $u = 0.008$ (vgl. Abbildung 10.23) wählen. Diese Werte wurden auch für die VaR Berechnung mit der POT Methode in Tabelle 10.4 und Tabelle 10.8 verwendet.

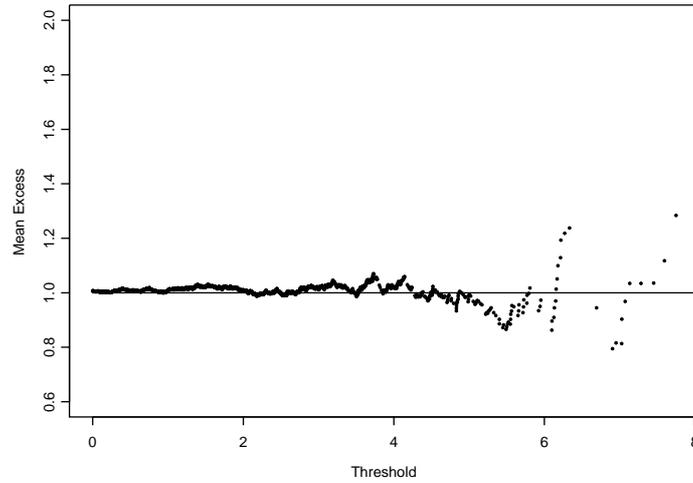


Abbildung 3.2:
Die empirische Exzessfunktion simulierter Daten ($n = 10000$) (gepunktet) im Vergleich zur theoretischen Exzessfunktion (Linie) einer Standard Exponentialverteilung.

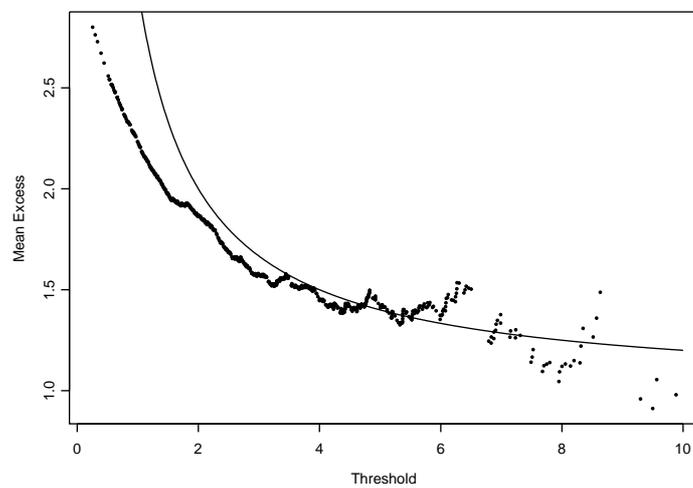


Abbildung 3.3:
Die empirische Exzessfunktion simulierter Daten ($n = 10000$) (gepunktet) im Vergleich zur theoretischen Exzessfunktion (Linie) einer Gamma-Verteilung mit Parametern $\beta = 1$ und $\alpha > 0$.

Kapitel 4

Der extreme Index

Der extreme Index ist eine Größe, die die Beziehung zwischen der Abhängigkeitsstruktur von Daten und ihrem extremen Verhalten beschreibt. Bei der Schätzung des Index besteht das gleiche Problem wie bei der Schätzung von kleinen Quantilen, nämlich das Problem der kleinen Stichproben.

Um zu erklären, was der extreme Index mit der Clusterbildung zu tun hat, betrachtet man die klassische Extremwerttheorie. Die Extremwerttheorie beschreibt das Verhalten der grössten Werte einer Stichprobe bzw. der Exzedenten über einer hohen Schwelle. Bei einer Stichprobe X_1, \dots, X_n , die iid verteilt ist, betrachtet man die partiellen Maxima

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

Die Anzahl der Beobachtungen, die die Schwelle u_n überschreiten, ist

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

Diese Beobachtungen sind natürlich binomialverteilt mit Parameter n und der Wahrscheinlichkeit $P(X_i > u_n)$.

Betrachtet man nun

$$nP(X_i > u_n) = \tau \in (0, \infty), \quad (4.1)$$

so konvergiert nach dem klassischen Satz von Poisson die Verteilung von $\#\{i : X_i > u_n; i = 1, \dots, n\}$ gegen eine Poissonverteilung mit Parameter τ .

Satz 4.1. (*Satz von Poisson*)

Sei $X \sim \text{bin}(n, p_n)$ verteilt und $np_n \xrightarrow{n \rightarrow \infty} \tau$, dann gilt

$$\lim_{n \rightarrow \infty} P(X_n = k) = \lim_{n \rightarrow \infty} \text{bin}(k, p_n) = \frac{\tau^k}{k!} e^{-\tau} = \text{Poi}(\tau)$$

Beweis. vgl. Behnen und Neuhaus [1] (Lemma 3.3.1 S. 118)

□

Markiert man in dem Intervall $[0, n]$ die Punkte $\{i : X_i > u_n; i = 1, \dots, n\}$, so wird das Intervall immer größer, die markierten Punkte werden immer seltener, da die Schwelle u_n mit n zunimmt. Die Schwelle u_n findet man dadurch, indem man die Anzahl der Exzedenten vorgibt.

Deswegen bekommt man eine bessere Darstellung, wenn man das Intervall auf $[0, 1]$ normiert. Ein Exzedent X_i von u_n wird dann also nicht bei i , sondern bei i/n auftreten.

Definition 4.1. (zeitnormierter Punktprozess)

Die Anzahl der Exzedenten kann man mit

$$N_n((a, b]) = \#\{i/n \in (a, b] : X_i > u_n, i = 1, \dots, n\} \quad \text{für } (a, b] \subset (0, 1] \quad (4.2)$$

beschreiben. N_n definiert einen Punktprozess auf dem Intervall $(0, 1]$. Diesen nennt man den *zeitnormierten Punktprozess der Exzedenten*.

Wenn man die Daten in k Blöcke der Größe r aufteilt, kann man für jeden Block dessen Maximum mit

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

definieren. Jetzt kann analog zu (4.2) der zeitnormierte Punktprozess der Clusterexzedenten definiert werden.

Definition 4.2. (zeitnormierter Punktprozess der Clusterexzedenten)

Die Folge N_n^C

$$N_n^C((a, b]) = \#\{ir/n \in (a, b] : X_i > u_n, i = 1, \dots, k\} \quad \text{für } (a, b] \subset (0, 1]$$

ist der *zeitnormierte Punktprozess der Clusterexzedenten*.

Diese Folge N_n^C von Punktprozessen konvergiert für $n \rightarrow \infty$ gegen einen Poissonprozess mit Intensität $\theta\tau$ für ein $\theta \in (0, 1]$. Also sind die Cluster unabhängig.

Dieser Parameter θ heißt *extremaler Index* der Zeitreihe. Er spielt eine wesentliche Rolle bei der Modellierung und Interpretation für die Tail- und Quantilschätzung.

Definition 4.3. (extremaler Index)

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

$$\lim_{n \rightarrow \infty} n P(X_i > u_n) = \tau, \quad (4.3)$$

$$\lim_{n \rightarrow \infty} P(M_n \leq u_n) = \exp(-\theta\tau), \quad (4.4)$$

gelten, dann heißt θ der *extremale Index* der Folge (X_n) .

Äquivalent zu den Bedingungen der Definition 4.3 sind folgende Aussagen:

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

$$\text{Mit } M_n = \max(X_1, \dots, X_n) \quad \text{und} \quad \widetilde{M}_n = \max(\widetilde{X}_1, \dots, \widetilde{X}_n),$$

wobei (\widetilde{X}_n) wieder die assoziierten unabhängigen Folgen von Zufallsvariablen sind und F sei die dazugehörige Verteilungsfunktion.

Nun hilft uns der Grundsatz der Extremwerttheorie, der Folgendes besagt:

Satz 4.2. (Satz von Fisher und Tippett)

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

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

dann ist H eine der folgenden Verteilungsfunktionen:

$$\begin{aligned}\text{Fréchet : } \quad \Phi_\alpha(x) &= \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \quad \alpha > 0. \\ \text{Weibull : } \quad \Psi_\alpha(x) &= \begin{cases} \exp(-(-x)^\alpha), & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0. \\ \text{Gumbel : } \quad \Lambda(x) &= \exp(-e^{-x}), \quad x \in \mathbb{R}.\end{aligned}$$

Beweis. Eine Skizze des Beweises findet man in Embrechts, Klüppelberg und Mikosch [5] (Theorem 3.2.3, S. 121f). \square

Um sich die Verteilungen aus Satz 4.2 vorstellen zu können, gibt Abbildung 4.1 einen Überblick.

Die Funktion $H(x)$ wird auch Fisher-Tippett Funktion genannt und hat folgende Gestalt:

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

Für $\xi > 0$ ergibt sich die Fréchet Verteilung, für $\xi < 0$ die Weibull- und für $\xi = 0$ die Gumbel Verteilung. Satz 4.2 besagt eben, dass die asymptotische Verteilung der Maxima

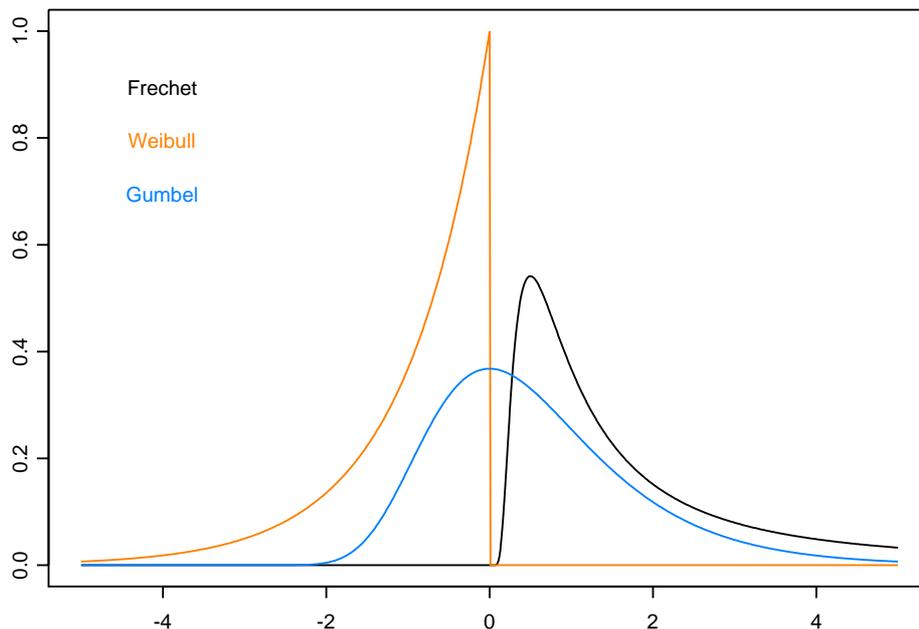


Abbildung 4.1:

Dichten der Fréchet, Weibull und Gumbel Verteilung. Bei Fréchet und Weibull wurde $\beta = 1$ gewählt.

eine dieser drei Verteilungen ist.

Da uns Satz 4.2 also die Grenzverteilung der Maxima liefert, hilft dieser sehr viel bei den Schätzmethoden des extremalen Index. Es gilt: $\lim P(M_n \leq a_n x + b_n) = \lim F^n(a_n x + b_n)$.

Mit der Aussage $P(M_n \leq u_n) \approx P(\tilde{M}_n \leq u_n)^\theta$ (siehe (5.4)) und Satz 4.2 kann man die Tabelle 4.1 erstellen.

Verteilung	Verteilungsfkt.	$\lim_{n \rightarrow \infty} P(\tilde{M}_n \leq u_n)$	$\lim_{n \rightarrow \infty} P(M_n \leq u_n)$	τ
Fréchet	$\exp(-x^{-\alpha}), x > 0$	$\exp(-x^{-\alpha})$	$(\exp(-x^{-\alpha}))^\theta$	$x^{-\alpha}$
Weibull	$\exp(-(-x)^\alpha), x \leq 0$	$\exp(-(-x)^\alpha)$	$(\exp(-(-x)^\alpha))^\theta$	$(-x)^\alpha$
Gumbel	$\exp(-e^{-x}), x \in \mathbb{R}$	$\exp(-e^{-x})$	$(\exp(-e^{-x}))^\theta$	e^{-x}

Tabelle 4.1: *Übersicht über die asymptotische Verteilung der Maxima*

θ liegt im Intervall $(0, 1]$. $\theta = 0$ ist ein pathologischer Fall, der für uns nicht in Betracht kommt. $\theta > 0$ gilt, da dies im praktischen Sinne liegt. Also bleibt nur noch $\theta \leq 1$ zu zeigen. Dies ist jedoch einfach.

Es ist klar, dass

$$\begin{aligned} P(M_n \leq u_n) &= 1 - P\left(\bigcup_{i=1}^n \{X_i > u_n\}\right) \geq 1 - n \bar{F}(u_n) \\ \Leftrightarrow P(M_n \leq u_n) &\geq 1 - n \bar{F}(u_n) \\ \Leftrightarrow \lim_{n \rightarrow \infty} P(M_n \leq u_n) &\geq \lim_{n \rightarrow \infty} (1 - n \bar{F}(u_n)) \\ \Leftrightarrow e^{-\theta\tau} &\geq 1 - \tau \end{aligned}$$

Für $\tau > 1$ ist die Aussage trivial, da die Exponentialverteilung immer größer Null ist. Für $\tau \in (0, 1)$ betrachte folgende Grenzwerte:

$$\begin{aligned} \theta &\leq \lim_{\tau \rightarrow 0} -\frac{\ln(1 - \tau)}{\tau} \stackrel{v\text{Hopital}}{=} \lim_{\tau \rightarrow 0} \frac{1}{1 - \tau} = 1 \\ \theta &\leq \lim_{\tau \rightarrow 1} -\frac{\ln(1 - \tau)}{\tau} = \infty \end{aligned}$$

Da die Aussage $e^{-\theta\tau} \geq 1 - \tau$ für alle τ gelten muß, gilt also $\theta \leq 1$.

Demnach ist $\theta \in (0, 1]$.

Interpretation von θ

Wir erinnern uns an die Definition des Punktprozesses der Exzedenten (4.2).

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

wobei die X_i nicht unabhängig sein müssen.

Sei Y_n die Zufallsvariable, die die Clustergröße beschreibt. Mit Hilfe des Konvergenzverhaltens von N_n^C gegen einen Poissonprozess mit Intensität $\theta\tau$, weiß man dass N_n für eine stationäre Zeitreihe in Verteilung gegen einen zusammengesetzten Poissonprozess konvergiert (unter bestimmten Eigenschaften (vgl. Embrechts, Klüppelberg und Mikosch [5] Abschnitt 4.4), der wie folgt aussieht.

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

wobei $\sum_{n=1}^0 Y_n = 0$ und N^C der Grenzprozess der Punktprozessfolge N_n^C , also ein Poissonprozess mit Intensität $\theta\tau$, ist.

Also gilt folgendes Lemma.

Lemma 4.1. (*Interpretation von θ*)

θ kann als Kehrwert des Erwartungswertes interpretiert werden. Also gilt:

$$E(Y_1) = 1/\theta \quad (4.5)$$

Beweis. Es gilt

$$\tau = \theta \tau E(Y_1) \quad (4.6)$$

wegen

$$\begin{aligned} \tau &\stackrel{(4.3)}{=} \lim_{n \rightarrow \infty} n P(X > u_n) \stackrel{Def. E(x)}{=} \lim_{n \rightarrow \infty} E(N_n(0, 1]) = E(N(0, 1]) = \\ &= E \left(\sum_{n=1}^{N^C(1) - N^C(0)} Y_n \right) \stackrel{N^C(0)=0}{=} E \left(\sum_{n=1}^{N^C(1)} Y_n \right) = E(N^C(1) Y_1) = \\ &= E(N^C(1)) E(Y_1) \stackrel{N^C \sim Poi(\theta \tau)}{=} \theta \tau E(Y_1) \end{aligned}$$

Mit (4.6) ergibt sich demnach:

$$E(Y_1) = 1/\theta.$$

□

Also ist die **mittlere Clustergröße** gerade $1/\theta$ (unter schwachen Bedingungen).

Kapitel 5

Alte Schätzmethoden des extremalen Index

Es gibt verschiedene Möglichkeiten, den extremalen Index zu schätzen. Dazu gehören die Blockmethode, run Methode oder das Interpretieren als den Kehrwert des Erwartungswertes, die in diesem Kapitel kurz beschrieben werden. Das asymptotische Verhalten, wie z.B. die Konsistenz, des run und des Blockschätzers können bei Hsing [10, 11], Smith und Weissman [22] und Weissman und Novak [23] nachgelesen werden.

5.1 Blockmethode

Die vielleicht einfachste Möglichkeit ist die Blockmethode.

Man erhält aus der Definition des extremalen Index im Kapitel 4 auch folgende Aussagen:

$$\lim_{n \rightarrow \infty} n \bar{F}(u_n) = \tau, \quad (5.1)$$

$$\lim_{n \rightarrow \infty} P(\tilde{M}_n \leq u_n) = \exp(-\tau), \quad (5.2)$$

$$\lim_{n \rightarrow \infty} P(M_n \leq u_n) = \exp(-\theta \tau). \quad (5.3)$$

M_n sei wieder $\max(X_1, \dots, X_n)$ und \tilde{M}_n sei $\max(\tilde{X}_1, \dots, \tilde{X}_n)$. Auch hier seien wieder $(X_i)_{i=1}^n$ eine Folge von stationären Beobachtungen und $(\tilde{X}_i)_{i=1}^n$ wieder die assoziierte unabhängige Folge.

Mit diesen Aussagen läßt sich leicht nachweisen, dass

$$P(M_n \leq u_n) \approx \exp(-\theta \tau) = (\exp(-\tau))^\theta \approx P^\theta(\tilde{M}_n \leq u_n) = F^{\theta n}(u_n), \quad (5.4)$$

wobei F die Verteilungsfunktion der Folge $(X_i)_{i=1}^n$ sei.

Lemma 5.1. (*θ als Grenzwert*)

Es sei (5.1), (5.2), (5.3) und (5.4) erfüllt. Dann läßt sich der Grenzwert

$$\lim_{n \rightarrow \infty} \frac{\ln P(M_n \leq u_n)}{n \ln F(u_n)} = \theta \quad (5.5)$$

berechnen.

Beweis.

$$\begin{aligned} (1) \quad & P(M_n \leq u_n) \rightarrow \exp(-\theta\tau) \\ & \Rightarrow \ln P(M_n \leq u_n) \rightarrow \ln(\exp(-\theta\tau)) = -\theta\tau \\ (2) \quad & F(u_n)^n \stackrel{(5.4)}{=} P(\widetilde{M}_n \leq u_n) \rightarrow \exp(-\tau) \\ & \Rightarrow n \ln F(u_n) = \ln F(u_n)^n \rightarrow \ln \exp(-\tau) = -\tau \\ (1) + (2) \quad & \Rightarrow \frac{P(M_n \leq u_n)}{F(u_n)^n} = \frac{-\theta\tau}{-\tau} = \theta \end{aligned}$$

□

$$\frac{N}{n} = \frac{1}{n} \sum_{i=1}^n I_{\{I_i > u_n\}} =: \hat{\theta}_n^{(empirisch)}(u_n) \quad (5.6)$$

ist eine empirische Methode den Tail von $\overline{F}(u_n)$ zu schätzen.

Aus Embrechts, Klüppelberg und Mikosch [5] (S. 211, 419) wissen wir:

$$P(M_n \leq u_n) \approx P^k(M_{\lfloor n/k \rfloor} \leq u_n). \quad (5.7)$$

Diese Aussage ist die Basis der Blockmethode. Im ersten Schritt der Blockmethode wird der Datensatz in k Blöcke der Größe r aufgeteilt, also ist $n = rk$. Damit man keine Probleme mit der ganzzahligen Größe der Blöcke bekommt, wählt man $r = \lfloor n/k \rfloor$. Somit sieht unser Datensatz folgendermaßen aus:

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

Das Maximum für jeden Block sei:

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

Mit diesen Teilergebnissen wollen wir nun $P(M_n \leq u_n)$ anschauen und θ schätzen.

Lemma 5.2. ($P(M_n \leq u_n)$ mit der Blockmethode)

Sei K die Anzahl der Blöcke, in denen einer oder mehrere Exzedenten auftreten, d.h. $K = \#\{\text{Anzahl der „Cluster-Exzedenten“}\}$.

Dann gilt:

$$P(M_n \leq u_n) \approx \left(1 - \frac{K}{k}\right)^k.$$

Beweis. Aus (5.6), (5.7) folgt:

$$\begin{aligned} P(M_n \leq u_n) &= P(\max_{1 \leq i \leq k} M_r^{(i)} \leq u_n) \stackrel{(5.7)}{\approx} P^k(M_r \leq u_n) \approx \\ &\stackrel{(5.6)}{\approx} \left(\frac{1}{k} \sum_{i=1}^k I_{\{M_r^{(i)} \leq u_n\}}\right)^k = \\ &= \left(1 - \frac{1}{k} \sum_{i=1}^k I_{\{M_r^{(i)} \leq u_n\}}\right)^k = \\ &\stackrel{(5.6)}{=} \left(1 - \frac{K}{k}\right)^k \end{aligned}$$

□

Lemma 5.3. (Schätzung von θ mit der Blockmethode)

Sei $N = \#\{\text{Exzedenten von } u_n\}$ und K die Anzahl der Blöcke, in denen einer oder mehrere Exzedenten auftreten, d.h. $K = \#\{\text{Anzahl der „Cluster-Exzedenten“}\}$.

Dann läßt sich θ durch die Blockmethode wie folgt schätzen:

$$\hat{\theta}_n^{(Block)}(u_n) = \frac{1 \ln(1 - K/k)}{r \ln(1 - N/n)}. \quad (5.8)$$

Beweis.

$$\hat{\theta}_n^{(Block)}(u_n) = \frac{\ln \hat{P}(M_n \leq u_n)}{n \ln \hat{F}(u_n)} = \frac{\ln(1 - K/k)^k}{\ln(1 - N/n)^n} \stackrel{n=kr}{=} \frac{k \ln(1 - K/k)}{kr \ln(1 - N/n)} = \frac{1 \ln(1 - K/k)}{r \ln(1 - N/n)}$$

□

Aus (5.6) und einer Taylorentwicklung folgt:

$$\hat{\theta}_n^{(empirisch)}(u_n) = \frac{K}{N} = \frac{1}{r} \frac{K}{k} \frac{kr}{N} \stackrel{n=kr}{=} \frac{1}{r} \frac{K}{k} \frac{n}{N} = \frac{1}{r} \frac{K/k}{N/n} \approx \hat{\theta}_n^{(Block)}(u_n).$$

Bei der Blockmethode wird ein Cluster identifiziert, falls das Maximum in einem Block über der Schwelle u liegt. Sobald es Blöcke gibt, in denen mindestens eine Beobachtung über der Schwelle u liegt, werden alle Exzedenten in diesem Block zu einem Cluster zusammengefasst (vgl. Borkovec und Klüppelberg [2], S. 11). Zur Veranschaulichung in Abbildung 5.1 ein theoretisches Beispiel.

Clusterbildung mit der Blockmethode

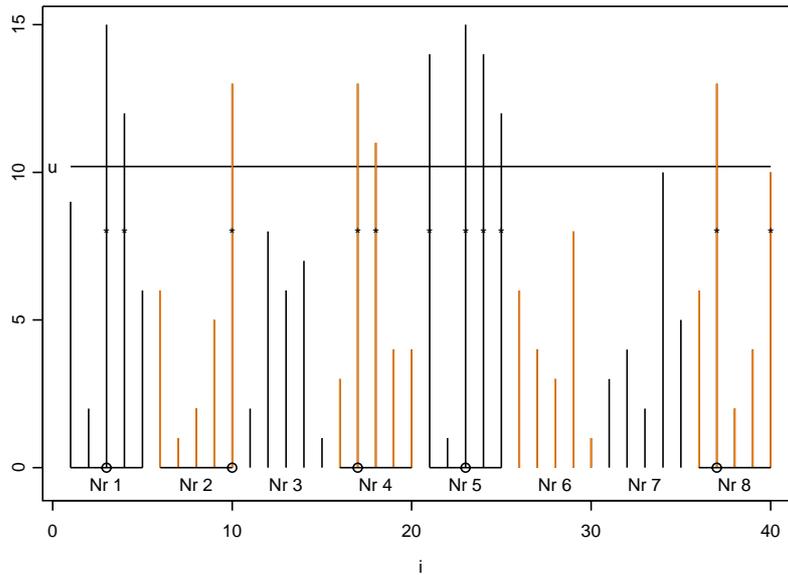


Abbildung 5.1:

Theoretisches Beispiel für die Blockmethode. Es wurde die Blockgröße $r=5$ gewählt und die Blöcke sind durch die verschiedenen Farben erkenntlich gemacht. Die Punkte kennzeichnen die Maxima, die oberhalb der Schwelle u liegen. Somit befindet sich im Block Nr. 1, 2, 4, 5 und 8 ein Cluster, die mit „“ gekennzeichnet sind.*

5.2 Der extreme Index als Kehrwert des Erwartungswertes

In Kapitel 4 wurde schon gezeigt, dass der extreme Index als Kehrwert des Erwartungswertes dargestellt werden kann (vgl. (4.5))

Damit bekommt man auf der Basis der Blockmethode einen Schätzer für θ , der folgendermaßen aussieht:

$$\begin{aligned} \hat{\theta}_n^{(\text{Kehrwert})}(u_n) &= \frac{\text{„Anzahl der Blockmaxima größer } u\text{“}}{\text{„Anzahl der Exzedenten“}} = \\ &= \frac{\sum_{i=1}^k I_{\{M_r^{(i)} > u_n\}}}{\sum_{(i=1)}^n I_{\{X_i > u_n\}}} = \frac{K}{N}, \end{aligned}$$

wobei I die Indikatorfunktion ist und $M_r^{(i)} = \max(X_{(i-1)r+1}, \dots, X_{ir})$.

5.3 run Methode

Wenn man das asymptotisches Verhalten von $P(M_n \leq c_n)$ für $n \rightarrow \infty$ untersucht, kommt man auf das Ergebnis des Theorems 5.1 (vgl. O'Brien [21]).

Sei $M_{i,j} = \max(X_{i+1}, \dots, X_j)$ und $M_i = M_{0,i} = \max(X_1, \dots, X_n)$.

Definition 5.1. (asymptotische Unabhängigkeit von Maxima (AUM))

Eine stationäre Folge (X_n) hat eine asymptotische Unabhängigkeit von Maxima (AUM) relativ zur Folge (c_n) , wenn es eine Folge (q_n) mit $q_n = o(n)$ gibt, so dass

$$\alpha_n := \max |P[M_i \leq c_n, M_{i+q, i+q+j} \leq c_n] - P[M_i \leq c_n]P[M_j \leq c_n]| \rightarrow 0$$

wobei das Maximum über alle $i \geq q$ und $j \geq q$ mit $i + q + j \leq n$ gebildet wird.

Theorem 5.1. (asymptotisches Verhalten von $P[M_n \leq c_n]$)

Sei (X_n) eine Folge, die AUM hat, wobei (c_n) eine Folge von reellen Zahlen sei. (q_n) und (α_n) seien wie in Definition 5.1 definiert. (p_n) sei eine Folge von positiven Integern, so dass

$$p = o(n), \quad n \alpha = o(p) \quad \text{und} \quad q = o(p).$$

Wenn entweder

$$\liminf F^n(c_n) > 0$$

oder

$$\liminf P[M_{1,p} \leq c_n | X_1 > c_n] > 0$$

gilt, dann gelten

$$P[M_n \leq c_n] - (F(c_n))^{nP[M_{1,p} \leq c_n | X_1 > c_n]} \rightarrow 0$$

und

$$P[M_n \leq c_n] - \exp\{-nP[X_1 > c_n, M_{1,p} \leq c_n]\} \rightarrow 0$$

Beweis. vgl. O'Brien [21] (S. 282ff) □

Wegen Theorem 5.1 gilt

$$\begin{aligned} P(M_n \leq u_n) &= (F(u_n))^{nP(M_{2,s} \leq u_n | X_1 > u_n)} + o(1) \\ &= \exp\{-nP(X_1 > u_n, M_{2,s} \leq u_n)\} + o(1), \end{aligned}$$

wobei $n \bar{F}(u_n) \rightarrow \tau$, $M_{2,s} = \max(X_2, \dots, X_s)$, $s = s(n)$, und $s/n \rightarrow 0$.

Nach der Definition des extremalen Index gilt zusätzlich

$$P(M_n \leq u_n) \approx \exp(-\theta \tau).$$

Dies kann auch interpretiert werden als

$$P(M_n \leq u_n) = \exp(-\theta \tau) + o(1).$$

Mit den oben genannten Eigenschaften läßt sich Folgendes herleiten:

$$\lim_{n \rightarrow \infty} \theta_n(u_n, s(n)) = \lim_{n \rightarrow \infty} P(M_{2,s} \leq u_n | X_1 > u_n) = \theta \quad (5.9)$$

Hier kann also θ als Grenzverteilung interpretiert werden. Die bedingte Wahrscheinlichkeit $\theta_n(s(n), u_n)$ ist ein Maß für eine Tendenz zu großen Cluster. O'Briens Schätzung von θ beruht auf „runs“:

$$\hat{\theta}_n^{(run)}(u_n, s(n)) = \frac{\text{„Anzahl der Cluster“}}{\text{„Anzahl der Exzedenten“}} = \frac{\sum_{i=1}^{n-r} I_{A_{i,n}}}{\sum_{i=1}^n I_{\{X_i > u_n\}}}, \quad (5.10)$$

wobei $A_{i,n} = \{X_i > u_n, X_{i+1} \leq u_n, \dots, X_{i+r} \leq u_n\}$ und I die Indikatorfunktion ist.

Ein Cluster beginnt bei der run Methode beim ersten Exzedenten. Ein Cluster endet, wenn mindestens s aufeinander folgende Daten unterhalb der Schwelle u liegen. Das nächste Cluster beginnt dann wieder beim ersten Exzedenten nach den mindestens s Daten, die unterhalb der Schwelle lagen. Hierzu auch eine theoretische Veranschaulichung:

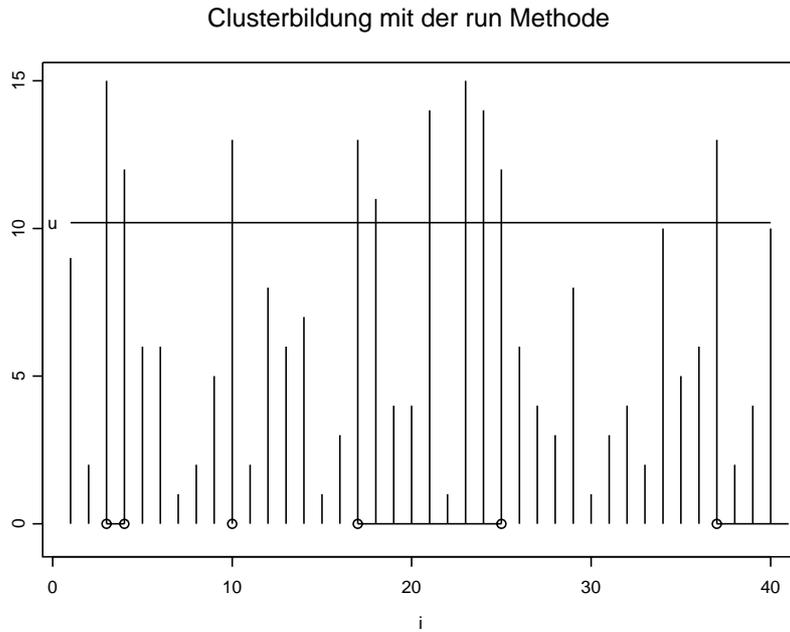


Abbildung 5.2:

Theoretisches Beispiel für die run Methode mit der run Länge $s = 5$. Wie immer ist natürlich u der Schwellenwert. Es wurden die gleichen Daten wie bei Abbildung 5.1 verwendet. Die Punkte kennzeichnen den Beginn bzw. das Ende eines Clusters. Ein Cluster wird dadurch kenntlich gemacht, dass die Daten mit einer Linie verbunden sind. Also sind die Daten mit i von 3 - 4, 10, 17 - 25, 36 - ..., Cluster. Wie der Datenpunkt mit $i = 10$ zeigt, kann ein Cluster im Extremfall auch nur aus einem Datenpunkt bestehen.

Kapitel 6

Neue Methode 1: Theoretische Momenten Methode

Ein Problem bei den bisherigen Methoden ist die Wahl der Clusterparameter. Die Cluster hängen von der Laufzeit und anderen Parametern ab, auf die die Methoden sehr sensibel reagieren. Es wurden Methoden entwickelt, die den Einfluß dieser Parameter vermindern sollen.

Eine dieser Methoden bedient sich der theoretischen Momente der Grenzverteilung. Auf diesem Weg bekommt man ein automatisches Clusteridentifizierungsschema, das nicht auf die Wahl von Hilfsparametern angewiesen ist. Zusätzlich ermöglicht diese Methode die Berechnung von Konfidenzintervallen mit Hilfe eines Bootstrap Prozesses.

6.1 Asymptotische Verteilung

Sei $(X_n)_{n \geq 1}$ der Prozess der Beobachtungen, die den Schwellenwert u überschreiten. $T(u)$ sei die zum minimalen Zeitabstand zwischen Exzedenten zugehörige Zufallsvariable.

Definition 6.1. (kleinster Zeitabstand)

Sei $M_{s,n} = \max(X_s, \dots, X_n)$, dann ist

$$\min\{n \geq 1, X_{n+1} > u\}, \quad X_1 > u \quad (6.1)$$

der *kleinste Zeitabstand* zwischen Exzedenten.

(6.1) kann auch folgendermaßen geschrieben werden:

$$P\{T(u) = n\} = P(M_{1,n} \leq u, X_{n+1} > u | X_1 > u), \quad n \geq 1$$

bzw.

$$P\{T(u) > n\} = P(M_{1,n+1} \leq u | X_1 > u), \quad n \geq 1.$$

Lemma 6.1. (*asympt. Verteilung*)

Die asymptotische Verteilung von $\overline{F}(u)T(u)$ ist im Fall von iid verteilten Zufallsvariablen eine Standard Exponentialverteilung.

Beweis. Im Fall von iid verteilten Zufallsvariablen gilt:

$$P\{T(u) > n\} = F(u)^n, \quad n \geq 1, \quad (6.2)$$

so dass

$$P\{\overline{F}(u)T(u) > x\} = \exp \left[\lfloor x/\overline{F}(u) \rfloor \log\{F(u)\} \right] \quad \text{für } x > 0. \quad (6.3)$$

(6.3) ist leicht nachzurechnen, da

$$\begin{aligned} P\{\overline{F}(u)T(u) > x\} &= P\{T(u) > \lfloor x/\overline{F}(u) \rfloor\} \stackrel{(6.2)}{=} F(u)^{\lfloor x/\overline{F}(u) \rfloor} = \exp \left[\log \left\{ F(u)^{\lfloor x/\overline{F}(u) \rfloor} \right\} \right] \\ &= \exp \left[\lfloor x/\overline{F}(u) \rfloor \log\{F(u)\} \right]. \end{aligned}$$

Sei ω der Endpunkt der Verteilung und $\log(1 + \epsilon) \approx \epsilon$, für $\epsilon \rightarrow 0$, dann gilt

$$\lim_{u \rightarrow \omega^+} [P\{\overline{F}(u)T(u) > x\}] = \exp(-x)$$

wegen

$$\lim_{u \rightarrow \omega^+} [P\{\overline{F}(u)T(u) > x\}] = \lim_{u \rightarrow \omega^+} \exp \left[\underbrace{\left\lfloor \frac{x}{\overline{F}(u)} \right\rfloor}_{\rightarrow 1 - (1+\epsilon)} \underbrace{\log\{F(u)\}}_{\substack{\rightarrow 1+\epsilon \\ \rightarrow \epsilon}} \right] = \text{„} \exp \left[\frac{x}{-\epsilon} \epsilon \right] \text{“} = \exp(-x)$$

Also ist die **asymptotische Verteilung** von $\overline{F}(u)T(u)$ eine **Standard Exponentialverteilung** und der Punktprozess der Exzedenten ist ein Poissonprozess.

□

Wir betrachten nun den generellen Fall $\theta \in [0, 1]$. Für $1 \leq k \leq l$ sei $F_{k,l}(u)$ die σ -Algebra, die durch die Ereignisse $\{X_i > u, k \leq i \leq l\}$ erzeugt wird. Damit kann man sich die gemischten Koeffizienten

$$\alpha_{n,q}(u) = \max_{l \leq k \leq n-q} \sup |P(B|A) - P(B)|$$

definieren, mit $A \in F_{1,k}(u)$, $P(A) > 0$ und $B \in F_{k+q,n}(u)$.

Theorem 6.1. (*asymptotische Verteilung*)

Sei r_n positiv und u_n der Schwellenwert, $n \geq 1$, so dass

$$r_n \rightarrow \infty,$$

$$r_n \bar{F}(u_n) \rightarrow \tau$$

und

$$P(M_{r_n} \leq u_n) \rightarrow \exp(-\theta \tau),$$

für $\tau \in (0, \infty)$ und $\theta \in [0, 1]$.

