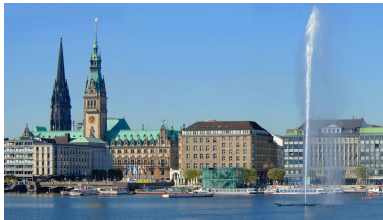


Quantitatives Risikomanagement 2

SS 2026

Univ.-Prof. Dr. Michael Merz
Universität Hamburg

Lehrstuhl für BWL, insb. Mathematik und Statistik in den Wirtschaftswissenschaften



Kapitel 7

Extremwerttheorie

Abschnitt 7.1

Einleitung

7. Extremwerttheorie

7.1 Einleitung

In Kapitel 4 wurden verschiedene **Schadenanzahl-** und **Schadenhöhenverteilungen** betrachtet, die zur stochastischen Modellierung eines breiten Spektrums von Risiken geeignet sind. Im darauffolgenden Kapitel 5 wurden dann stochastische Modelle für den Gesamtschaden eines Portfolios/Kollektivs von Einzelrisiken vorgestellt. Diese **Gesamtschadenmodelle** erlauben es unter der Annahme der stochastischen Unabhängigkeit, die Schadenanzahl und die Einzelschadenhöhen getrennt voneinander zu modellieren.

Bei der Anpassung einer Verteilung an die Schadenanzahl und die Einzelschadenhöhen werden i.d.R. **konventionelle Methoden**, wie die z.B. die Momentenmethode und die Maximum-Likelihood-Methode, verwendet. Eine hinreichend große Datenmenge vorausgesetzt, bieten diese Verfahren i.A. eine gute Anpassung an den gegebenen Datensatz. Allerdings basiert diese Anpassung vor allem auf **Beobachtungen im Zentrum der Verteilung**, was häufig zu einer schlechten Anpassung **im äußersten rechten Tail der Verteilung** führen kann, wo die sehr seltenen, aber potenziell katastrophalen Extremereignisse zu finden sind.

Selbst wenn die angepasste Verteilung insgesamt zufriedenstellend mit den gegebenen Daten übereinstimmt, erfasst sie den äußersten rechten Tail einer Verteilung, der außerhalb des Bereichs der verfügbaren Daten liegt, oftmals nicht adäquat. In diesen Fällen ist es häufig sinnvoll, die herkömmlichen Verfahren um Methoden aus der sog. **Extremwerttheorie** zu ergänzen.

7. Extremwerttheorie

7.1 Einleitung

Die **Extremwerttheorie** ist der Zweig der Statistik, der sich mit der Prognose **extrem-er Ereignisse** und der Schätzung ihrer Eintrittswahrscheinlichkeit beschäftigt. Extrem bedeutet hierbei, dass das Ereignis weitgehend oder sogar vollständig außerhalb des Bereichs bisheriger Beobachtungen liegt, wie es z.B. bei

- a) Rekorden,
- b) Maxima/Minima oder
- c) Überschreitungen großer Schwellenwerte

der Fall ist.

Ursprung: Am 1.2.1953 kam es in den Niederlanden in Folge eines Sturms zur Überflutung mehrerer Deiche und zu mehr als 1800 Toten. Die niederländische Regierung beauftragte daher eine Kommission mit der Untersuchung der benötigten Deichhöhe, so dass eine solche Überflutung im Durchschnitt nur alle 10.000 Jahre stattfindet.



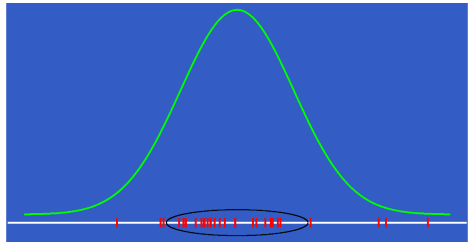
Problem: Nur Daten aus den letzten ca. 100 Jahren waren verfügbar!

7. Extremwerttheorie

7.1 Einleitung

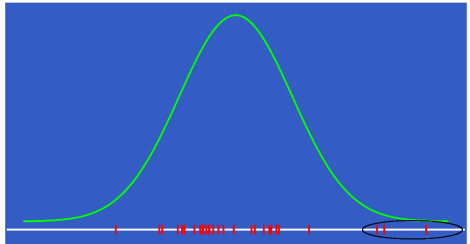
Klassische Statistik:

Untersuchung des **zentralen Bereichs** der Beobachtungen, d.h. durchschnittliche/übliche Ereignisse, wie z.B. Mittelwert, Median. Extreme Ereignisse werden oft als „Ausreißer“ bei der Analyse ausgeschlossen.



Extremwerttheorie:

Untersuchung des **dezentralen Bereichs** der Beobachtungen, d.h. extreme Ereignisse, wie z.B. Maxima oder Überschreitungen extremer Schwellenwerte. Der zentrale Bereich der Beobachtungen ist für die Analyse relativ wertlos.



7. Extremwerttheorie

7.1 Einleitung

Die Extremwerttheorie ist eine weit entwickelte Theorie mit Anwendungen z.B. in der

- **Finanzmathematik** (z.B. Wahrscheinlichkeit eines Börsencrashes),
- **Versicherungsmathematik** (z.B. extreme Schadenhöhen, Häufigkeit von Naturkatastrophen oder Man-made-Katastrophen),
- **Hydrologie** (z.B. ausreichende Dammhöhe, Höhe von Monsterwellen),
- **Meteorologie** (z.B. extreme Windstärken oder Regenfälle),
- **Umwelttechnik** (z.B. Grenzwerte bei Wasser- oder Luftverschmutzungen),
- **Qualitätskontrolle** (z.B. maximale Lebensdauer von Produkten) und
- **Sportwissenschaft** (z.B. Bestimmung der menschlichen Leistungsgrenzen)

Im Folgenden werden die beiden folgenden klassischen Extremwertmethoden und ihr Zusammenhang vorgestellt:

- a) **Block-Maxima-Methode:** Methode zur Modellierung der größten Beobachtung in einer großen Stichprobe von i.i.d. Beobachtungen.
- b) **Peak-over-Threshold-Methode:** Methode zur Modellierung von Beobachtungen, die einen hohen Threshold (Schwellenwert) in einer großen Stichprobe von i.i.d. Beobachtungen überschreiten.

7. Extremwerttheorie

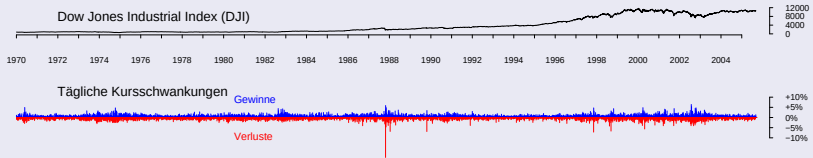
7.1 Einleitung

Beispiel (Börsencrash „Schwarzer Montag“)

Die untere Abbildung zeigt die zeitliche Entwicklung S_t des Dow Jones Index (wichtigster US-amerikanischer Aktienindex bestehend aus den 30 größten US-Unternehmen) und die zugehörigen Tagesrenditen (daily returns)

$$X_t = \frac{S_t - S_{t-1}}{S_{t-1}}$$

im Zeitraum von 1.1.1970 bis 31.10.2005.



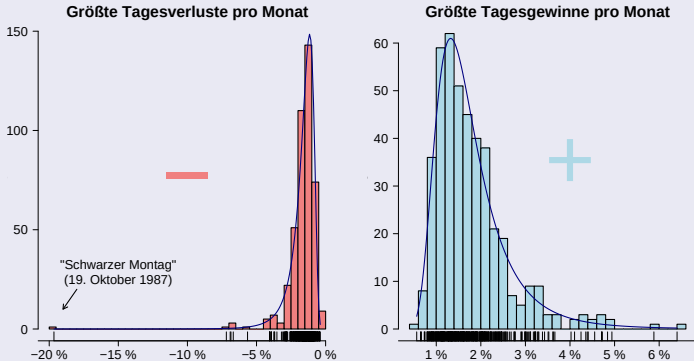
Gut zu erkennen ist der „Schwarze Montag“ am 19.10.1987, an dem der Dow Jones an einem Tag um 22,6% einbrach. Dieser Börsencrash stellt den größten prozentualen Einbruch innerhalb eines Tages in der Geschichte des Dow Jones Index dar.

7. Extremwerttheorie

7.1 Einleitung

Beispiel (Fortsetzung)

Werden die größten Tagesgewinne und Tagesverluste innerhalb eines Monats jeweils als extreme Ereignisse betrachtet, dann liegen die Daten für 430 Monate vor und man erhält für die Tagesgewinne und Tagesverluste die folgenden beiden Histogramme und angepassten Extremwertverteilungen:



7. Extremwerttheorie

7.1 Einleitung

Beispiel (Fortsetzung)

U.a. ist folgendes zu erkennen:

- Starke Kurssprünge nach unten sind extremer und treten häufiger auf als nach oben.
- Die angepassten Extremwertverteilungen weisen eine sehr gute Anpassung auf.
- Mit Hilfe der angepassten Extremwertverteilungen können z.B. Aussagen darüber gemacht werden, wie wahrscheinlich ein Tagesverlust von mehr als 5% im nächsten Monat ist oder mit welchem maximalen Tagesverlust man alle 10 Jahre rechnen muss.