Wenn ein positives $q_n = o(r_n)$ existiert, so dass $\alpha_{cr_n, q_n}(u_n) = o(1)$, für alle $c > 0$, dann gilt:

$$P\{\bar{F}(u_n)T(u_n) > t\} \rightarrow \theta \exp(-\theta t), \quad \text{für } t > 0.$$

Beweis. vgl. Ferro und Segers [8] (S. 555, Appendix A, Theorem 1) □

Wenn man die Konvergenz in Verteilung betrachtet, kommt man wegen Theorem 6.1 auf $\bar{F}(u)T(u) \xrightarrow{d} T_\theta$, für $u \rightarrow w^+$, wobei T_θ die Zufallsvariable ist, die sich auf die Mischungsverteilung

$$(1 - \theta)\epsilon_0 + \theta\mu_\theta \tag{6.4}$$

bezieht. Dabei ist ϵ_0 das Diracmaß an der Stelle 0 und μ_θ steht für eine Exponentialverteilung mit dem Erwartungswert $1/\theta$.

θ hat demnach zwei Bedeutungen. Zum einen ist θ der Anteil der Interexzedentenzeiten ungleich Null und zum anderen ist θ der Kehrwert des Erwartungswertes der Interexzedentenzeiten ungleich Null.

6.2 Schätzung des extremalen Index

Seien X_1, \dots, X_n Zufallsvariablen und u ein hoher Schwellenwert. Sei

$$N = N_n(u) = \sum_{i=1}^n I(X_i > u)$$

die Summe aller Exzedenten, die größer sind als der Schwellenwert u . Zusätzlich seien $1 \leq S_1 \leq \dots \leq S_N \leq n$ die Zeiten, zu denen Exzedenten auftreten. Somit sind $T_i = S_{i+1} - S_i$ für $i = 1, \dots, N - 1$.

6.2.1 Intervallschätzer

Lemma 6.2. (zweites Moment von T)

Das zweite Moment von T mit der Mischungsverteilung (6.4) ist $E(T^2) = 2/\theta$

Beweis. Das zweite Moment einer Zufallsvariabel X mit $X \sim \exp(\theta)$ ist.

$$\begin{aligned}
 E(x^2) &= \int_0^\infty x^2 \theta \exp(-\theta x) dx = \\
 &= \theta \left\{ \left[x^2 \exp(-\theta x) \left(-\frac{1}{\theta} \right) \right]_0^\infty - \int_0^\infty 2x \exp(-\theta x) \left(-\frac{1}{\theta} \right) dx \right\} = \\
 &= 2 \int_0^\infty x \exp(-\theta x) dx = -\frac{2}{\theta} \left\{ [x \exp(-\theta x)]_0^\infty - \int_0^\infty \exp(-\theta x) dx \right\} = \\
 &= \frac{2}{\theta} \int_0^\infty \exp(-\theta x) dx = \frac{2}{\theta} \left[\exp(-\theta x) \left(-\frac{1}{\theta} \right) \right]_0^\infty \\
 &= \frac{2}{\theta^2}
 \end{aligned}$$

Sei nun Y eine Zufallsvariable mit der Verteilungsfunktion ϵ_0 und X eine Zufallsvariable mit der Verteilungsfunktion μ_θ . Wie allgemein für Mischungsverteilungen bekannt, läßt sich das zweite Moment von T wie folgt berechnen.

$$E(T^2) = (1 - \theta)E(Y^2) + \theta E(X^2) = \theta \frac{2}{\theta^2} = \frac{2}{\theta}$$

wegen $E(Y^2) = 0$. □

Lemma 6.3. (erster Momentenschätzer)

Mit Lemma 6.2 bekommt man den ersten Schätzer für θ . Dafür sei $\hat{F}_n(u)$ ein Schätzer für $\bar{F}(u)$. Somit ist der Momentenschätzer für θ

$$\hat{\theta}_n^{(1)}(u) = \frac{2(N-1)}{\hat{F}^2(u) \sum_{i=1}^{N-1} T_i^2} \quad (6.5)$$

Beweis.

$$\frac{2}{\hat{\theta}_n^{(1)}} = \frac{1}{N-1} \sum_{i=1}^{N-1} \hat{F}^2(u) T_i^2 \quad \Leftrightarrow \quad \hat{\theta}_n^{(1)} = \frac{2(N-1)}{\sum_{i=1}^{N-1} \hat{F}^2(u) T_i^2} = \frac{2(N-1)}{\hat{F}^2(u) \sum_{i=1}^{N-1} T_i^2}$$

□

Wählt man als Schätzung für $\bar{F}(u)$ zum Beispiel $\hat{F}_n(u) = \frac{N}{n}$, ergibt sich für $\hat{\theta}_n^{(1)}(u)$

$$\hat{\theta}_n^{(1)}(u) = \frac{2n^2(N-1)}{N^2 \sum_{i=1}^{N-1} T_i^2} .$$

Das erste Moment von T ist 1, da $(1 - \theta)E(Y) + \theta E(X) = 0 + \theta \frac{1}{\theta} = 1$, wobei Y wieder eine Zufallsvariable vom Diracmaß ϵ_0 ist und $X \sim \exp(\theta)$ (aus der Mischungsverteilung (6.4)). Somit ist $\text{var}(T) = \nu^2 = E(T^2) - E(T)^2 = E(T^2) - 1$. Damit kann man folgende Beziehung herleiten:

$$1 + \nu^2 = E(T^2) = \frac{E(T^2)}{E(T)^2} = \frac{2}{\theta} \quad (6.6)$$

Die Interexzedenzzeiten sind zerstreut, mit Clustern an den Grenzen, genau dann, wenn $\theta < 1$. Diese Beziehung veranlaßt einen dazu, einen anderen Schätzer für θ zu finden.

Lemma 6.4. (zweiter Schätzer für θ)

Mit (6.6) bekommt man den folgenden Schätzer für θ :

$$\hat{\theta}_n^{(2)}(u) = \frac{2 \left(\sum_{i=1}^{N-1} T_i \right)^2}{(N-1) \sum_{i=1}^{N-1} T_i^2}$$

Beweis.

$$\begin{aligned} (1) \quad \hat{E}(T^2) &= \frac{2}{\hat{\theta}_n} = \frac{1}{N-1} \sum_{i=1}^{N-1} \hat{F}_n^2(u) T_i^2 \\ (6.5) \Rightarrow \hat{\theta}_n &= \frac{2(N-1)}{\hat{F}_n^2(u) \sum_{i=1}^{N-1} T_i^2} \\ (2) \quad \hat{E}(T) &= \frac{1}{N-1} \sum_{i=1}^{N-1} \hat{F}_n(u) T_i \\ (1) + (2) \Rightarrow \hat{\theta}_n^{(2)}(u) &= \frac{2 \hat{E}(T)^2}{\hat{E}(T^2)} = \frac{2(N-1) \hat{F}_n^2(u) \left(\sum_{i=1}^{N-1} T_i \right)^2}{(N-1)^2 \hat{F}_n^2(u) \sum_{i=1}^{N-1} T_i^2} = \frac{2 \left(\sum_{i=1}^{N-1} T_i \right)^2}{(N-1) \sum_{i=1}^{N-1} T_i^2} \end{aligned}$$

□

Der Vorteil dieses Schätzers ist es, dass man $\bar{F}(u)$ nicht schätzen muß.

Eine weitere Verbesserung des Schätzers ist möglich, indem man die folgende Approximation der Mischungsverteilung (6.4) betrachtet. Setzt man im Beweis von Theorem 6.1 $r_n = n$, so erhält man für die Verteilung der Interexzedenzzeiten

$$P(T(u) > n) = \theta F(u)^{n\theta} + o(1).$$

Mit dieser Beziehung kann ein Schätzer für θ bestimmt werden. Sei T eine Zufallsvariable, die folgende Verteilung hat:

$$P(T > n) = \theta p^{n\theta} \quad \text{für } n \geq 1 \quad (6.7)$$

wobei $\theta \in (0, 1]$ und $p \in (0, 1]$.

Lemma 6.5. (Folgerungen aus der Verteilung (6.7))

Betrachtet man die Wahrscheinlichkeit (6.7) der Approximation der gemischten Grenzverteilung (6.4), so erhält man

$$\begin{aligned}\frac{2 E(T)^2}{E(T^2)} &= \frac{2 (1 - (1 - \theta) p^\theta)^2}{2 \theta p^\theta + \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2} \\ &= \theta + \theta \left(2 - \frac{3\theta}{2} \right) (1 - p) + O\{(1 - p)^2\}.\end{aligned}$$

Beweis. Hier soll nur eine Beweisskizze dargestellt werden, da einige Zwischenergebnisse später noch benötigt werden. Der vollständige Beweis befindet sich im Anhang A.1.

$$(1) \quad E(T - 1) = E(T) - 1$$

$$\Rightarrow E(T) = E(T - 1) + 1$$

$$(2) \quad E\left(\frac{T(T-1)}{2}\right) = E\left(\frac{T^2 - T}{2}\right) = \frac{1}{2} (E(T^2) - E(T))$$

$$\Rightarrow E(T^2) = 2 E\left(\frac{T(T-1)}{2}\right) + E(T)$$

$$(3) \quad E(T - 1) = \sum_{n=1}^{\infty} P(T > n) = \sum_{n=1}^{\infty} \theta p^{n\theta} = \theta \sum_{n=0}^{\infty} p^{(n+1)\theta} = \theta p^\theta \sum_{n=0}^{\infty} p^{n\theta} =$$

$$\stackrel{\text{geo.Reihe}}{=} \theta p^\theta (1 - p^\theta)^{-1}$$

$$(4) \quad E\left(\frac{T(T-1)}{2}\right) = \sum_{n=1}^{\infty} n P(T > n) = \sum_{n=1}^{\infty} n \theta p^{n\theta} = \theta \sum_{n=1}^{\infty} n p^{n\theta} =$$

$$\stackrel{\text{geo.Reihe}}{=} \theta \left[(1 - p^\theta)^{-1-1} \sum_{j=1}^1 a_j^{(1)} (p^\theta)^j \right] \stackrel{a_1^{(1)}=1}{=} \theta p^\theta (1 - p^\theta)^{-2}$$

$$(1) - (4) \Rightarrow \frac{2 E(T)^2}{E(T^2)} =$$

$$\begin{aligned}&= \frac{2 (\theta p^\theta (1 - p^\theta)^{-1} + 1)^2}{2 \theta p^\theta (1 - p^\theta)^{-2} + \theta p^\theta (1 - p^\theta)^{-1} + 1} = \\ &= \frac{2 (\theta^2 p^{2\theta} (1 - p^\theta)^{-2} + 2 \theta p^\theta (1 - p^\theta)^{-1} + 1)}{2 \theta p^\theta (1 - p^\theta)^{-2} + \theta p^\theta (1 - p^\theta)^{-1} + 1} = \\ &= \frac{2 (\theta^2 p^{2\theta} + 2 \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2)}{2 \theta p^\theta + \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2} = \\ &= \frac{2 (\theta p^\theta + (1 - p^\theta))^2}{2 \theta p^\theta + \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2} = \\ &= \frac{2 (1 - (1 - \theta) p^\theta)^2}{2 \theta p^\theta + \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2}\end{aligned}$$

□

Definition 6.2. (bias erster Ordnung)

Sei $R(X)$ die Schätzstatistik und $q(\theta)$ eine Funktion des zu schätzenden Parameters θ , dann ist

$$bias_{(erster\ Ordnung)} = E(R(X)) - q(\theta)$$

unter Vernachlässigung der quadratischen Terme der *bias erster Ordnung*.

Der Bias erster Ordnung für $\hat{\theta}_n^{(2)}(u)$ bzgl. einer Schwelle u und für die Verteilung (6.7) ist $\theta(2 - 3\theta/2)\bar{F}(u)$. Dies folgt aus Lemma 6.5.

$$\begin{aligned} bias_{(erster\ Ordnung)} &= E\left(\frac{2\left(\sum_{i=1}^{N-1} T_i\right)^2}{(N-1)\sum_{i=1}^{N-1} T_i^2}\right) - \theta \approx \frac{2\left(\sum_{i=1}^{N-1} E(T_i)\right)^2}{(N-1)\sum_{i=1}^{N-1} E(T_i^2)} - \theta = \\ &= \frac{2(N-1)^2 E(T_1)^2}{(N-1)^2 E(T_1^2)} - \theta \stackrel{Lemma\ 6.5}{=} \theta + \theta\left(2 - \frac{3\theta}{2}\right)(1-p) - \theta = \\ &= \theta\left(2 - \frac{3\theta}{2}\right)(1-p) = \theta\left(2 - \frac{3\theta}{2}\right)\bar{F}(u) \end{aligned}$$

Die letzte Folgerung gilt, weil $\theta F(u)^{n\theta} = \theta p^{n\theta} \Rightarrow F(u) = p \Rightarrow 1 - F(u) = (1-p)$.

Ferner gilt die Beziehung

$$\frac{2(E(T-1))^2}{E((T-1)(T-2))} = \theta, \quad (6.8)$$

da

$$\begin{aligned} \frac{2(E(T-1))^2}{E((T-1)(T-2))} &= \frac{2[\theta p^\theta(1-p^\theta)^{-1}]^2}{E(T^2) - 3E(T) + 2} \stackrel{\text{Beweis Lemma 6.5(1)+(2)}}{=} \\ &= \frac{2\theta^2 p^{2\theta}(1-p^\theta)^{-2}}{2\theta p^\theta(1-p^\theta)^{-2} + \theta p^\theta(1-p^\theta)^{-1} + 1 - 3\theta p^\theta(1-p^\theta)^{-1} - 3 + 2} = \\ &= \frac{2\theta^2 p^{2\theta}(1-p^\theta)^{-2}}{2\theta p^\theta(1-p^\theta)^{-2} - 2\theta p^\theta(1-p^\theta)^{-1}} = \frac{\theta p^\theta(1-p^\theta)^{-2}}{(1-p^\theta)^{-2} - (1-p^\theta)^{-1}} = \\ &= \frac{\theta p^\theta}{1 - (1-p^\theta)} = \\ &= \theta \end{aligned}$$

Damit kommt man zum Schätzer

$$\hat{\theta}_n^{(3)}(u) = \frac{2\left(\sum_{i=1}^{N-1} (T_i - 1)\right)^2}{(N-1)\sum_{i=1}^{N-1} (T_i - 1)(T_i - 2)},$$

dessen Bias erster Ordnung 0 ist.

Der Bias erster Ordnung für $\hat{\theta}_n^{(3)}(u)$ läßt sich wie folgt verifizieren:

$$\begin{aligned}
bias_{(erster\ Ordnung)} &= E \left(\frac{2 \left(\sum_{i=1}^{N-1} (T_i - 1) \right)^2}{(N-1) \sum_{i=1}^{N-1} (T_i - 1)(T_i - 2)} \right) - \theta \approx \\
&\approx \frac{2 \left(\sum_{i=1}^{N-1} E(T_i - 1) \right)^2}{(N-1) \sum_{i=1}^{N-1} E((T_i - 1)(T_i - 2))} - \theta = \\
&= \frac{2(N-1)^2 (E(T_1 - 1))^2}{(N-1)(N-1)E((T_1 - 1)(T_1 - 2))} - \theta = \\
&= \frac{2((E(T_1 - 1))^2)}{E((T_1 - 1)(T_1 - 2))} - \theta \stackrel{(6.8)}{=} \\
&= \theta - \theta = 0
\end{aligned}$$

Die kleinsten Interexzedentzeiten werden vom Grenzmodell (6.4) so wie 0 modelliert. Der Schätzer $\hat{\theta}_n^{(3)}(u)$ sichert, dass der Einfluß der kleinsten Interexzedentzeiten tatsächlich 0 ist. Die großen Interexzedentzeiten sind so gut wie nicht betroffen. Der Schätzer $\hat{\theta}_n^{(3)}(u)$ kann Werte annehmen, die größer als 1 sind, und ist für Interexzedentzeiten, die nicht größer als 2 sind, nicht definiert. Um diese Fälle zu vermeiden, definiert man sich den

Intervallschätzer für den extremalen Index θ :

$$\theta_n^{(Intervall)}(u) = \begin{cases} \min(1, \hat{\theta}_n^{(2)}(u)), & \text{wenn } \max\{T_i : 1 \leq i \leq N-1\} \leq 2, \\ \min(1, \hat{\theta}_n^{(3)}(u)), & \text{wenn } \max\{T_i : 1 \leq i \leq N-1\} > 2. \end{cases} \quad (6.9)$$

Die Konsistenz des Intervallschätzers wird im Theorem 6.2 geklärt. Dabei erinnern wir uns, dass für einen positiven Integer m , die Folge $(X_n)_{n \geq 1}$ genau dann m -unabhängig ist, wenn für alle positiven Integer k , die σ -Algebra $\sigma(X_i : 1 \leq i \leq k)$ und die $\sigma(X_i : i \geq m+k)$ unabhängig sind.

Theorem 6.2. (*Konsistenz des Intervallschätzers*)

Sei r_n positiv und u_n der Schwellenwert, $n \geq 1$, so dass

$$\begin{aligned}
r_n &= o(n) \\
r_n &\rightarrow \infty \\
r_n \overline{F}(u_n) &\rightarrow \tau
\end{aligned}$$

und

$$P(M_{r_n} \leq u_n) \rightarrow \exp(-\theta \tau),$$

für $\tau \in (0, \infty)$ und $\theta \in (0, 1]$. Wenn $(X_n)_{n \geq 1}$ m -unabhängig ist, gilt:

$$\tilde{\theta}_n(u_n) \xrightarrow{p} \theta$$

Beweis. vgl. Ferro und Segers (2002) [7] (S. 12ff, Appendix B, Theorem 3.1) □

6.2.2 Maximum Likelihood Schätzer

Man kann den extremalen Index mit Maximum Likelihood schätzen. Die Schätzer werden unter der Annahme konstruiert, dass die Interexzedentenzeiten unabhängig sind. Dann gibt es einen Schätzer bzgl. des Grenzmodells (6.4) und einen bzgl. des Modells mit der Verteilung (6.7) für den extremalen Index.

a) Sei $t_i = \frac{NT_i}{n}$, dann ergibt sich der MLE des Grenzmodells (6.4) aus:

$$l(\theta) = 2(N - 1) \log(\theta) - \theta \sum_{i=1}^{N-1} t_i$$

wegen

$$\begin{aligned} l(\theta) &= \sum_{i=1}^{N-1} [\log((1 - \theta)^{I(t_i=0)}) + \log(\{\theta^2 \exp(-\theta t_i)\}^{I(t_i>0)})] = \\ &= \sum_{i=1}^{N-1} \log[(1 - \theta)^{I(t_i=0)} \{\theta^2 \exp(-\theta t_i)\}^{I(t_i>0)}] \stackrel{T_i>0 \cong t_i>0}{=} \\ &= \sum_{i=1}^{N-1} \log[\theta^2 \exp(-\theta t_i)] = \\ &= \sum_{i=1}^{N-1} [\log(\theta^2) + \log(\exp(-\theta t_i))] = \\ &= \sum_{i=1}^{N-1} [2 \log(\theta) - \theta t_i] = \\ &= 2(N - 1) \log(\theta) - \theta \sum_{i=1}^{N-1} t_i, \quad \text{mit } 0 \leq \theta \leq 1. \end{aligned}$$

Demnach ist der MLE von θ das $\min(1, \frac{2}{\bar{t}})$, wobei \bar{t} der Erwartungswert von t_i ist. Dieser Schätzer konvergiert für wachsende n gegen 1. Dies ist ein Fehler des Modells, das alle Interexzedentenzeiten der exponentialen Komponente der Mischungsverteilung zuteilt. Dieser Fehler kann behoben werden, indem man bestimmte kleinste Exzedenten zu einer Gruppe verbindet und mit Null bewertet. Dazu braucht man aber wieder einen Hilfsparameter. Die Wahl ist zufällig und kann signifikante Auswirkungen auf den Schätzer haben und somit ist die Methode wieder so sensibel wie die Blockmethode oder run Methode.

b) Der MLE des Modells mit der Verteilung (6.7) ergibt sich aus:

$$l(\theta) = m_1 \log(1 - \theta p^\theta) + (N - 1 - m_1) (\log(\theta) + \log(1 - p^\theta)) + \theta \log(p) \sum_{i=1}^{N-1} (T_i - 1), \quad (6.10)$$

wobei m_i die Anzahl der Interexzedentenzeiten ist, die Länge i haben. Der MLE kann durch numerische Methoden gefunden werden.

Die Formel (6.10) gilt, da

$$\begin{aligned}
l(\theta) &= \log \left\{ P(T=1)^{m_1} \prod_{i=2}^{\infty} P(T=i)^{m_i} \right\} = \\
&= \log \left\{ (1 - P(T > 1))^{m_1} \prod_{i=2}^{\infty} [P(T > i-1) - P(T > i)]^{m_i} \right\} = \\
&= \log \left\{ (1 - \theta p^\theta)^{m_1} \prod_{i=2}^{\infty} [\theta p^{(i-1)\theta} - \theta p^{i\theta}]^{m_i} \right\} = \\
&= \log \left\{ (1 - \theta p^\theta)^{m_1} \prod_{i=2}^{\infty} [\theta p^{i\theta} (p^{-\theta} - 1)]^{m_i} \right\} = \\
&= \log(1 - \theta p^\theta)^{m_1} + \sum_{i=2}^{\infty} \log \theta^{m_i} + \sum_{i=2}^{\infty} \log p^{m_i i \theta} + \sum_{i=2}^{\infty} \log (p^{-\theta} - 1)^{m_i} = \\
&= m_1 \log(1 - \theta p^\theta) + \log \theta \sum_{i=2}^{\infty} m_i + \log(p^\theta) \sum_{i=2}^{\infty} m_i i + \sum_{i=2}^{\infty} m_i \log \left(\frac{1 - p^\theta}{p^\theta} \right) = \\
&= m_1 \log(1 - \theta p^\theta) + \log \theta \sum_{i=2}^{\infty} m_i + \log(p^\theta) \sum_{i=2}^{\infty} m_i i + \sum_{i=2}^{\infty} m_i \log(1 - p^\theta) - \sum_{i=2}^{\infty} m_i \log(p^\theta) = \\
&= m_1 \log(1 - \theta p^\theta) + (\log \theta + \log(1 - p^\theta)) \sum_{i=2}^{\infty} m_i + \log(p^\theta) \left[\sum_{i=2}^{\infty} (m_i i - m_i) \right]
\end{aligned}$$

Mit $\sum_{i=2}^{\infty} (m_i i - m_i) = \sum_{i=1}^{N-1} (T_i) - m_1 - (N - 1 - m_1) = \sum_{i=1}^{N-1} (T_i) - (N - 1)$ folgt:

$$l(\theta) = m_1 \log(1 - \theta p^\theta) + (N - 1 - m_1) (\log \theta + \log(1 - p^\theta)) + \theta \log(p) \sum_{i=1}^{N-1} (T_i - 1)$$

6.3 Bootstrap Prozess

Wie schon erwähnt, benötigen die bisherigen Methoden einige Hilfsparameter zur Bestimmung der Cluster. Dies soll nun mit Hilfe der Grenzverteilung (6.4) umgangen werden. Die Grenzverteilung unterstützt einen Bootstrap Prozess, der es ermöglicht Konfidenzintervalle der Schätzer zu berechnen, um die Schätzungenauigkeiten zu ermitteln. Man nutzt den Grenzprozess der Exzedentenanzahl aus, der ein Poissonprozess ist, und unterscheidet die Abhängigkeit zwischen den einzelnen Clustern und dem abhängigen Auftreten der Exzedenten innerhalb eines Clusters.

Man betrachtet N **Zeiten** $S_1 \leq \dots \leq S_N$, zu denen Exzedenten auftreten, und die **Interexzedentenzeiten** $T_i = S_{i+1} - S_i$, für $i = 1, \dots, N - 1$.

Demnach kann man laut Ferro und Segers [8] (S. 551) annehmen, dass die

$$C - 1 = \lfloor \theta N \rfloor \tag{6.11}$$

längsten Interexzedentenzeiten nahezu unabhängig sind.

Im folgendem Abschnitt seien **IEZ** die Interexzedentenzeiten, **ICZ** die Interclusterzeiten und **IntraCZ** seien die Intraclusterzeiten.

Sei $T_{(C)}$ die C -te längste **IEZ**
und
 T_{i_j} die j -te **IEZ**, die $T_{(C)}$ überschreitet.

Dann ist $\{T_{i_j}\}_{j=1}^C$ eine Gruppe von nahezu unabhängigen **ICZ**'s.

$$\text{Sei } \tau_j = \{T_{i_{j-1}+1}, \dots, T_{i_j}\},$$

wobei $i_0 = 0$, $i_C = N$ und $\tau_j = \emptyset$, wenn $i_j = i_{j-1} + 1$.

Dann ist $\{\tau_j\}_{j=1}^C$ eine Sammlung von nahezu unabhängigen **IntraCZ**'s. Weiter ist mit jedem τ_j ein $C_j = \{X_k : k \in S_j\}$ verbunden, wobei $S_j = \{S_{i_{j-1}+1}, \dots, S_{i_j}\}$ ist.

Durch diese Interpretation kann man den Prozess in C Cluster aufteilen, wobei das j -te Cluster die Exzedenten von C_j enthält. Laut Ferro und Segers [8] (S. 551) ist dies dasselbe wie die run Methode mit einer Laufzeit von $T_{(C)}$. In der Praxis wird θ durch $\hat{\theta}$ geschätzt.

Eine Übersicht über diese Bezeichnungen und Berechnungen der Clusteridentifizierung sind in Abbildung 6.1 dargestellt.

Jetzt soll eine Clusterfunktion H (z.B. die „excess height statistic“ $Z_n = \sum(X_j - u_n)^+$ vgl. Leadbetter [16] (S. 257)) hergeleitet werden, so dass man für jedes Cluster die Werte $\{H\}_{j=1}^C$ erhält, die verwendet werden können, um eine Schätzung \tilde{H} für die Wahrscheinlichkeit H zu finden. Man kann z.B. H durch $\tilde{H} = C^{-1}(H_1 + \dots + H_C)$ schätzen. Den Bootstrap Algorithmus kann man dazu nutzen, ein Konfidenzintervall für \tilde{H} und $\hat{\theta}$ zu bekommen.

Bootstrap Algorithmus

- Resample mit dem Ersetzen von $C - 1$ ICZ's von $\{T_{i_j}\}_{j=1}^{C-1}$.
- Resample mit dem Ersetzen von C Gruppen von IntraCZ's (einige davon werden leer sein) und assoziiere damit die Exzedenten von $\{(\tau_j, C_j)\}_{j=1}^C$.

- Benutze diese IEZ's und ICZ's um eine Bootstrap Replikation des Prozesses zu beschreiben
- Berechne N für die Bootstrap Replikation und schätze θ
- Berechne \overline{H} der clusterbildenden Bootstrap Replikation

Man erhält nach B Bootstrap Prozessen die Schätzungen $(\hat{\theta}_1, \dots, \hat{\theta}_B)$ und $(\overline{H}_{(1)}, \dots, \overline{H}_{(B)})$. Damit kann man die Verteilung des Originals annähern. Das empirische α - und das $(1-\alpha)$ -Quantil jedes Samples definieren ein $(1 - 2\alpha)$ -Konfidenzintervall.

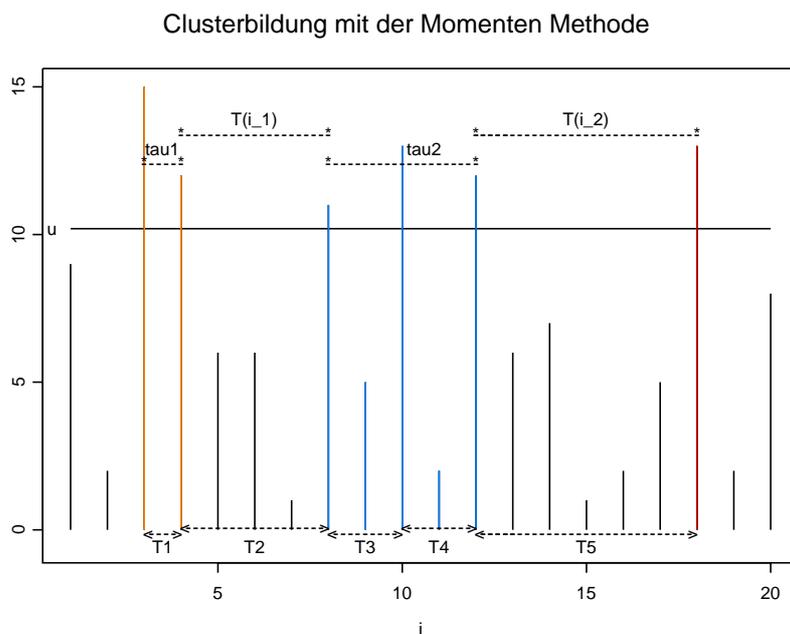


Abbildung 6.1:

Theoretisches Beispiel für die Momenten Methode. Die farbigen Linien seien die Cluster. Wenn man $C = 3$ wählt, dann sucht man jetzt die drittgrößte IEZ. $T_C = T_3 = 2$. Also sind T_2 und T_5 die zwei größten IEZ's, die T_C überschreiten. Demnach ist $i_1 = 2$ und $i_2 = 5$. Mit diesen Werten lassen sich dann die τ_j berechnen. $\tau_1 = \{T_{i_0+1}, \dots, T_{i_1-1}\} = \{T_{0+1}, \dots, T_{2-1}\} = \{T_1, \dots, T_1\} = T_1$ und $\tau_2 = \{T_{i_1+1}, \dots, T_{i_2-1}\} = \{T_3, \dots, T_4\}$. In der Praxis wird C nicht gewählt, sondern mit (6.11) berechnet. Dies ist dasselbe Ergebnis wie bei der run Methode mit run Länge $s = T_C = 2$.

Kapitel 7

Neue Methode 2: Kombination der subject-matter Methode und der run Methode

Schon der Intervallschätzer aus Kapitel 6 (von Ferro und Segers) ist stabiler bzgl. der Wahl der Hilfsparameter als die Schätzer der run Methode oder der Blockmethode, jedoch nur für hohe Schwellenwerte u . Die Methode, die in diesem Kapitel dargestellt wird, ist generell stabiler. Das Ziel ist es, die subject-matter Methode, die im letzten Absatz kurz beschrieben wird, mit der run Methode zu kombinieren, um ein neues Schema zur Clusteridentifikation zu erhalten.

Es ist sehr wichtig für die Schätzung von θ ein gutes Schema der Clusteridentifikation zu haben. Dazu betrachtet man die Prozesse um die Exzedenten, entweder mit einem assoziierten autoregressiven Verhalten für extreme Abhängigkeiten (vgl. Abschnitt 7.1) oder mit einem Volatilitätsmechanismus (vgl. Abschnitt 7.2). Diese Abhängigkeitstypen charakterisieren einen Unterraum der Prozesse, und man glaubt, dass diese Typen praktische Situationen sehr gut wiedergeben und eine suffiziente Basis bilden, um bessere Annäherungen zu liefern. Demnach spielt der Prozess, der um die extremen Werte stattfindet, eine große Rolle.

Sowohl Block- als auch die run Methode sind eher naiv, weil sie eben nur grobe Informationen über den Prozess, der um die extremen Werte stattfindet, verwenden. Die run Methode ignoriert z.B. die $s - 1$ aufeinander folgenden Beobachtungen nach einem Exzedenten und setzt voraus, dass diese keine Exzedenten sind. Es könnte jedoch sein, dass der Prozess knapp unter der Schwelle u für die $s - 1$ Werte verbleibt und nach wenigen lags wieder übersteigt. Dann würden die alten Methoden ein Cluster beenden, obwohl die $s - 1$ Werte nur knapp unter der Schwelle gelegen waren. Die alten Methoden berücksichtigen also die „Bahn“ bzw. den „Pfad“ des Prozesses nicht.

Bisher gibt es vier Möglichkeiten um Cluster zu identifizieren: Blockmethode, run Methode, fixierte Längen über lokalen Ausreißern und niedrige Schwellen. Die vierte Möglichkeit macht man sich in der neuen Methode zu Nutze. In der Subject-matter Methode für Fi-

nanzzeitreihen werden diese Zeitreihen in ihre Volatilitäts- und Residuenkomponenten aufgeteilt. Dabei werden die drei Typen der Volatilitäts-Abhängigkeits-Modelle ARCH, SV und GARCH verwendet.

Die Kombination der Struktur der subject-matter Methode und der run Methode führt zu einem neuen Schema der Clusterbildung.

7.1 Zwei-Schwellen Index

Die Erweiterung durch eine zweite Schwelle beendet ein Cluster, wenn ein genügend kleiner Wert des Prozesses zwischen zwei Exzedenten gefunden wird.

Sei $\omega_F = \inf\{x : F(x) > 0\}$ und $\omega^F = \sup\{x : F(x) < 1\}$. Die zwei Schwellenwerte seien u und c mit $w_F \leq c \leq u < w^F$ und

$$L_{i,j}^X = \min(X_i, \dots, X_j) \text{ und } M_{i,j}^X = \max(X_i, \dots, X_j) \quad \text{für } i < j. \quad (7.1)$$

Die Schätzung für θ ist gegeben durch

$$\theta_X^{(zwei-Schwellen)}(u, s, c) = P \left\{ M_{2,s}^X \leq u \text{ oder } \bigcup_{i=2}^{s-1} T_{i,c,u} | X_1 > u \right\} \quad (7.2)$$

mit

$$T_{i,c,u} = \{c < L_{2,i-1}^X, M_{2,i-1}^X < u, X_i \leq c\}.$$

Die Formel (7.2) kann man auch schreiben als

$$\theta_X^{(zwei-Schwellen)}(u, s, c) = P(c < L_{2,s}^X, M_{2,s}^X \leq u | X_1 > u) + \sum_{i=2}^s P(T_{i,s,c} | X_1 > u). \quad (7.3)$$

Es ist klar, dass $\theta_X^{(zwei-Schwellen)}(u, s, c) \geq \theta^{(run)}(u, s)$. Weiter sieht man an der Formel (7.2), dass sowohl das Kriterium der run Methode einfließt ($M_{2,s} \leq u$), aber zusätzlich wird ein Cluster identifiziert, falls der Prozess unter den Schwellenwert c fällt.

Wenn man die Abbildung 7.1 mit Abbildung 5.2 vergleicht, sieht man sofort, dass die Anzahl der Cluster gestiegen ist und die Cluster kleiner geworden sind. Das liegt eben an der zusätzlich eingeführten zweiten Grenze, die den Ablauf des Prozesses berücksichtigt.

Für $s = 2$ ist die Schwelle c überflüssig, also ist $\theta_X^{zwei-Schwellen}(u, 2, c) = \theta^{(run)}(u, 2)$. Liegt die Schwelle c nahe an der Schwelle u , so entspricht die Zwei-Schwellen Methode ungefähr der run Methode mit $s = 2$, da ein Cluster schnell beendet wird. Jedoch eine Schwelle c , die sich eher ω_F annähert, wird im Vergleich zur run Methode fast keine zusätzlichen Cluster ergeben. Demnach gilt, dass

$$\begin{aligned} \lim_{c \rightarrow u^+} \theta_X^{(zwei-Schwellen)}(u, s, c) &= \theta^{(run)}(u, 2) \text{ und} \\ \lim_{c \rightarrow \omega_F^-} \theta_X^{(zwei-Schwellen)}(u, s, c) &= \theta^{(run)}(u, s). \end{aligned}$$

Mit dieser Wahrscheinlichkeit und den Ergebnissen von O'Brien [21] ist klar, dass für $n \rightarrow \infty$ und $c_n \rightarrow \omega_F$ gilt $\theta_X^{(zwei-Schwellen)}(u_n, s_n, c_n) = \theta$.

7.2 Volatilitäts-Schwellen Index

Ähnlich wie beim Zwei-Schwellen Index wird beim Volatilitätsindex das Cluster beendet, wenn die lokale Varianz genügend klein ist. Dabei wird angenommen, dass die größten Werte von der lokalen Varianz des Prozesses beeinflusst werden.

Wir gehen also nun davon aus, dass die Cluster von der Volatilität des Prozesses gebildet werden. Dann hat der Prozess die Form

$$X_t = \sigma_t Z_t,$$

mit $E(Z_t) = 0$, $Z_t \sim iid$, $E(Z) = 0$, $Var(Z_t) = \sigma_t^2$. Für solch einen Prozess tendiert X_{t+1} dazu, in den niedrigen Tail genauso oft wie in den fetteren Tail zu fallen, wobei X_t ein Exzedent ist. Also ist der Schätzer $\theta_X^{(zwei-Schwellen)}$ nicht so geeignet. Hier gibt der Volatilitätsmechanismus bessere Kenntnisse über den Pfad des Prozesses der extremen Ereignisse wieder. Asymptotische Annäherungen zu diesem Problem findet man unter de Haan, Resnick, Rootzén und de Vries [4], Breidt und Davis [3] und Mikosch und Stărică [20].

Um eine asymptotische Annäherung dieses Problems zu bekommen, untersucht man das Quadrat des Prozesses. Dennoch liefert der Schätzer $\theta_{X^2}(u, s, c)$ keine gute Annäherung, wenn kleine Werte in Perioden mit hoher Volatilität auftreten. Also ist der Volatilitäts-Schätzer von θ

$$\theta_\sigma^{(Volatilität)}(u, s, c) = P \left(M_{2,s}^X \leq u \text{ oder } \bigcup_{i=2}^s V_{i,c,u} | X_1 > u \right) \quad (7.4)$$

mit $V_{i,c,u} = \{c < L_{2,i-1}^\sigma, M_{2,i}^X \leq u, \sigma_i \leq c\}$ für $i = 2, \dots, s$.