Neben der **20-80-Regel** besitzt die Extremwerttheorie für die Versicherungsmathematik eine große Bedeutung, da

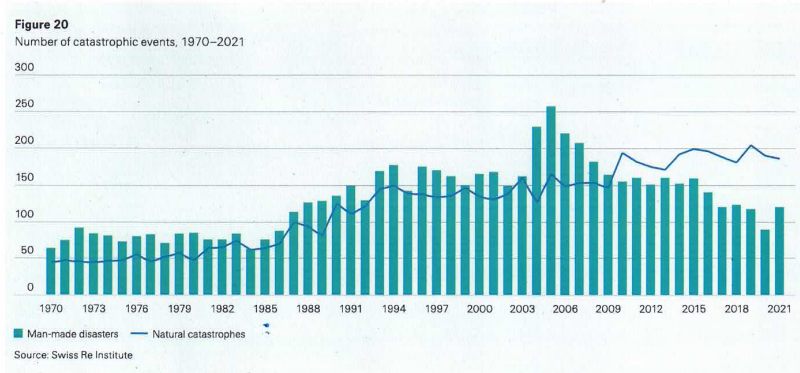
- a) die Anzahl an Natur- und Man-made-Katastrophen sowie
- b) die jährliche Gesamtschadenhöhe

aufgrund klimatischer Veränderungen, zunehmender Wertekonzentration in für Naturgefahren exponierten Lagen (z.B. Küsten/Flüssen/Gebirgen), Terrorismus, technischer Katastrophen (z.B. nukleare Katastrophen, Unglücke in der Luftfahrt, Seefahrt und im Bergbau) einen **ansteigenden Trend** aufweisen.

7. Extremwerttheorie

7.1 Einleitung

Anzahl Man-made- und Naturkatastrophen weltweit 1970 – 2021 (Quelle: Swiss Re)



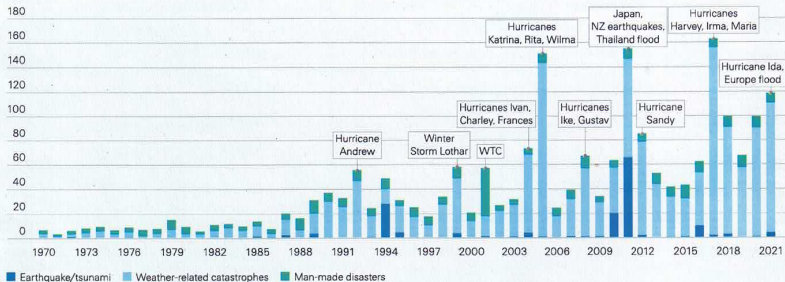
7. Extremwerttheorie

7.1 Einleitung

Von Naturkatastrophen verursachte versicherte Schäden in Mrd. US-Dollar weltweit
1970 – 2021 (Quelle: Swiss Re)

Figure 22

Insured catastrophe losses, 1970–2021, in USD billion at 2021 prices

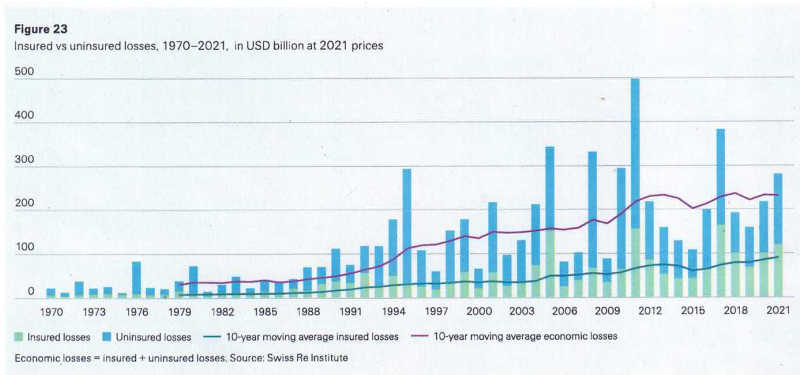


Source: Swiss Re Institute

7. Extremwerttheorie

7.1 Einleitung

Versicherte und nicht versicherte Schäden in Mrd. US-Dollar weltweit 1970 – 2021
(Quelle: Swiss Re)



Abschnitt 7.2

Wiederholung: Zentraler Grenzwertsatz

7. Extremwerttheorie

7.2 Wiederholung: Zentraler Grenzwertsatz

Zentraler Grenzwertsatz von Lindeberg-Lévy:

Die Verteilung der **standardisierten Summenvariable**

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\text{Var}(S_n)}} \quad \text{mit} \quad S_n := \sum_{i=1}^n X_i$$

mit

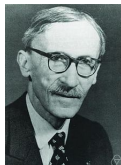
$$\text{i.i.d. } X_1, \dots, X_n \sim F_X \quad \text{und} \quad \text{Var}(X_i) < \infty$$

konvergiert für $n \rightarrow \infty$ gegen die **Standardnormal-Verteilung** Φ . D.h. es gilt

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\text{Var}(S_n)}} \leq x \right) = \Phi(x) \quad \text{für } x \in \mathbb{R}.$$

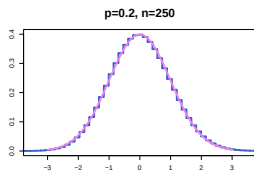
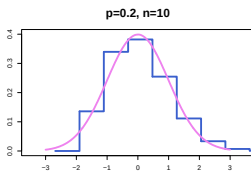
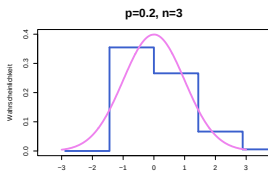
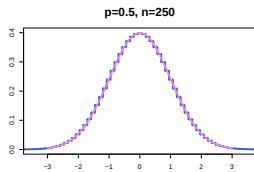
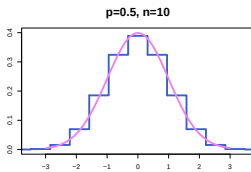
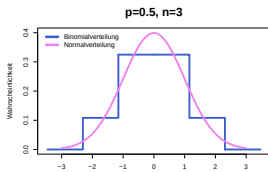
Speziell für $F_X = \text{Bernoulli}(p)$ mit $p \in (0, 1)$ erhält man daraus den **zentralen Grenzwertsatz von Moivre-Laplace**, d.h.

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \leq x \right) = \Phi(x) \quad \text{für } x \in \mathbb{R}.$$



7. Extremwerttheorie

7.2 Wiederholung: Zentraler Grenzwertsatz



Beobachtungen:

- Die Grenzverteilung Φ ist unabhängig von F_X .
- Der zentrale Grenzwertsatz rechtfertigt für die Summenvariable $S_n = \sum_{i=1}^n X_i$ bei großem n die Annahme einer Normalverteilung.
- Für die Formulierung des zentralen Grenzwertsatzes wird die Summenvariable S_n **standardisiert**.

Abschnitt 7.3

Verteilung des Maximums

7. Extremwerttheorie

7.3 Verteilung des Maximums

Es seien $X_1, X_2, \dots$ stochastisch unabhängige und identisch-verteilte Zufallsvariablen (z.B. Schäden, Verluste) mit der Verteilungsfunktion F_X . Im Folgenden werden mit

$$M_n := \max \{X_1, \dots, X_n\}$$

und

$$M_N := \max \{X_1, \dots, X_N\}$$

deren **Maximum** bei einer **deterministischen Anzahl n** bzw. einer **zufallsabhängigen Anzahl N** von Schäden/Verlusten X_i betrachtet.

Der folgende Satz macht eine Aussage über die **Verteilung** von M_n und M_N :

Satz (Verteilung von M_n und M_N)

Es seien $X_1, \dots, X_n$ unabhängige und identisch-verteilte Zufallsvariablen mit der Verteilungsfunktion F_X und N eine von $X_1, \dots, X_n$ stochastisch unabhängige $\mathbb{N}_0$ -wertige Zufallsvariable. Dann gilt:

a) $F_{M_n}(x) = F_X(x)^n$

b) $F_{M_N}(x) = g_N(F_X(x))$ mit $g_N(t) := \mathbb{E}[t^N] = \sum_{n=0}^{\infty} t^n \mathbb{P}(N = n)$ für $t \in [0, 1]$

7. Extremwerttheorie

7.3 Verteilung des Maximums

Beweis: Zu a): Es gilt:

$$\begin{aligned}F_{M_n}(x) &= \mathbb{P}(M_n \leq x) = \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) \\&= \prod_{j=1}^n \mathbb{P}(X_j \leq x) \\&= F_X(x)^n\end{aligned}$$

Zu b): Mit Aufgabenteil a) erhält man:

$$\begin{aligned}F_{M_N}(x) &= \mathbb{P}(M_N \leq x) = \mathbb{P}\left(\bigcup_{n=0}^{\infty} \{M_n \leq x \text{ und } N = n\}\right) \\&= \sum_{n=0}^{\infty} \mathbb{P}(M_n \leq x \text{ und } N = n) \\&= \sum_{n=0}^{\infty} \mathbb{P}(M_N \leq x | N = n) \mathbb{P}(N = n) \\&= \sum_{n=0}^{\infty} F_X(x)^n \mathbb{P}(N = n) = g_N(F_X(x))\end{aligned}$$

■

Der letzte Satz besagt, dass F_{M_n} und F_{M_N} **exakt berechnet** werden können, falls F_X und n bzw. F_X und F_N bekannt sind.

7. Extremwerttheorie

7.3 Verteilung des Maximums

Beispiel (Verteilungsfunktion von M_n mit $X_i \sim \text{Exp}(\theta)$)

Die Zufallsvariablen $X_1, \dots, X_n$ erfüllen die Voraussetzungen des letzten Satzes mit der Verteilungsfunktion $F_X(x) = 1 - e^{-\theta x}$ für $x > 0$ und $F_X(x) = 0$ für $x \leq 0$ mit $\theta > 0$. Dann besitzt das Maximum M_n nach Aussage a) des letzten Satzes die Verteilungsfunktion

$$\begin{aligned} F_{M_n}(x) &= (F_X(x))^n \\ &= (1 - e^{-\theta x})^n \end{aligned}$$

für $x \geq 0$.

7. Extremwerttheorie

7.3 Verteilung des Maximums

Beispiel (Verteilungsfunktion von M_N mit $X_i \sim \text{Exp}(\theta)$ und $N \sim \Pi(\lambda)$)

Die Zufallsvariablen $X_1, \dots, X_N$ und N erfüllen die Voraussetzungen des letzten Satzes mit der Verteilungsfunktion F_X des letzten Beispiels und der Wahrscheinlichkeitsfunktion $f_N(n) = e^{-\lambda} \frac{\lambda^n}{n!}$ für $n \in \mathbb{N}_0$ mit $\lambda > 0$. Dann folgt für $t \in [0, 1]$:

$$\begin{aligned} g_N(t) &= \mathbb{E}[t^N] = \sum_{n=0}^{\infty} t^n e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(t\lambda)^n}{n!} \\ &= e^{-\lambda} e^{t\lambda} = e^{-\lambda(1-t)} \end{aligned} \quad (1)$$

Zusammen mit Aussage b) des letzten Satzes erhält man damit für die Verteilungsfunktion von M_N für $x \geq 0$:

$$\begin{aligned} F_{M_N}(x) &= g_N(F_X(x)) = e^{-\lambda(1-F_X(x))} \\ &= e^{-\lambda e^{-\theta x}} \\ &= \exp\left(-\exp\left(-\frac{x - \frac{1}{\theta} \ln(\lambda)}{1/\theta}\right)\right) \end{aligned}$$

7. Extremwerttheorie

7.3 Verteilung des Maximums

Beispiel (Verteilungsfunktion von M_N mit $X_i \sim \text{Par}(\alpha, \theta)$ und $N \sim \Pi(\lambda)$)

Die Zufallsvariablen $X_1, \dots, X_n$ und N erfüllen die Voraussetzungen des letzten Satzes mit der Verteilungsfunktion $F_X(x) = 1 - \left(\frac{\theta}{\theta+x}\right)^\alpha$ für $x > 0$ und $F_X(x) = 0$ für $x \leq 0$ mit $\alpha, \theta > 0$ bzw. der Wahrscheinlichkeitsfunktion $f_N(n) = e^{-\lambda} \frac{\lambda^n}{n!}$ für $n \in \mathbb{N}_0$ mit $\lambda > 0$. Für die Verteilungsfunktion von M_N erhält man mit (1) für $x \geq 0$:

$$\begin{aligned} F_{M_N}(x) &= g_N(F_X(x)) = e^{-\lambda(1-F_X(x))} = \exp\left(-\lambda\left(\frac{\theta}{\theta+x}\right)^\alpha\right) \\ &= \exp\left(-\left(\frac{\theta+x}{\theta\lambda^{1/\alpha}}\right)^{-\alpha}\right) \\ &= \exp\left(-\left(1 + \frac{\theta+x-\theta\lambda^{1/\alpha}}{\theta\lambda^{1/\alpha}}\right)^{-\frac{1}{1/\alpha}}\right) \\ &= \exp\left(-\left(1 + \frac{1}{\alpha} \frac{x - (\theta\lambda^{1/\alpha} - \theta)}{\theta/\alpha\lambda^{1/\alpha}}\right)^{-\frac{1}{1/\alpha}}\right) \end{aligned}$$

7. Extremwerttheorie

7.3 Verteilung des Maximums

In der Regel ist jedoch F_X nicht bekannt, so dass eine exakte Berechnung von F_{M_n} **nicht möglich** ist. Der naheliegende Ansatz F_X durch eine Schätzung $\hat{F}_X$ zu ersetzen und dann F_{M_n} mittels

$$\hat{F}_{M_n} := \hat{F}_X^n$$

zu schätzen, ist jedoch ebenfalls nicht praktikabel, da bei diesem Ansatz die Schätzfehler bei der Schätzung von F_X durch $\hat{F}_X$ potenziert werden. In den meisten Fällen ist es daher besser für hinreichend große n mit Hilfe der extremen Beobachtungen für X direkt eine Approximation für F_{M_n} zu bestimmen.

Dabei ist jedoch folgendes zu beachten: Mit $z := \inf_{x \in \mathbb{R}} \{F_X(x) = 1\}$ folgt aus Aussage a) des letzten Satzes:

$$\lim_{n \rightarrow \infty} F_{M_n}(x) = \begin{cases} 0 & \text{für } x < z \\ 1 & \text{für } x \geq z \end{cases}$$

D.h. F_{M_n} konvergiert für $n \rightarrow \infty$ gegen eine **degenerierte Verteilungsfunktion** mit der Punktmasse 1 an der Stelle $x = z$. Also gegen die Verteilungsfunktion der degenerierten Zufallsvariable $X = z$ $\mathbb{P}$ -f.s.

7. Extremwerttheorie

7.3 Verteilung des Maximums

Aus diesem Grund wird analog zum klassischen Grenzwertsatz für die Summenvariable S_n (vgl. Abschnitt 7.2) nicht die Verteilungsfunktion von M_n für $n \rightarrow \infty$ approximiert, sondern die Verteilungsfunktion der **standardisierten Zufallsvariablen**

$$M_n^* = \frac{M_n - a_n}{b_n}$$

für geeignete Konstanten $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ und $(b_n)_{n \in \mathbb{N}} \subseteq (0, \infty)$ approximiert.

Abschnitt 7.4

Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippett

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippett

Für die Extremwerttheorie ist die **verallgemeinerte Extremwertverteilung** von zentraler Bedeutung:

Definition (Verallgemeinerte Extremwertverteilung)

Die verallgemeinerte Extremwertverteilung besitzt die Verteilungsfunktion

$$F_{\xi}(x; \mu, \sigma) := \begin{cases} \exp\left(-\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}\right) & \text{für } \xi \neq 0 \text{ und } 1 + \xi \frac{x-\mu}{\sigma} > 0 \\ \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right) & \text{für } \xi = 0 \text{ und } x \in \mathbb{R} \end{cases}.$$

Für $\xi > 0$ gilt definitionsgemäß $F_{\xi}(x; \mu, \sigma) = 0$ für $1 + \xi \frac{x-\mu}{\sigma} \leq 0$ und im Falle von $\xi < 0$ gilt definitionsgemäß $F_{\xi}(x; \mu, \sigma) = 1$ für $1 + \xi \frac{x-\mu}{\sigma} \leq 0$. Bei $\mu \in \mathbb{R}$ handelt es sich um einen Lokalisations- und bei $\sigma > 0$ um einen Skalierungsparameter.

Die Abbildung $\xi \mapsto F_{\xi}(x; \mu, \sigma)$ ist auch **an der Stelle $\xi = 0$ stetig**. D.h. es gilt

$$\lim_{\xi \rightarrow 0} F_{\xi}(x; \mu, \sigma) = F_0(x; \mu, \sigma).$$

Dies folgt aus

$$\lim_{\xi \rightarrow 0} (1 + \xi y)^{-1/\xi} = \lim_{\xi \rightarrow 0} \left(1 + \frac{y}{1/\xi}\right)^{-1/\xi} = e^{-y} \quad \text{für alle } y \in \mathbb{R}.$$

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Der **Shape-Parameter (Tail-Index)** ξ bestimmt den Typ der Extremwertverteilung:

$\xi = 0$ **Gumbel-Verteilung** mit nach oben und unten unbeschränktem Träger (benannt nach dem deutsch-amerikanischen Mathematiker E. J. GUMBEL (1891–1966)).



$\xi > 0$ **Fréchet-Verteilung** mit nach unten beschränktem und nach oben unbeschränktem Träger (benannt nach dem französischen Mathematiker M. R. FRÉCHET (1878–1973)).



$\xi < 0$ **(Verallgemeinerte) Weibull-Verteilung** mit nach unten unbeschränktem und nach oben beschränktem Träger. Diese Weibull-Verteilung entspricht im Wesentlichen einer verschobenen und gespiegelten Weibull-Verteilung gemäß Abschnitt 4.11.