Mit der Definition (7.1) können wir die vorhergehende Wahrscheinlichkeit (7.4) umschreiben in

$$\theta_\sigma^{(Volatilität)}(u, s, c) = P \{c < L_{2,s}^\sigma, M_{2,s}^X \leq u\} + \sum_{i=2}^s P\{V_{i,c,u} | X_1 > u\}.$$

Wieder wie beim Zwei-Schwellen Index gilt auch hier $\theta_\sigma^{(Volatilität)}(u, s, c) \geq \theta^{(run)}(u, s)$ und es können folgende Grenzwerte gebildet werden:

$$\begin{aligned} \lim_{c \rightarrow \infty} \theta_\sigma^{(Volatilität)}(u, s, c) &= \theta^{(run)}(u, 2) \text{ und} \\ \lim_{c \rightarrow 0} \theta_\sigma^{(Volatilität)}(u, s, c) &= \theta^{(run)}(u, s). \end{aligned}$$

Mit dieser Wahrscheinlichkeit und $c_n \rightarrow 0$ $n \rightarrow \infty$ gilt $\theta_\sigma^{Volatilität}(u_n, s_n, c_n) = \theta$ für jeden Prozess.

Das einzige Problem ist, dass man angenommen hat, dass der Volatilitätsprozess bekannt ist. Der muß in der Realität natürlich geschätzt werden.

Zur Veranschaulichung ein theoretisches Beispiel zur Zwei-Schwellen Methode:

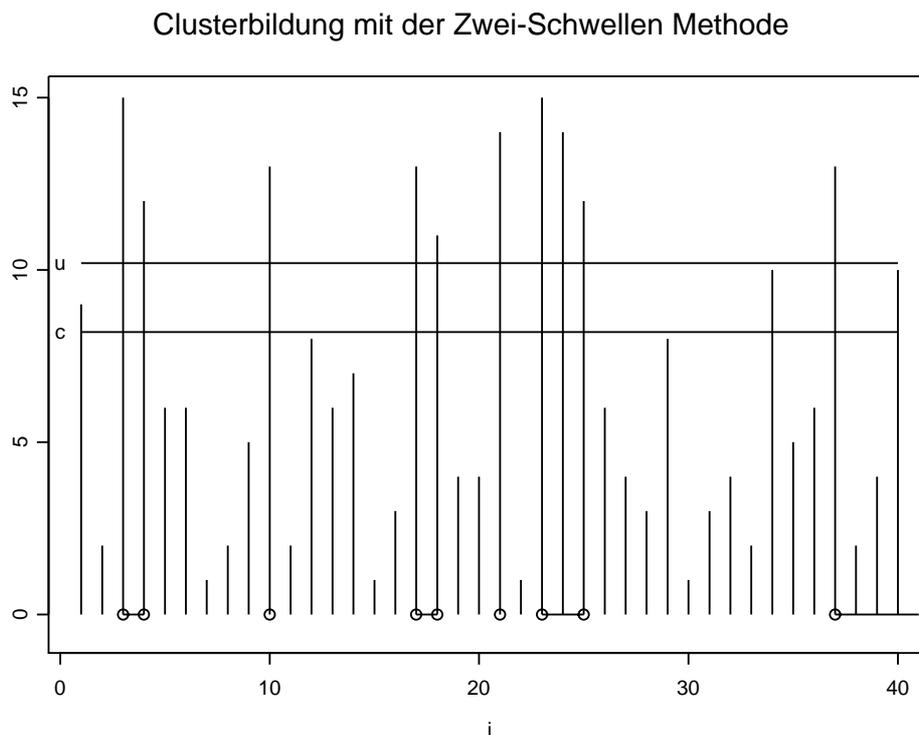


Abbildung 7.1:

Theoretisches Beispiel für die Zwei-Schwellen Methode mit der run Länge $s = 5$. u ist der obere Schwellenwert und c die zusätzliche untere Grenze. Es wurden die gleichen Daten wie bei Abbildung 5.1 und Abbildung 5.2 verwendet. Die Punkte kennzeichnen den Beginn bzw. das Ende eines Clusters. Ein Cluster wird dadurch kenntlich gemacht, dass die Daten mit einer Linie verbunden sind. Also sind die Daten mit i von 3 - 4, 10, 17 - 18, 21, 23 - 25, 36 - ..., Cluster. Wie der Datenpunkt mit $i = 10$ zeigt, kann ein Cluster im Extremfall auch nur aus einem Datenpunkt bestehen.

Verwendet man die Volatilitäts-Schwellen Methode muß zuerst der Volatilitätsprozess geschätzt werden und dann ist die Vorgehensweise die gleiche wie bei der Zwei-Schwellen Methode.

7.3 Schätzung des extremalen Index θ

Diese Schätzung richtet sich nach dem Schätzer für θ von Smith und Weisman [22], der $\hat{\theta}$ als einen empirischen Schätzer der Wahrscheinlichkeit definiert.

Sei $W_{n,i} = I_{\{X_i > u\}}$, die Anzahl der Exzedenten $N_n(u_n) = \sum_{i=1}^n W_{n,i}$ und die Anzahl der Cluster

$$C_n(u_n) = \sum_{i=1}^{n-(s_n-1)} W_{n,i} (1 - W_{n,i+1}) \cdots (1 - W_{n,i+(s_n-1)}).$$

Mit (7.3) kann man die Schätzung für θ durch

$$\hat{\theta}_{X,n}^{(\text{zwei-Schwellen})}(u_n, s(n), c_n) = \frac{C_{X,n}(u_n)}{N_n(u_n)} \quad (7.5)$$

bestimmen, wobei

$$C_{X,n}(u_n) = \sum_{i=1}^{n-(s_n-1)} W_{n,i} \left(\prod_{j=1}^{s_n-1} B_{n,i+j}^X + \sum_{k=1}^{s_n-1} S_{n,i+k}^X \prod_{r=1}^{k-1} B_{n,i+r}^X \right),$$

mit $B_{n,i}^X = I_{\{c_n < X_i \leq u_n\}}$ und $S_{n,i}^X = I_{\{X_i \leq c_n\}}$

Der Schätzer für $\theta_\sigma^{(\text{Volatilität})}$ ist demnach gegeben durch

$$\hat{\theta}_{\sigma,n}^{(\text{Volatilität})}(u_n, s(n), c_n) = \frac{C_{\sigma,n}(u_n)}{N_n(u_n)}, \quad (7.6)$$

wobei

$$C_{\sigma,n}(u_n) = \sum_{i=1}^{n-(s_n-1)} W_{n,i} \left(\prod_{j=1}^{s_n-1} B_{n,i+j}^\sigma + \sum_{k=1}^{s_n-1} S_{n,i+k}^\sigma \prod_{r=1}^{k-1} B_{n,i+r}^\sigma \right),$$

mit $B_{n,i}^\sigma = I_{\{X_i \leq u_n, \hat{\sigma}_i \geq c_n\}}$ und $S_{n,i}^\sigma = I_{\{X_i \leq u_n, \hat{\sigma}_i \leq c_n\}}$.

Kapitel 8

Simulationsstudie

8.1 Unabhängige Prozesse

Hat man nun X_1, \dots, X_n iid verteilte Zufallsvariablen mit Verteilungsfunktion F , so ist $\theta = 1$ (siehe Laurini und Tawn [14], S. 198).

Lemma 8.1. (*Extremaler Index θ eines unabhängigen Prozesses mit der Zwei-Schwellen Methode*)

Mit $\theta^{(run)} \stackrel{(5,9)}{=} P(M_{2,s}^X \leq u) = \{F(u)\}^{s-1}$ und (7.3), folgt

$$\begin{aligned} \theta_X^{(zwei-Schwellen)} &= P\{c < L_{2,s}^X, M_{2,s}^X \leq u | X_1 > u\} + \sum_{i=2}^s P\{T_{i,c,u} | X_1 > u\} = \\ &= \frac{F(c) + \{F(u) - F(c)\}^{s-1} \{1 - F(u)\}}{1 - \{F(u) - F(c)\}} \end{aligned}$$

Beweis.

$$\begin{aligned} &P\{c < L_{2,s}^X, M_{2,s}^X \leq u | X_1 > u\} + \sum_{i=2}^s P\{T_{i,c,u} | X_1 > u\} \\ &= P\{c < L_{2,s}^X, M_{2,s}^X \leq u | X_1 > u\} + \sum_{i=2}^s P\{c < L_{2,i-1}^X, M_{2,i-1}^X < u, X_i \leq c | X_1 > u\} \\ &= \{F(u) - F(c)\}^{s-1} + \sum_{i=2}^s \{F(u) - F(c)\}^{i-2} F(c) = \\ &= \{F(u) - F(c)\}^{s-1} + \sum_{i=0}^{s-2} \{F(u) - F(c)\}^i F(c) = \\ &= \{F(u) - F(c)\}^{s-1} + F(c) \frac{1 - \{F(u) - F(c)\}^{s-1}}{1 - \{F(u) - F(c)\}} = \end{aligned}$$

$$\begin{aligned}
&= \frac{\{F(u) - F(c)\}^{s-1}\{1 - (F(u) - F(c))\} + F(c) - F(c)\{F(u) - F(c)\}^{s-1}}{1 - \{F(u) - F(c)\}} = \\
&= \frac{\{F(u) - F(c)\}^{s-1} - \{F(u) - F(c)\}^s + F(c) - F(c)\{F(u) - F(c)\}^{s-1}}{1 - \{F(u) - F(c)\}} = \\
&= \frac{F(c) + \{F(u) - F(c)\}^{s-1}\{1 - F(u) + F(c) - F(c)\}}{1 - \{F(u) - F(c)\}} = \\
&= \frac{F(c) + \{F(u) - F(c)\}^{s-1}\{1 - F(u)\}}{1 - \{F(u) - F(c)\}}
\end{aligned}$$

□

Weiter ist klar, dass für ein festes c , $c \rightarrow \omega^F$ und $c < u$

$$1 - \theta^{(run)}(u, s) \sim (s - 1)\{1 - F(u)\} \text{ für } u \rightarrow (\omega^F)^+$$

und für ein festes c

$$1 - \theta_X^{(zwei-Schwellen)}(u, s, c) \sim \left[\frac{1 - \{1 - F(c)\}^{s-1}}{F(c)} \right] \{1 - F(u)\}.$$

Für c , so dass $F(c) \rightarrow 1$ mit $\frac{1-F(u)}{1-F(c)} \rightarrow c'$, wobei $0 \leq c' \leq 1$, dann gilt:

$$1 - \theta_X^{(zwei-Schwellen)}(u, s, c) \sim \{1 - F(u)\}$$

Somit folgt zum Schluß in diesem Fall von $\theta = 1$

$$\theta^{(run)}(u, s) \leq \theta_X^{(zwei-Schwellen)}(u, s, c) \leq \theta$$

unabhängig von der Wahl von u, s und c .

8.2 ARMAX Prozess

Wir betrachten den ARMAX(1) Prozess, der definiert ist durch

Definition 8.1. (ARMAX(1))

Sei $0 < \alpha < 1$, $\epsilon_t \sim iid$ und $E(\epsilon_t) = 0$, dann ist

$$X_t = \max(\alpha X_{t-1}, \epsilon_t)$$

ein ARMAX(1) Prozess.

Nun nehmen wir für die Verteilungsfunktion $F_\epsilon = \exp\{-(1 - \alpha)/x\}$ an, denn dann ist (X_t) unit Fréchet verteilt ($\Rightarrow \Phi_1(x) = \exp(-x^{-1})$, für $x > 0$).

Sei

$$\xi_n = \max_{k \geq 0} \rho^k \eta_{n-k},$$

wobei $0 < \rho < 1$ und (η_n) eine iid Folge von Zufallsvariablen mit Dichte $\exp(-1/x)$ ist. Dann ist der extreme Index $\theta = 1 - \rho$ (vgl. Leadbetter [15] (S. 302/303)).

Beim ARMAX(1) Prozess gilt $k = 1$, $\rho = \alpha$ und damit ist der extreme Index $\theta = 1 - \alpha$.

Analytisch ist $\theta^{(run)}(u, s)$ leicht zu berechnen.

$$\begin{aligned} \theta^{(run)}(u, s) &= P(M_{2,s} \leq u | X_1 > u) = P(X_s \leq u, \dots, X_2 \leq u | X_1 > u) = \\ &= P(X_s \leq u | X_{s-1} \leq u, \dots, X_2 \leq u, X_1 > u) P(X_{s-1} \leq u, \dots, X_2 \leq u | X_1 > u) = \\ &= P(X_s \leq u | X_{s-1} \leq u) P(X_{s-1} \leq u, \dots, X_2 \leq u | X_1 > u) \stackrel{X_s \leq u \Rightarrow \alpha X_{s-1} \leq u \Rightarrow}{=} \\ &= P(\epsilon_s \leq u) P(X_{s-1} \leq u, \dots, X_2 \leq u | X_1 > u) = \\ &= \{F_\epsilon(u)\} P(X_{s-1} \leq u, \dots, X_2 \leq u | X_1 > u) = \\ &= \dots \\ &= \{F_\epsilon(u)\}^{s-2} P(X_2 \leq u | X_1 > u) = \{F_\epsilon(u)\}^{s-2} P(\alpha X_1 \leq u, \epsilon \leq u | X_1 > u) = \\ &= P(\alpha X_1 < u | X_1 > u) \{F_\epsilon(u)\}^{s-1} \\ &= \frac{P(\alpha X_1 < u, X_1 > u)}{P(X_1 > u)} \{F_\epsilon(u)\}^{s-1} = \\ &= \frac{P(X_1 < u/\alpha, X_1 > u)}{1 - P(X_1 \leq u)} (\exp\{-(1 - \alpha)/u\})^{(s-1)} = \\ &= \frac{P(u < X_1 < u/\alpha)}{1 - P(X_1 \leq u)} \exp\{-(s-1)(1 - \alpha)/u\} = \\ &= \frac{\exp(-(u/\alpha)^{-1}) - \exp(-u^{-1})}{1 - \exp(-u^{-1})} \exp\{-(s-1)(1 - \alpha)/u\} = \\ &= \frac{\exp(-(\alpha/u)) - \exp(-1/u)}{1 - \exp(-1/u)} \exp\{-(s-1)(1 - \alpha)/u\} \\ &\approx (1 - \alpha)[1 - \{(s-1)(1 - \alpha) + \alpha/2\}/u] \text{ f\u00fcr gro\u00dfe } u \end{aligned} \tag{8.1}$$

Also ist der relative Fehler

$$\begin{aligned} \frac{\theta^{(run)}(u, s) - \theta}{\theta} &\sim \frac{(1 - \alpha)[1 - \{(s-1)(1 - \alpha) + \alpha/2\}/u] - (1 - \alpha)}{1 - \alpha} \\ &= 1 - \{(s-1)(1 - \alpha) + \alpha/2\}/u - 1 \\ &= -\delta(s-1)/u \end{aligned}$$

mit $\delta(k) = k(1 - \alpha) + \alpha/2$.

Sei, wie bekannt, $o(g(x))$ und $O(g(x))$ wie folgt definiert:

Definition 8.2. (klein o und groß O)

Wenn $f(h) = O(g(h))$, dann gilt

$$|f(h)| \leq C|g(h)|.$$

Wenn $f(h) = o(g(h))$, dann gilt

$$\lim_{h \rightarrow 0, h \neq 0} \frac{f(h)}{g(h)} = 0$$

Für $c = o(u)$ gilt, dass

$$\frac{\theta_X^{(zwei-Schwellen)}(u, s, c) - \theta}{\theta} \sim \frac{\theta^{(run)}(u, s) - \theta}{\theta}$$

für $u \rightarrow \infty$.

Für den Fall, dass $c = O(u)$, definieren wir uns eine Variable J , so dass $\alpha^J \leq c' \leq \alpha^{J-1}$ mit $c = uc'$ und $0 < c' < 1$. Demnach ist $X_1 > u$, $X_2 < u$, $X_J > c$ und $X_{J+2} < c$.

Für $J \geq s - 1$ gilt für das approximative Verhalten von $\theta_X^{(zwei-Schwellen)}(u, s, c)$ dasselbe wie bei $c = o(u)$.

Für $J < s - 1$ hängen die Grenzen für $\theta_X^{(zwei-Schwellen)}(u, s, c) - \theta^{(run)}(u, s)$ davon ab, ob $X_{J+1} < c$ oder $X_{J+1} > c$. Dann ist $\theta_X^{(zwei-Schwellen)}(u, s, c) - \theta^{(run)}(u, s)$ gegeben durch die Wahrscheinlichkeit, dass X_2 die Schwelle u unterschreitet, X_{J+1} oder X_{J+2} unterschreitet die Schwelle c und ein X_i mit $i < s$ überschreitet wieder die Schwelle u . Die zwei Möglichkeiten, dass X_{J+1} oder X_{J+2} kleiner als die Schwelle c sind, bestimmen die folgenden zwei Grenzen:

$$-(1 - \alpha)^2(s - J) \leq \lim_{u \rightarrow \infty} u \{ \theta_X^{(zwei-Schwellen)}(u, s, c) - \theta^{(run)}(u, s) \} \leq -(1 - \alpha)^2(s - J - 1) \quad (8.2)$$

Mit (8.1) und den Grenzen aus (8.2) ergibt sich:

$$-\delta(J + 1) \leq \lim_{u \rightarrow \infty} u \left\{ \frac{\theta_X^{(zwei-Schwellen)} - \theta}{\theta} \right\} \leq -\delta(J)$$

In Abbildung 8.1 wurde ein ARMAX(1) Prozess der Länge $n=10000$ mit $\alpha = 0.3$ simuliert. Somit ist der wahre Wert für $\theta = 0.7$, was der horizontalen Linie bei 0.7 entspricht.

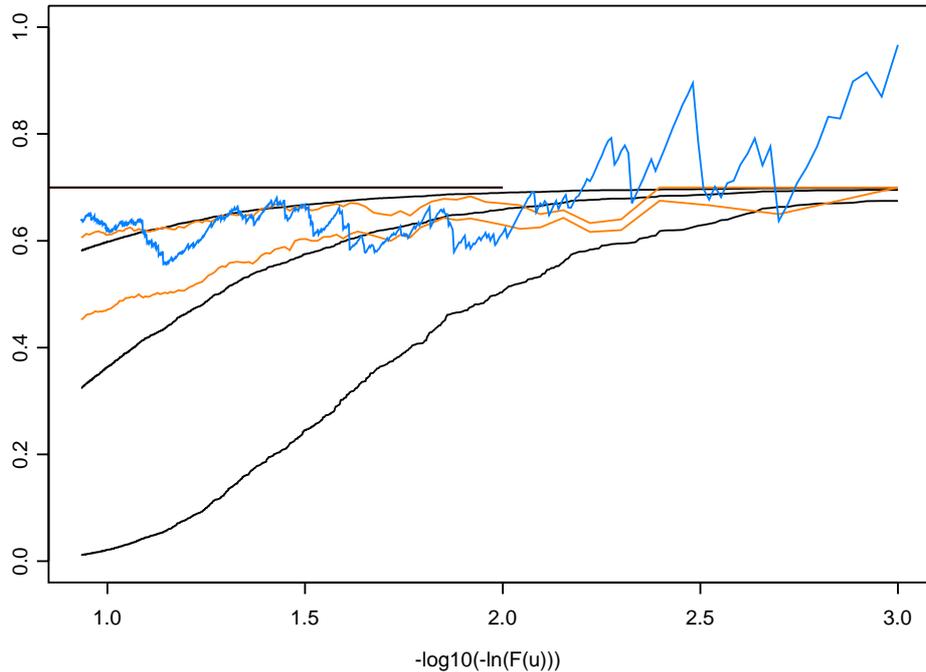


Abbildung 8.1:

10000 Daten eines $ARMAX(1)$ Prozesses mit $\alpha = 0.3$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_X^{(zwei-Schwellen)}$ (orange) für $s = 50$ und für c mit $F(c) = 0.75$ und $F(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet. $\theta = 0.7$ ist durch die horizontalen Linie kenntlich gemacht.

Für die obere Schwelle u wurde mit der Funktion *threshold* (vgl. Anhang A.2) das 89% bis 99.9%-Quantil berechnet und in einer $-\log_{10}(-\ln(F(u)))$ Skala abgebildet.

$\theta^{(run)}$ wurde für $s = 3, 10$ und 50 berechnet. Dabei wurde die nicht approximierte Formel (8.1) verwendet. Man sieht sehr schön, dass die run Methode wieder sehr sensibel auf die Wahl von s reagiert hat. Die Zwei-Schwellen Methode nähert sich schnell und stabil dem wahren Wert von θ an, dabei wurde zur Schätzung von θ die Formel (7.5) angewendet mit der run Länge $s = 50$ und c mit $F(c) = 0.75$ und $F(c) = 0.25$, wobei F die Verteilungsfunktion der Daten darstellt. Somit ist also c der Wert des 75%- und des 25%-Quantils der 10000 simulierten Daten des $ARMAX(1)$ Prozesses. Der Intervallschätzer schwank zwar sehr, aber ist trotzdem stabiler als die run Methode und ist nur von der Wahl von u abhängig. Der Grund für die größeren Sprünge der Werte ab $-\log_{10}(-\ln(F(u))) = 2$ liegt an der mit dem Logarithmus transformierten x-Achse.

8.3 AR Prozess

Im Folgenden soll ein AR(1) Prozess untersucht werden.

Definition 8.3. (AR(1) Prozess)

Sei

$$X_t = \alpha X_{t-1} + Z_t \text{ mit } Z_t \stackrel{iid}{\sim} N(0, 1).$$

Dann heißt X_t AR(1) Prozess.

In Abbildung 8.2 wurde ein AR(1) Prozess der Länge $n=10000$ mit $\alpha = 0.3$ simuliert. Die Wahl von u ist dieselbe wie im ARMAX(1) Prozess.

Bei der Berechnung von $\theta^{(run)}$ mit (5.10) wurde $s = 3, 10$ und 50 gewählt. Die Zwei-Schwellen Methode reagiert wie immer weniger auf die Wahl der Hilfsparameter als die Schätzer der run Methode. Für $\theta_X^{(zwei-Schwellen)}$ wurde wieder (7.5) mit der run Länge

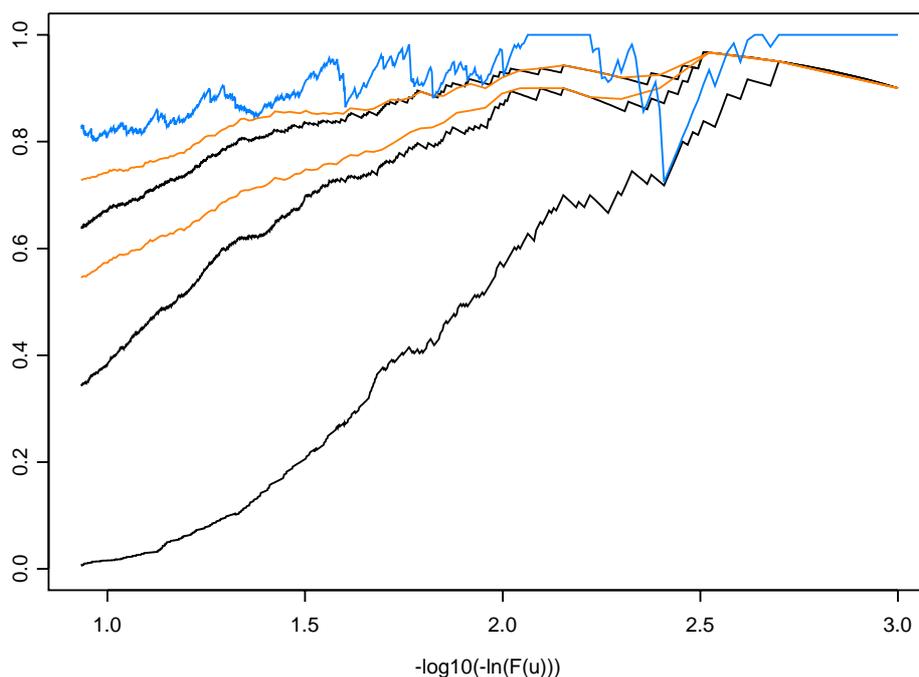


Abbildung 8.2:

10000 Daten eines AR(1) Prozesses mit $\alpha = 0.3$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_X^{(zwei-Schwellen)}$ (orange) für $s = 50$ und für c mit $F(c) = 0.75$ und $F(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet.

$s = 50$ und c mit $F(c) = 0.75$ und $F(c) = 0.25$ angewenden, wobei F die Verteilungsfunktion der Daten darstellt. Somit ist also c der Wert des 75%- und des 25%-Quantils der 10000 simulierten Daten des AR(1) Prozesses. Der Intervallschätzer verhält sich ebenfalls wieder stabiler als die run Methode, jedoch schwankt er etwas mehr als der Schätzer der Zwei-Schwellen Methode.

8.4 ARCH Prozess

Wir betrachten nun den ARCH(1) Prozess, der definiert ist durch

Definition 8.4. (ARCH(1) Prozess)

Mit

$$X_t = \sigma_t Z_t \text{ mit } Z_t \stackrel{iid}{\sim} N(0, 1),$$

wobei

$$\sigma_t^2 = \beta + \lambda X_{t-1}^2$$

ist X_t ein *ARCH(1) Prozess*.

β und λ sind positive Parameter, damit die Stationarität gesichert ist. de Haan, Resnick und Rootzén [4] haben den extreme Index in diesem Fall durch einen random walk, der auf Monte Carlo Methoden beruht, hergeleitet. Jedoch nur λ beeinflusst unseren extremen Index. Hier betrachtet man dann natürlich $\theta_\sigma^{(Volatilität)}(u, s, c)$. Wählt man $\lambda = 0.9$ und $\beta = 0.5$ ist der Wert für $\theta = 0.612$ (vgl. Laurini und Tawn [14] (S. 202) und Embrechts, Klüppelberg und Mikosch [5](S. 480)).

In Abbildung 8.3 wurde ein ARCH(1) Prozess der Länge $n=10000$ mit $\lambda = 0.9$ und $\beta = 0.5$ simuliert. Der wahre Wert für $\theta = 0.612$ wurde von Embrechts, Klüppelberg und Mikosch [5] berechnet. Was die Wahl von u angeht, siehe ARMAX(1) Prozess.

$s = 3, 10$ und 50 wurde bei der Schätzung von $\theta^{(run)}$ in (5.10) eingesetzt. Es ist zu beachten, dass bei einem ARCH Prozess wegen dem Volatilitätsmechanismus die Volatilitäts-Schwellen Methode zu wählen ist. Dazu wird die Formel (7.6) verwendet. Die Schätzung für $\theta_\sigma^{(Volatilität)}$ wurde mit der run Länge $s = 50$ durchgeführt und c mit $F_\sigma(c) = 0.75$ und $F_\sigma(c) = 0.25$ angewenden, wobei F_σ die Verteilungsfunktion des geschätzten Volatilitätsprozesses σ_t ist. c ist also hier der Wert des 75%- und des 25%-Quantils der 10000 simulierten σ_t . Der Intervallschätzer konvergiert ebenso gegen den wahren Wert von θ und schwankt um 0.612. Der Grund für die größeren Sprünge der Werte ab $-\log_{10}(-\ln(F(u))) = 2$ liegt wieder an der mit dem Logarithmus transformierten x-Achse.

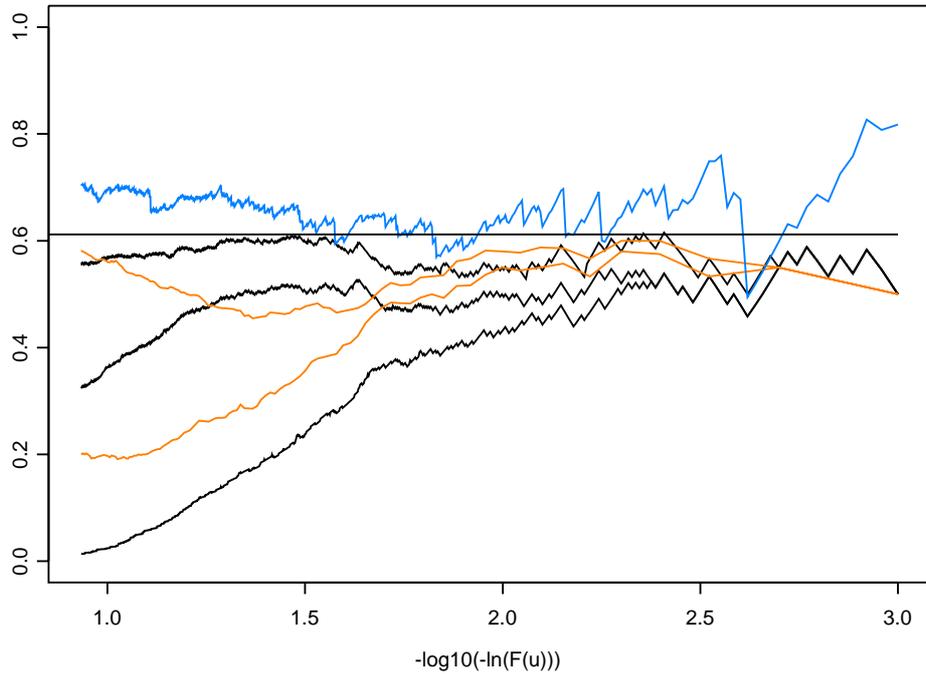


Abbildung 8.3:

10000 Daten eines $ARCH(1)$ Prozesses mit $\lambda = 0.9$ und $\beta = 0.5$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_\sigma^{(Volatilität)}$ (orange) für $s = 50$ und für c mit $F_\sigma(c) = 0.75$ und $F_\sigma(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet. $\theta = 0.612$ ist durch die horizontalen Linie kenntlich gemacht.

8.5 GARCH Prozess

Der $GARCH(1,1)$ Prozess, den wir untersuchen, hat folgende Form

Definition 8.5. (GARCH Prozess)

Sei

$$X_t = \sigma_t Z_t \text{ mit } Z_t \stackrel{iid}{\sim} N(0, 1),$$

wobei

$$\sigma_t^2 = \beta + \lambda X_{t-1}^2 + \gamma \sigma_{t-1}^2.$$

Dann heißt X_t $GARCH(1,1)$ Prozess.

Ein $GARCH(1,1)$ Prozess ist ein Mitglied der Klasse der verallgemeinerten ARCH Prozesse. Er hat eine ähnliche Struktur wie ein SV Prozess (vgl. Laurini und Tawn [14] (S. 203f)). Damit die Stationarität gesichert ist, müssen β, λ und γ bestimmte Eigenschaften

haben. Mikosch und Stărică [20] haben den extremalen Index des Prozesses (X_t^2) hergeleitet. Jedoch läßt sich θ mit Hilfe von Monte Carlo Methoden berechnen (vgl. Laurini und Tawn [14] (S. 205f)). Wie schon erwähnt, beeinflusst β bei einem ARCH Prozess nicht die Berechnung von θ . So ist der extremale Index bei einem GARCH Prozess eine Funktion von λ und γ . Den größten Einfluss jedoch hat der Parameter λ . Für $\beta = 0.005$, $\lambda = 0.2$ und $\gamma = 0.7$ ergibt sich für θ der Wert 0.722, der in der Abbilung 8.4 wieder durch eine horizontale Linie eingezeichnet wurde.

In Abbildung 8.4 wurde ein GARCH(1,1) Prozess der Länge $n=10000$ mit $\lambda = 0.2$, $\gamma = 0.7$ und $\beta = 0.005$ simuliert. Der wahre Wert für θ ist somit 0.722. Auch hier wurde u mit der Funktion *threshold* (vgl. Anhang A.2) das 89% bis 99.9%-Quantil verwendet und in einer $-\log_{10}(-\ln(F(u)))$ Skala abgebildet, was wieder die stark auseinander gezogenen Wert ab $x = 2$ erklärt.

Zur Schätzung von θ mit der run Methode wurde die run Länge $s = 3, 10$ und 50 abgebildet. Bei einem GARCH Prozess ist wegen dem Volatilitätsmechanismus wieder die

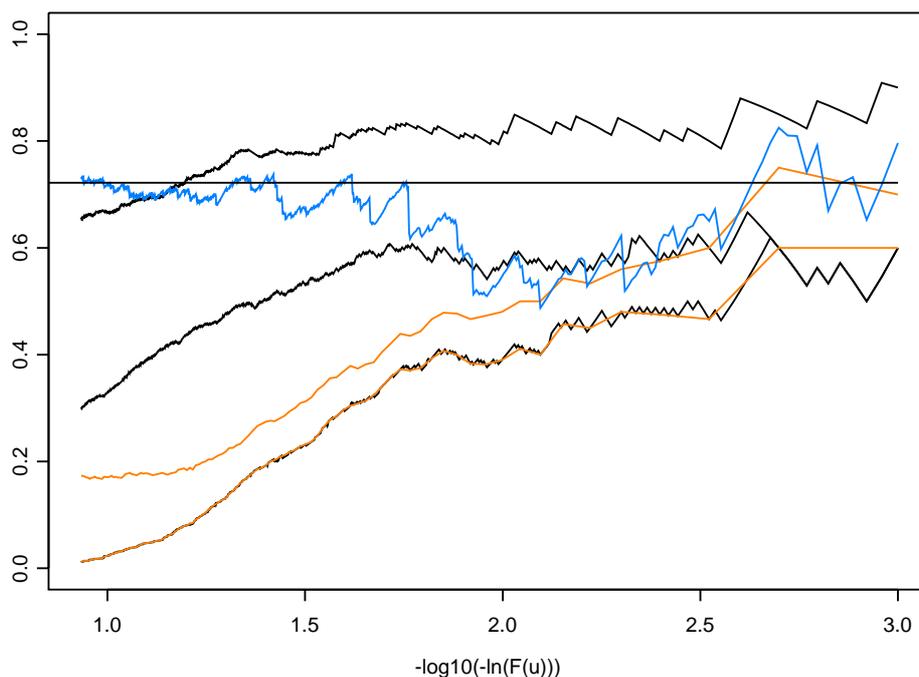


Abbildung 8.4:

10000 Daten eines GARCH(1,1) Prozesses mit $\lambda = 0.2$, $\gamma = 0.7$ und $\beta = 0.005$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_{\sigma}^{(Volatilität)}$ (orange) für $s = 50$ und für c mit $F_{\sigma}(c) = 0.75$ und $F_{\sigma}(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet. $\theta = 0.612$ ist durch die horizontalen Linie kenntlich gemacht.

Volatilitäts-Schwellen Methode besser. $\theta_\sigma^{(Volatilität)}$ wurde wie beim ARCH(1) Prozess mit (7.6) für $s = 50$ durchgeführt. c wurde mit $F_\sigma(c) = 0.75$ und $F_\sigma(c) = 0.25$ berechnet, wobei wie beim ARCH Prozess F_σ die Verteilungsfunktion des geschätzten Volatilitätsprozesses σ_t ist. Somit ist c der Wert des 75%- und des 25%-Quantils der 10000 simulierten σ_t . Der Intervallschätzer schwank zwar sehr stark um den wahren Wert von θ , aber da er nur von der Wahl von u abhängt, ist die Schätzung trotzdem stabiler als bei der run Methode. Der run Schätzer mit $s = 50$ und der Zwei-Schwellen Schätzer für $F_\sigma(c) = 0.25$ und $s = 50$ fallen fast zusammen.

Kapitel 9

Vergleich der Methoden

Mit den Schätzern $\hat{\theta}^{(run)}$ sowie $\hat{\theta}^{(Block)}$ kann man unabhängige Cluster identifizieren. Die run Methode ist die bevorzugte Methode. Jedoch sind beide Methoden eher ein naiver Weg θ zu schätzen. Für jede Methode werden zwei Parameter gebraucht. Zum einen die Schwelle $u = u_n$, die Größe der Blocks $r = r_n$ und die run Länge $s = s_n$ (vgl. (5.9)). Beide Methoden sind sehr sensibel bezüglich der Wahl von u und die run Methode ist stark von der Wahl von s abhängig. Beim Zwei-Schwellen Index und Volatilitätsindex wurde festgestellt, dass die Sensibilität der Schätzung für θ gegenüber der Wahl von u und der Laufzeit s reduziert werden kann.

An Abbildung 9.1 kann man sehen, dass die mit der run Methode berechneten Werte von θ für $s = 3, 10$ und 50 sehr variieren. Jedoch die mit der Zwei-Schwellen Methode berechneten Werte für $s = 3$ und $s = 10$ liegen sehr nahe beieinander. Die Methode reagiert also nicht so sensibel auf die Wahl von s . Berechnet man z.B. zusätzlich noch die Werte für $s = 50$ mit der Zwei-Schwellen Methoden fallen die Linien für $s = 10$ und $s = 50$ fast zusammen, wie man an Abbildung 9.2 sieht.

Für die Tabellen 9.1, 9.2 und 9.3 wurden die negativen Log Returns von S&P 500 (vgl. Kapitel 10) untersucht. Die Werte für θ wurden bei der Momenten Methode mit (6.9), bei der run Methode mit (5.10) und bei der Zwei-Schwellen Methode mit (7.5) verwendet.