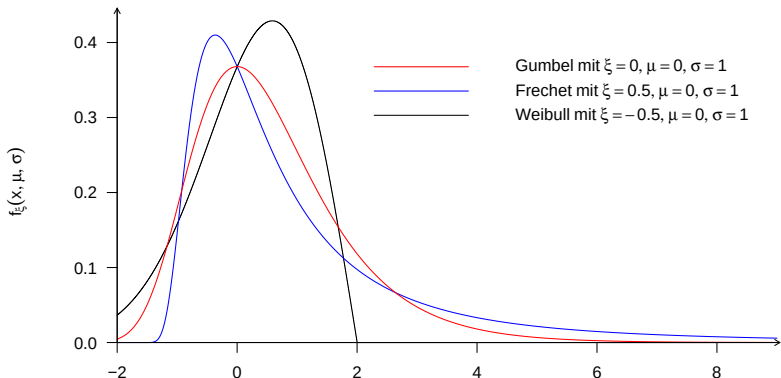


7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippett

Die **Dichte** der verallgemeinerten Extremwertverteilung ist gegeben durch:

$$f_{\xi}(x; \mu, \sigma) := \begin{cases} \frac{1}{\sigma} \exp \left(- \left(1 + \xi \frac{x-\mu}{\sigma} \right)^{-1/\xi} \right) \left(1 + \xi \frac{x-\mu}{\sigma} \right)^{-\frac{1+\xi}{\xi}} & \text{für } \xi \neq 0 \\ \frac{1}{\sigma} \exp \left(- \exp \left(- \frac{x-\mu}{\sigma} \right) - \frac{x-\mu}{\sigma} \right) & \text{für } \xi = 0 \end{cases} \quad (2)$$



7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Bemerkungen:

- Für $\mu = 0$ und $\sigma = 1$ erhält man die **standardisierte verallgemeinerte Extremwertverteilung**.
- Für $X \sim F_{\xi}(x; 0, 1)$ gilt $Y = \sigma X + \mu \sim F_{\xi}(y; \mu, \sigma)$.
- Für zwei Zufallsvariablen X und Y mit einer verallgemeinerten Extremwertverteilung und **identischem Shape-Parameter** ξ gibt es Konstanten $\sigma > 0$ und $\mu \in \mathbb{R}$, so dass gilt

$$Y \stackrel{d}{=} \sigma X + \mu.$$

- Die Weibull-Verteilung ist aufgrund ihres nach rechts beschränkten Trägers zur Modellierung von Maxima im Quantitativen Risikomanagement **oftmals ungeeignet**, da Einzelschadenhöhenverteilungen i.d.R. einen nach rechts unbeschränkten Träger besitzen. Sie kann jedoch von Nutzen sein, wenn Maxima von beschränkten Verlusten betrachtet werden.

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Bemerkungen (Fortsetzung):

- Für $\xi > 0$ gehört die verallgemeinerte Extremwertverteilung zur Klasse der **subexponentiellen Verteilungen** (vgl. Abschnitt 4.1). Grundsätzlich gilt: Umso größer $\xi > 0$ ist, desto mehr Maße liegt auf dem rechten Tail. Wie man zeigen kann, existiert das **k -te Moment** für $k < \frac{1}{\xi}$. D.h. es gilt

$$\mathbb{E}[X^k] = \infty \quad \text{für } k \geq \frac{1}{\xi}. \quad (3)$$

Eine Verteilungsfunktion F besitzt die **Max-stability-Eigenschaft**, wenn es für alle $n \in \mathbb{N}$ Konstanten $a_n \in \mathbb{R}$ und $b_n > 0$ gibt, so dass

$$F^n(x) = F(b_n x + a_n) \quad \text{für alle } x \in \mathbb{R}$$

gilt. D.h. die Verteilungsfunktion F ist **bzgl. Maximumsbildung abgeschlossen**.

Der folgende Satz besagt, dass die verallgemeinerte Extremwertverteilung die einzige Verteilung mit dieser Eigenschaft ist:

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Satz (Charakterisierung der Max-stability-Eigenschaft)

Eine Verteilungsfunktion F besitzt genau dann die Max-stability-Eigenschaft, wenn sie eine verallgemeinerte Extremwertverteilung ist.

Beweis: Ist F zum Beispiel eine Gumbel-Verteilung, dann gilt

$$F^n(x) = \exp\left(-n \exp\left(-\frac{x-\mu}{\sigma}\right)\right) = \exp\left(-\exp\left(-\frac{x-(\mu+\sigma \ln(n))}{\sigma}\right)\right).$$

Dies ist jedoch wieder eine Gumbel-Verteilung mit verändertem Lokalisationsparameter. Völlig analog zeigt man, dass auch die Fréchet- und Weibull-Verteilung die Max-stability-Eigenschaft besitzen. Für den Nachweis, dass eine Verteilungsfunktion mit der Max-stability-Eigenschaft stets eine verallgemeinerte Extremwertverteilung ist, werden Hilfsmittel aus der Funktionalanalysis benötigt (siehe z.B. LEADBETTER ET AL. (1983)). ■



Bei den in den letzten beiden Beispielen in Abschnitt 7.3 resultierenden Verteilungsfunktionen für M_N handelt es sich offensichtlich um eine **Gumbel-** bzw. **Fréchet-Verteilung** mit $\xi = \frac{1}{\alpha}$. Wie der folgende nach den englischen Statistikern R. A. FISHER (1890–1962) und L. H. C. TIPPETT (1902–1985) bezeichnete **Grenzwertsatz für Maxima** besagt, ist dies kein Zufall:



7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippett

Satz (Fisher-Tippett (1928))

Die Zufallsvariablen $X_1, \dots, X_n$ seien stochastisch unabhängig und identisch-verteilt und $M_n = \{X_1, \dots, X_n\}$. Ferner existieren Konstanten $b_n > 0$ und $a_n \in \mathbb{R}$ für alle $n \in \mathbb{N}$ sowie eine nicht-degenerierte Verteilungsfunktion G mit

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{M_n - a_n}{b_n} \leq x \right) = G(x) \quad (4)$$

für alle Stetigkeitsstellen x von G (d.h. $\left(\frac{M_n - a_n}{b_n} \right)_{n \in \mathbb{N}}$ konvergiert in Verteilung gegen eine Zufallsvariable mit der Verteilungsfunktion G). Dann ist G eine verallgemeinerte Extremwertverteilung $F_\xi(x; \mu, \sigma)$.

Beweisskizze: Für einen formalen Beweis siehe z.B. LEADBETTER ET AL. (1983). Betrachtet wird

$$\begin{aligned} M_{kn} &:= \max \{X_1, \dots, X_n, X_{n+1}, \dots, X_{2n}, \dots, X_{(k-1)n+1}, \dots, X_{kn}\} \\ &= \max \left\{ \max \{X_1, \dots, X_n\}, \max \{X_{n+1}, \dots, X_{2n}\}, \dots, \max \{X_{(k-1)n+1}, \dots, X_{kn}\} \right\}. \end{aligned}$$

Ferner sei angenommen, dass $\frac{M_n - a_n}{b_n}$ für geeignete Konstanten $b_n > 0$ und $a_n \in \mathbb{R}$ für $n \rightarrow \infty$ die nicht-degenerierte Grenzverteilung G besitzt. D.h. es gelte (4).

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Beweisskizze (Fortsetzung): Für ein hinreichend großes $n \in \mathbb{N}$ und alle $k \in \mathbb{N}$ folgt dann

$$\mathbb{P}(M_{kn} \leq x) = \mathbb{P}\left(\frac{M_{kn} - a_{nk}}{b_{nk}} \leq \frac{x - a_{nk}}{b_{nk}}\right) \approx G\left(\frac{x - a_{nk}}{b_{nk}}\right).$$

Ferner ist M_{kn} das Maximum von k Zufallsvariablen, die alle die gleiche Verteilung wie M_n besitzen. Mit Teil a) des Satzes in Abschnitt 7.3 folgt daher

$$\begin{aligned}\mathbb{P}(M_{kn} \leq x) &= \mathbb{P}\left(\frac{M_{kn} - a_n}{b_n} \leq \frac{x - a_n}{b_n}\right) \\ &= \left(\mathbb{P}\left(\frac{M_n - a_n}{b_n} \leq \frac{x - a_n}{b_n}\right)\right)^k \\ &\approx G^k\left(\frac{x - a_n}{b_n}\right).\end{aligned}$$

Die Verteilungen G und G^k stimmen somit bis auf einen Lokalisations- und Skalierungsparameter überein. D.h. die Verteilung G besitzt die Max-stability-Eigenschaft und ist daher gemäß des letzten Satzes eine verallgemeinerte Extremwertverteilung. ■

Der Satz von Fisher-Tippett ermöglicht die **Approximation** der Verteilung von M_n ohne Kenntnis der Verteilung F_X .

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Beispiel (Grenzverteilung von M_n mit $X_i \sim \text{Exp}(\lambda)$)

Die Zufallsvariablen $X_1, \dots, X_n$ seien stochastisch unabhängig und identisch-verteilt mit der Verteilungsfunktion

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{für } x > 0 \\ 0 & \text{sonst} \end{cases}$$

mit $\lambda > 0$. Mit Teil a) des Satzes in Abschnitt 7.3 und den Standardisierungskonstanten $a_n = \frac{1}{\lambda} \ln(n)$ und $b_n = 1$ folgt dann für $x \in \mathbb{R}$:

$$\begin{aligned} \mathbb{P}\left(\frac{M_n - a_n}{b_n} \leq x\right) &= \mathbb{P}(M_n \leq b_n x + a_n) \\ &= \mathbb{P}\left(M_n \leq x + \frac{1}{\lambda} \ln(n)\right) \\ &= \left(1 - e^{-\lambda(x + \frac{1}{\lambda} \ln(n))}\right)^n = \left(1 - \frac{e^{-\lambda x}}{n}\right)^n \end{aligned}$$

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Beispiel (Fortsetzung)

Es gilt somit:

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{M_n - a_n}{b_n} \leq x \right) &= \lim_{n \rightarrow \infty} \left(1 - \frac{e^{-\lambda x}}{n} \right)^n \\ &= e^{-e^{-\lambda x}} \\ &= F_0(x; 0, 1/\lambda)\end{aligned}$$

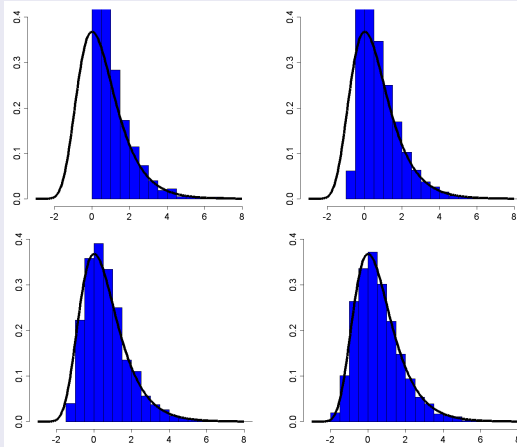
D.h. die Verteilung des standardisierten Maximums $\frac{M_n - a_n}{b_n}$ konvergiert für $n \rightarrow \infty$ gegen eine Gumbel-Verteilung mit den Parametern $\mu = 0$ und $\sigma = \frac{1}{\lambda}$.

Die folgende Abbildung zeigt die Konvergenz der simulierten Verteilung von $\frac{M_n - a_n}{b_n}$ für wachsendes $n = 1, 2, 5, 10$ und $\lambda = 1$ gegen die $F_0(x; 0, 1/\lambda)$ -Verteilung (d.h. Gumbel-Verteilung mit den Parametern $\mu = 0$ und $\sigma = \frac{1}{\lambda}$):

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Beispiel (Fortsetzung)



7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Definition (Maximum domain of attraction)

Der Maximum domain of attraction (MDA) einer verallgemeinerten Extremwertverteilung G ist die Menge bestehend aus allen Verteilungen F_X mit der Eigenschaft, dass für $M_n = \max \{X_1, \dots, X_n\}$ mit geeigneten Konstanten $a_n \in \mathbb{R}$ und $b_n > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{M_n - a_n}{b_n} \leq x \right) = G(x)$$

gilt (Kurzschreibweise: $F_X \in \text{MDA}(G)$).

Verteilungen F_X aus dem gleichen MDA besitzen somit als Grenzverteilung für das standardisierte Maximum die gleiche verallgemeinerte Extremwertverteilung G . Verteilungen (mit nicht-degenerierter Grenzverteilung) können folglich in **drei Klassen** eingeteilt werden:

- Gumbel-Verteilung als Grenzverteilung für das standardisierte Maximum
- Fréchet als Grenzverteilung für das standardisierte Maximum
- Weibull als Grenzverteilung für das standardisierte Maximum

Der folgende Satz besagt, dass die Zugehörigkeit von F_X zu einem bestimmten MDA ausschließlich durch das Verhalten des **rechten Tails** von F_X bestimmt wird:

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Satz (MDA bei Tail-Äquivalenz)

Zwei Verteilungen F_X und F_Y gehören genau dann zum gleichen Maximum domain of attraction (MDA), wenn sie Tail-äquivalent sind, d.h. wenn es eine Konstante $0 < c < \infty$ gibt mit

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_X(x)}{\bar{F}_Y(x)} = c.$$

Beweis: Siehe z.B. EMBRECHTS ET AL. (2003), Seiten 132–133. ■

Beispiel (Tailverhalten einer Exponential- und Pareto-Verteilung)

- a) Eine $\text{Exp}(\lambda)$ -Verteilung besitzt die Verteilungsfunktion $F_X(x) = 1 - e^{-\lambda x}$ für $x > 0$. Es gilt somit

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_X(x)}{re^{-\lambda x}} = \lim_{x \rightarrow \infty} \frac{e^{-\lambda x}}{re^{-\lambda x}} = \frac{1}{r}.$$

D.h. Verteilungsfunktionen mit einem rechten Tail der asymptotischen Form $re^{-\lambda x}$ gehören zum MDA einer Gumbel-Verteilung.

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Beispiel (Fortsetzung)

- b) Eine $\text{Par}(\alpha, \lambda)$ -Verteilung hat die Verteilungsfunktion $F_X(x) = 1 - \left(\frac{\lambda}{\lambda+x}\right)^\alpha$ für $x > 0$. Es gilt somit

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_X(x)}{rx^{-\alpha}} = \lim_{x \rightarrow \infty} \frac{\left(\frac{\lambda+x}{\lambda}\right)^{-\alpha}}{rx^{-\alpha}} = \lim_{x \rightarrow \infty} \frac{\left(1 + \frac{x}{\lambda}\right)^{-\alpha}}{rx^{-\alpha}} = \lim_{x \rightarrow \infty} \frac{\left(\frac{1}{x} + \frac{1}{\lambda}\right)^{-\alpha}}{r} = \frac{\lambda^\alpha}{r}.$$

D.h. Verteilungsfunktionen mit einem rechten Tail der asymptotischen Form $rx^{-\alpha}$ gehören zum MDA einer Fréchet-Verteilung.

Bemerkungen:

- Die Fréchet-, Pareto-, Cauchy-, Burr-, Student-, Loggamma- und F -Verteilung sowie stable distributions (mit Exponent $\alpha < 2$) gehören zum MDA einer **Fréchet-Verteilung**. Wegen (3) kann eine Verteilung, deren Momente alle existieren, nicht zum MDA einer Fréchet-Verteilung gehören.
- Zum MDA einer **Weibull-Verteilung** gehören neben der Weibull-Verteilung z.B. die Beta- und Gleichverteilung.

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippet

Bemerkungen (Fortsetzung):

- Die Gumbel-, Exponential-, Normal-, Gamma-, χ^2 -, Lognormal-, Weibull- und Inverse Gauss-Verteilung gehören zum MDA einer **Gumbel-Verteilung**. Da die Momente einer Gumbel-Verteilung alle existieren, besitzen Verteilungen im MDA einer Gumbel-Verteilung Momente beliebiger Ordnung. Dennoch weisen Verteilungen im MDA einer Gumbel-Verteilung zum Teil sehr unterschiedliches Tailverhalten auf. Zum Beispiel hat die Inverse Gauss-Verteilung einen deutlich schwereren rechten Tail als die Normalverteilung.

In manchen Fällen können mit Hilfe des folgenden Satzes die **Standardisierungskonstanten** $a_n \in \mathbb{R}$ und $b_n > 0$ in (4) berechnet werden:

Satz (Charakterisierung des MDA)

Eine Verteilungsfunktion F_X gehört zum Maximum of attraction (MDA) von $F_\xi(x; 0, 1)$ genau dann, wenn gilt:

$$\lim_{n \rightarrow \infty} n(1 - F_X(b_n x + a_n)) = -\ln(F_\xi(x; 0, 1)) = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{für } \xi \neq 0 \\ \exp(-x) & \text{für } \xi = 0 \end{cases}$$

Beweis: Siehe z.B. EMBRECHTS ET AL. (2003), Seite 128.

7. Extremwerttheorie

7.4 Verallgemeinerte Extremwertverteilung und Satz von Fisher-Tippett

Für ein Beispiel zur Anwendung dieses Satzes siehe Übungsaufgaben.

Für die Anwendung des Satzes von Fisher-Tippett ist jedoch die Kenntnis geeigneter Standardisierungskonstanten a_n und b_n **nicht erforderlich**. Denn für hinreichend große $n \in \mathbb{N}$ gilt nämlich

$$\mathbb{P}\left(\frac{M_n - a_n}{b_n} \leq x\right) \approx F_\xi(x; \mu, \sigma) \implies \mathbb{P}(M_n \leq x) \approx F_\xi(x; \tilde{\mu}, \tilde{\sigma})$$

mit $\tilde{\mu} := b_n \mu + a_n$ und $\tilde{\sigma} := b_n \sigma$ (vgl. Übungsaufgaben). Da jedoch bei einer konkreten Anwendung der Lokalisations- und Skalenparameter der verallgemeinerten Extremwertverteilung sowieso aus den Daten geschätzt werden müssen, spielt es bei der Anpassung einer verallgemeinerten Extremwertverteilung an einen gegebenen Datensatz keine Rolle, wenn die Konstanten a_n und b_n nicht bekannt sind.