An den Ergebnissen aus Tabelle 9.2 und Tabelle 9.3 kann man sehen, dass für kleine s das Kriterium der run Methode überwiegt. Jedoch für mittelmäßige s und große c als auch für große s überwiegt das Kriterium der Zwei-Schwellen Methode und identifiziert fast alle Cluster. Was jedoch leicht nachvollziehbar ist. Denn ist s ziemlich klein, dann werden schon alleine durch die run Bedingung sehr viele Cluster gefunden und die zusätzlich Bedingung überschneidet sich mit diesen. Ist jedoch s groß, dann entdeckt die run Methode nicht so viele Cluster, denn es wird schwieriger s aufeinander folgende Daten zu finden, die sich unterhalb der Schwelle u befinden. Jetzt findet also die zweite Bedingung dazwischen einige Cluster, die die run Methode „übersieht“. Betrachtet man die Werte für $s = 2$ ist $\theta^{(run)}$ ungefähr $\theta^{(zwei-Schwellen)}$. Nähert sich c der Schwelle u , so wird ein Cluster sehr schnell beendet und die Zwei-Schwellen Methode entspricht ungefähr der run Methode mit run Länge $s = 2$. Nimmt jedoch die untere Schwelle c einen sehr niedrigen

Wert an, so ergeben sich fast dieselben Werte wie der run Methode unabhängig von der Wahl von s . Dies verifiziert die Grenzaussagen aus Kapitel 7.1.

Der Intervallschätzer ist in der Tabelle 9.1 abgebildet. Die Werte für $\hat{\theta}^{(Intervall)}$ schwanken, wenn die Schwelle u steigt. Der Intervallschätzer ist stark von der Wahl von u abhängig, da damit die run Länge $s = T(C)$ (vgl. Kapitel 6.3) berechnet wird.

Momenten Methode	$u = 75\%$	$u = 80\%$	$u = 95\%$	$u = 99\%$
	0.9958843	0.9245999	0.4723417	0.2729402

Tabelle 9.1: θ geschätzt mit der Momenten - Methode für verschiedene s und u

run Methode	$s = 2$	$s = 3$	$s = 10$	$s = 50$	$s = 100$
$u = 95\%$	0.8618182	0.7963636	0.4545455	0.09454545	0.03272727
$u = 99\%$	0.8909091	0.8545455	0.6545455	0.2909091	0.1818182

Tabelle 9.2: θ geschätzt mit der run Methode für verschiedene s und u

Zwei-Schwellen		$s = 2$	$s = 3$	$s = 10$	$s = 50$	$s = 100$
$u = 95\%$	$c = 1\%$	0.8654545	0.8	0.5090909	0.2145455	0.1781818
	$c = 25\%$	0.8909091	0.8472727	0.8036364	0.8	0.8
	$c = 75\%$	0.9018182	0.8981818	0.8981818	0.8981818	0.8981818
	$c = 90\%$	0.9127273	0.9127273	0.9127273	0.9127273	0.9127273
$u = 99\%$	$c = 1\%$	0.8909091	0.8545455	0.7090909	0.5272727	0.4727273
	$c = 25\%$	0.8909091	0.8727273	0.8727273	0.8727273	0.8727273
	$c = 75\%$	0.8909091	0.8909091	0.8909091	0.8909091	0.8909091
	$c = 90\%$	0.9090909	0.9090909	0.9090909	0.9090909	0.9090909

Tabelle 9.3: θ geschätzt mit der Zwei-Schwellen Methode für verschiedene s , u und c

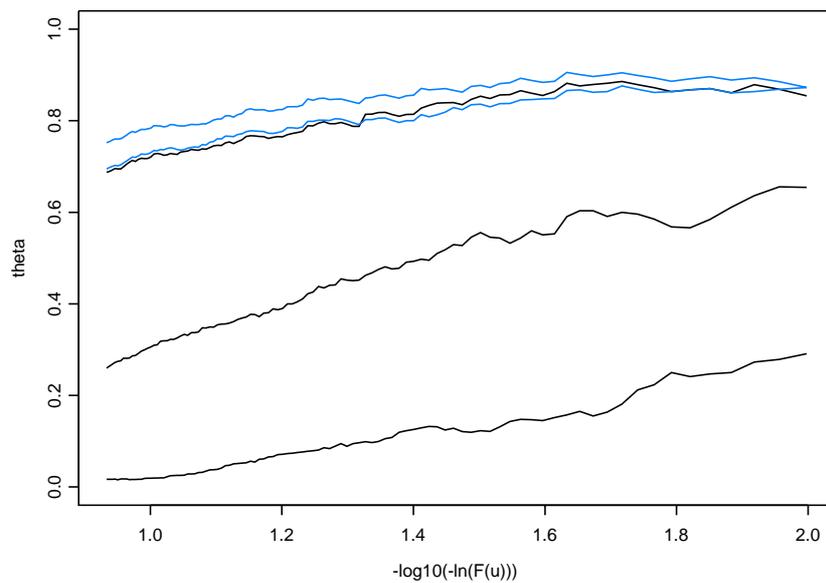


Abbildung 9.1:
 $\hat{\theta}_n^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten) und $\hat{\theta}_{X,n}^{(zweischwellen)}$ (blau) für $s = 3$ und 10 (von oben nach unten) dargestellt.

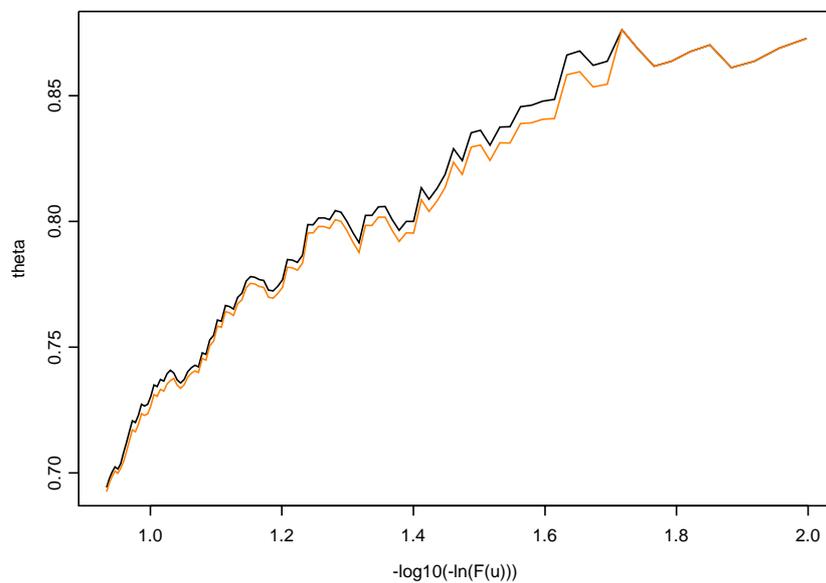


Abbildung 9.2: $\hat{\theta}_{X,n}^{(zweischwellen)}$ abgebildet für $s = 10$ (schwarz) und für $s = 50$ (orange).

Kapitel 10

Datenanalyse

10.1 S&P 500

Wenn man nur einen Index will, um die Wirtschaft der U.S. zu messen, verwendet man den S&P 500. Als der führende Index deckt der S&P 500 80% des U.S. Marktes ab. Er umfasst mehr als 100 verschiedene Industriegruppen mit 50 Stocks. Der S&P 500 ist nur einer der S&P Indexreihe und kann gut zur Konstruktion von Portfolios verwendet werden. Der Index wird von einem S&P 500 Index Komitee, einer Gruppe von Standard&Poors Wirtschafts- und Indexanalysten, aufrecht erhalten. Das Ziel dieser Gruppe ist, dass sich der S&P 500 an der Spitze als führender Index hält. Zusätzlich hat sie natürlich Aufgaben wie die Liquidität zu wahren, usw. Weitere Informationen findet man dazu auch unter der Internetadresse [12]

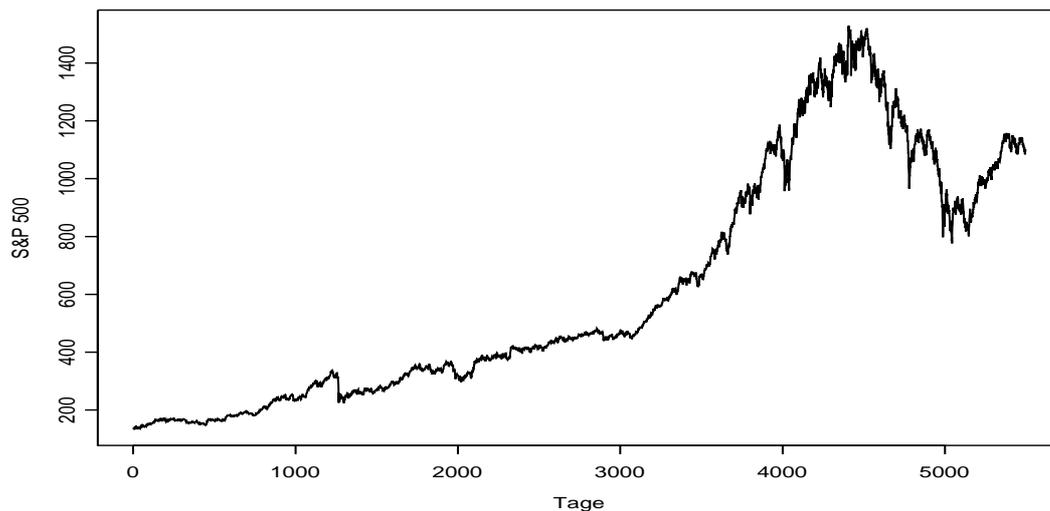


Abbildung 10.1: *Tägliche Schlußdaten des S&P 500 - 1982 bis 2004*

Der untersuchte Datensatz S&P 500 hat 5497 Daten, die in Abbildung 10.1 dargestellt sind. Die negativen Log Returns $X_i = -(\log P_i - \log P_{i-1})$ werden in Abbildung 10.2 aufgezeichnet. Diese haben dementsprechend $n = 5496$ Daten. Es wurden die täglichen Schlußdaten des S&P 500 gewählt und davon die negativen Log Returns gebildet. Folgende Tabelle gibt einige interessante Werte der negativen Log Returns wieder.

n	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
5496	-0.2288678	-0.0046772	0.0004752	0.0003764	0.0056934	0.0870853
Schiefe	Kurtosis					
-1.924438	42.06168					

Tabelle 10.1:

Übersicht über die wichtigsten Kenngrößen der negativen Log Returns des S&P 500.

Die Abbildung 10.3 bildet die Autokorrelationsfunktion und ein Histogramm der negativen Log Returns ab. Da die Autokorrelationsfunktion für lags ≥ 1 im wesentlichen innerhalb der Grenzen liegt, kann man von unkorrelierten Daten ausgehen. Am Wert der Schiefe, der kleiner Null ist, sieht man, dass die Daten eher linksschief angeordnet sind. Die Kurtosis, die viel größer als Null ist, beweist, dass die Daten sehr „spitz“ verteilt sind.

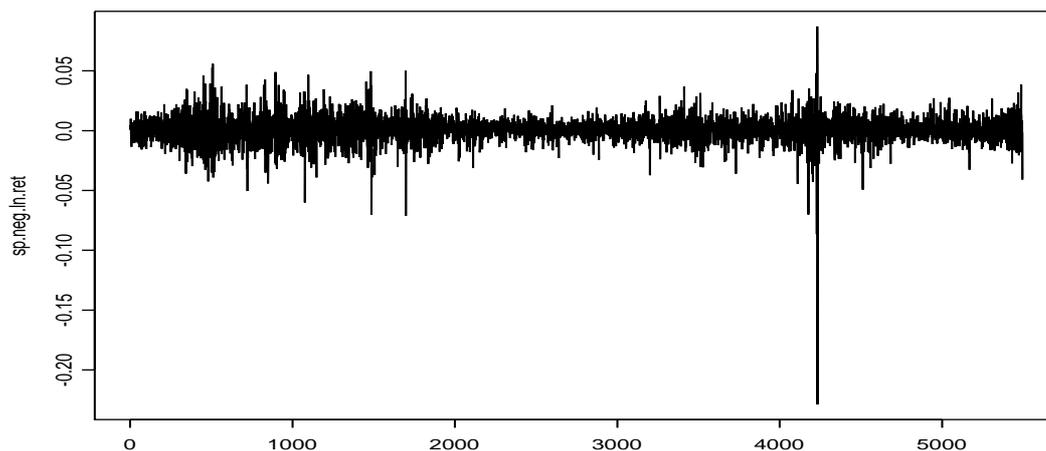


Abbildung 10.2:

Negative Log Returns der täglichen Schlußdaten des S&P 500 - 1982 bis 2004.

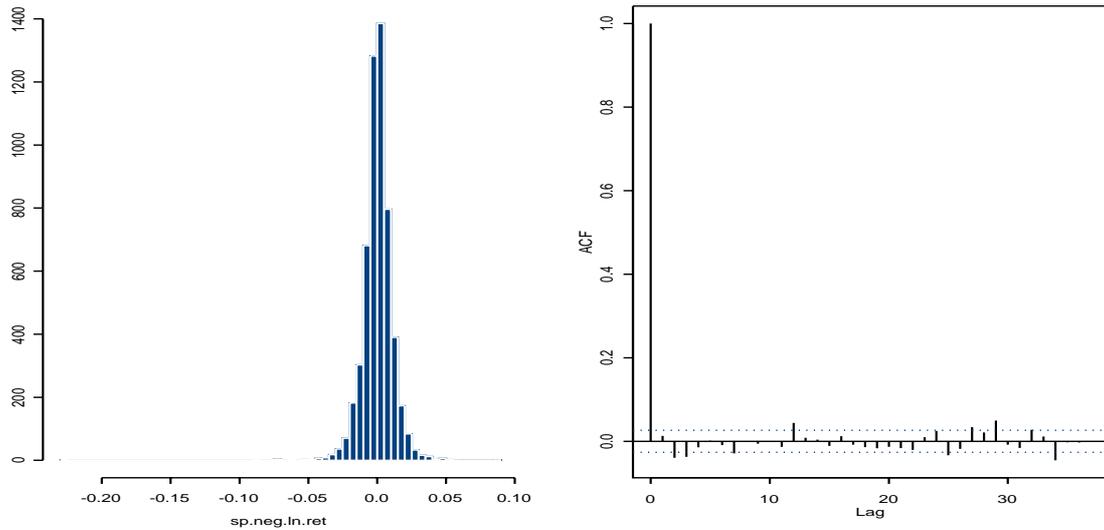


Abbildung 10.3:
Autokorrelationsfunktion und Histogramm der negativen Log Returns des S&P 500 - 1982 bis 2004.

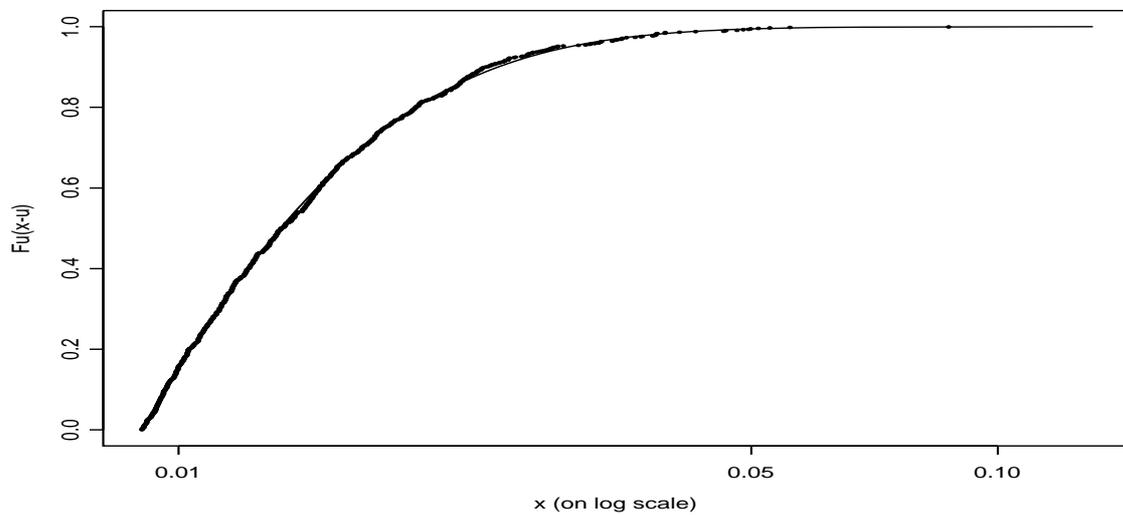


Abbildung 10.4:
Anpassung der negativen Log Returns des S&P 500 an eine GPD (Standard POT Methode). ξ und β wurden durch einen Maximum Likelihood Schätzer geschätzt und sind in der Tabelle 10.2 zu finden.

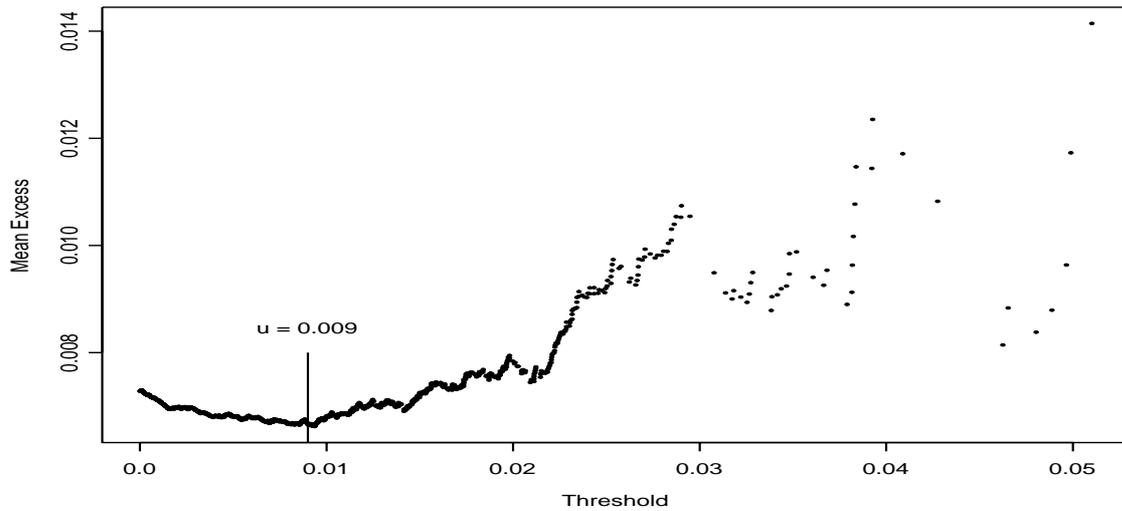


Abbildung 10.5:
Die Exzessfunktion der positiven Daten der negativen Log Returns des S&P 500.

S&P 500	u	$\hat{\xi}$	$\hat{\beta}$	n	N_u
	0.009	0.0991	0.00601	5496	854

Tabelle 10.2:
Der Schwellenwert u , der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Anzahl der Daten n , die Anzahl der Exzedenten N_u und die Maximum Likelihood Schätzer für ξ und β des Datensatzes S&P 500. (Standard POT Methode)

S&P 500	u	s	c	$\hat{\xi}$	$\hat{\beta}$	n	N_u^C	$\hat{\theta}$
run Methode	0.009	5		0.0289	0.00809	5496	369	0.43208
Zwei-Schwellen Methode	0.009	5	-0.00468	0.045	0.00746	5496	564	0.66042

Tabelle 10.3:
Der Schwellenwert u , der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Schwelle c , für die das 25%-Quantil genommen wurde, run Länge s , die Anzahl der Daten n , die Anzahl der Clustermaxima N_u^C , die Schätzer der run bzw. Zwei-Schwellen Methode und die Maximum Likelihood Schätzer für ξ und β des Datensatzes S&P 500 (POT Methode für stationäre Zeitreihen).

S&P 500	empirisch	Standard POT	POT für stat. Zeitreihen (run Methode)	POT für stat. Zeitreihen (Zwei-Schwellen Methode)
VaR	-0.01585	-0.01584	-0.01795	-0.01730
DRC	0.01623	0.01621	0.01833	0.01768

Tabelle 10.4:

POT Methoden angewandt auf die negativen Log Returns des Datensatzes S&P 500.

Man sieht, in Abbildung 10.4 die Anpassung der Exzedenten an die GPD. Die Parameter der GPD wurden mit Maximum Likelihood geschätzt. Der Schwellenwert u wurde mit Hilfe der Exzessfunktion ermittelt (vgl. Abbildung 10.5). Wir erinnern uns, dass die beste Wahl für u der Wert ist, ab dem die Exzessfunktion linear steigt. Im Fall der negativen Log Returns steckt das Risiko in den Werten größer Null, also werden bei der Anpassung an die GPD automatisch die richtigen Daten verwendet. Die geschätzten Parameter findet man in der Tabelle 10.2. Die Funktion *mepplot* wählt nicht automatisch nur die Werte größer 0 aus, also muß man bei der Anwendung dieser Funktion aufpassen. An der Tabelle 10.4 sieht man die Ergebnisse der Schätzung durch die Standard POT Methode. Für die Schätzung von VaR muß bei den negativen Log Returns das 95%-Quantil geschätzt werden, da dort das Risiko steckt. Mit den geschätzten Parametern kann dann das 95%-Quantil der Verteilung mit (2.4) berechnet werden und somit erhält man das DRC, das korrigiert um den Erwartungswert den VaR ergibt. Was sofort auffällt, ist, dass der VaR durch die Standard POT Methode viel höher geschätzt wird als durch die empirische Methode, d.h. das der VaR weiter links liegt. Es werden also mehr Daten als Risiko gewertet.

In Tabelle 10.4 wurden auch die Ergebnisse der Schätzung des VaR mit der POT Methode für stationäre Zeitreihen zusammengefasst. Dabei muß zuerst θ geschätzt werden. Dazu wurden die run und die Zwei-Schwellen Methode verwendet. Es hat sich gezeigt, dass die Wahl von $s = 5$ geeignet ist. Die Ergebnisse der Schätzungen der Parameter findet man in der Tabelle 10.3. Das 95%-Quantil wird bei der POT Methode für stationäre Zeitreihen mit (2.7) berechnet. Wieder fällt auf, dass der VaR höher kalkuliert wird.

Die Schätzung des extremalen Index mit den verschiedenen Methoden sieht man in Abbildung 10.6. Für die run Methode wurde die Formel (5.10), für die Zwei-Schwellen Methode die Formel (7.5) und für die Momenten Methode die Formel (6.9). Die Ergebnisse der Blockmethode findet man in der Abbildung 10.7, die auch für verschiedene Schwellenwerte u angewandt wurde.

Die Abbildungen 10.8 bis 10.18 zeigen die Ergebnisse des Declustering der Daten mit Hilfe der verschiedenen Methoden. Bleibt die Schwelle c und u gleich und nur s ändert sich, kann man beobachten, dass bei fast allen Methoden weniger Cluster identifiziert werden. Nur die Zwei-Schwellen Methode findet fast immer die gleichen Cluster, wenn die Schwelle

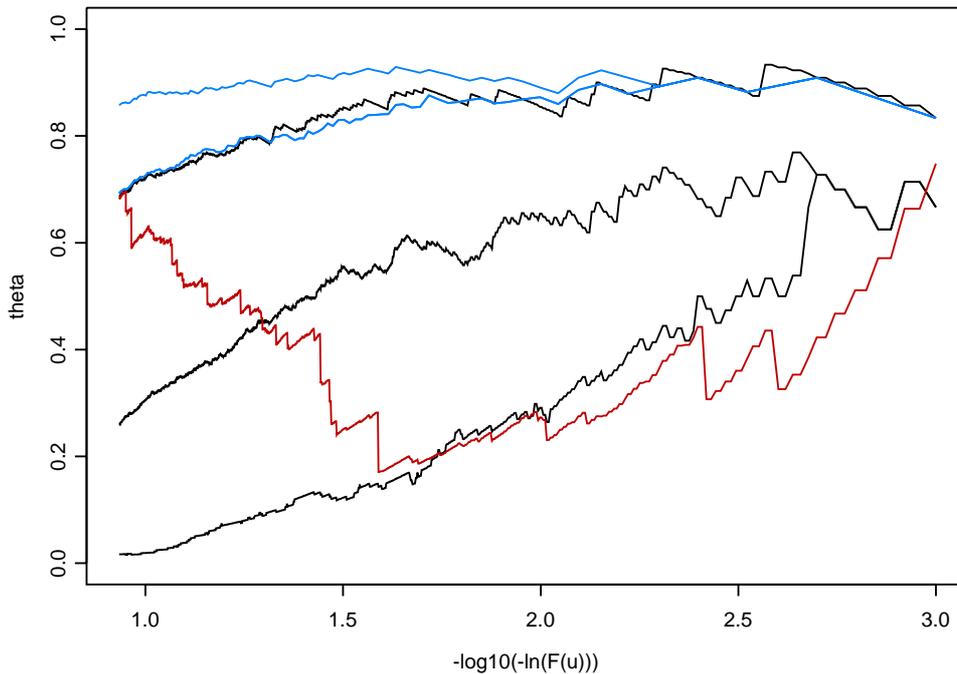


Abbildung 10.6:

θ geschätzt für die negativen Log Returns des S&P 500 mit der run Methode (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), mit der Zwei-Schwellen Methode für $s = 50$ und c mit $F(c) = 0.75, 0.25$ (von oben nach unten) und der Momenten Methode. u entspricht dem 89%- bis 99.9%-Quantil des Datensatzes und ist in einer $-\log_{10}(\ln(F(u)))$ Skala abgebildet.

u hoch ist. Die Cluster werden weniger mit steigendem u , da auch die Exzedentenanzahl sinkt. Wenn der Schwellenwert u sehr hoch ist und die run Länge s sehr niedrig ist, bringt die run Methode fast das gleiche Ergebnis wie die Zwei-Schwellen Methode (vgl. Abbildung 10.16). Man kann auch beobachten, dass die Clusteranzahl stark abnimmt, falls s und c klein sind und konstant bleiben und nur der Schwellenwert u größer wird (siehe Abbildung 10.15, 10.16). Ändert sich die run Länge nicht viel, so wie in Abbildung 10.11 und 10.12, ist das Ergebnis fast das gleiche. Abbildung 10.17 verifiziert nochmal die Aussage, dass sich das Ergebnis der Zwei-Schwellen Methode für große c der run Methode mit niedriger run Länge s annähert. Allgemein sieht man, dass die Zwei-Schwellen Methode im Vergleich zu anderen Methoden sehr viele Cluster identifiziert. Da die Momentenmethode nur vom Schwellenwert u abhängt, sind die Ergebnisse für $u=95\%$ und $u=99\%$ immer das gleiche. Deshalb wird das Resultat der Momenten Methode in der Abbildung 10.18 alleine dargestellt. Man bemerkt sofort, dass bei einer kleinen Erhöhung von u (95%- auf 99%-Quantil), die run Länge stark ansteigt. Die run Länge wird bei der Momenten Methode berechnet (siehe (6.11)) und ist nicht frei wählbar.

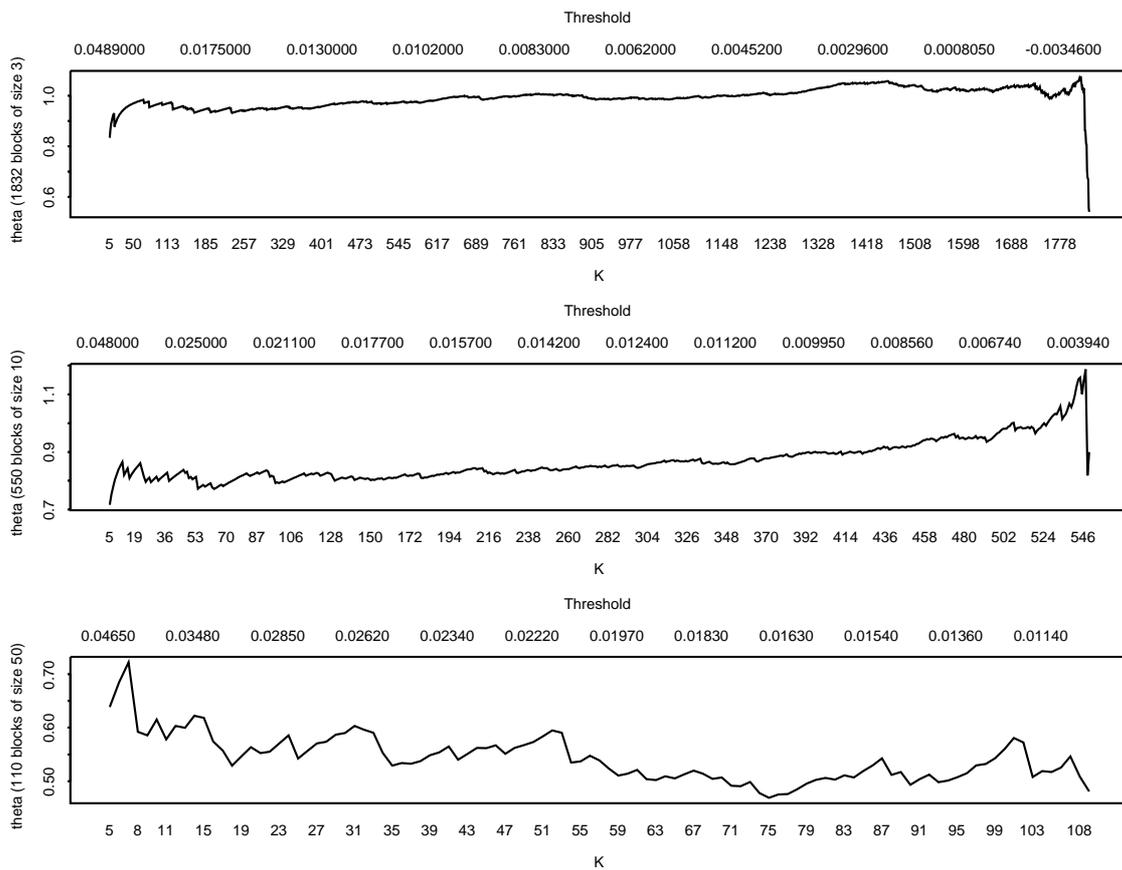


Abbildung 10.7:
 θ geschätzt für die negativen Log Returns des S&P 500 mit der Blockmethode. Dabei wurde die Blockgröße 3, 10 und 50 (von oben nach unten) gewählt.

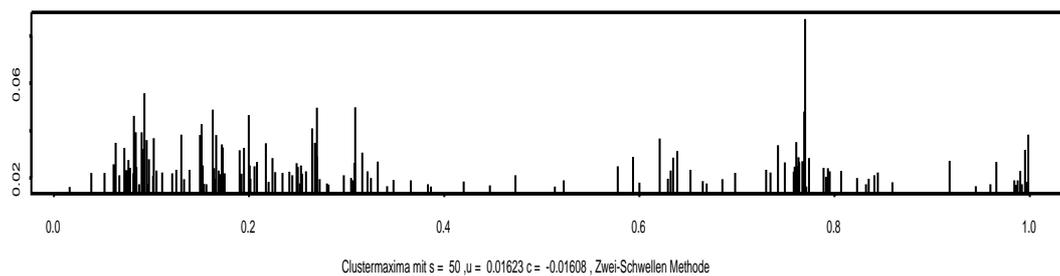
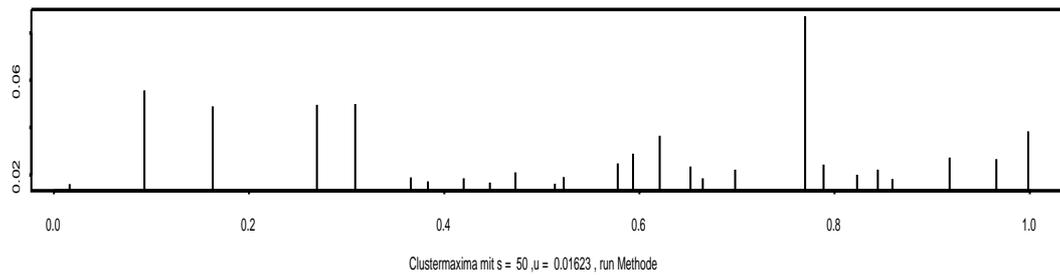
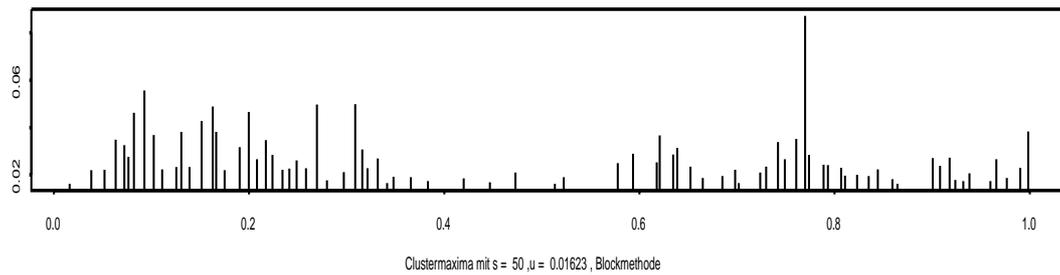
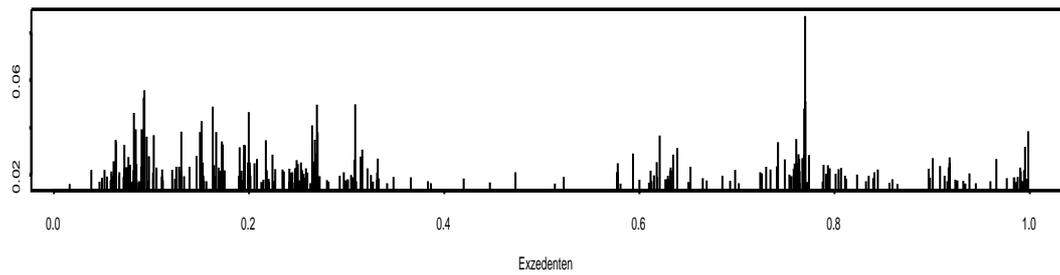


Abbildung 10.8:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 50$.

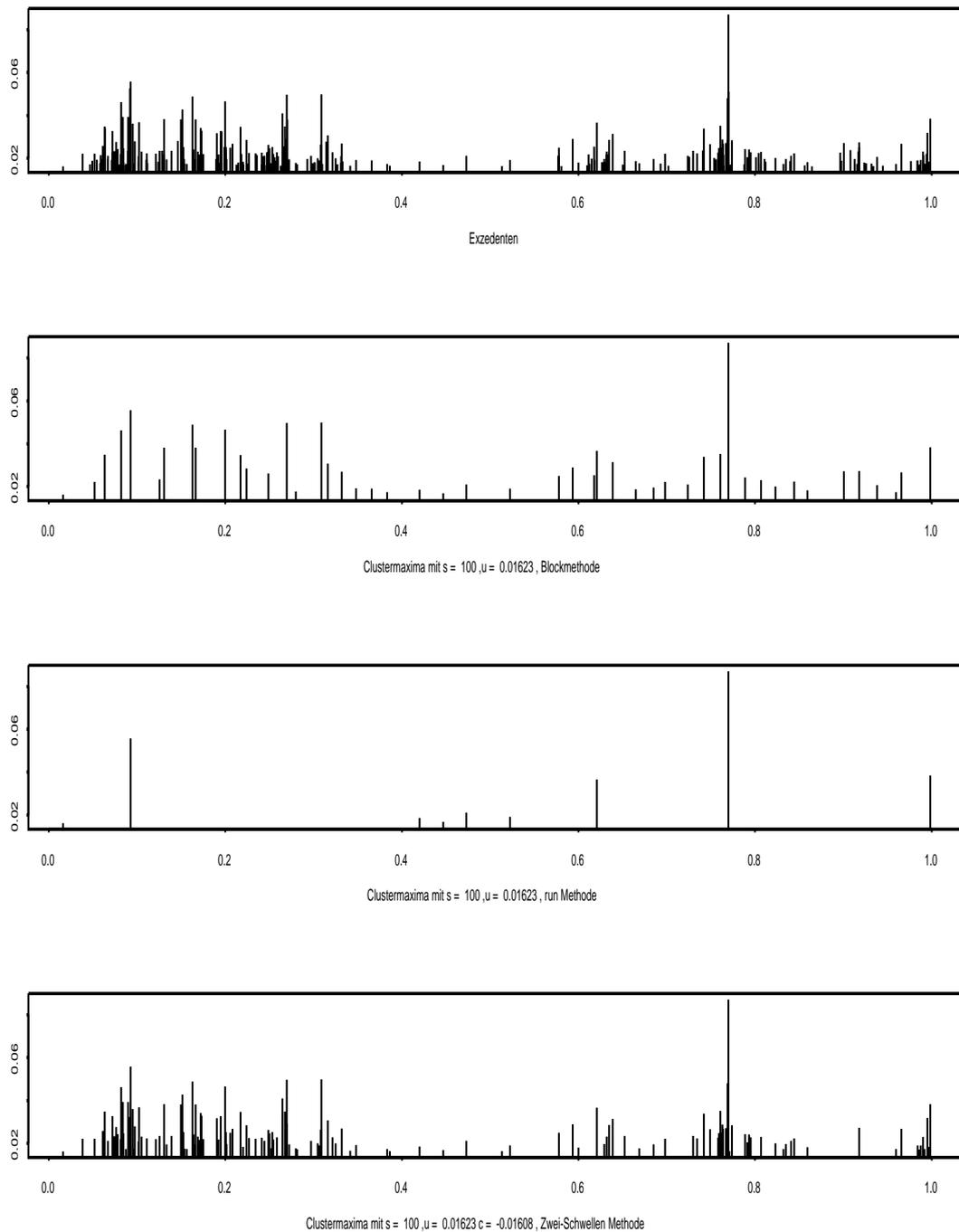


Abbildung 10.9:

Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 100$.