Abschnitt 7.5

Block-Maxima-Methode

7. Extremwerttheorie

7.5 Block-Maxima-Methode

Gegeben seien n stochastisch unabhängige und identisch-verteilte Zufallsvariablen

$$X_1, \dots, X_n \sim F_X \in \text{MDA}(F_\xi(x; \mu, \sigma)). \quad (5)$$

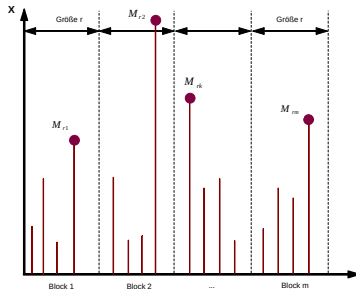
Bei der **Block-Maxima-Methode** werden zur Anpassung einer verallgemeinerten Extremwertverteilung an die Grenzverteilung G von

$$M_n = \max \{X_1, \dots, X_n\} \quad \text{für } n \gg 1$$

die Beobachtungen (5) in m Blöcke der gleichen Größe r unterteilt (z.B. Jahres-, Monats-, Quartals-Daten) und die zugehörigen m **Block-Maxima**

$$M_{r1}, \dots, M_{rm} \quad (6)$$

ermittelt.



7. Extremwerttheorie

7.5 Block-Maxima-Methode

Sind

$$z_1, \dots, z_m \quad (7)$$

die **Realisierungen** der m Block-Maxima (6) und ist $f_\xi(z; \mu, \sigma)$ die Dichte der verallgemeinerten Extremwertverteilung $F_\xi(z; \mu, \sigma)$ (vgl. (2)), dann erhält man nach kurzer Umformung für die **Log-Likelihoodfunktion** der m Beobachtungen (7) den Ausdruck

$$\begin{aligned} \ln(L(\xi, \mu, \sigma; z_1, \dots, z_m)) &= \sum_{k=1}^m \ln(f_\xi(z_k; \mu, \sigma)) \\ &= \begin{cases} -m \ln(\sigma) - (1 + \xi) \sum_{k=1}^m y_k - \sum_{k=1}^m \exp(-y_k) & \text{für } \xi \neq 0 \\ -m \ln(\sigma) - \sum_{k=1}^m y_k - \sum_{k=1}^m \exp(-y_k) & \text{für } \xi = 0 \end{cases} \end{aligned} \quad (8)$$

für $z_k \in \mathbb{R}$ mit $1 + \xi \left(\frac{z_k - \mu}{\sigma} \right) > 0$ und

$$y_k := \begin{cases} \frac{1}{\xi} \ln \left(1 + \xi \frac{z_k - \mu}{\sigma} \right) & \text{für } \xi \neq 0 \\ \frac{z_k - \mu}{\sigma} & \text{für } \xi = 0 \end{cases}. \quad (9)$$

Durch Maximierung von (8) mittels numerischer Methoden (z.B. Newton-Raphson) erhält man **ML-Schätzer** $\hat{\xi}, \hat{\mu}, \hat{\sigma}$ für ξ, μ, σ .

Bemerkungen:

- Bei der Wahl der Blockgröße r und der Blockanzahl m besteht ein **Zielkonflikt**:
 - Ein großes r (d.h. große Blöcke) führt zu einer **besseren Approximation** der Verteilung von M_n durch eine verallgemeinerte Extremwertverteilung. Andererseits stehen dann aber u.U. nur wenige Block-Maxima für die ML-Schätzung von ξ, μ und σ zur Verfügung.
 - Ein großes m (d.h. viele Blöcke) führt zu vielen Block-Maxima und somit zu einer **geringeren Varianz** bei der ML-Schätzung von ξ, μ und σ . Die verallgemeinerte Extremwertverteilung liefert dann aber u.U. keine zufriedenstellende Approximation für die Verteilung von M_n .

Eine **pragmatische Lösung** besteht oftmals darin, Blöcke zu verwenden, die mit einem Jahr oder Quartal korrespondieren.

- Bei der Maximierung von (8) müssen die **Nebenbedingungen** $\sigma > 0$ und $1 + \xi \frac{z_k - \mu}{\sigma} > 0$ berücksichtigt werden, weshalb die gewöhnlichen **Regularitätsbedingungen**, welche die guten Eigenschaften von ML-Schätzern sicherstellen, verletzt sind. Es kann jedoch gezeigt werden, dass die ML-Schätzer für $\xi > -0,5$ **konsistent** und **asymptotisch Normal-verteilt** sind. Dieser Fall deckt die meisten finanz- und versicherungswirtschaftlichen Anwendungen ab.

Bemerkungen (Fortsetzung):

- Die Einteilung der vorhandenen Beobachtungen in m Blöcke ergibt sich bei **Zeitreihendaten** (z.B. Aktienpreise, Wasserstände) auf natürliche Weise. Wenn die Beobachtungen keine natürliche Ordnung aufweisen, können sie **zufällig** in Blöcke aufgeteilt werden.
- Häufig werden **grafische Methoden** oder die **Momentenmethode** zur Berechnung von geeigneten Startwerten für die ML-Methode verwendet.
- Die Standardabweichungen der ML-Schätzungen für die Parameter μ , σ und ξ sind selbst bei umfangreichen Datensätzen **häufig relativ groß**. Dies gilt besonders für die ML-Schätzung des Shape-Parameters ξ . Dies ist problematisch, da die ML-Schätzung für den Parameter ξ angibt, welche der drei verallgemeinerten Extremwertverteilungen am besten geeignet ist.

7. Extremwerttheorie

7.5 Block-Maxima-Methode

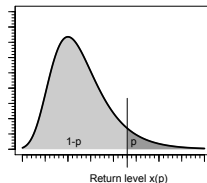
Wiederkehrwert und Wiederkehrperiode

Mit Hilfe der angepassten verallgemeinerten Extremwertverteilung $F_{\hat{\xi}}(x; \hat{\mu}, \hat{\sigma})$ kann

- a) für eine vorgegebene Überschreitungswahrscheinlichkeit $p \in (0, 1)$ der zugehörige **Wiederkehrwert (return level)** $\hat{x}(p)$ bzw.
- b) für einen vorgegebenen Wiederkehrwert $\hat{x}(p)$ die zugehörige **Wiederkehrperiode (return period)** T

ermittelt werden. Dies erfolgt durch Auflösen der Gleichung

$$\mathbb{P}(M_n > \hat{x}(p)) \approx 1 - F_{\hat{\xi}}(\hat{x}(p); \hat{\mu}, \hat{\sigma}) = p$$



nach $\hat{x}(p)$. Man erhält dann:

$$\hat{x}(p) = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left(1 - [-\ln(1-p)]^{-\hat{\xi}} \right) & \text{für } \hat{\xi} \neq 0 \\ \hat{\mu} - \hat{\sigma} \ln[-\ln(1-p)] & \text{für } \hat{\xi} = 0 \end{cases} \quad (10)$$

Beim Wiederkehrwert $\hat{x}(p)$ handelt es sich um eine ML-Schätzung für den Wert, der vom Maximum mit der Wahrscheinlichkeit p überschritten, d.h. der durchschnittlich alle $T = \frac{1}{p}$ Zeiteinheiten übertroffen, wird.

7. Extremwerttheorie

7.5 Block-Maxima-Methode

Mit der **Delta-Methode** zur Normalverteilungsapproximation differenzierbarer Transformationen asymptotisch normalverteilter Zufallsvektoren kann ein Schätzer für die Varianz des Quantil-Schätzers $\hat{x}(p)$ ermittelt werden, der eine differenzierbare Funktion des asymptotisch normalverteilten Zufallsvektors $(\hat{\xi}, \hat{\mu}, \hat{\sigma})$ ist (vgl. Abschnitt 8.3).

Man erhält dann

$$\widehat{\text{Var}}(\hat{x}(p)) \approx \text{grad } \hat{x}(p)^T \Sigma \text{ grad } \hat{x}(p), \quad (11)$$

wobei Σ die **(asymptotische) Varianz-Kovarianz-Matrix** von $(\hat{\xi}, \hat{\mu}, \hat{\sigma})$ und

$$\begin{aligned} \text{grad } \hat{x}(p) &= \begin{pmatrix} \frac{\partial \hat{x}(p)}{\partial \hat{\xi}} \\ \frac{\partial \hat{x}(p)}{\partial \hat{\mu}} \\ \frac{\partial \hat{x}(p)}{\partial \hat{\sigma}} \end{pmatrix} \\ &= \begin{pmatrix} \hat{\sigma} \hat{\xi}^{-2} (1 - [-\ln(1-p)]^{-\hat{\xi}}) - \hat{\sigma} \hat{\xi}^{-1} [-\ln(1-p)]^{-\hat{\xi}} \ln(-\ln(1-p)) \\ 1 \\ -\hat{\xi}^{-1} (1 - [-\ln(1-p)]^{-\hat{\xi}}) \end{pmatrix} \end{aligned} \quad (12)$$

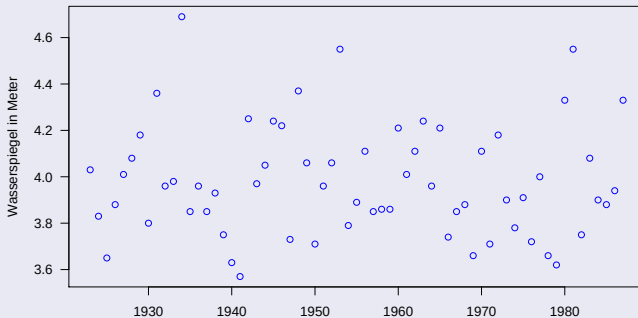
der **Gradient** von $\hat{x}(p)$ ist.