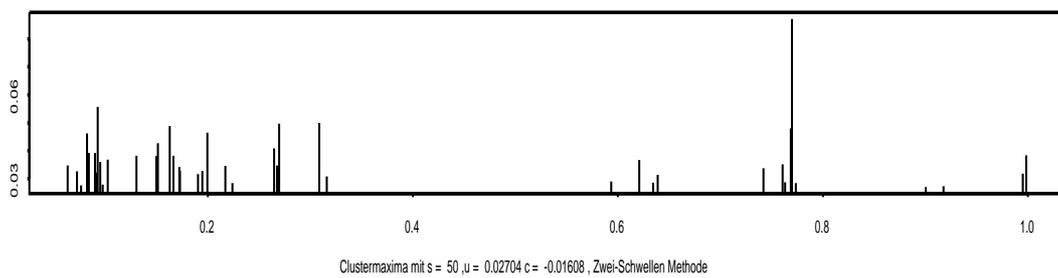
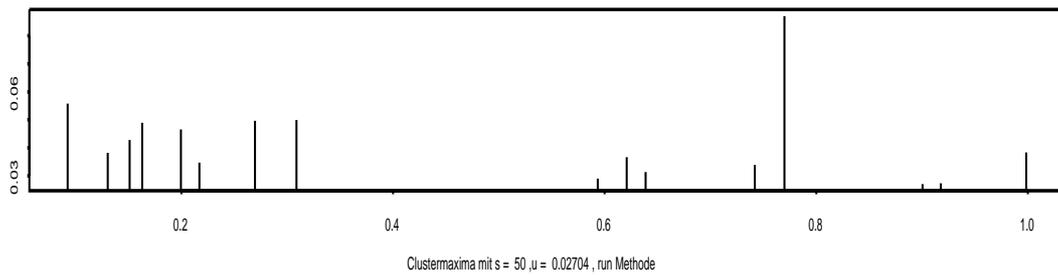
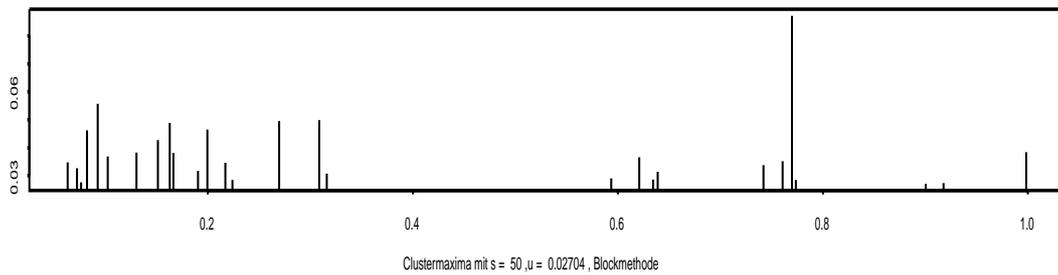
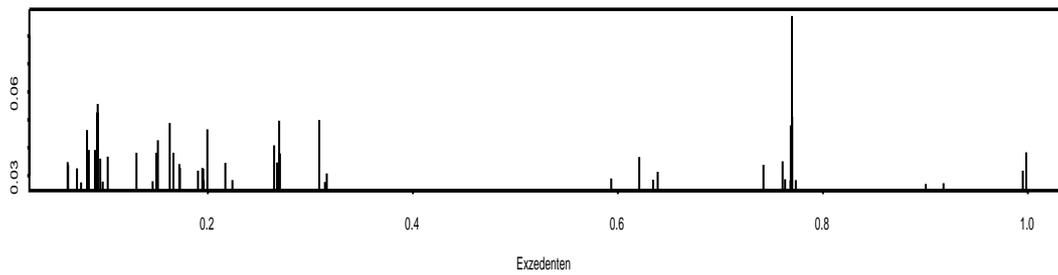


Abbildung 10.10:

Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 50$.

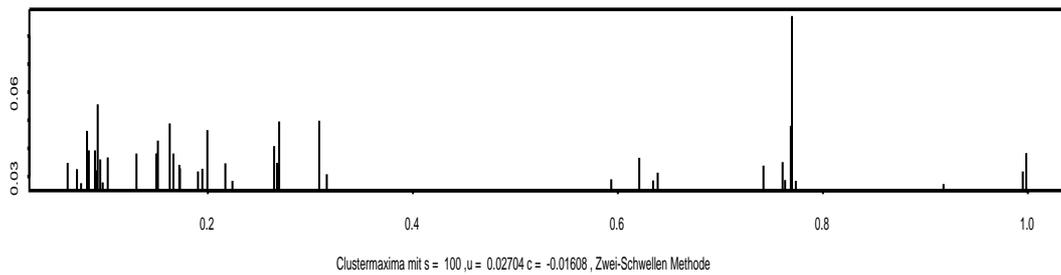
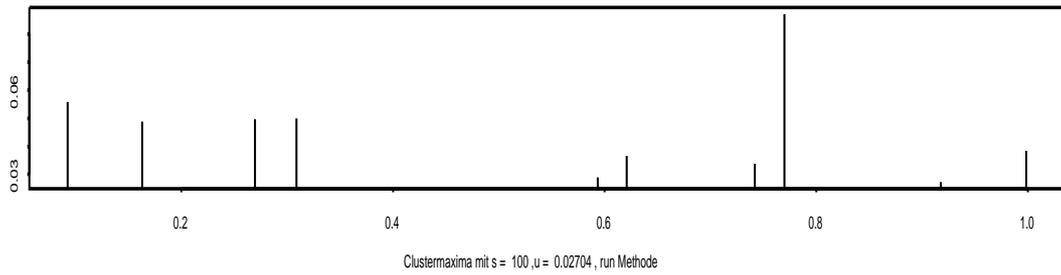
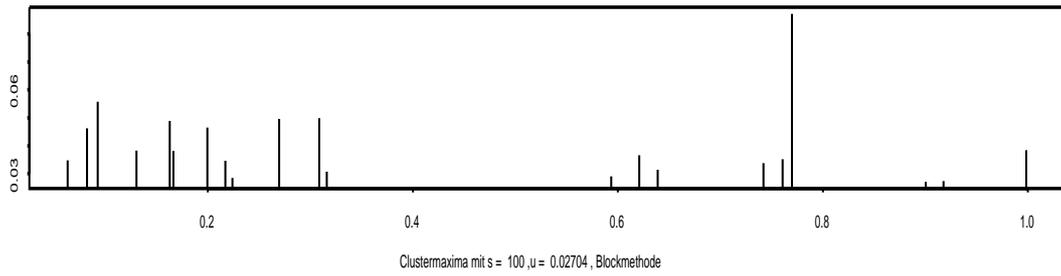
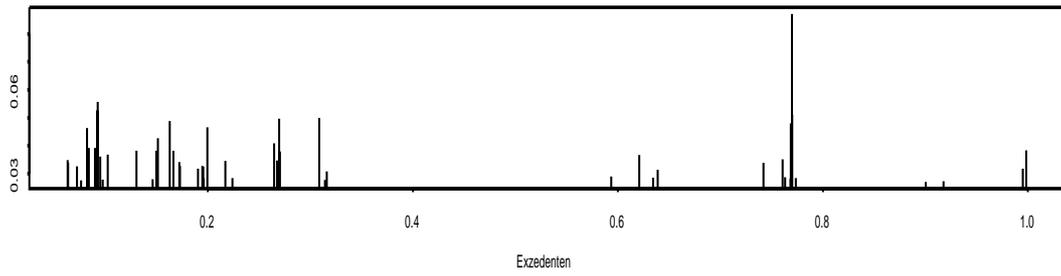


Abbildung 10.11:

Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 100$.

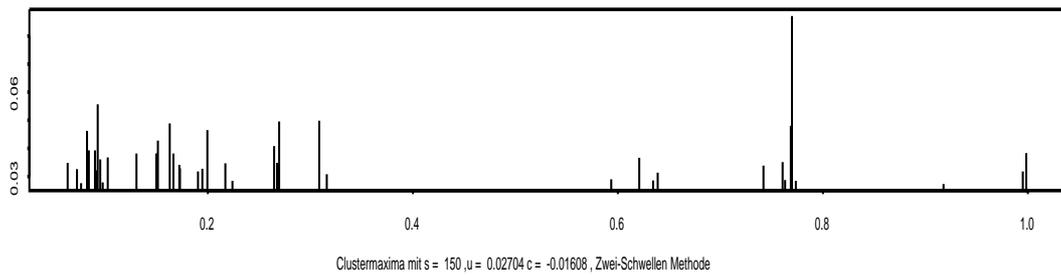
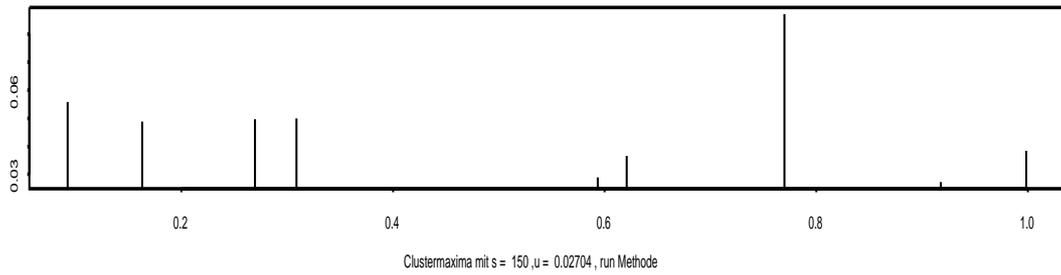
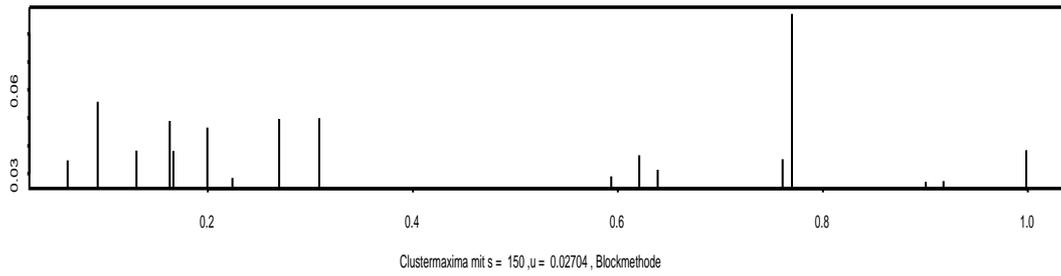
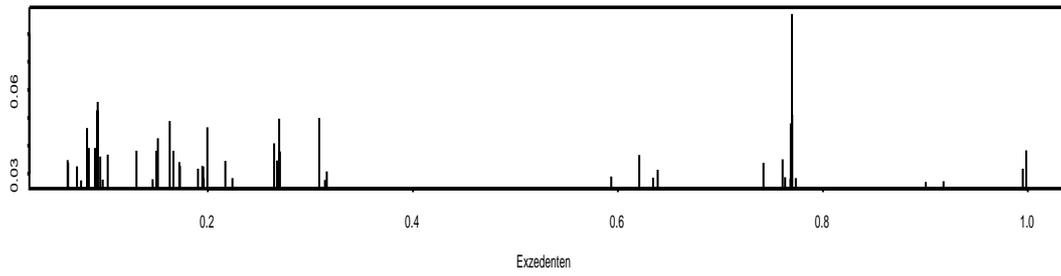


Abbildung 10.12:

Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 150$.

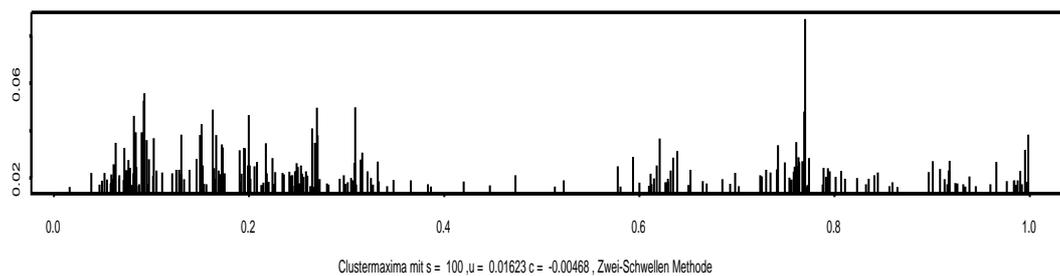
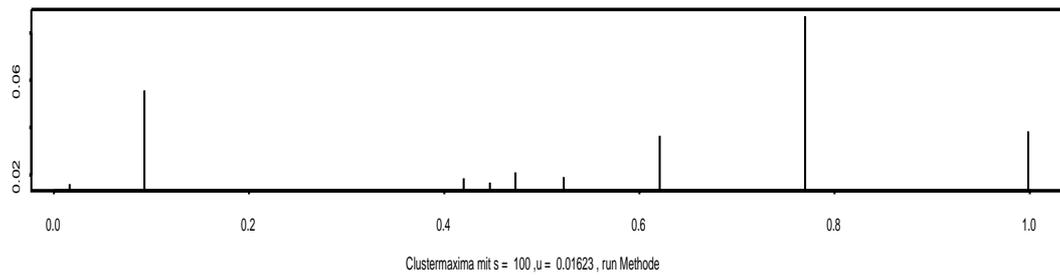
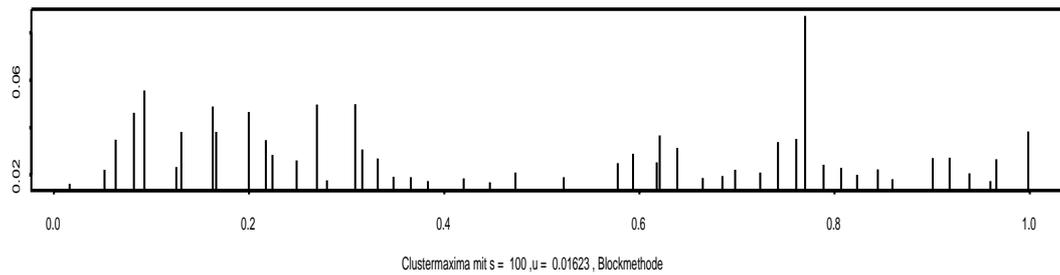
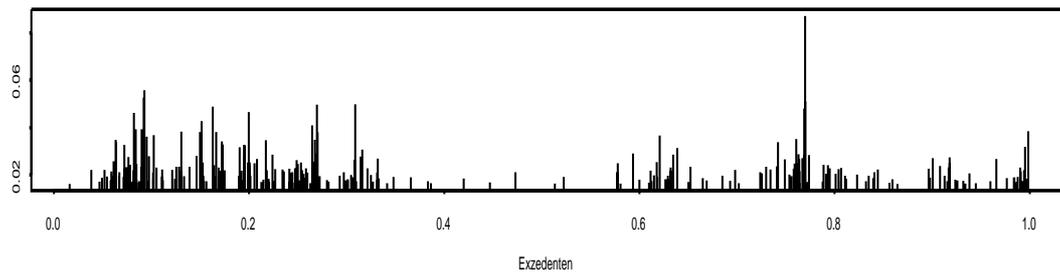


Abbildung 10.13:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 25\%$ -Quantil und $s = 100$.

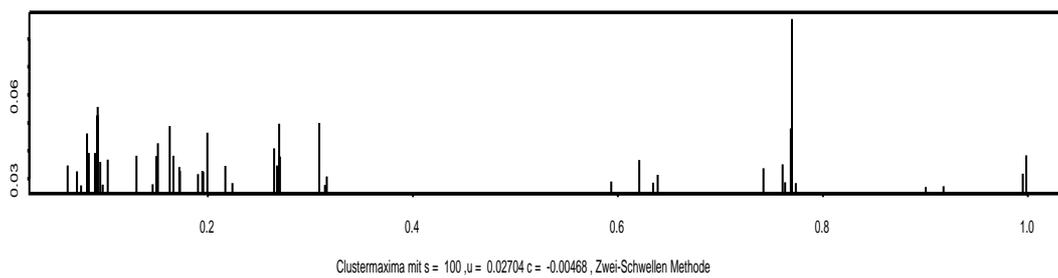
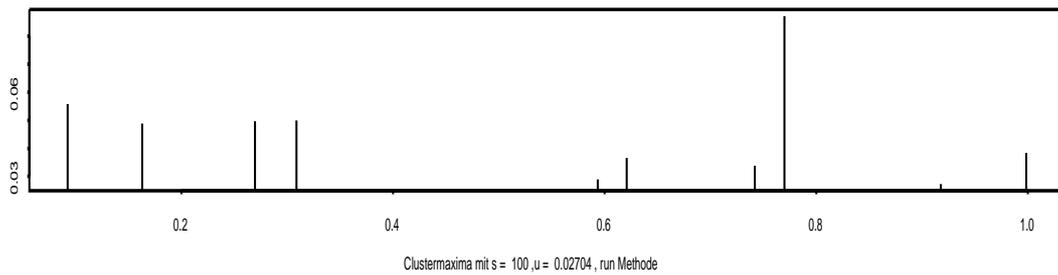
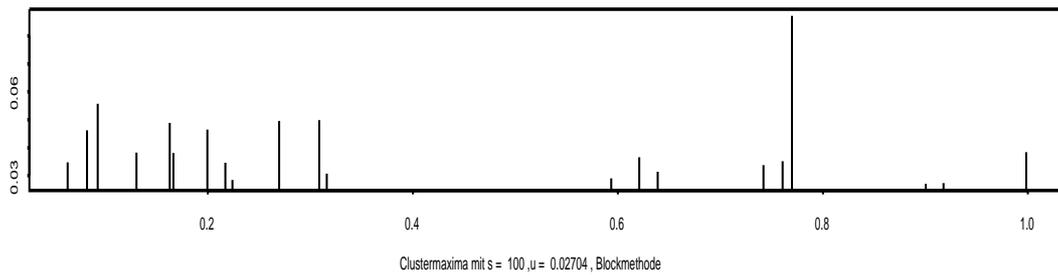
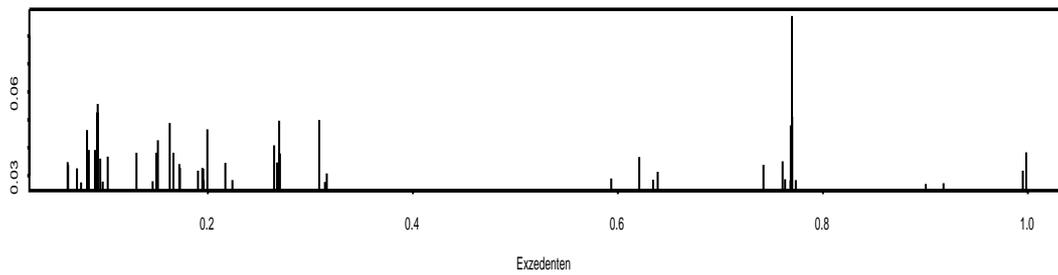


Abbildung 10.14:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 25\%$ -Quantil und $s = 100$.

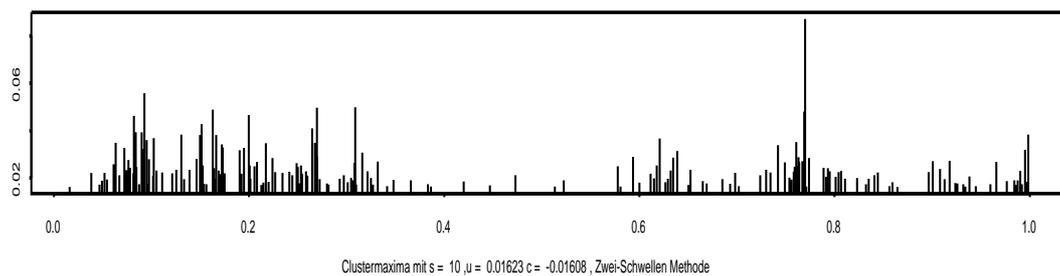
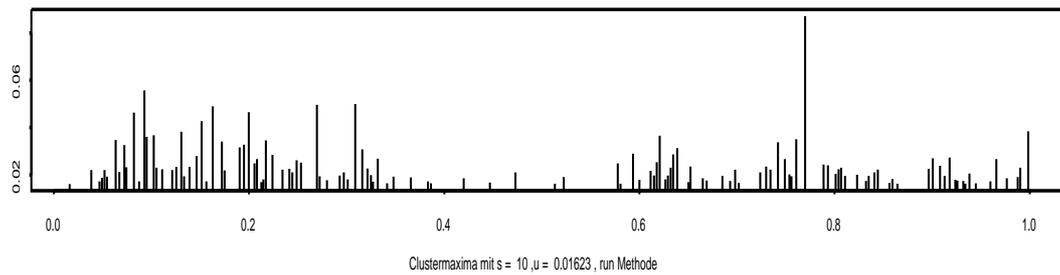
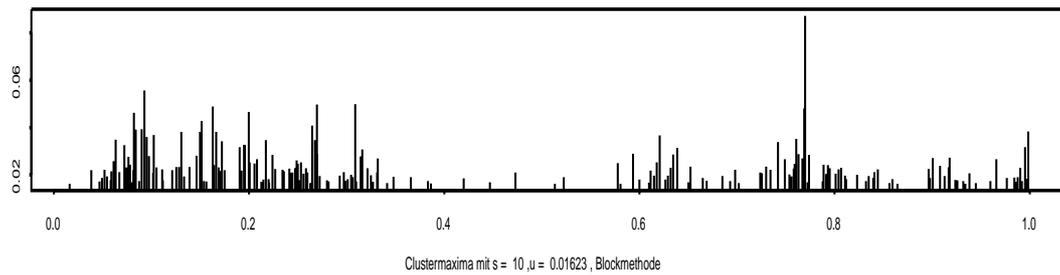
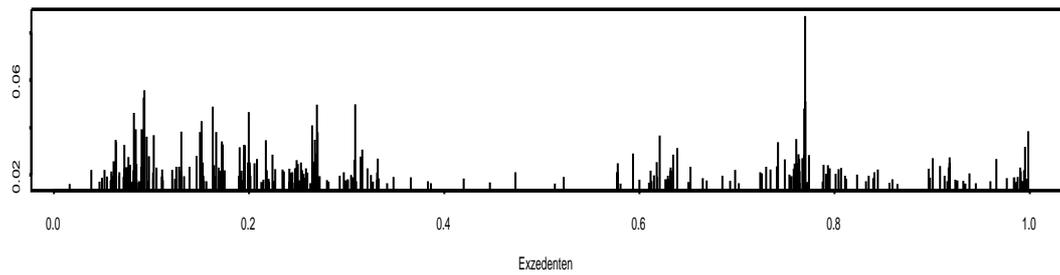


Abbildung 10.15:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 10$.

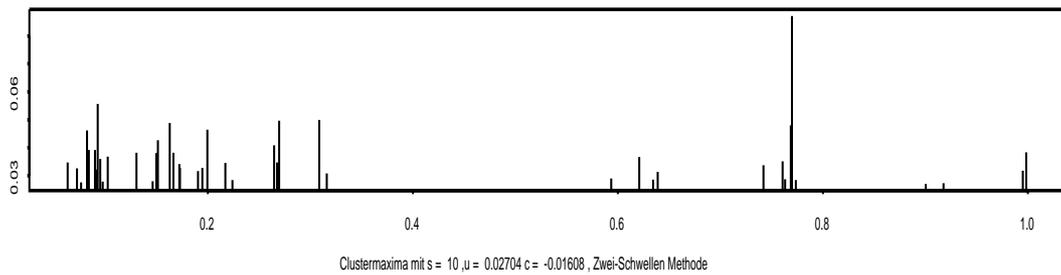
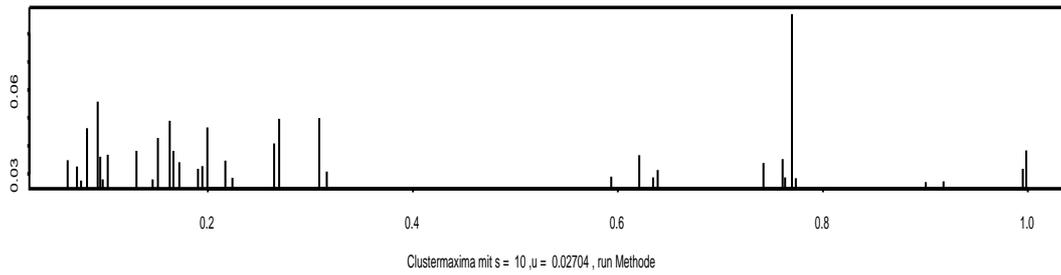
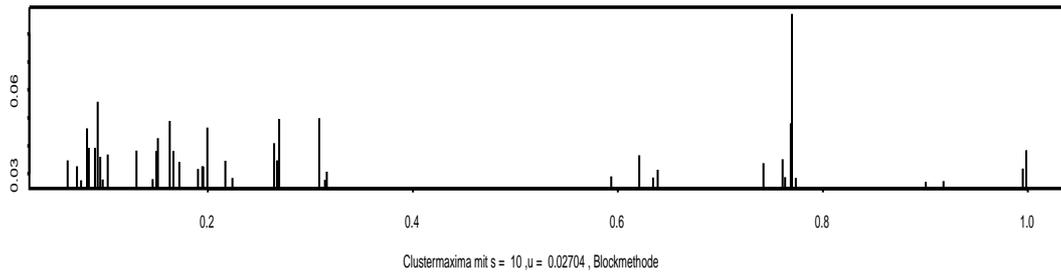
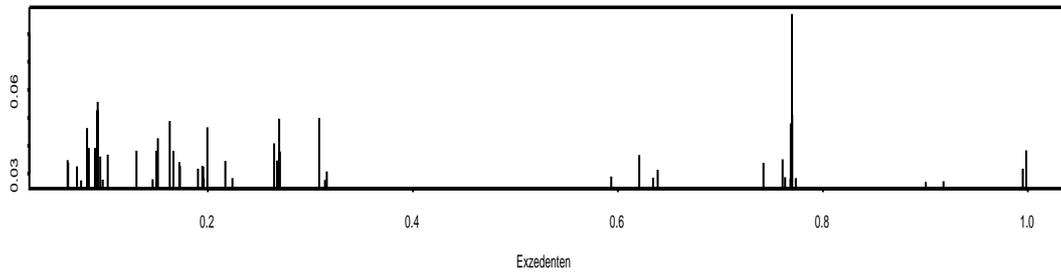


Abbildung 10.16:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 10$.

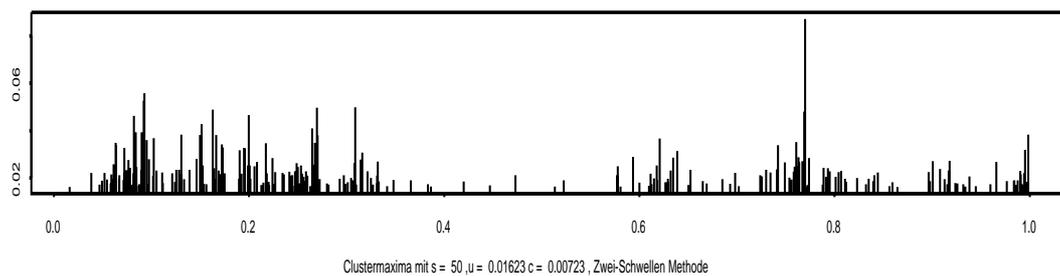
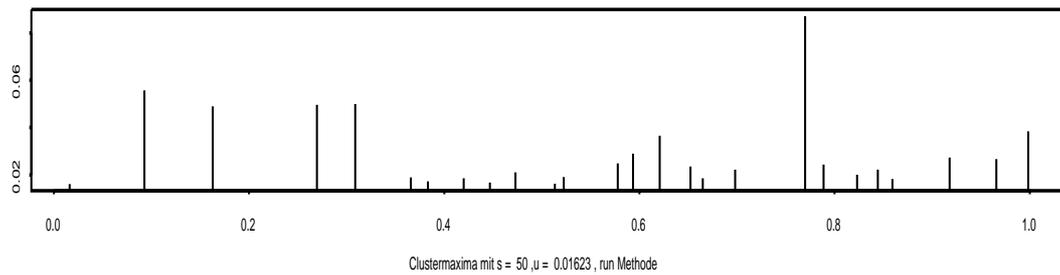
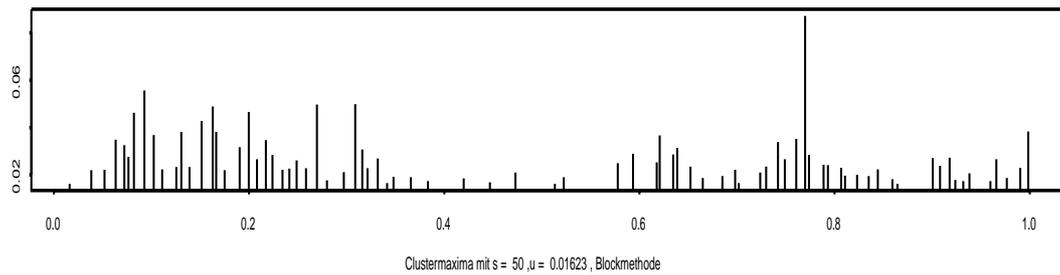
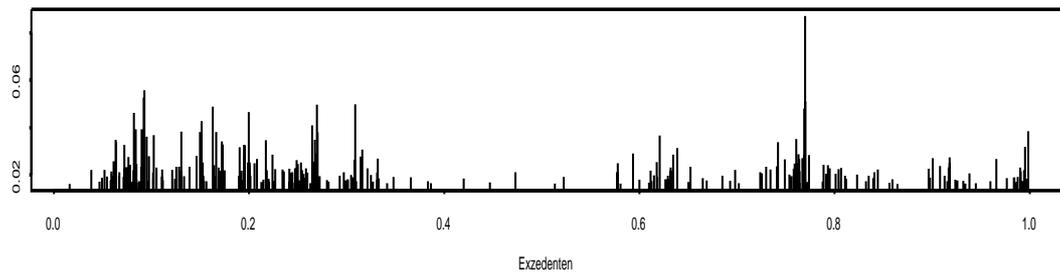


Abbildung 10.17:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 80\%$ -Quantil und $s = 50$.

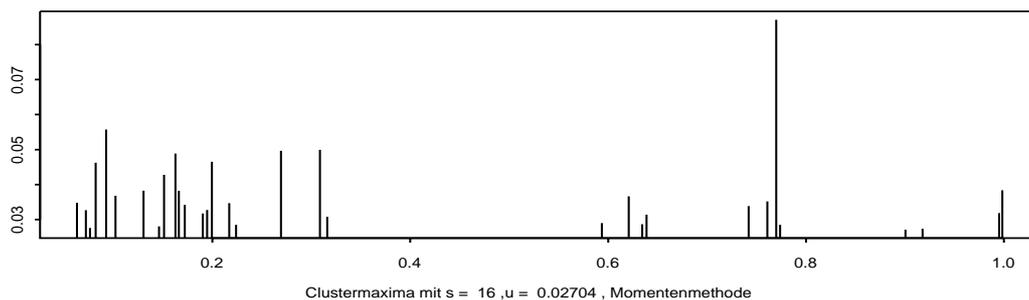
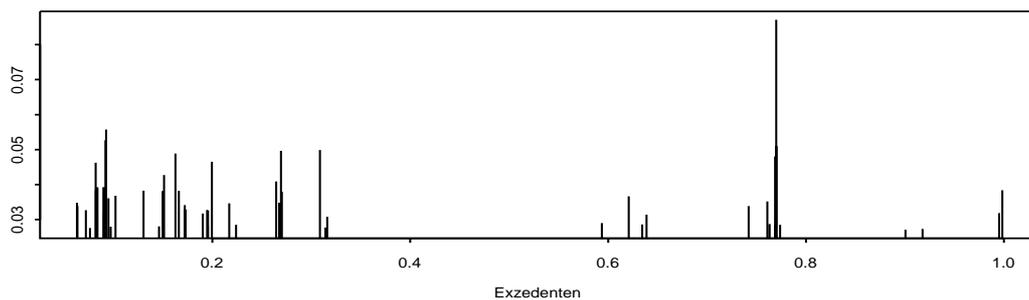
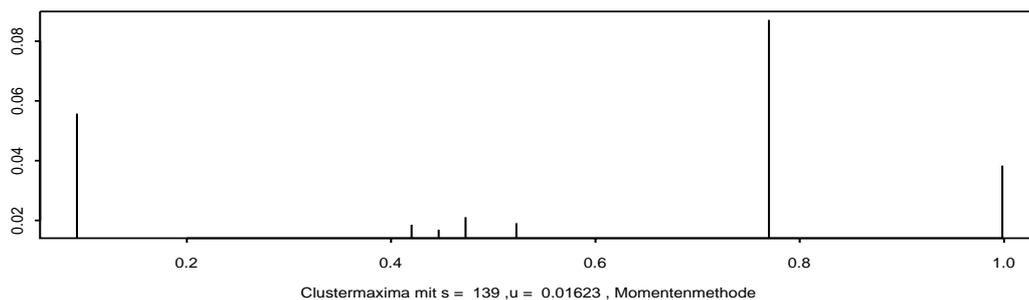
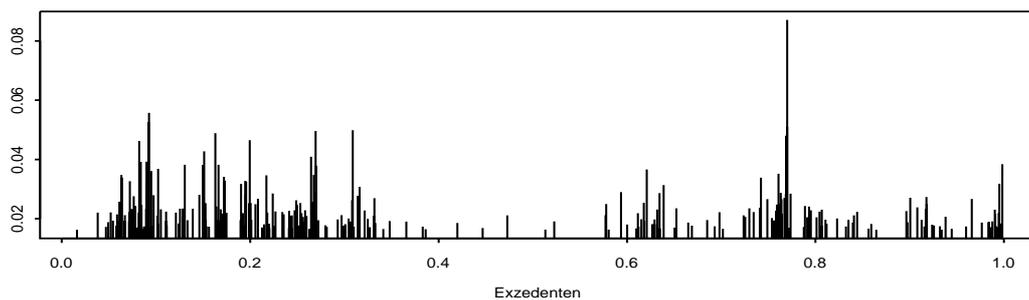


Abbildung 10.18:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Momentenmethode. Die gewählten Werte waren $u = 95\%$ -Quantil (oberen zwei Graphen) und $u = 99\%$ -Quantil (unteren zwei Graphen).

10.2 FTSE 100

Der untersuchte Datensatz FTSE 100 hat 5139 Daten, die in Abbildung 10.19 dargestellt sind. Da alle Berechnungen mit den negativen Log Returns $X_i = -(\log P_i - \log P_{i-1})$ durchgeführt werden, deshalb wurden diese in Abbildung 10.20 aufgezeichnet. Es ist klar, dass diese Daten nur noch die Länge $n = 5138$ haben. Es wurden die täglichen Schlußdaten des FTSE 100 gewählt und davon die negativen Log Returns gebildet. Folgende Tabelle gibt einige interessante Werte der Daten der negativen Log Returns wieder.

n	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
5138	-0.1302860	-0.0054545	0.0005720	0.0002687	0.0063944	0.0759697
Schiefe	Kurtosis					
-0.5377097	7.825408					

Tabelle 10.5:

Übersicht über die wichtigsten Kenngrößen der negativen Log Returns des FTSE 100.

Die Abbildung 10.21 bildet die Autokorrelationsfunktion und ein Histogramm der Daten ab. Da auch hier die Autokorrelationsfunktion für alle lags ≥ 1 im wesentlichen innerhalb der Grenzen liegt, kann man von unkorrelierten Daten ausgehen. Da die Schiefe kleiner Null und die Kurtosis größer Null ist, sind die Daten wieder linksschief und sind eher „spitz“ angeordnet.

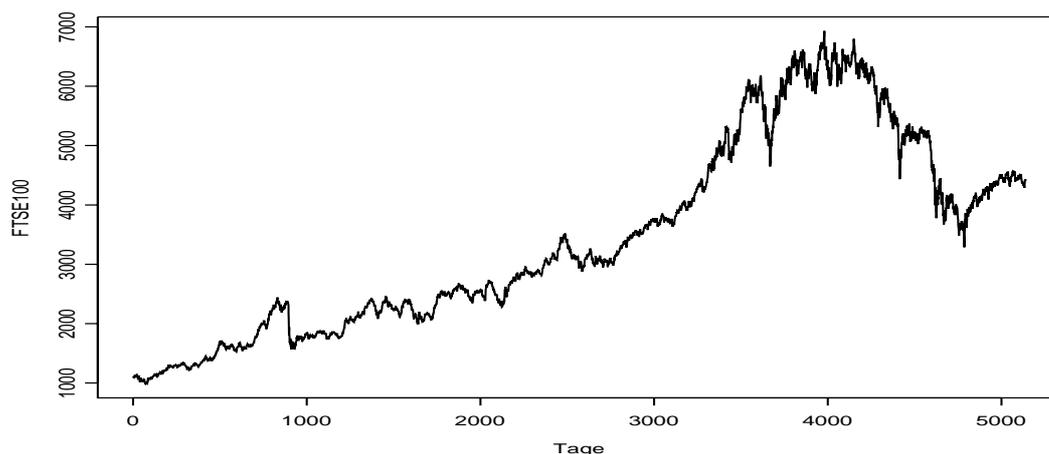


Abbildung 10.19: *Tägliche Schlußdaten des FTSE 100 - 1984 bis 2004.*

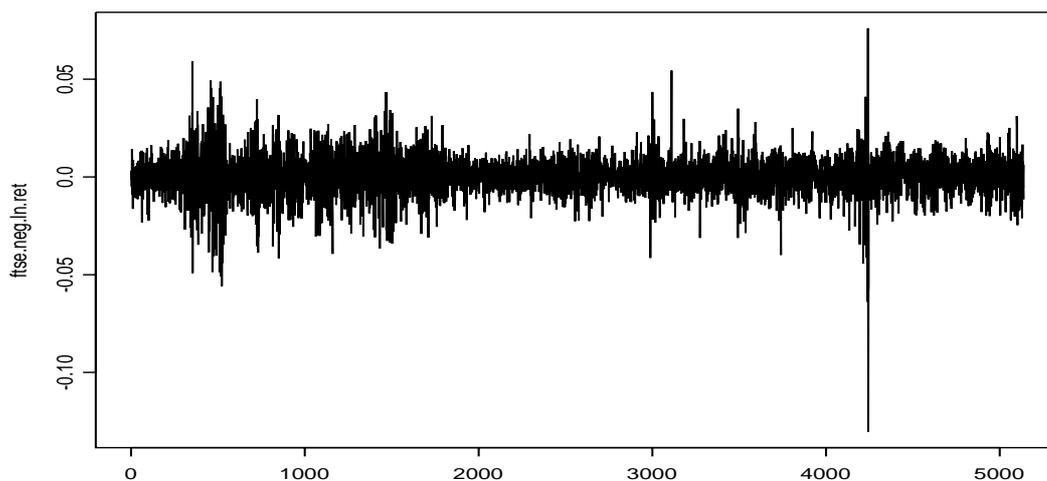


Abbildung 10.20: *Negative Log Returns der täglichen Schlußdaten des FTSE 100.*

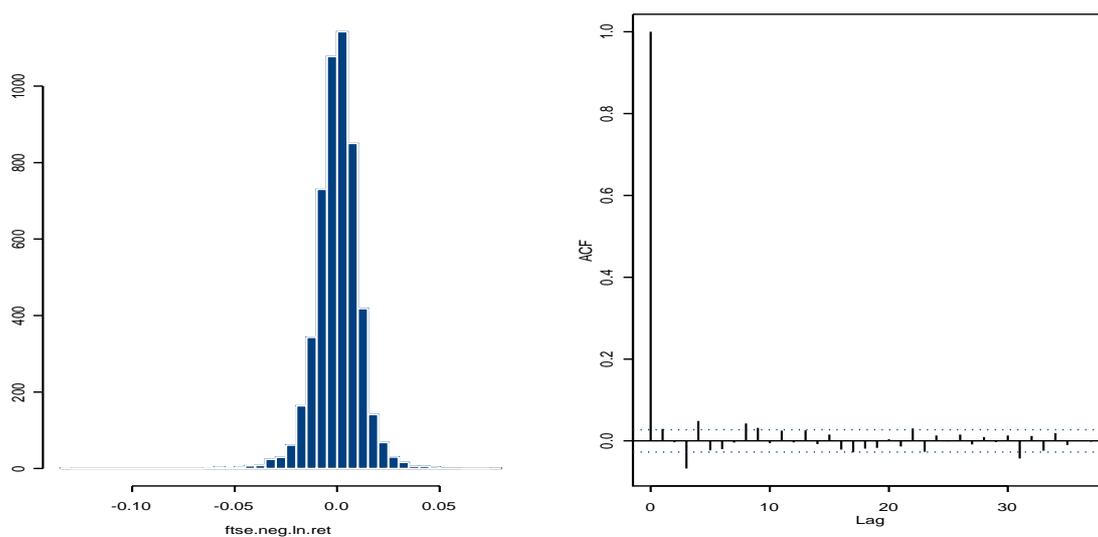


Abbildung 10.21: *Autokorrelationsfunktion und Histogramm der negativen Log Returns des FTSE 100*

Für die negativen Log Returns wird die POT Methode und POT Methode für stationäre Zeitreihen angewandt. Die Ergebnisse der geschätzten Parameter und der Wert für den VaR findet man in den Tabellen 10.6, 10.7 und 10.8. Der Wert für die Schwelle u wurde mit der Exzessfunktion bestimmt und kann in der Abbildung 10.23 abgelesen werden. In der Abbildung 10.6 wurden die Exzesse bzgl. der Schwelle u der negativen Log Returns an

eine GPD (Standard POT Methode) angepasst. Wieder wird bei der POT Methode das 95%-Quantil geschätzt und damit das DRC berechnet. Der Wert des VaR ist das DRC korrigiert um den Erwartungswertes. Auch beim Datensatz FTSE 100 wird bei den POT Methoden höher angesetzt als bei der empirischen Methode. Für die Schätzer für ξ und β wurden bei der Standard POT Methode der Maximum Likelihood der Exzedenten und bei der POT Methode für stationäre Zeitreihen der ML der Maxima benutzt.

Die Schätzung von θ mit den verschiedenen Methoden sieht man in Abbildung (10.24), dabei wurde für die run Methode (5.10), für die Zwei-Schwellen Methode (7.5) und für die Momenten Methode (6.9) verwendet. In Abbildung 10.25 werden die Resultate der Schätzung mit der Blockmethode abgebildet.

Die Abbildungen 10.26 bis 10.36 zeigen die Ergebnisse des Declustering der Daten mit Hilfe der verschiedenen Methoden, die in dieser Arbeit beschrieben wurden. Die Auswertung dieser Abbildungen entspricht den Ergebnissen des Datensatzes S&P 500.

	u	$\hat{\xi}$	$\hat{\beta}$	n	N_u
FTSE 100	0.008	0.13	0.00514	5138	1015

Tabelle 10.6:

Der Schwellenwert u , der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Anzahl der Daten n , die Anzahl der Exzedenten N_u und die Maximum Likelihood Schätzer für ξ und β des Datensatzes FTSE 100. (Standard POT Methode).

FTSE 100	u	s	c	$\hat{\xi}$	$\hat{\beta}$	n	N_u^C	$\hat{\theta}$
run Methode	0.008	5		0.081	0.0069	5138	345	0.33990
Zwei-Schwellen Methode	0.008	5	-0.00545	0.0411	0.0072	5138	582	0.5734

Tabelle 10.7:

Der Schwellenwert u , der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Schwelle c , für die das 25%-Quantil genommen wurde, run Länge s , die Anzahl der Daten n , die Anzahl der Clustermaxima N_u^C , die Schätzer der run bzw. Zwei-Schwellen Methode und die Maximum Likelihood Schätzer für ξ und β des Datensatzes FTSE 100 (POT Methode für stationäre Zeitreihen).

FTSE 100	empirisch	Standard POT	POT für stat. Zeitreihen (run Methode)	POT für stat. Zeitreihen (Zwei-Schwellen Methode)
VaR	-0.01516	-0.01546	-0.01776	-0.01791
DRC	0.01543	0.01573	0.01803	0.01818

Tabelle 10.8:

POT Methoden angewandt auf die negativen Log Returns des Datensatzes FTSE 100.

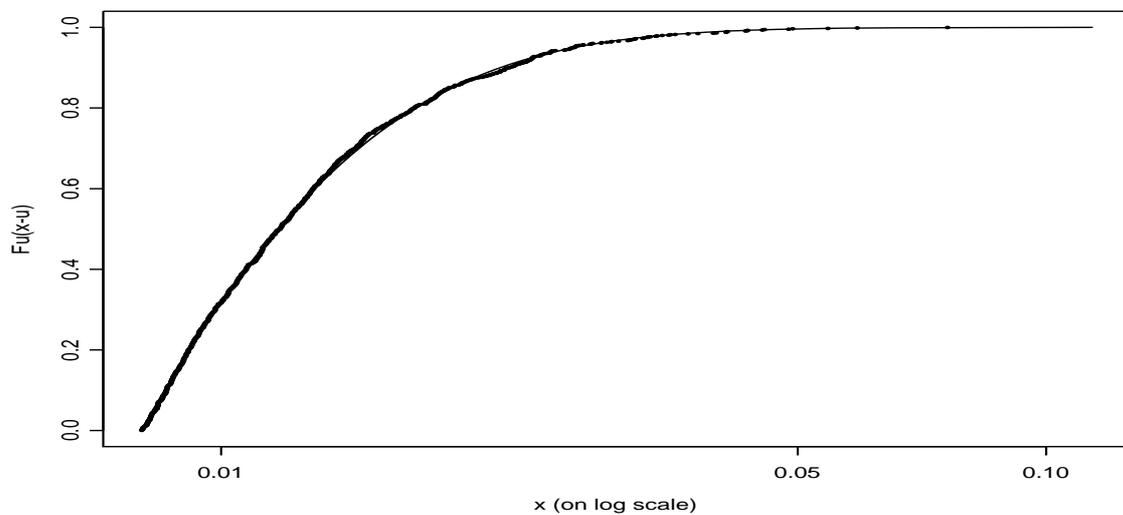


Abbildung 10.22:

Anpassung der negativen Log Returns des FTSE 100 an eine GPD. ξ und β wurden durch einen Maximum Likelihood Schätzer geschätzt und sind in de Tabelle 10.6 zu finden.

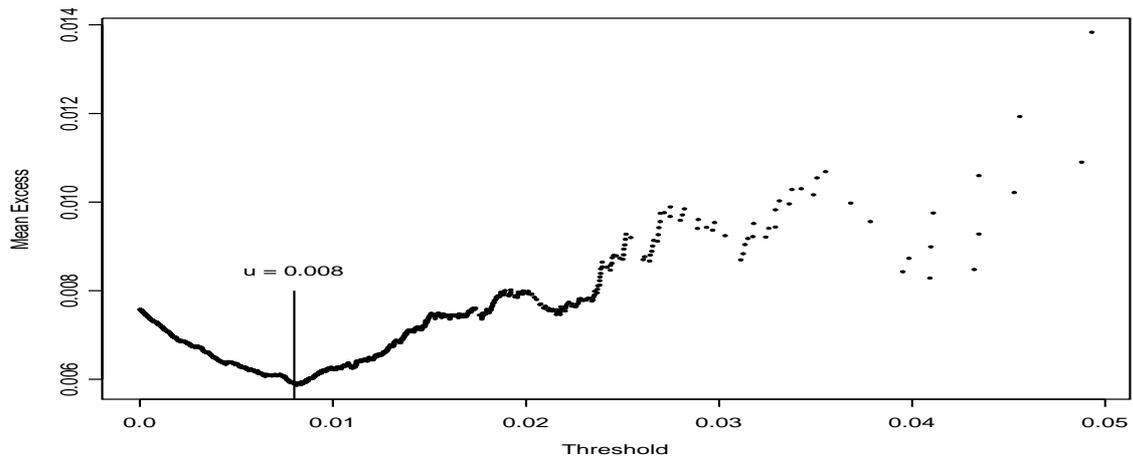


Abbildung 10.23:
 Die Exzessfunktion der positiven Daten der negativen Log Returns des FTSE 100.

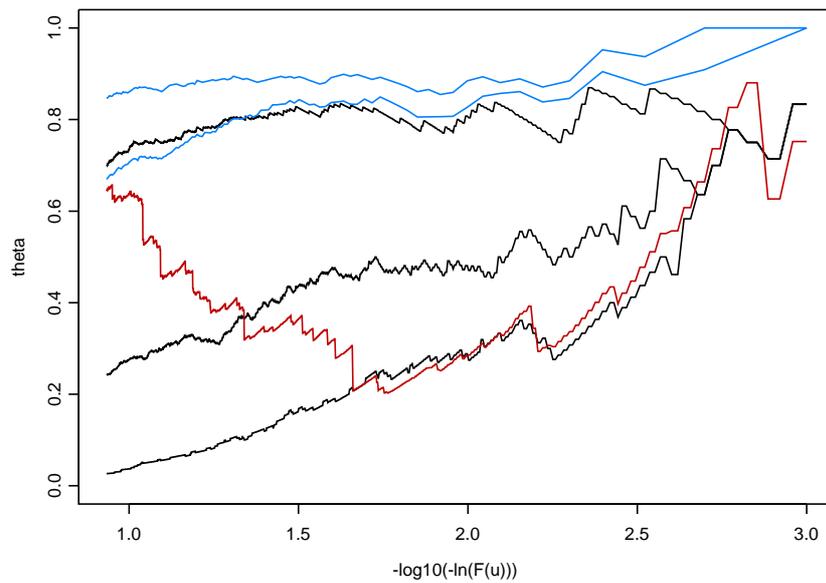


Abbildung 10.24:
 θ geschätzt für den Datensatz FTSE 100 mit der run Methode (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), mit der Zwei-Schwellen Methode für $s = 50$ und c mit $F(c) = 0.75, 0.25$ (von oben nach unten) und der Momenten Methode. u entspricht dem 89%- bis 99.9%-Quantil des Datensatzes und ist in einer $-\log_{10}(\ln(F(u)))$ Skala abgebildet.

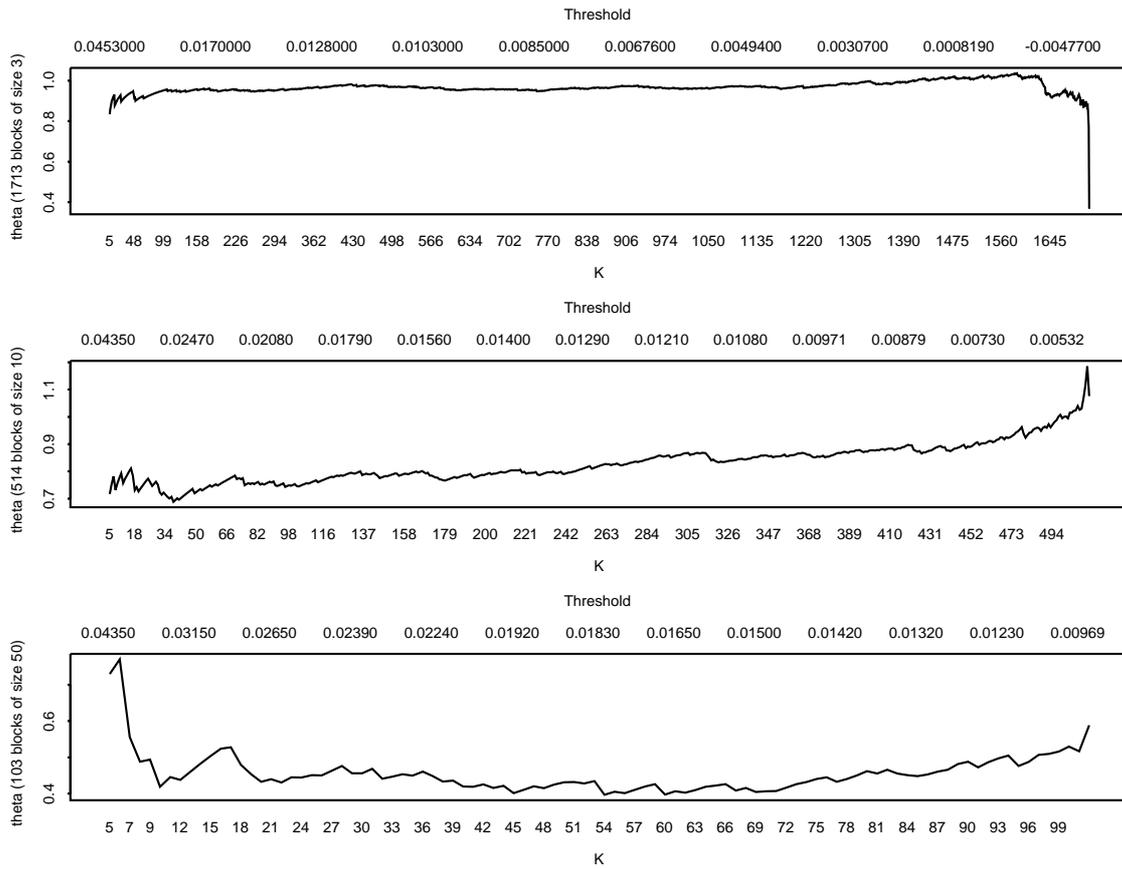


Abbildung 10.25:
 θ geschätzt für den Datensatz FTSE 100 mit der Blockmethode. Dabei wurde die Blockgröße 3, 10 und 50 (von oben nach unten) gewählt.

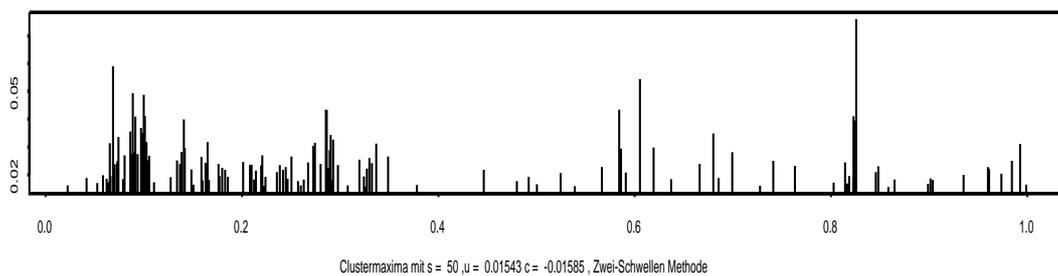
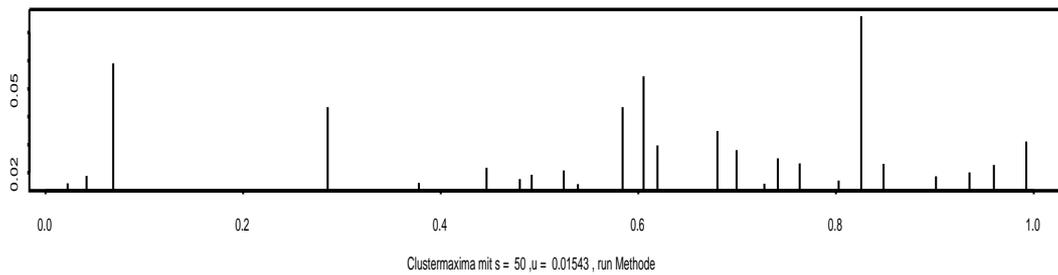
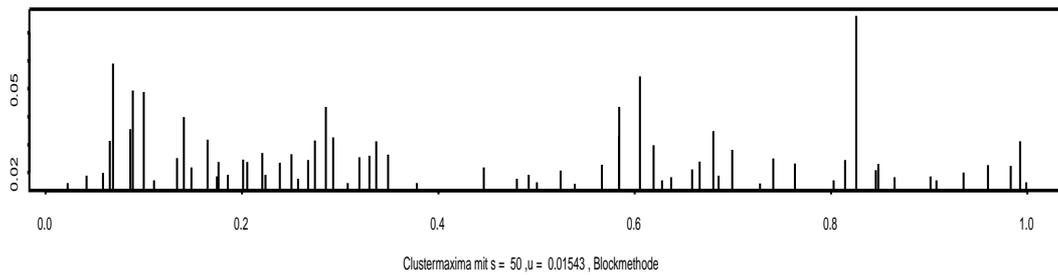
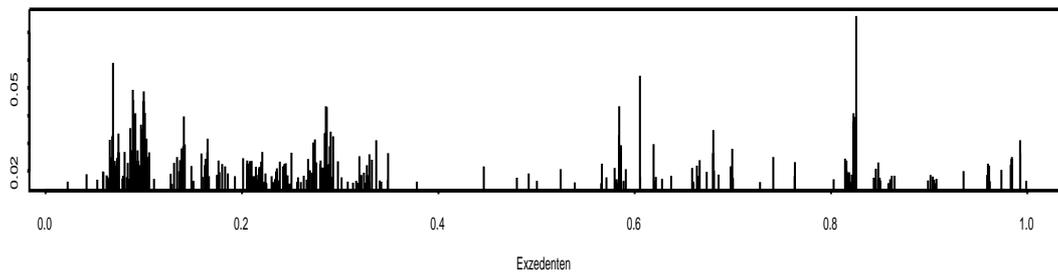


Abbildung 10.26:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 50$.

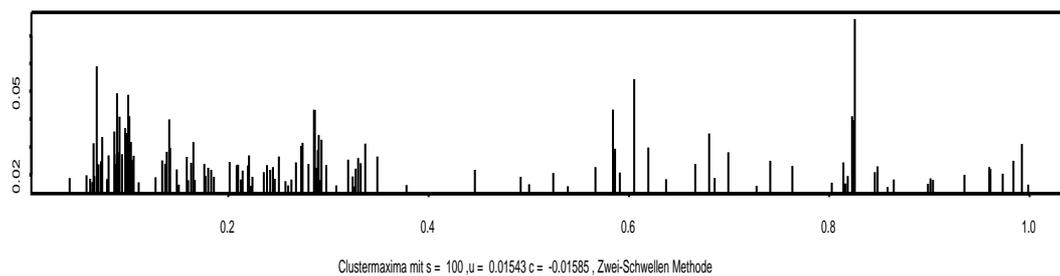
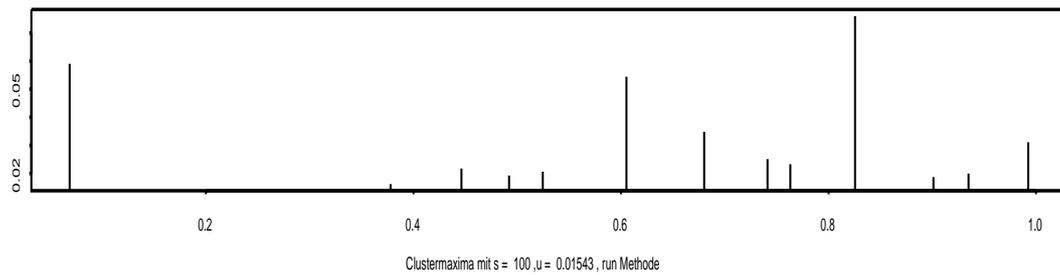
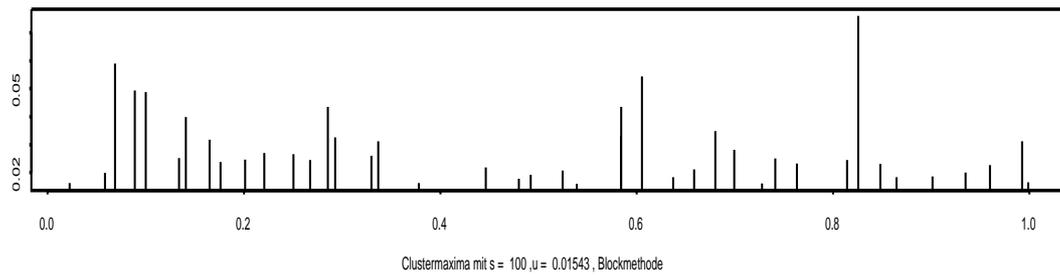
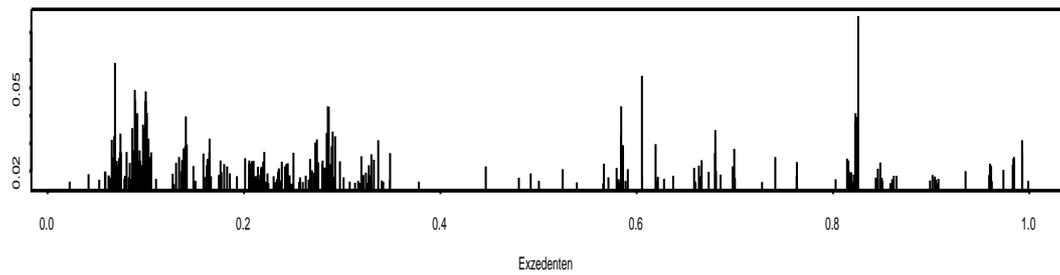


Abbildung 10.27:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 100$.

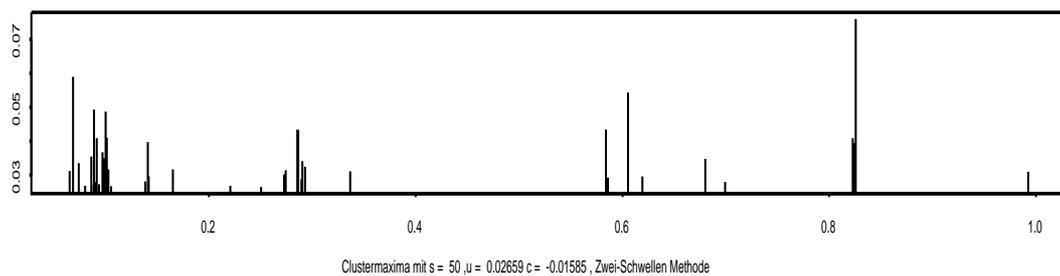
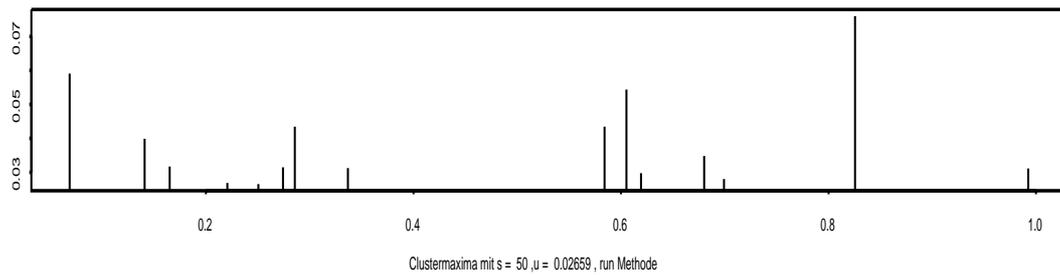
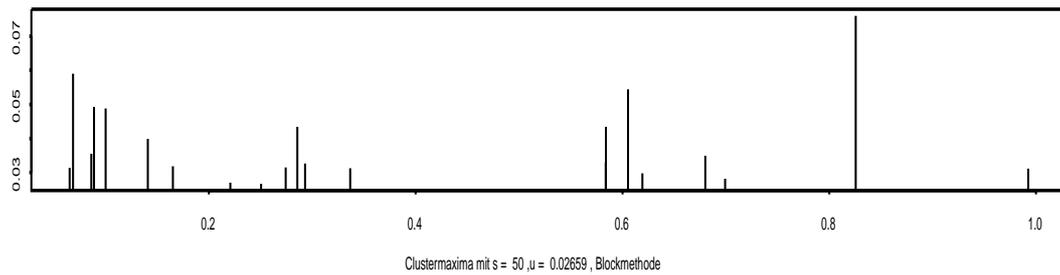
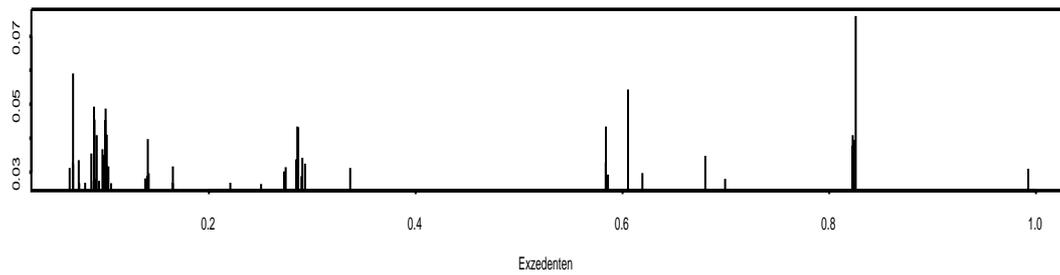


Abbildung 10.28:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 50$.

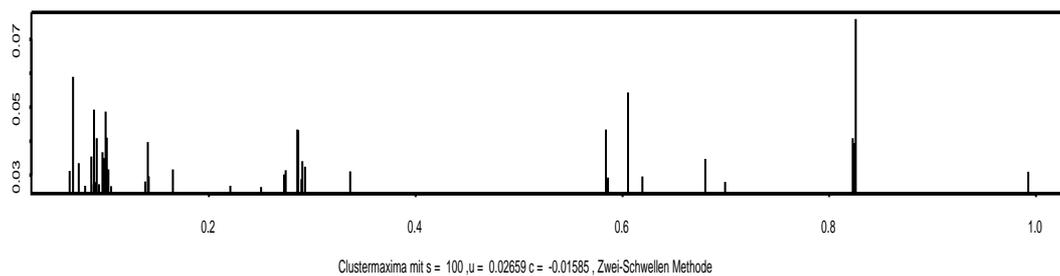
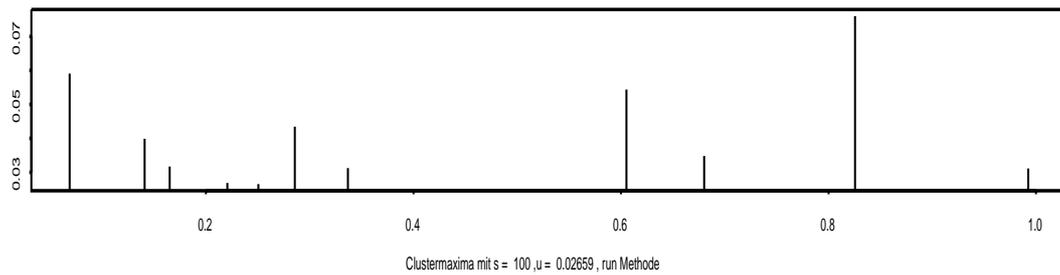
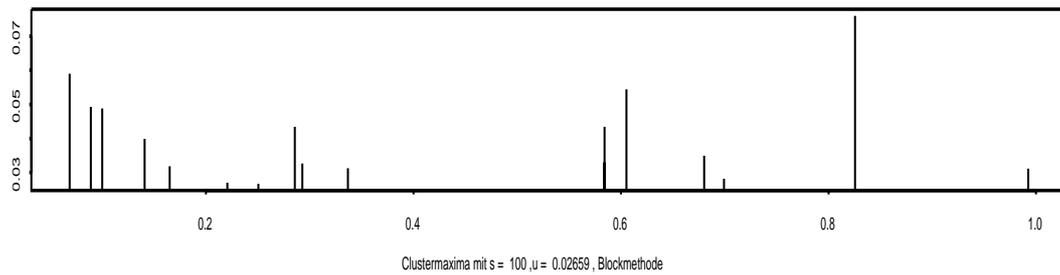
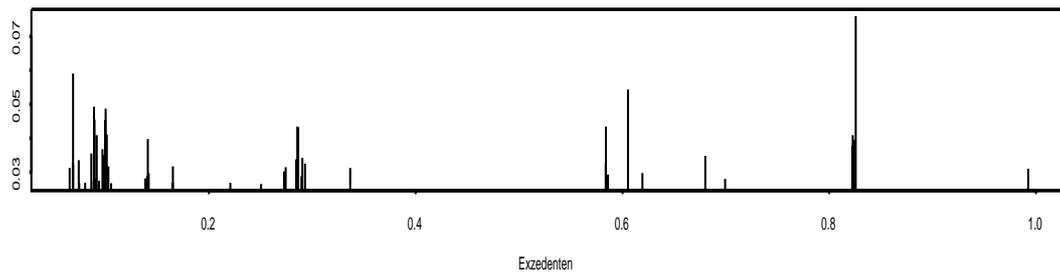


Abbildung 10.29:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 100$.

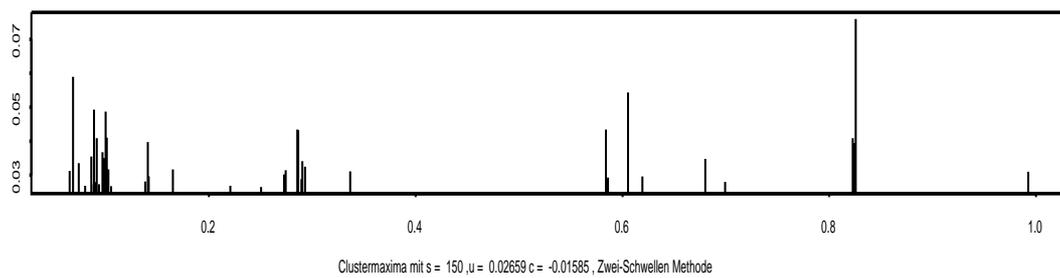
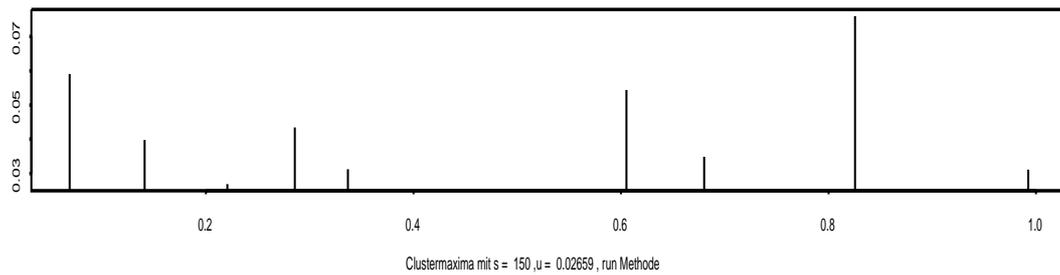
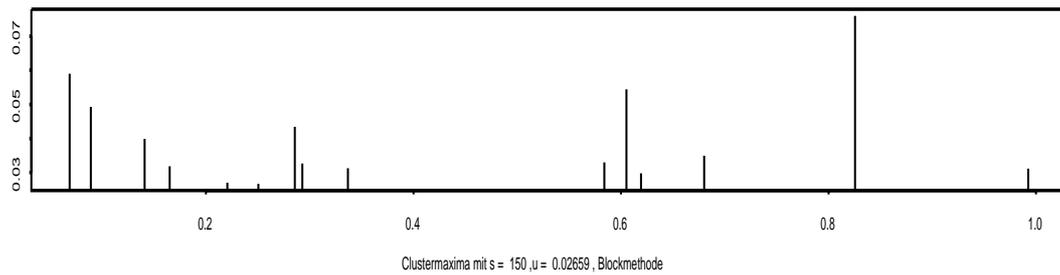
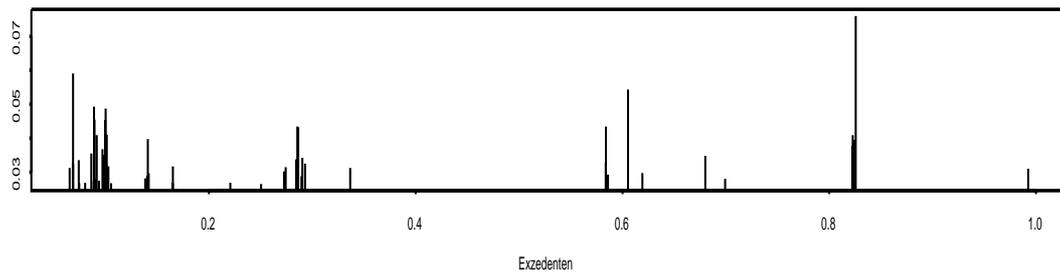


Abbildung 10.30:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 150$.

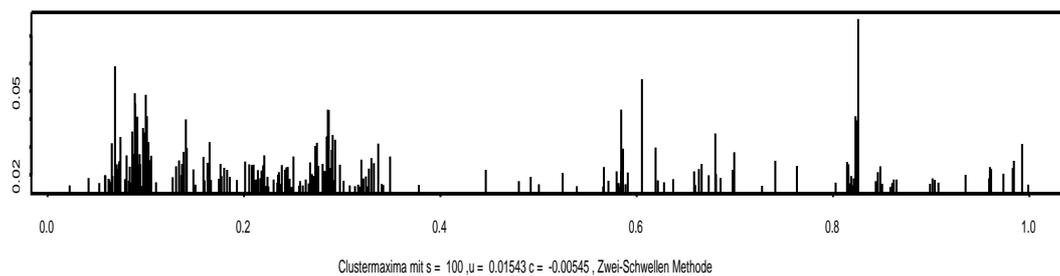
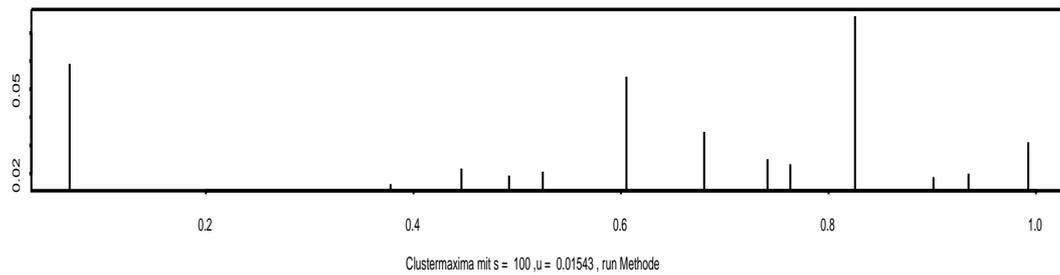
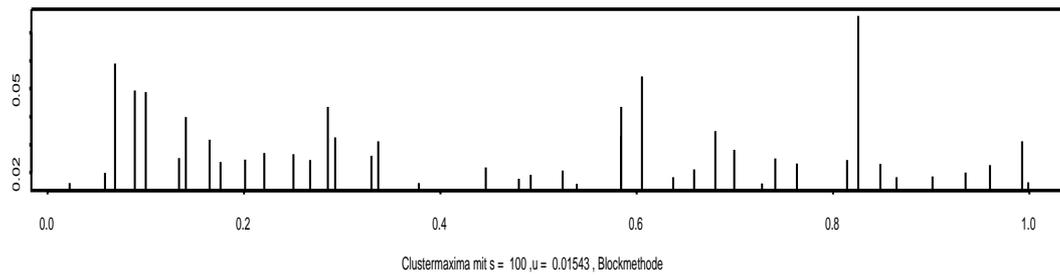
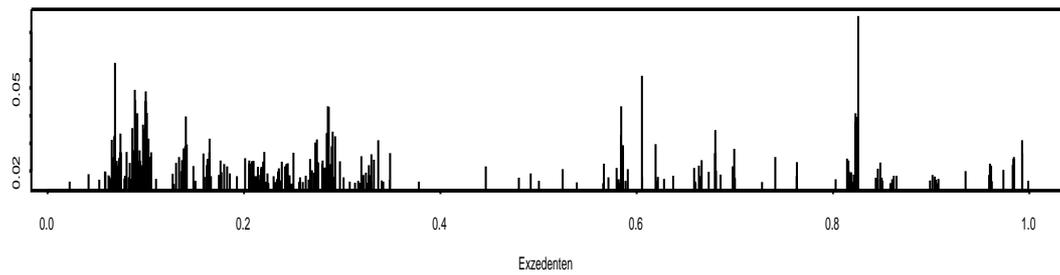


Abbildung 10.31:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 25\%$ -Quantil und $s = 100$.