7. Extremwerttheorie

7.5 Block-Maxima-Methode

Beispiel (Anwendung der Block-Maxima-Methode)

Die untenstehende Abbildung zeigt die jährlichen Maxima für den Wasserspiegel bei Port Pirie (Süd-Australien) im Zeitraum 1923–1987. Es ist kein zeitlicher Trend zu erkennen. Im Folgenden wird eine ML-Schätzung für den maximalen Wasserspiegel berechnet, der alle 10 Jahre bzw. alle 100 Jahre (Jahrhundertereignis) auftreten kann, unter der Voraussetzung, dass keine Klimaveränderungen oder andere Veränderungen stattfinden (Zeitstabilitätshypothese).



7. Extremwerttheorie

7.5 Block-Maxima-Methode

Beispiel (Fortsetzung)

Hierzu wird an die 65 Jahres-Maxima eine verallgemeinerte Extremwertverteilung $F_{\xi}(x; \mu, \sigma)$ angepasst. Dies liefert für die Modellparameter μ, σ, ξ die folgenden ML-Schätzungen und Standardabweichungen:

Parameter	ML-Schätzung	Standardabweichung	
ξ	-0,050	0,09825633	nicht signifikant
μ	3,87	0,02793211	hoch signifikant
σ	0,198	0,02024610	hoch signifikant

Die Parameterschätzung für ξ unterscheidet sich nicht signifikant von 0. Für ξ kann daher die ML-Schätzung

$$\hat{\xi} = 0 \text{ (Gumbel)} \quad \text{oder} \quad \hat{\xi} = -0,050 \text{ (Weibull)}$$

verwendet werden.

Beispiel (Fortsetzung)

Mit $\hat{\xi} = -0,050$ resultiert die Weibull-Verteilung

$$\hat{F}_{-0,050}(x; 3,87, 0,198) = \begin{cases} \exp\left(-\left(1 - 0,050 \cdot \frac{x-3,87}{0,198}\right)^{-1/-0,050}\right) & \text{für } x < 7,83 \\ 1 & \text{für } x \geq 7,83 \end{cases}.$$

Für die 10-Jahres und die 100-Jahres Wiederkehrwerte erhält man mit (10) die ML-Schätzungen

$$\hat{x}(0,1) = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left(1 - [-\ln(1 - 0,1)]^{-\hat{\xi}}\right) = 4,30$$

bzw.

$$\hat{x}(0,01) = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left(1 - [-\ln(1 - 0,01)]^{-\hat{\xi}}\right) = 4,69$$

Beispiel (Fortsetzung)

Mit (11) erhält man für die Varianzen dieser geschätzten Wiederkehrwerte

$$\widehat{\text{Var}}(\hat{x}(0,1)) = 0,00303$$

bzw.

$$\widehat{\text{Var}}(\hat{x}(0,01)) = 0,02502.$$

Damit erhält man für die ML-Schätzungen $\hat{x}(0,1)$ und $\hat{x}(0,01)$ die approximativen 95%-Konfidenzintervalle

$$4,30 \pm 1,96\sqrt{0,00303} = [4,19; 4,41]$$

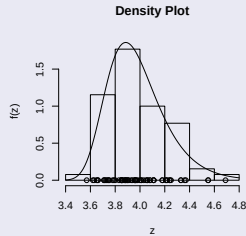
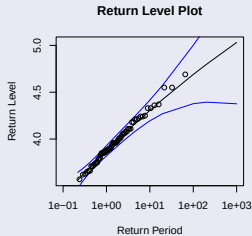
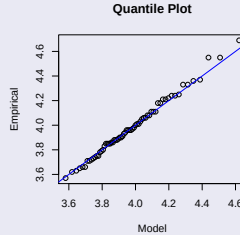
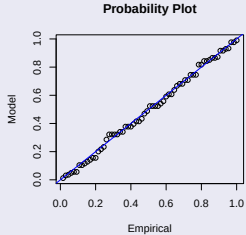
bzw.

$$4,69 \pm 1,96\sqrt{0,02502} = [4,38; 5,00].$$

7. Extremwerttheorie

7.5 Block-Maxima-Methode

Beispiel (Fortsetzung)

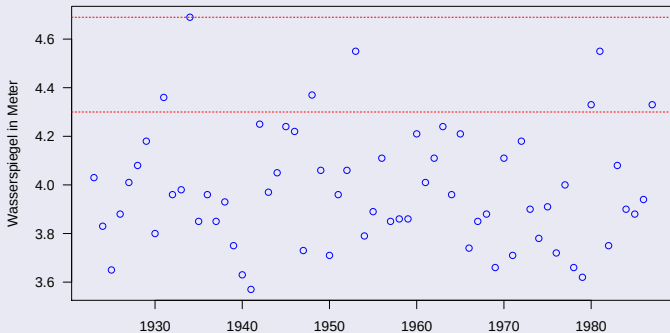


7. Extremwerttheorie

7.5 Block-Maxima-Methode

Beispiel (Fortsetzung)

Es ist zu erkennen, dass der Probability Plot und der Quantile Plot nahezu linear verlaufen. Ferner ist die Dichte der angepassten verallgemeinerten Extremwertverteilung $F_{-0,050}(x; 3,87, 0,198)$ konsistent zum Histogramm mit den Beobachtungen. Dies sind Indizien dafür, dass die angepasste verallgemeinerte Extremwertverteilung einen guten Fit für die 65 Jahres-Maxima liefert.



7. Extremwerttheorie

7.5 Block-Maxima-Methode

Beispiel (Fortsetzung)

R-Code:

```
1  ###
2  # Anpassung einer verallgemeinerten EWV an den Wasserspiegel bei Port Pirie mittels
3  # Block-Maxima-Methode
4  ###
5
6  library(ismev)
7  data(portpirie)
8  par(mfrow=c(1,1),cex=2,mar=c(2,5,1,1))
9  plot(portpirie, ylab="Wasserspiegel in Meter", xlab="Jahr", col=4,las=1)
10
11  fit<-gev.fit(portpirie[,2])
12  fit$mle
13  fit$se
14  fit$cov
15  gev.diag(fit)
16  par(mfrow=c(1,1),cex=2,mar=c(2,5,1,1))
17  plot(portpirie, ylab="Wasserspiegel in Meter", xlab="Jahr", col=4,las=1)
18  abline(h=c(4.30,4.69), col=2, lty=2)
```

Bemerkung: Mit Hilfe eines **Likelihood-Quotienten-Tests** kann untersucht werden, ob $\xi = 0$ gilt. Dieser Test basiert auf der Teststatistik

$$\begin{aligned} L &:= -2 \ln \left(\frac{L(\xi, \hat{\mu}, \hat{\sigma}; z_1, \dots, z_m)}{L(\hat{\xi}, \hat{\mu}, \hat{\sigma}; z_1, \dots, z_m)} \right) \\ &= -2 \left(\ln(L(\xi, \hat{\mu}, \hat{\sigma}; z_1, \dots, z_m)) - \ln(L(\hat{\xi}, \hat{\mu}, \hat{\sigma}; z_1, \dots, z_m)) \right), \end{aligned}$$

welche unter der Nullhypothese

$$H_0 : \xi = 0$$

asymptotisch χ^2 -verteilt ist mit einem Freiheitsgrad. Zum Beispiel beträgt das 95%-Quantil einer χ^2_1 -Verteilung 3,84. Bei einem Signifikanzniveau von $\alpha = 5\%$ würde man daher $H_0 : \xi = 0$ zugunsten von $H_1 : \xi \neq 0$ verwerfen, wenn

$$L > 3,84$$

gilt.

Abschnitt 7.6

Verteilung des Excess

7. Extremwerttheorie

7.6 Verteilung des Excess

Die **Block-Maxima-Methode** besitzt die folgenden **Schwächen**:

- Sie geht mit vorhandenen Beobachtungen „verschwenderisch“ um. Da aus jedem Block nur der größte Wert verwendet wird, kann es vorkommen, dass extreme Beobachtungen bei der Schätzung der Parameter nicht berücksichtigt werden.
- Auf der anderen Seite ist es möglich, dass nicht so extreme Beobachtungen bei der Parameterschätzung verwendet werden, weil ein Block nur aus kleineren Beobachtungen besteht.
- Das Maximum ist in vielen Anwendungen nicht von so großem Interesse wie die Verteilung der Überschreitung $X - u | X > u$ für einen großen Threshold u .

In der Praxis wird daher oft die sog. **POT-Methode (Peak-over-Threshold-Methode)** verwendet, bei der alle Daten oberhalb eines vorgegebenen **Schwellenwerts (Thresholds)**

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

berücksichtigt werden.