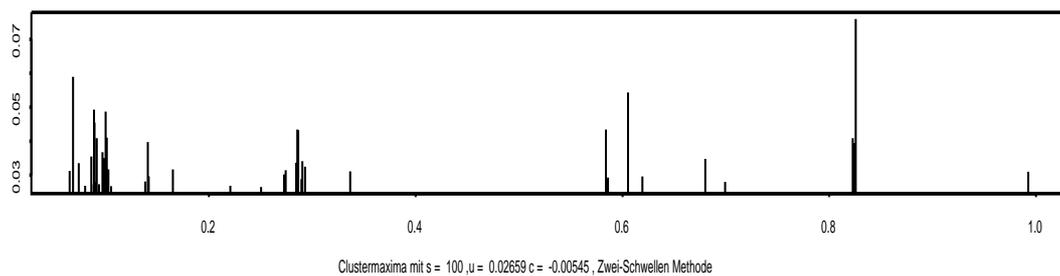
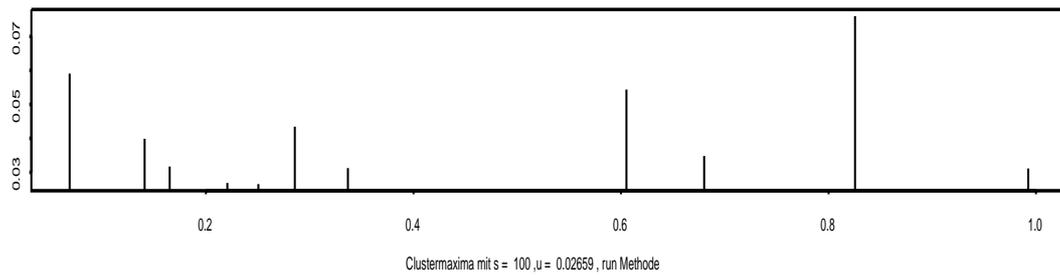
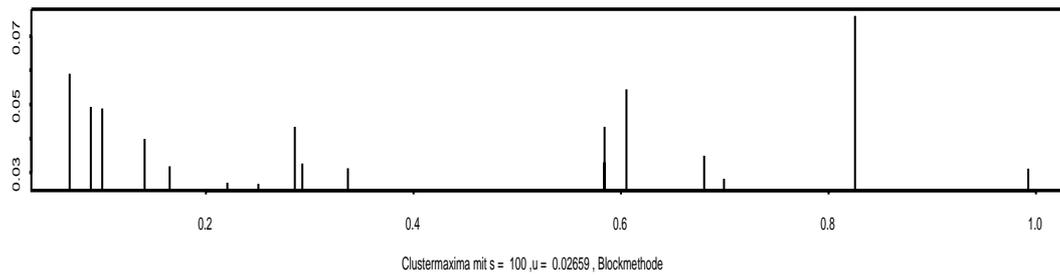
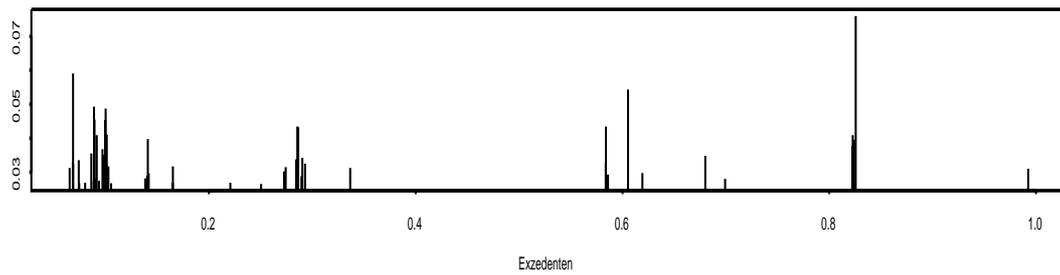


Abbildung 10.32:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 25\%$ -Quantil und $s = 100$.

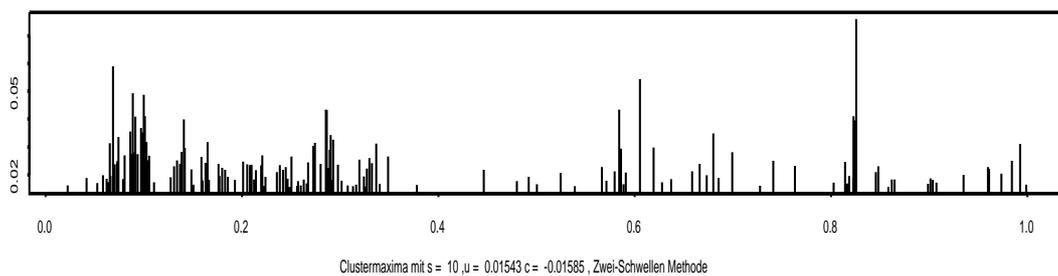
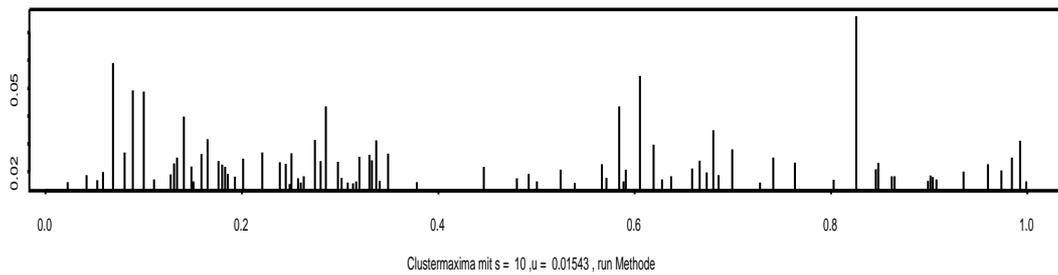
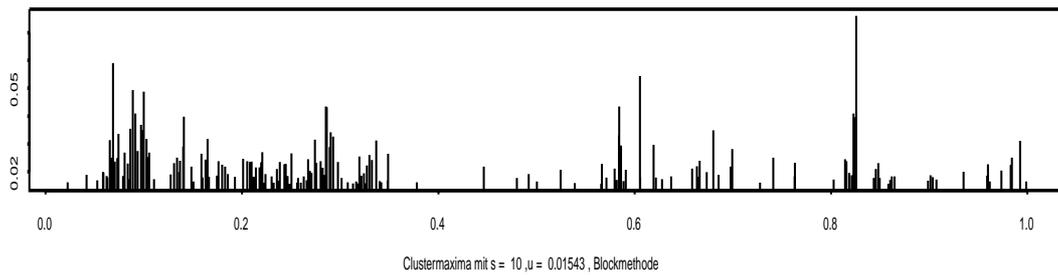
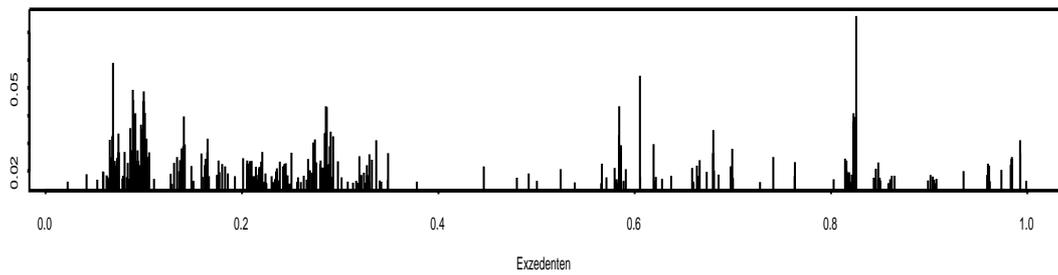


Abbildung 10.33:

Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 10$.

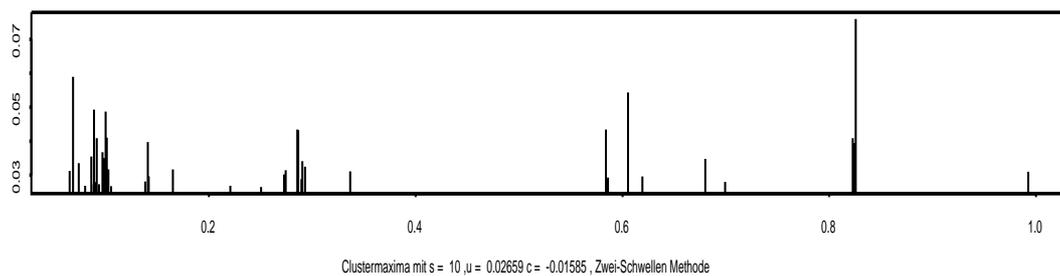
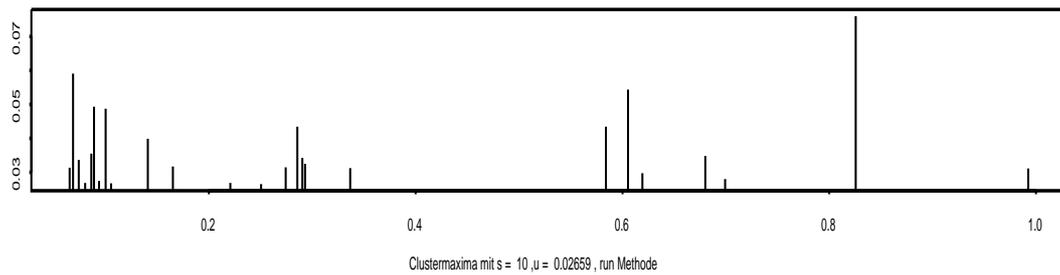
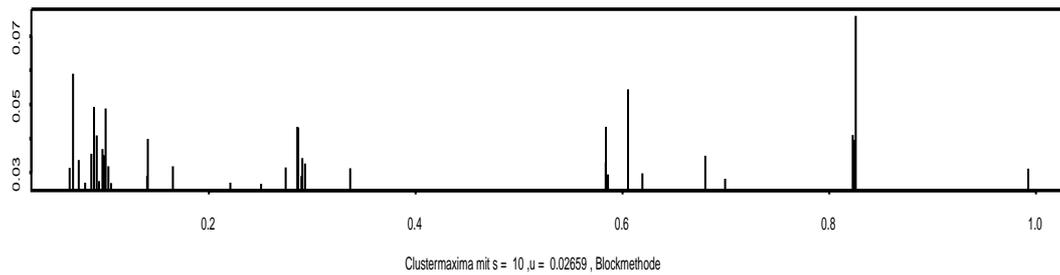
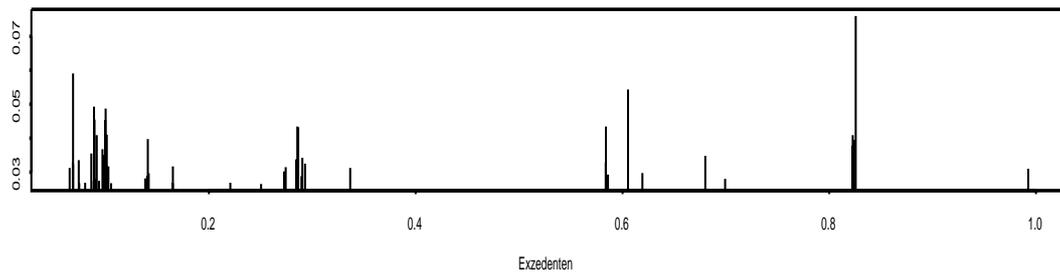


Abbildung 10.34:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 10$.

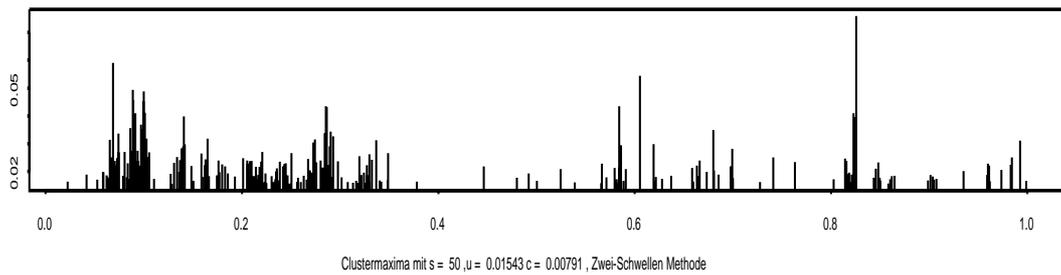
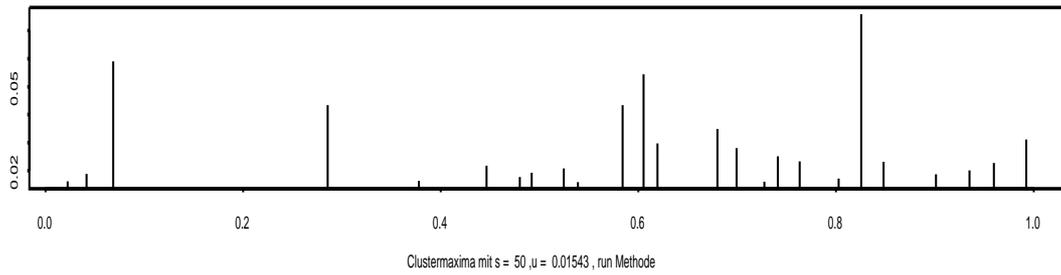
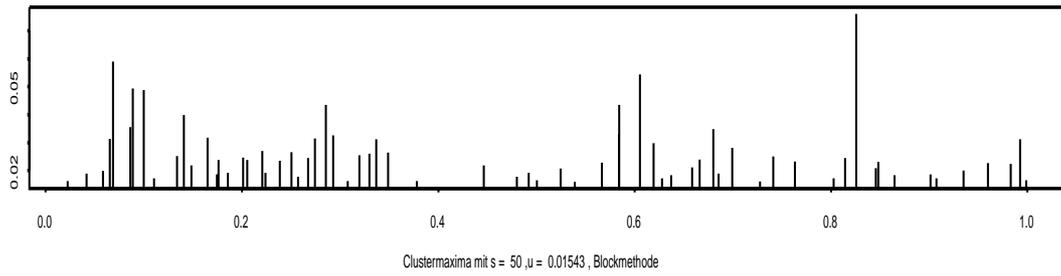
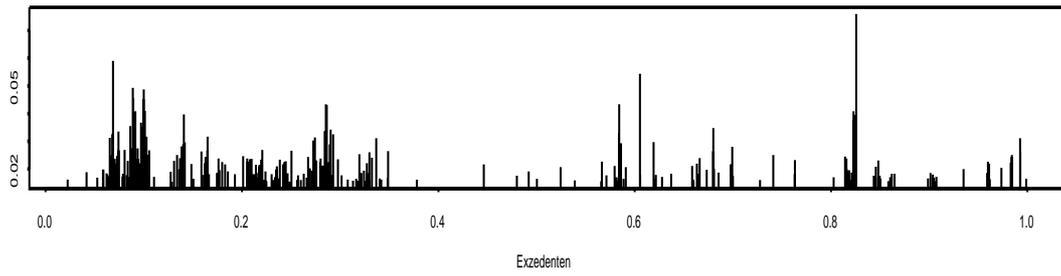


Abbildung 10.35:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 80\%$ -Quantil und $s = 50$.

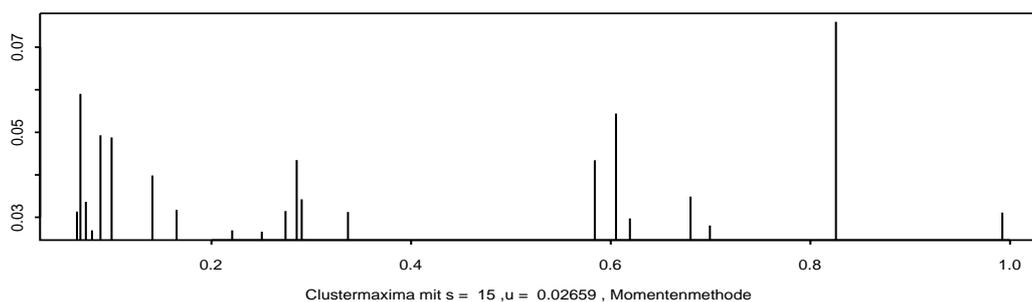
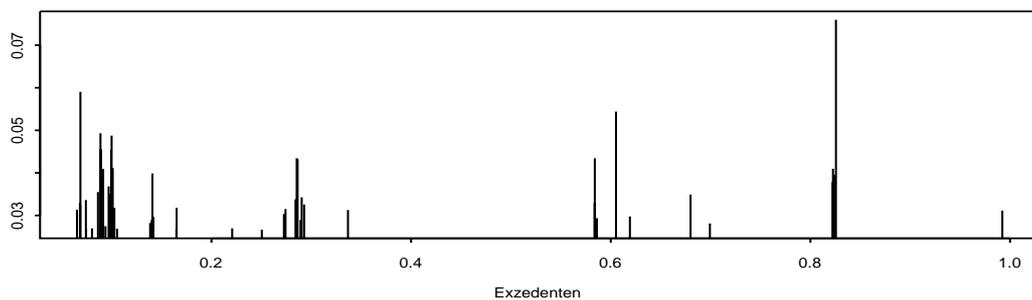
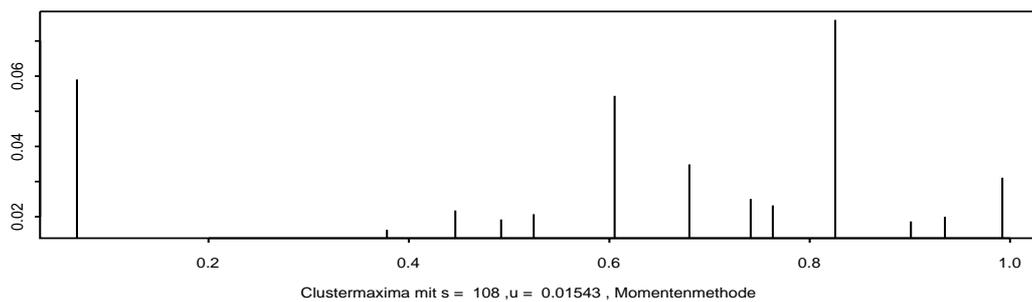
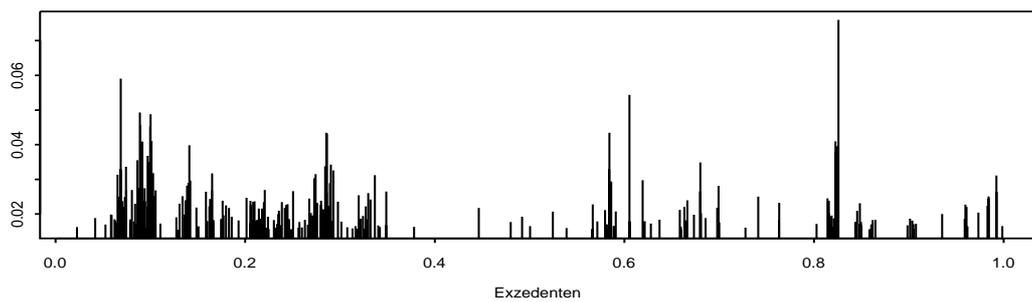


Abbildung 10.36:
Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch Momenten Methode. Die gewählten Werte waren $u = 95\%$ -Quantil (oberen zwei Graphen) und $u = 99\%$ -Quantil (unteren zwei Graphen).

Kapitel 11

Abschliessende Bemerkungen

Alle Berechnungen und Graphen wurden mit Hilfe des Programms SPlus und dem Paket EVIS hergestellt. Informationen zu EVIS findet man unter [19]. Für die Untersuchung der Datensätze wurden auch eigene Funktionen geschrieben, die im Anhang zu finden sind.

Allgemein wurde gesehen, dass die Zwei-Schwellen Methode und Volatilitäts-Schwellen Methode eine der stabilsten Methoden sind. D.h. dass diese Methoden weniger sensibel auf die Wahl von Hilfsparametern reagieren. Sie berücksichtigen den Pfad der Daten. Der Intervallschätzer ist nur von u abhängig und somit ist nur ein Parameter zu wählen. Dies vereinfacht die Sache in der Hinsicht, da es oft schwierig ist, die richtige run Länge s oder untere Schwelle c zu wählen.

In dieser Arbeit wurde ein ausführlicher Vergleich der alten und neuen Methoden durchgeführt und an zwei Datensätzen veranschaulicht. Dabei wurden die theoretischen Tatsachen verifiziert. Auch die Simulationsstudie soll dazu dienen, die Theorie zu veranschaulichen. Vor allem ging es darum, den extremalen Index θ so gut wie möglich zu schätzen und Cluster zu identifizieren.

Anhang A

A.1 Beweis Lemma 6.5

Die Aussage von Lemma 6.5 war: Betrachtet man die Wahrscheinlichkeit 6.7 der Approximation der gemischten Grenzverteilung (6.4), so erhält man

$$\theta \stackrel{(6.6)}{=} \frac{2E(T)^2}{E(T^2)} = \frac{2(1 - (1 - \theta)p^\theta)^2}{2\theta p^\theta + \theta p^\theta(1 - p^\theta) + (1 - p^\theta)^2} = \theta + \theta \left(2 - \frac{3\theta}{2}\right) (1 - p) + O\{(1 - p)^2\}.$$

Beweis. Es gilt $P(T > n) = \theta p^{n\theta}$

$$\begin{aligned} (1) \quad & E(T - 1) = E(T) - 1 \\ & \Rightarrow E(T) = E(T - 1) + 1 \\ (2) \quad & E\left(\frac{T(T - 1)}{2}\right) = E\left(\frac{T^2 - T}{2}\right) = \frac{1}{2}(E(T^2) - E(T)) \\ & \Rightarrow E(T^2) = 2E\left(\frac{T(T - 1)}{2}\right) + E(T) \\ (3) \quad & E(T - 1) = E(T) - P(T > 0) = \sum_{n=1}^{\infty} nP(T = n) - P(T > 0) = \\ & = \sum_{n=1}^{\infty} n[P(T > n - 1) - P(T > n)] - P(T > 0) = \\ & = \sum_{n=1}^{\infty} nP(T > n - 1) - \sum_{n=1}^{\infty} nP(T > n) - P(T > 0) = \\ & = \sum_{n=0}^{\infty} (n + 1)P(T > n) - \sum_{n=1}^{\infty} nP(T > n) - P(T > 0) = \\ & = P(T > 0) + \sum_{n=1}^{\infty} (n + 1)P(T > n) - \sum_{n=1}^{\infty} nP(T > n) - P(T > 0) = \\ & = \sum_{n=1}^{\infty} P(T > n) = \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} \theta p^{n\theta} = \theta \sum_{n=0}^{\infty} p^{(n+1)\theta} = \theta p^\theta \sum_{n=0}^{\infty} p^{n\theta} \stackrel{geo.Reihe}{=} \theta p^\theta (1 - p^\theta)^{-1} \\
(4) \quad E\left(\frac{T(T-1)}{2}\right) &= \frac{1}{2} [E(T^2) - E(T)] = \frac{1}{2} \left[\sum_{n=1}^{\infty} n^2 P(T=n) - \sum_{n=1}^{\infty} n P(T=n) \right] = \\
&= \frac{1}{2} \left[\sum_{n=1}^{\infty} n^2 (P(T > n-1) - P(T > n)) - \sum_{n=1}^{\infty} n (P(T > n-1) - P(T > n)) \right] = \\
&= \frac{1}{2} \left[\sum_{n=0}^{\infty} (n^2 + 2n + 1) P(T > n) - \sum_{n=1}^{\infty} n^2 P(T > n) - \sum_{n=0}^{\infty} (n+1) P(T > n) + \right. \\
&\quad \left. + \sum_{n=1}^{\infty} n P(T > n) \right] = \\
&= \frac{1}{2} \left[P(T > 0) + \sum_{n=1}^{\infty} (n^2 + 2n + 1) P(T > n) - \sum_{n=1}^{\infty} n^2 P(T > n) - P(T > 0) - \right. \\
&\quad \left. - \sum_{n=1}^{\infty} (n+1) P(T > n) + \sum_{n=1}^{\infty} n P(T > n) \right] = \\
&= \frac{1}{2} \left[\sum_{n=1}^{\infty} (n^2 + 2n + 1 - n^2 - n - 1 + n) P(T > n) \right] = \frac{1}{2} \left[\sum_{n=1}^{\infty} 2n P(T > n) \right] = \\
&= \sum_{n=1}^{\infty} n P(T > n) = \\
&= \sum_{n=1}^{\infty} n \theta p^{n\theta} = \theta \sum_{n=1}^{\infty} n p^{n\theta} \stackrel{geo.Reihe}{=} \theta \left[(1 - p^\theta)^{-1-1} \sum_{j=1}^1 a_j^{(1)} (p^\theta)^j \right] \stackrel{a_1^{(1)}=1}{=} \\
&= \theta p^\theta (1 - p^\theta)^{-2}
\end{aligned}$$

$$\begin{aligned}
(1) - (4) \Rightarrow \frac{2E(T)^2}{E(T^2)} &= \frac{2(\theta p^\theta (1 - p^\theta)^{-1} + 1)^2}{2\theta p^\theta (1 - p^\theta)^{-2} + \theta p^\theta (1 - p^\theta)^{-1} + 1} = \\
&= \frac{2(\theta^2 p^{2\theta} (1 - p^\theta)^{-2} + 2\theta p^\theta (1 - p^\theta)^{-1} + 1)}{2\theta p^\theta (1 - p^\theta)^{-2} + \theta p^\theta (1 - p^\theta)^{-1} + 1} = \\
&= \frac{2(\theta^2 p^{2\theta} + 2\theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2)}{2\theta p^\theta + \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2} = \\
&= \frac{2(\theta p^\theta + (1 - p^\theta))^2}{2\theta p^\theta + \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2} = \\
&= \frac{2(1 - (1 - \theta)p^\theta)^2}{2\theta p^\theta + \theta p^\theta (1 - p^\theta) + (1 - p^\theta)^2}
\end{aligned}$$

Jetzt wird eine lineare Approximation von p^θ in $p = 1$ durchgeführt. Informationen zur linearen Approximation findet man in Königsberger [13] (S. 139f). Sei $f(p) = p^\theta$.

$$\Rightarrow f'(p) = (p^\theta)' = \theta p^{\theta-1}$$

Für $p = 1$ gilt:

$$f'(1) = \theta.$$

Demnach sieht die lineare Approximation von p^θ in $p = 1$ folgendermaßen aus:

$$\begin{aligned} p^\theta &\approx f(1) - f'(1)(1-p) = 1 - \theta(1-p) \\ \Rightarrow 1 - p^\theta &\approx \theta(1-p) \end{aligned} \tag{A.1}$$

Die lineare Approximation von $g(p) := \frac{2E(T)^2}{E(T^2)}$ in $p = 1$ hat die Gestalt:

$$g(p) = \frac{2E(T)^2}{E(T^2)} \approx g(1) - d(1-p) + O((1-p)^2),$$

wobei $d = g'(1)$ sei.

Mit der linearen Approximation von p^θ in $p = 1$ (A.1) gilt:

$$\begin{aligned} g(p) &= \frac{2(1 - (1-\theta)p^\theta)^2}{2\theta p^\theta + \theta p^\theta(1-p^\theta) + (1-p^\theta)^2} \approx \\ &\approx \frac{2(1 - (1-\theta)[1 - \theta(1-p)])^2}{2\theta[1 - \theta(1-p)] + \theta[1 - \theta(1-p)][\theta(1-p)] + (\theta(1-p))^2} = \\ &= \frac{2(1 - 2(1-\theta)(1 - \theta(1-p)) + (1-\theta)^2(1 - \theta(1-p))^2)}{2\theta - 2\theta^2(1-p) + \theta^2(1-p) - \theta^3(1-p)^2 + \theta^2(1-p)^2} = \\ &= \frac{2(1 - 2(1-\theta - \theta(1-p) + \theta^2(1-p)) + (1-\theta)^2(1 - 2\theta(1-p) + \theta^2(1-p)^2))}{2\theta + (1-p)(-\theta^2) + (1-p)^2(-\theta^3 + \theta^2)} = \\ &= \frac{2(-1 + 2\theta + (1-p)2(\theta - \theta^2) + (1-\theta)^2 - 2\theta(1-\theta)^2(1-p) + \theta^2(1-\theta)^2(1-p)^2)}{2\theta + (1-p)(-\theta^2) + (1-p)^2(-\theta^3 + \theta^2)} = \\ &= \frac{-2 + 4\theta + 4(1-p)(\theta - \theta^2 - \theta(1-\theta)^2) + 2(1-\theta)^2 + 2\theta^2(1-\theta)^2(1-p)^2}{2\theta + (1-p)(-\theta^2) + (1-p)^2(-\theta^3 + \theta^2)} = \\ &= \frac{2\theta^2 + 4(1-p)(\theta^2 - \theta^3) + 2\theta^2(1-\theta)^2(1-p)^2}{2\theta + (1-p)(-\theta^2) + (1-p)^2(-\theta^3 + \theta^2)} = \\ &= \frac{2\theta + 4(1-p)\theta(1-\theta) + 2\theta(1-\theta)^2(1-p)^2}{2 + (1-p)(-\theta) + (1-p)^2\theta(1-\theta)} \\ &\Rightarrow g(1) = \theta \end{aligned}$$

$$\begin{aligned}
g'(p) &= \left(\frac{2\theta}{2 + (1-p)(-\theta) + (1-p)^2\theta(1-\theta)} \right)' + \left(\frac{4(1-p)\theta(1-\theta)}{2 + (1-p)(-\theta) + (1-p)^2\theta(1-\theta)} \right)' + \\
&+ \left(\frac{2\theta(1-\theta)^2(1-p)^2}{2 + (1-p)(-\theta) + (1-p)^2\theta(1-\theta)} \right)' \stackrel{z:=2+(1-p)(-\theta)+(1-p)^2\theta(1-\theta)}{=} \\
&= \frac{-2\theta(\theta - 2(1-p)(\theta - \theta^2))}{z^2} + \\
&+ \frac{z(-4)(\theta - \theta^2) - 4(1-p)(\theta - \theta^2)(\theta - 2(1-p)(\theta - \theta^2))}{z^2} + \\
&+ \frac{z(-4\theta(1-\theta)^2(1-p)) - (2\theta(1-\theta)^2(1-p)^2)(\theta - 2(1-p)(\theta - \theta^2))}{z^2}
\end{aligned}$$

Für $p = 1$ gilt $z = 2 + (1-1)(-\theta) + (1-1)^2\theta(1-\theta) = 2$

$$\begin{aligned}
\Rightarrow d &= \frac{-2\theta(\theta - 2(1-1)(\theta - \theta^2))}{4} + \\
&+ \frac{2(-4)(\theta - \theta^2) - 4(1-1)(\theta - \theta^2)(\theta - 2(1-1)(\theta - \theta^2))}{4} + \\
&+ \frac{2(-4\theta(1-\theta)^2(1-1)) - (2\theta(1-\theta)^2(1-1)^2)(\theta - 2(1-1)(\theta - \theta^2))}{4} = \\
&= \frac{-2\theta^2}{4} + \frac{-8(\theta - \theta^2)}{4} + \frac{0}{4} = \\
&= -\frac{1}{2}\theta^2 - 2\theta + 2\theta^2 = -2\theta + \frac{3}{2}\theta^2 = \\
&= -\theta \left(2 - \frac{3\theta}{2} \right)
\end{aligned}$$

Demnach gilt für die lineare Approximation:

$$\frac{2E(T)^2}{E(T^2)} \approx \theta + \theta \left(2 - \frac{3\theta}{2} \right) (1-p) + O((1-p)^2)$$

□

A.2 Verwendete Funktionen in SPlus

A.2.1 Voreingestellte Funktionen von EVIS

Alle Berechnungen und Graphen wurden mit Hilfe des Programms SPlus und dem Paket EVIS hergestellt. Informationen zu EVIS findet man unter [19]. Dabei wurde zur Erstellung der Exzessfunktion die Funktion *meplot*, zur Anpassung an die GPD die Funktion *gpd* und für den QQplot in den decluster Funktionen *qplot* verwendet. Neben den von EVIS voreingestellten Funktionen habe ich die Funktionen in den nächsten Kapitel entwickelt.

A.2.2 Berechnung der Schwellenwerte

Berechnung der oberen Schwellen u

Neben der Funktion *findthresh* aus EVIS, die eine Schwelle u dadurch findet, indem man die Anzahl der Exzedenten angibt, berechnet die Funktion *threshold* die Schwellen bzgl. der Quantile der Daten. Dabei gibt man für *serie* den Datensatz, für *beginn/ende* das Quantil, mit dem man starten/enden will, *schritt* gibt die Schrittlänge vor. Dadurch bekommt man einen Vektor von mehreren Schwellen u bzgl. der Quantile. Gibt man z.B. `threshold(danish,89,99,0.1)` bekommt man vom 89%-Quantil bis 99%-Quantil im 0.1 Schritt die Quantile des Datensatzes *danish*, der auch in EVIS enthalten ist, und diese kann man für die Schwellenwerte u verwenden.

```
threshold_function(serie,beginn,ende,schritt)
{
  thres_seq(beginn,ende,schritt)/100
  q_quantile(serie,thres)

  q
}
```

Berechnung der unteren Schwelle c

Dementsprechend berechnet die Funktion *untere.gre* die untere Schwelle c mit den Quantilen der Daten, die man untersuchen will. *serie*, *beginn*, *ende* und *schritt* sind wie bei der Funktion *threshold* zu interpretieren.

```
untere.gre_function(serie,beginn,ende,schritt)
{
  unten_seq(beginn,ende,schritt)/100
  q2_quantile(serie,unten)

  q2
}
```

A.2.3 Simulationsstudie

ARMAX(1) Prozess

Mit dieser Funktion kann man n Daten eines ARMAX(1) Prozesses simulieren. *alpha* steht für α aus der Definition des ARMAX(1) Prozesses (Definition 8.1).

```
armax <- function(n,alpha,start=0)
{
  X <- rep(0,n)
  X[1] <- start

  epsilon <- rep(0,n)
  uni_runif(n)

  for (i in 1:n)
    {epsilon[i] <- (alpha-1)/log(uni[i])}

  for (t in 2:n)
```

```

    {X[t] <- max(alpha*X[t-1],epsilon[t-1])}
  X
}

```

ARMAX theta run

Hiermit kann man θ eines ARMAX(1) berechnen auf der Basis der Formel (8.1). *alpha* steht natürlich für α , *run* für die run Länge s und *u* für die Schwelle u aus derselben Formel (8.1).

```

theta.armax.run <- function(alpha,run,u)
{
  theta_rep(0,length(u))

  for ( i in 1:(length(u)))
  {
    theta[i] <- ((exp(-alpha/u[i])-exp(-1/u[i]))/(1-exp(-1/u[i]))) *
                *exp(-(run-1)*(1-alpha)/u[i])
  }

  theta
}

```

AR Prozess

ar.1 simuliert n Daten AR(1) Prozesses. *alpha* steht für α in der Definition des AR(1) Prozesses (Definition 8.3).

```

ar.1 <- function(n,alpha,start=0)
{
  X <- rep(0,n)
  X[1] <- start

  Z_rnorm(n,0,1)

  for (t in 2:n)
  {X[t] <- alpha*X[t-1]+Z[t]}

  X
}

```

ARCH(1) Prozess

Mit dieser Funktion kann man n Daten eines ARCH(1) Prozesses simulieren. *lambda* steht für λ und *beta* für β aus der Definition des ARCH(1) Prozesses (Definition 8.4).

```

arch <- function(n,lambda,beta,start=0)
{
  X <- rep(0,n)
  X[1] <- start

  Z <- rnorm(n,0,1)
  sigma <- rep(0,n)

```

```

for (t in 2:n)
{
sigma[t] <- sqrt(beta+lambda*X[t-1]^2)
X[t] <- sigma[t]*Z[t]
}

return(X,sigma)
}

```

GARCH(1,1) Prozess

garch simuliert n Daten eines GARCH(1,1) Prozesses, wobei *beta* steht für β , *lambda* für λ und *gamma* für γ aus der Definition des GARCH(1,1) Prozesses (Definition 8.5).

```

garch <- function(n,beta,lambda,gamma,start=0)
{
X <- rep(0,n)
X[1] <- start

Z <- rnorm(n,0,1)
sigma <- rep(0,n)

for (t in 2:n)
{
sigma[t] <- sqrt(beta+lambda*X[t-1]^2+gamma*sigma[t-1]^2)
X[t] <- sigma[t]*Z[t]
}

return(X,sigma)
}

```

A.2.4 Clusteridentifizierung mit den verschiedenen Methoden

Bei den Funktionen *cluster.** ist zu beachten, dass sie nur dann ein richtiges Ergebnis liefern, falls alle Clustermaxima unterschiedlich sind. Wenn gleiche Maxima vorhanden sind, findet die Funktion *match* nicht die richtigen Positionen der Maxima im Datensatz.

Identifizierung der Cluster mit der Blockmethode

Für *serie* gibt man den Datensatz ein, den man untersuchen will, für *obere.grenze* gibt man die obere Schwelle u ein und für r die Blockgröße. Die Schätzung wird mit der Formel (5.8) berechnet. Wählt man $plot=T$ werden die Exzedenten (links oben) und der zugehörige QQ-Plot (rechts oben), als auch die Maxima der Cluster (links unten) und der zugehörige QQ-Plot (rechts unten) geplottet.

```

cluster.block_function(serie,obere.grenze,r,plot=T)
{
times <- (1:length(serie))/length(serie)
daten <- signalSeries(serie,times)
daten.exce <- daten[daten>obere.grenze]

nblock <- (length(daten) %/% r)+1

```

```

grouping <- rep(1:nblock, rep(r,nblock))[1:length(serie)]
block.max <- tapply(as.numeric(serie),grouping,max)
block.gap <- block.max>obere.grenze

b <- {}
for (i in 1:nblock) {if(block.gap[i]==1) {b <- c(b,i)}}

blockseries <- daten[match(block.max[b],daten)]

if(plot)
{
  cat("Declustering Graph...\n")
  cat(paste("Daten wurden reduziert von ", length(as.numeric(daten.exce)),
"auf", length(as.numeric(blockseries)), "\n"))
  par(mfrow = c(2., 2.))
  plot(positions(daten.exce),as.numeric(daten.exce), type="h",xlab="Exzedenten",
ylab=" ",xlim=c(0,1),ylim=c(0,max(daten.exce)))
  gaps <- as.numeric(diff(positions(daten.exce)))
  qplot(gaps)
  plot(positions(blockseries),as.numeric(blockseries),type="h",xlab="Blockmaxima",
ylab=" ",xlim=c(0,1),ylim=c(0,max(daten.exce)))
  newgaps <- as.numeric(diff(positions(blockseries)))
  qplot(newgaps)
  par(mfrow = c(1, 1))
}

blockseries
}

```

Identifizierung der Cluster mit der run Methode

Mit der Funktion *cluster.run* werden die Cluster mit der run Methode identifiziert. *serie* und *obere.grenze* haben dieselbe Bedeutung wie bei der Funktion *cluster.block*. Die run Länge wird bei *run* eingegeben. Wählt man *plot=T*, ergeben sich dieselben Plots wie bei *cluster.block*. Die Funktion sucht sich die Exzedenten, berechnet die Interexzedentenzeiten in *gaps* und untersucht in *longgaps*, welche *gaps* größer als die run Länge sind. Mit der Funktion *tapply* bildet man die Maxima der Gruppen, die durch die *cumsum* gebildet werden. Die Positionen der Maxima werden mit *match* gesucht.

```

cluster.run <- function(serie,obere.grenze,run,plot=T)
{
  times <- (1:length(serie))/length(serie)
  daten <- signalSeries(serie,times)
  daten.exce <- daten[daten>obere.grenze]
  r <- run/length(serie)
  times2 <- positions(daten.exce)
  gaps <- as.numeric(diff(times2))
  longgaps <- gaps>r
  if(sum(longgaps)<=1.)
  stop("Der Cluster-Parameter s ist zu gross.")

  cluster.run.meth <- c(0.,cumsum(longgaps))
  cluster.max <- tapply(as.numeric(daten.exce),cluster.run.meth,max)
  newseries <- daten.exce[match(cluster.max,daten.exce)]
}

```

```

if(plot)
{
  cat("Declustering Graph...\n")
  cat(paste("Daten wurden reduziert von ", length(as.numeric(daten.exce)),
"auf", length(as.numeric(newseries)), "\n"))
  par(mfrow = c(2., 2.))
  plot(positions(daten.exce),as.numeric(daten.exce),type="h",xlab="Exzedenten",
ylab=" ",xlim=c(0,1),ylim=c(0,max(daten.exce)))
  qplot(gaps)
  plot(positions(newseries),as.numeric(newseries),type="h",xlab="Clustermaxima",
ylab=" ",xlim=c(0,1),ylim=c(0,max(daten.exce)))
  newgaps <- as.numeric(diff(positions(newseries)))
  qplot(newgaps)
  par(mfrow = c(1, 1))
}

```

```

newseries
}

```

Identifizierung der Cluster mit der Momenten Methode

Die Funktion *cluster.intervall* benutzt für die Identifizierung der Cluster die Momenten Methode. Für *serie* und *obere.grenze* setzt man wie immer den Datensatz und die obere Schwelle *u* ein. *plot=T* liefert diesleben Plots wie bei der Funktion *cluster.block*.

```

cluster.intervall <- function(serie,obere.grenze,plot=T)
{
  daten <- signalSeries(serie)
  daten.exce <- daten[daten>obere.grenze]
  times <- positions(daten.exce)
  gaps <- as.numeric(diff(times))
  N <- length(as.numeric(daten.exce))

  if (max(gaps) <= 2)
    {schatzer <- min(1,(2*(sum(gaps))^2)/((N-1)*(sum(gaps^2))))}
  else
    {schatzer <- min(1,(2*(sum(gaps)-(N-1))^2)/((N-1)*(sum(gaps^2)-sum(3*gaps)+2*(N-1))))}

  C.Wert <- trunc(schatzer * N) + 1

  cluster.run(serie,obere.grenze,C.Wert,plot=T)
}

```

Identifizierung der Cluster mit der Zwei-Schwellen Methode

Die Funktion *cluster.zwei.schwellen* identifiziert die Cluster mit der Zwei-Schwellen Methode. *serie* steht für den Datensatz, *obere.grenze* für die Schwelle *u*, *untere.grenze* für die zweite Schwelle *c* und *run* für die run Länge *s*. *plot=T* erzeugt wieder die Plots wie bei *cluster.block*. Die Funktion sucht sich die Exzedenten, berechnet die Interexzedentenzeiten in *gaps* und untersucht in *longgaps1*, welche *gaps* größer als die run Länge sind. In *longgaps2* werden die Daten geprüft, ob sie unter die Schwelle *c* fallen. *longgaps1+longgaps2* liefert also die Clusterenden. Damit keine Cluster doppelt gezählt werden, wird *longgaps* überprüft, ob eine 2 vorkommt und auf 1 zurückgesetzt. Das restliche Vorgehen ist wie bei der run Methode.

```

cluster.zwei.schwellen <- function(serie, obere.grenze, untere.grenze, run, plot=T)
{
  times <- (1:length(serie))/length(serie)
  daten <- signalSeries(serie, times)
  daten.exce <- daten[daten>obere.grenze]

  r <- run/length(serie)
  times2 <- positions(daten.exce)
  gaps <- as.numeric(diff(times2))
  longgaps1 <- gaps>r

  cluster.pos <- times2*length(serie)
  number <- length(longgaps1)
  longgaps2_rep(0, length(longgaps1))

  for(i in 1:number)
  {
    daten.neu <- as.numeric(daten[cluster.pos[i]:cluster.pos[i+1]])
    test <- daten.neu<(untere.grenze)
    if(sum(test)==0){longgaps2[i] <- 0}
    else {longgaps2[i] <- 1}
  }

  longgaps <- longgaps1+longgaps2
  for(i in 1:number)
  {
    if(longgaps[i]==2) {longgaps[i] <- 1}
  }

  if(sum(longgaps)<=1.)
  stop("Der Cluster-Parameter s ist zu gross oder die untere.grenze ist zu niedrig.")

  cluster.zwei.schw <- c(0., cumsum(longgaps))
  cluster.max <- tapply(as.numeric(daten.exce), cluster.zwei.schw, max)

  newseries <- daten.exce[match(cluster.max, daten.exce)]

  if(plot)
  {
    cat("Declustering Graph...\n")
    cat(paste("Daten wurden reduziert von ", length(as.numeric(daten.exce)),
    "auf", length(as.numeric(newseries)), "\n"))
    par(mfrow = c(2., 2.))
    plot(positions(daten.exce), as.numeric(daten.exce), type="h", xlab="Exzedenten",
    ylab=" ", xlim=c(0,1), ylim=c(0,max(daten.exce)))
    qplot(gaps)
    plot(positions(newseries), as.numeric(newseries), type="h", xlab="Clustermaxima",
    ylab=" ", xlim=c(0,1), ylim=c(0,max(daten.exce)))
    newgaps <- as.numeric(diff(positions(newseries)))
    qplot(newgaps)
    par(mfrow = c(1, 1))
  }
  newseries
}

```

Identifizierung der Cluster mit Volatilitäts-Schwellen Methode

cluster.vola funktioniert genauso wie *cluster.zwei.schwellen* nur, dass die zweite Bedingung nicht die zweite Schwelle *c* ist, sondern die Volatilitätsbedingung.

```
cluster.vola <- function(serie, obere.grenze, untere.grenze, run, plot=T)
{
  times <- (1:length(serie))/length(serie)
  daten <- signalSeries(serie, times)
  daten.exce <- daten[daten>obere.grenze]

  r <- run/length(serie)
  times2 <- positions(daten.exce)
  gaps <- as.numeric(diff(times2))
  longgaps1 <- gaps>r

  cluster.pos <- times2*length(serie)
  number <- length(longgaps1)
  longgaps2 <- rep(0, length(longgaps1))

  for(i in 1:number)
  {
    daten.neu <- as.numeric(daten[cluster.pos[i]:cluster.pos[i+1]])
    test <- stdev(daten.neu)<(untere.grenze)
    if(sum(test)==0){longgaps2[i] <- 0}
    else {longgaps2[i] <- 1}
  }

  longgaps <- longgaps1+longgaps2

  for(i in 1:number)
  {
    if(longgaps[i]==2) {longgaps[i] <- 1}
  }

  if(sum(longgaps)<=1.)
  stop("Der Cluster-Parameter s ist zu gross oder die untere.grenze ist zu niedrig.")

  cluster.zwei.schw <- c(0., cumsum(longgaps))
  cluster.max <- tapply(as.numeric(daten.exce), cluster.zwei.schw, max)
  newseries <- daten.exce[match(cluster.max, daten.exce)]

  if(plot)
  {
    cat("Declustering Graph...\n")
    cat(paste("Daten wurden reduziert von ", length(as.numeric(daten.exce)),
              "auf", length(as.numeric(newseries)), "\n"))
    par(mfrow = c(2., 2.))
    plot(positions(daten.exce), as.numeric(daten.exce), type="h", xlab="Exzedenten",
              ylab=" ", xlim=c(0,1), ylim=c(0,max(daten.exce)))
    qplot(gaps)
    plot(positions(newseries), as.numeric(newseries), type="h", xlab="Clustermaxima",
              ylab=" ", xlim=c(0,1), ylim=c(0,max(daten.exce)))
    newgaps <- as.numeric(diff(positions(newseries)))
    qplot(newgaps)
```

```

    par(mfrow = c(1, 1))
  }

newseries
}

```

A.2.5 Berechnung von θ mit den verschiedenen Methoden

Auch hier gilt, dass die Funktionen *theta.** nur dann ein richtiges Ergebnis liefern, falls alle Clustermaxima unterschiedlich sind, da diese Funktionen auf die *cluster.** zugreifen.

Schätzung von θ mit der Blockmethode

$\theta_n^{(Block)}$ wurde mit der voreingestellten Funktion *exindex* aus dem EVIS Paket berechnet.

Schätzung von θ mit der run Methode

Mit *theta.run* wird der Schätzer für θ mit Hilfe der run Methode (vgl. (5.10)) berechnet. Die Eingabewerte *serie*, *obere.grenze* und *run* sind wie immer der zu untersuchende Datensatz, der Schwellenwert *u* und die run Länge *s*.

```

theta.run_function(serie, obere.grenze, run)
{
  exce <- length(obere.grenze)

  for(n in 1:length(obere.grenze))
  {
    exce[n] <- length(serie[serie > obere.grenze[n]])
  }

  clus <- rep(0, length(obere.grenze))

  for(i in 1:length(obere.grenze))
  {
    cluster.best <- cluster.run(serie, obere.grenze[i], run, plot = F)
    clus[i] <- length(cluster.best)
  }

  theta.r <- rep(0, length(obere.grenze))

  for(j in 1:length(obere.grenze))
  {
    theta.r[j] <- clus[j]/exce[j]
  }

  theta.r
}

```

Schätzung von θ mit der Momenten Methode

theta.intervall berechnet $\hat{\theta}$ mit dem Intervallschätzer (6.9). Die Eingabewerte *serie* und *obere.grenze* sind der zu untersuchende Datensatz und die Schwelle *u*.

```

theta.intervall <- function(serie, obere.grenze)

```

```

{
  theta <- rep(0,length(obere.grenze))

  for (i in 1:length(obere.grenze))
  {
    daten <- signalSeries(serie)
    daten.exce <- daten[daten>obere.grenze[i]]
    times <- positions(daten.exce)
    gaps <- as.numeric(diff(times))
    N <- length(as.numeric(daten.exce))

    if (max(gaps) <= 2)
      {theta[i] <- min(1, (2*(sum(gaps))^2)/((N-1)*(sum(gaps^2))))}
    else
      {theta[i] <- min(1, (2*(sum(gaps)-(N-1))^2)/((N-1)*(sum(gaps^2)-sum(3*gaps)+2*(N-1))))}
  }
theta
}

```

Schätzung von θ mit der Zwei-Schwellen Methode

Die Funktion *theta.zwei.schwellen* schätzt θ mit der Zwei-Schwellen Methode (vgl. (7.5)). *serie*, *obere.grenze*, *untere.grenze* und *run* stehen für den zu untersuchende Datensatz, den Schwellenwert u , die zweite Schwelle c und die run Länge s .

```

theta.zwei.schwellen_function(serie, obere.grenze, untere.grenze, run)
{
  exce <- length(obere.grenze)

  for(n in 1:length(obere.grenze))
  {
    exce[n] <- length(serie[serie > obere.grenze[n]])
  }

  clus <- rep(0, length(obere.grenze))

  for(i in 1:length(obere.grenze))
  {
    cluster.best<-cluster.zwei.schwellen(serie,obere.grenze[i],untere.grenze,run,plot=F)
    clus[i] <- length(cluster.best)
  }

  theta.zwei <- rep(0, length(obere.grenze))

  for(j in 1:length(obere.grenze))
  {
    theta.zwei[j] <- clus[j]/exce[j]
  }
theta.zwei
}

```

Schätzung von θ mit der Volatilitäts-Schwellen Methode

Die Funktion `theta.vola` schätzt θ mit dem Volatilitätsschätzer (7.6). `serie`, `obere.grenze`, `untere.grenze` und `run` stehen für den zu untersuchende Datensatz, den Schwellenwert u , die Volatilitätsschwelle c und die run Länge s .

```
theta.vola_function(serie, obere.grenze, untere.grenze, run)
{
  exce <- length(obere.grenze)

  for(n in 1:length(obere.grenze))
  {
    exce[n] <- length(serie[serie > obere.grenze[n]])
  }

  clus <- rep(0, length(obere.grenze))

  for(i in 1:length(obere.grenze))
  {
    cluster.best<-cluster.vola(serie,obere.grenze[i],untere.grenze,run,plot=F)
    clus[i] <- length(cluster.best)
  }

  theta.vol <- rep(0, length(obere.grenze))

  for(j in 1:length(obere.grenze))
  {
    theta.vol[j] <- clus[j]/exce[j]
  }

  theta.vol
}
```

A.2.6 Fréchet-, Weibull- und Gumbel-Verteilung

Fréchet Verteilungsfunktion

Diese Funktion gibt für verschieden Werte von x die Verteilungsfunktion der Fréchet Verteilung wieder. Dabei steht *alpha* für α aus der Definition der Fréchet Verteilung (vgl. Satz 4.2).

```
frechet <- function(x,alpha,plot=T)
{
  frechet <- rep(0,length(x))

  for ( i in 1:length(x))
  {
    if (x[i]<=0) frechet[i] <- 0
    else frechet[i] <- exp(-(x[i])^(-alpha))*(x[i])^(-alpha-1)*alpha
  }

  if(plot=T)
  plot(frechet,type="l")
}
```

Weibull Verteilungsfunktion

Diese Funktion gibt für verschieden Werte von x die Verteilungsfunktion der Weibull Verteilung wieder. Dabei steht *alpha* wieder für α aus der Definition der Weibull Verteilung (vgl. Satz 4.2).

```
weibull <- function(x,alpha,plot=T)
{
  weibull <- rep(0,length(x))

  for ( i in 1:length(x))
  {
    if (x[i]<=0) weibull[i] <- exp(-(-x[i])^(alpha))*alpha*(-x[i])^(alpha-1)
    else weibull[i] <- 0
  }

  if(plot=T)
  plot(weibull,type="l")
}
```

Gumbel Verteilungsfunktion

Diese Funktion gibt für verschieden Werte von x die Verteilungsfunktion der Gumbel Verteilung wieder. Formel der Gumbel Verteilung vgl. Satz 4.2.

```
gumbel.vert <- function(x,plot=T)
{
  gum <- exp(-(exp(-x)))*exp(-x)
  if(plot=T)
  plot(gum,type="l")
}
```

Vergleich der drei Verteilungen

Diese Funktion ermöglicht es für verschiedene Werte von x die Verteilungsfunktion der Fréchet, Weibull und Gumbel Verteilung wieder. Dabei steht *alpha* für α aus der Definition der Fréchet und der Weibull Verteilung (vgl. Satz 4.2).

```
vergleich.vert <- function(x,alpha,plot=T)
{
  frechet <- rep(0,length(x))

  for ( i in 1:length(x))
  {
    if (x[i]<=0) frechet[i] <- 0
    else frechet[i] <- exp(-(x[i])^(-alpha))*(x[i])^(-alpha-1)*alpha
  }

  weibull <- rep(0,length(x))

  for ( i in 1:length(x))
  {

    if (x[i]<=0) weibull[i] <- exp(-(-x[i])^(alpha))*alpha*(-x[i])^(alpha-1)
```

```

    else weibull[i] <- 0
  }

gum <- exp(-(exp(-x))*exp(-x))

if(plot=T)
{
  plot(x,frechet,type="l",ylim=c(0,1),xlab=" ",ylab=" ")
  lines(x,weibull,col=5)
  lines(x,gum,col=6)
}
}

```

A.2.7 GPD

GPD-Verteilungsfunktion

Mit dieser Funktion kann für verschieden Werte von x die Verteilungsfunktion der GPD berechnet. Dabei steht *daten* für die Werte x , die man berechnen will, und ξ und β für die Parameter der GPD Verteilungsfunktion (vgl. (2.1)).

```

gpd.vert <- function(daten,xi,beta)
{
  g <- rep(0,length(daten))

  if(xi > 0)
  {
    for (i in 1:length(daten))
      {if (daten[i] > 0) g[i] <- 1-(1+xi*daten[i]/beta)^(-1/xi)
      else g[i] <- 0}
  }

  if(xi == 0)
  {
    for (i in 1:length(daten))
      {if (daten[i] > 0) g[i] <- 1-exp(-daten[i]/beta)
      else g[i] <- 0}
  }

  if(xi < 0)
  {
    for (i in 1:length(daten))
      {if (daten[i] < 0) g[i] <- 0
      if (daten[i] > 0) g[i] <- 1-(1+xi*daten[i]/beta)^(-1/xi)
      if (daten[i] > -(beta/xi)) g[i] <- 1}
  }

  g
}

```

GPD Dichte

Mit dieser Funktion kann für verschieden Werte von x die Dichte der GPD berechnet werden. Dabei steht *daten* für die Werte x , die man berechnen will, und ξ und β für die Parameter der GPD Dichte (vgl. (2.5)).

```

gpd.dichte <- function(daten,xi,beta)
{
  dgpd <- rep(0,length(daten))

  if(xi > 0)
  {
    for (i in 1:length(daten))
      dgpd[i] <- (1/beta)*(1+xi*daten[i]/beta)^(-(1/xi)-1)
  }

  if(xi == 0)
  {
    for (i in 1:length(daten))
      dgpd[i] <- (1/beta)*exp(-daten[i]/beta)
  }

  if(xi < 0)
  {
    for (i in 1:length(daten))
    {
      if (daten[i] < (-beta/xi))
        dgpd[i] <- (1/beta)*(1+xi*daten[i]/beta)^(-(1/xi)-1)
    }
  }

  dgpd
}

```


Abbildungsverzeichnis

2.1	<i>Verteilungsfunktion der GPD für drei verschiedene ξ und $\beta = 1$</i>	4
2.2	<i>Dichte der GPD für verschiedene ξ und $\beta = 1$</i>	4
3.1	<i>Die empirische Exzessfunktion von verschiedenen Standardverteilungen unter Vernachlässigung der Restdaten.</i>	10
3.2	<i>Die empirische Exzessfunktion simulierter Daten ($n = 10000$) (gepunktet) im Vergleich zur theoretischen Exzessfunktion (Linie) einer Standard Exponentialverteilung.</i>	11
3.3	<i>Die empirische Exzessfunktion simulierter Daten ($n = 10000$) (gepunktet) im Vergleich zur theoretischen Exzessfunktion (Linie) einer Gammaverteilung mit Parametern $\beta = 1$ und $\alpha > 0$.</i>	11
4.1	<i>Dichten der Fréchet, Weibull und Gumbel Verteilung. Bei Fréchet und Weibull wurde $\beta = 1$ gewählt.</i>	16
5.1	<i>Theoretisches Beispiel für die Blockmethode. Es wurde die Blockgröße $r=5$ gewählt und die Blöcke sind durch die verschiedenen Farben erkenntlich gemacht. Die Punkte kennzeichnen die Maxima, die oberhalb der Schwelle u liegen. Somit befindet sich im Block Nr. 1, 2, 4, 5 und 8 ein Cluster, die mit „*“ gekennzeichnet sind.</i>	22
5.2	<i>Theoretisches Beispiel für die run Methode mit der run Länge $s = 5$. Wie immer ist natürlich u der Schwellenwert. Es wurden die gleichen Daten wie bei Abbildung 5.1 verwendet. Die Punkte kennzeichnen den Beginn bzw. das Ende eines Clusters. Ein Cluster wird dadurch kenntlich gemacht, dass die Daten mit einer Linie verbunden sind. Also sind die Daten mit i von 3 - 4, 10, 17 - 25, 36 - ..., Cluster. Wie der Datenpunkt mit $i = 10$ zeigt, kann ein Cluster im Extremfall auch nur aus einem Datenpunkt bestehen.</i>	24

6.1	<p><i>Theoretisches Beispiel für die Momenten Methode. Die farbigen Linien seien die Cluster. Wenn man $C = 3$ wählt, dann sucht man jetzt die drittgrößte IEZ. $T_C = T_3 = 2$. Also sind T_2 und T_5 die zwei größten IEZ's, die T_C überschreiten. Demnach ist $i_1 = 2$ und $i_2 = 5$. Mit diesen Werten lassen sich dann die τ_j berechnen. $\tau_1 = \{T_{i_0+1}, \dots, T_{i_1-1}\} = \{T_{0+1}, \dots, T_{2-1}\} = \{T_1, \dots, T_1\} = T_1$ und $\tau_1 = \{T_{i_1+1}, \dots, T_{i_2-1}\} = \{T_3, \dots, T_4\}$. In der Praxis wird C nicht gewählt, sondern mit (6.11) berechnet. Dies ist dasselbe Ergebnis wie bei der run Methode mit run Länge $s = T_C = 2$.</i></p>	36
7.1	<p><i>Theoretisches Beispiel für die Zwei-Schwellen Methode mit der run Länge $s = 5$. u ist der obere Schwellenwert und c die zusätzliche untere Grenze. Es wurden die gleichen Daten wie bei Abbildung 5.1 und Abbildung 5.2 verwendet. Die Punkte kennzeichnen den Beginn bzw. das Ende eines Clusters. Ein Cluster wird dadurch kenntlich gemacht, dass die Daten mit einer Linie verbunden sind. Also sind die Daten mit i von 3 - 4, 10, 17 - 18, 21, 23 - 25, 36 - ..., Cluster. Wie der Datenpunkt mit $i = 10$ zeigt, kann ein Cluster im Extremfall auch nur aus einem Datenpunkt bestehen.</i></p>	40
8.1	<p><i>10000 Daten eines ARMAX(1) Prozesses mit $\alpha = 0.3$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_X^{(zwei-Schwellen)}$ (orange) für $s = 50$ und für c mit $F(c) = 0.75$ und $F(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet. $\theta = 0.7$ ist durch die horizontalen Linie kenntlich gemacht.</i></p>	47
8.2	<p><i>10000 Daten eines AR(1) Prozesses mit $\alpha = 0.3$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_X^{(zwei-Schwellen)}$ (orange) für $s = 50$ und für c mit $F(c) = 0.75$ und $F(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet.</i></p>	48
8.3	<p><i>10000 Daten eines ARCH(1) Prozesses mit $\lambda = 0.9$ und $\beta = 0.5$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_\sigma^{(Volatilität)}$ (orange) für $s = 50$ und für c mit $F_\sigma(c) = 0.75$ und $F_\sigma(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet. $\theta = 0.612$ ist durch die horizontalen Linie kenntlich gemacht.</i></p>	50
8.4	<p><i>10000 Daten eines GARCH(1,1) Prozesses mit $\lambda = 0.2, \gamma = 0.7$ und $\beta = 0.005$ simuliert. Es sind $\theta^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), $\theta_\sigma^{(Volatilität)}$ (orange) für $s = 50$ und für c mit $F_\sigma(c) = 0.75$ und $F_\sigma(c) = 0.25$ (von oben nach unten) und $\theta^{(Intervall)}$ (blau) abgebildet. $\theta = 0.612$ ist durch die horizontalen Linie kenntlich gemacht.</i></p>	51

9.1	$\hat{\theta}_n^{(run)}$ (schwarz) für $s = 3, 10$ und 50 (von oben nach unten) und $\hat{\theta}_{X,n}^{(zweischwellen)}$ (blau) für $s = 3$ und 10 (von oben nach unten) dargestellt.	55
9.2	$\hat{\theta}_{X,n}^{(zweischwellen)}$ abgebildet für $s = 10$ (schwarz) und für $s = 50$ (orange).	55
10.1	Tägliche Schlußdaten des S&P 500 - 1982 bis 2004	57
10.2	Negative Log Returns der täglichen Schlußdaten des S&P 500 - 1982 bis 2004.	58
10.3	Autokorrelationsfunktion und Histogramm der negativen Log Returns des S&P 500 - 1982 bis 2004.	59
10.4	Anpassung der negativen Log Returns des S&P 500 an eine GPD (Standard POT Methode). ξ und β wurden durch einen Maximum Likelihood Schätzer geschätzt und sind in der Tabelle 10.2 zu finden.	59
10.5	Die Exzessfunktion der positiven Daten der negativen Log Returns des S&P 500.	60
10.6	θ geschätzt für die negativen Log Returns des S&P 500 mit der run Methode (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), mit der Zwei-Schwellen Methode für $s = 50$ und c mit $F(c) = 0.75, 0.25$ (von oben nach unten) und der Momenten Methode. u entspricht dem 89%- bis 99.9%-Quantil des Datensatzes und ist in einer $-\log_{10}(\ln(F(u)))$ Skala abgebildet.	62
10.7	θ geschätzt für die negativen Log Returns des S&P 500 mit der Blockmethode. Dabei wurde die Blockgröße 3, 10 und 50 (von oben nach unten) gewählt.	63
10.8	Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 50$	64
10.9	Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$ -Quantil, $c = 5\%$ -Quantil und $s = 100$	65
10.10	Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$ -Quantil, $c = 5\%$ -Quantil und $s = 50$	66

10.11	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 5\%$-Quantil und $s = 100$.</i>	67
10.12	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 5\%$-Quantil und $s = 150$.</i>	68
10.13	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 25\%$-Quantil und $s = 100$.</i>	69
10.14	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 25\%$-Quantil und $s = 100$.</i>	70
10.15	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 5\%$-Quantil und $s = 10$.</i>	71
10.16	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 5\%$-Quantil und $s = 10$.</i>	72
10.17	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 80\%$-Quantil und $s = 50$.</i>	73
10.18	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des S&P 500 durch die Momentenmethode. Die gewählten Werte waren $u = 95\%$-Quantil (oberen zwei Graphen) und $u = 99\%$-Quantil (unteren zwei Graphen).</i>	74
10.19	<i>Tägliche Schlußdaten des FTSE 100 - 1984 bis 2004.</i>	75
10.20	<i>Negative Log Returns der täglichen Schlußdaten des FTSE 100.</i>	76
10.21	<i>Autokorrelationsfunktion und Histogramm der negativen Log Returns des FTSE 100</i>	76
10.22	<i>Anpassung der negativen Log Returns des FTSE 100 an eine GPD. ξ und β wurden durch einen Maximum Likelihood Schätzer geschätzt und sind in de Tabelle 10.6 zu finden.</i>	78

10.23	<i>Die Exzessfunktion der positiven Daten der negativen Log Returns des FTSE 100.</i>	79
10.24	<i>θ geschätzt für den Datensatz FTSE 100 mit der run Methode (schwarz) für $s = 3, 10$ und 50 (von oben nach unten), mit der Zwei-Schwellen Methode für $s = 50$ und c mit $F(c) = 0.75, 0.25$ (von oben nach unten) und der Momenten Methode. u entspricht dem 89%- bis 99.9%-Quantil des Datensatzes und ist in einer $-\log_{10}(\ln(F(u)))$ Skala abgebildet.</i>	79
10.25	<i>θ geschätzt für den Datensatz FTSE 100 mit der Blockmethode. Dabei wurde die Blockgröße $3, 10$ und 50 (von oben nach unten) gewählt.</i>	80
10.26	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 5\%$-Quantil und $s = 50$.</i>	81
10.27	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 5\%$-Quantil und $s = 100$.</i>	82
10.28	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 5\%$-Quantil und $s = 50$.</i>	83
10.29	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 5\%$-Quantil und $s = 100$.</i>	84
10.30	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 5\%$-Quantil und $s = 150$.</i>	85
10.31	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 25\%$-Quantil und $s = 100$.</i>	86
10.32	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 25\%$-Quantil und $s = 100$.</i>	87
10.33	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FTSE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 5\%$-Quantil und $s = 10$.</i>	88

10.34	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FT-SE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 99\%$-Quantil, $c = 5\%$-Quantil und $s = 10$.</i>	89
10.35	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FT-SE 100 durch die Block-, run, Zwei-Schwellen Methode. Die gewählten Werte waren $u = 95\%$-Quantil, $c = 80\%$-Quantil und $s = 50$.</i>	90
10.36	<i>Vergleich der Clusteridentifizierung bei den negativen Log Returns des FT-SE 100 durch Momenten Methode. Die gewählten Werte waren $u = 95\%$-Quantil (oberen zwei Graphen) und $u = 99\%$-Quantil (unteren zwei Graphen).</i>	91

Tabellenverzeichnis

3.1	<i>empirische Exzessfunktion von verschiedenen Standardverteilungen</i>	9
4.1	<i>Übersicht über die asymptotische Verteilung der Maxima</i>	16
9.1	<i>θ geschätzt mit der Momenten - Methode für verschiedene s und u</i>	54
9.2	<i>θ geschätzt mit der run Methode für verschiedene s und u</i>	54
9.3	<i>θ geschätzt mit der Zwei-Schwellen Methode für verschiedene s, u und c . .</i>	54
10.1	<i>Übersicht über die wichtigsten Kenngrößen der negativen Log Returns des S&P 500.</i>	58
10.2	<i>Der Schwellenwert u, der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Anzahl der Daten n, die Anzahl der Exzedenten N_u und die Maximum Likelihood Schätzer für ξ und β des Datensatzes S&P 500. (Standard POT Methode)</i>	60
10.3	<i>Der Schwellenwert u, der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Schwelle c, für die das 25%-Quantil genommen wurde, run Länge s, die Anzahl der Daten n, die Anzahl der Clustermaxima N_u^C, die Schätzer der run bzw. Zwei-Schwellen Methode und die Maximum Likelihood Schätzer für ξ und β des Datensatzes S&P 500 (POT Methode für stationäre Zeitreihen).</i>	60
10.4	<i>POT Methoden angewandt auf die negativen Log Returns des Datensatzes S&P 500.</i>	61
10.5	<i>Übersicht über die wichtigsten Kenngrößen der negativen Log Returns des FTSE 100.</i>	75

10.6	<i>Der Schwellenwert u, der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Anzahl der Daten n, die Anzahl der Exzedenten N_u und die Maximum Likelihood Schätzer für ξ und β des Datensatzes FTSE 100. (Standard POT Methode).</i>	77
10.7	<i>Der Schwellenwert u, der mit Hilfe der Exzessfunktion (vgl. Kapitel 3) ermittelt wurde, die Schwelle c, für die das 25%-Quantil genommen wurde, run Länge s, die Anzahl der Daten n, die Anzahl der Clustermaxima N_u^C, die Schätzer der run bzw. Zwei-Schwellen Methode und die Maximum Likelihood Schätzer für ξ und β des Datensatzes FTSE 100 (POT Methode für stationäre Zeitreihen).</i>	77
10.8	<i>POT Methoden angewandt auf die negativen Log Returns des Datensatzes FTSE 100.</i>	78

Literaturverzeichnis

- [1] Behnen, K. und Neuhaus, G. (1984) *Grundkurs Stochastik - Eine integrierte Einführung in Wahrscheinlichkeitstheorie und Mathematische Statistik*. Teubner Verlag Stuttgart
- [2] Borkovec, M. und Klüppelberg, C. (2000) *Extremwerttheorie für Finanzzeitreihen - ein unverzichtbares Werkzeug in Risikomanagement*. In: Rudolph, B. und Johanning, I. (Hrsg.) *Handbuch Risikomanagement*, 219 - 241, Uhlenbruch Verlag Bad Soden
- [3] Breidt, F.J und Davis, R.A. (1998) Extremes of Stochastic Volatility Models *The Annals of Applied Probability* **8**, 664 -675
- [4] de Haan, L., Resnick, S.I., Rootzén, H. und de Vries, C.G. (1989) Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Processes and their Applications* **32**, 231 - 224
- [5] Embrechts, P., Klüppelberg, C. und Mikosch, T. (2003) *Modelling Extremal Events for Insurance and Finance*. Springer Verlag
- [6] Emmer, S., Klüppelberg, C. und Trüstedt, M. (1998) VaR - ein Maß für das extreme Risiko. *Solutions* **1**
- [7] Ferro, C.A.T. und Segers, J. (2002) Automatic declustering of extreme values via an estimator for the extremel index. *Technical Report 2002-025*, EURANDOM (erhältlich unter www.eurandom.nl)
- [8] Ferro, C.A.T. und Segers, J. (2003) Inference for cluster of extreme values. *Journal of the Royal Statistical Society B* **65**, Part2, 545 - 556
- [9] Gençay, R., Selçuk, F., Ulugülyağci, A. (2001) *EVIM: A Software Package for Extreme Value Analysis in Matlab* (erhältlich unter <http://www.essex.ac.uk/ccfea/seminarpapers/evim.pdf>)
- [10] Hsing, T. (1991) Estimating the parameters of rare events. *Stochastic Processes and their Applications* **37**, 117 - 139
- [11] Hsing, T. (1993) Extremal Index Estimation for a Weakly Dependent Stationary Sequence. *The Annals of Statistics* **21**, No.4, 2043 - 2071
- [12] Informationen zum S&P 500 (erhältlich unter <http://www.standardandpoors.com> und <http://www.indices.standardandpoors.com>)

- [13] Königsberger, K. (1999) *Analysis 1* Springer Verlag
- [14] Laurini, F. und Tawn, J.A. (2003) New estimators for the extremal index and other cluster characteristics. *Extremes* **6**, Preprint Lancaster University and University of Padua, 189 - 211
- [15] Leadbetter, M.R. (1983) Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **65**, 291 - 306
- [16] Leadbetter, M.R. (1995) On high-level exceedance modeling and tail-inference. *Journal of Statistical Planning and Inference* **45**, 247 - 260
- [17] Mathematica Notebook *Extreme Value Distribution*. (erhältlich unter <http://mathworld.wolfram.com/ExtremeValueDistribution.html>)
- [18] McNeil, A.J. (1998) Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory. *Paper der ETH Zentrum Zürich*
- [19] McNeil, A.J. (2001) *EVIS, Version 4* (erhältlich unter <http://www.math.ethz.ch/mcneil> und <http://www.math.ethz.ch/mcneil/ftp/>)
- [20] Mikosch, T. und Stărică, C. (2000) Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process. *The Annals of Statistics* **28**, 1427 - 1451
- [21] O'Brien, G.L. (1987) Extreme Values for stationary and Markov sequences. *The Annals of Probability* **15**, 281 - 291
- [22] Smith, R.L. und Weisman, I. (1994) Estimating the Extremal Index. *Journal of the Royal Statistical Society B* **56**, No.3, 515 - 528
- [23] Weisman, I. und Novak, S.Yu. (1998) On blocks and runs estimators of the extremal index. *Journal of Statistical Planning and Inference* **66**, 281 - 288