7. Extremwerttheorie

7.6 Verteilung des Excess

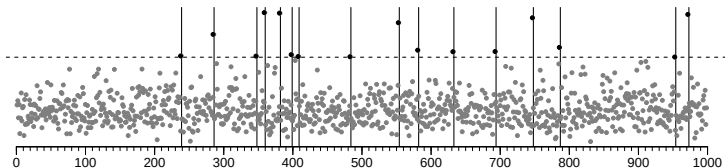
D.h. für $u < x_F$ wird die **bedingte Verteilung**

$$F_X^u(x) := \mathbb{P}(X - u \leq x | X > u) = \frac{\mathbb{P}(u < X \leq x + u)}{\mathbb{P}(X > u)} = \frac{F_X(x + u) - F_X(u)}{1 - F_X(u)} \quad (13)$$

untersucht. Diese Verteilung wird als **Excess-Verteilung von X über u** bezeichnet. Sie beschreibt die Verteilung des **Excess**

$$Y := X - u | X > u$$

von X über u , bedingt gegeben $X > u$.



7. Extremwerttheorie

7.6 Verteilung des Excess

Bemerkungen:

- Wenn die Verteilungsfunktion F_X bekannt ist, kann die Excess-Verteilung (13) **exakt** berechnet werden.
- Die Excess-Verteilung (13) wird z.B. für die Berechnung von Rückversicherungsprämien benötigt (vgl. Abschnitt 6.5).

Beispiel (Excess-Verteilung für $X \sim \text{Exp}(\lambda)$)

Die Zufallsvariable X besitze die Verteilungsfunktion $F_X(x) = 1 - e^{-\lambda x}$ für $x > 0$ und $F_X(x) = 0$ für $x \leq 0$ mit $\lambda > 0$. Dann gilt für $u < \infty$:

$$\begin{aligned} F_X^u(x) &= \frac{F_X(x+u) - F_X(u)}{1 - F_X(u)} \\ &= \frac{1 - e^{-\lambda(x+u)} - (1 - e^{-\lambda u})}{e^{-\lambda u}} = 1 - e^{-\lambda x} = F_X(x) \end{aligned}$$

D.h. X und der Excess $X - u$, bedingt gegeben $X > u$, besitzen dieselbe Verteilung. Dies ist die sog. Gedächtnislosigkeit der Exponential-Verteilung (vgl. Abschnitt 4.7).

Abschnitt 7.7

Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Wenn die Verteilungsfunktion F_X nicht bekannt ist, kann die Excess-Verteilung F_X^u nicht exakt berechnet werden. In diesem Fall ist man an einer **Approximation** der Excess-Verteilung für große u interessiert.

Hierbei besitzt die **verallgemeinerte Pareto-Verteilung** eine zentrale Bedeutung:

Definition (Verallgemeinerte Pareto-Verteilung)

Die verallgemeinerte Pareto-Verteilung besitzt die Verteilungsfunktion

$$\tilde{F}_\xi(x; \mu, \sigma) := \begin{cases} 1 - \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi} & \text{für } \xi \neq 0 \\ 1 - \exp\left(-\frac{x - \mu}{\sigma}\right) & \text{für } \xi = 0 \end{cases}$$

für $x > \mu$, falls $\xi \geq 0$ und $\mu < x < \mu - \frac{\sigma}{\xi}$, falls $\xi < 0$ gilt. Für $x \leq \mu$ gilt in diesen beiden Fällen definitionsgemäß $\tilde{F}_\xi(x; \mu, \sigma) = 0$. Im Falle von $\xi < 0$ setzt man ferner $\tilde{F}_\xi(x; \mu, \sigma) = 1$ für $x \geq \mu - \frac{\sigma}{\xi}$.

Bei $\mu \in \mathbb{R}$ handelt es sich um einen Lokalisations- und bei $\sigma > 0$ um einen Skalierungsparameter.

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Die Abbildung $\xi \mapsto \tilde{F}_\xi(x; \mu, \sigma)$ ist auch an der Stelle $\xi = 0$ stetig. D.h. es gilt

$$\lim_{\xi \rightarrow 0} \tilde{F}_\xi(x; \mu, \sigma) = \tilde{F}_0(x; \mu, \sigma).$$

Der **Shape-Parameter (Tail-Index)** ξ bestimmt den Typ der verallgemeinerten Pareto-Verteilung:

- $\xi = 0$ **Exponential-Verteilung** mit nach unten beschränktem und nach oben unbeschränktem Träger
- $\xi > 0$ **Pareto-Verteilung** mit nach unten beschränktem und nach oben unbeschränktem Träger (benannt nach dem italienischen Ingenieur, Soziologen und Ökonom V. F. D. PARETO (1848–1923))
- $\xi < 0$ **(Verallgemeinerte) Beta-Verteilung** mit nach unten und oben beschränktem Träger

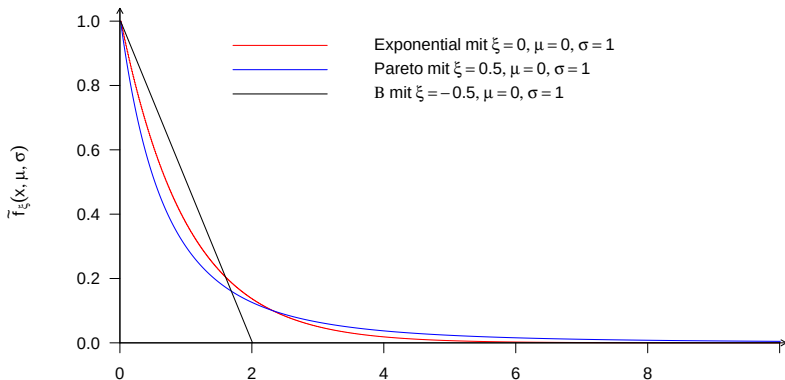


7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Die Dichte der verallgemeinerten Pareto-Verteilung ist gegeben durch:

$$\tilde{f}_{\xi}(x; \mu, \sigma) := \begin{cases} \frac{1}{\sigma} \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-\frac{1+\xi}{\xi}} & \text{für } \xi \neq 0 \\ \frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{für } \xi = 0 \end{cases} \quad (14)$$



7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Bemerkungen:

- Für $\mu = 0$ und $\sigma = 1$ erhält man die **standardisierte verallgemeinerte Pareto-Verteilung**.
- Für $X \sim \tilde{F}_\xi(x; 0, 1)$ gilt $Y = \sigma X + \mu \sim \tilde{F}_\xi(y; \mu, \sigma)$.
- Für zwei Zufallsvariablen X und Y mit einer verallgemeinerten Pareto-Verteilung und **identischem Shape-Parameter** ξ gibt es Konstanten $\sigma > 0$ und $\mu \in \mathbb{R}$, so dass gilt:

$$Y \stackrel{d}{=} \sigma X + \mu$$

- Analog zur Weibull-Verteilung ist die Beta-Verteilung wegen ihres nach rechts beschränkten Trägers oftmals zur Modellierung von Einzelschadenhöhen **nicht geeignet**.
- Für $\xi > 0$ gehört die verallgemeinerte Pareto-Verteilung zur Klasse der **sub-exponentiellen Verteilungen** (vgl. Abschnitt 4.1). Grundsätzlich gilt: Umso größer $\xi > 0$ ist, desto mehr Maße liegt auf dem rechten Tail.

Bemerkungen (Fortsetzung):

- Wie man zeigen kann, existiert das k -te Moment für $k < \frac{1}{\xi}$ und es gilt

$$\mathbb{E}[X^k] = \infty \quad \text{für } k \geq \frac{1}{\xi}. \quad (15)$$

D.h. der Erwartungswert existiert für $\xi < 1$ und ist dann gegeben durch

$$\mathbb{E}[X] = \mu + \frac{\sigma}{1 - \xi}. \quad (16)$$

(vgl. Übungsaufgaben).

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Zwischen der verallgemeinerten Pareto-Verteilung $\tilde{F}_\xi(x; \mu, \sigma)$ und der verallgemeinerten Extremwertverteilung $F_\xi(x; \mu, \sigma)$ besteht offensichtlich der **Zusammenhang**

$$\tilde{F}_\xi(x; \mu, \sigma) = 1 + \ln(F_\xi(x; \mu, \sigma))$$

für alle $x \in \mathbb{R}$ mit $F_\xi(x; \mu, \sigma) \geq \exp(-1)$. Darüber hinaus gilt der folgende Satz:

Satz (Zusammenhang zwischen $F_\xi(x; \mu, \sigma)$ und $\tilde{F}_\xi(x; \mu, \sigma)$)

Für alle $\xi \in \mathbb{R}$ gilt:

$$\tilde{F}_\xi(x; \mu_1, \sigma_1) \in \text{MDA}(F_\xi(x; \mu_2, \sigma_2))$$

Beweis: Siehe z.B. EMBRECHTS ET AL. (2003). ■

Eine Verteilungsfunktion F besitzt die **Excess-Stability-Eigenschaft**, wenn es für alle $u < x_F = \sup\{x \in \mathbb{R} : F_X(x) < 1\}$ Konstanten $a_u \in \mathbb{R}$ und $b_u > 0$ gibt, so dass

$$F^u(x) = F(b_u x + a_u) \quad \text{für alle } x \in \mathbb{R}$$

gilt. Man sagt dann, dass die Verteilungsfunktion F **bzgl. Thresholding abgeschlossen** ist.

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Satz (Charakterisierung der Excess-Stability-Eigenschaft)

Eine Verteilungsfunktion F besitzt genau dann die Excess-Stability-Eigenschaft, wenn sie eine verallgemeinerte Pareto-Verteilung $\tilde{F}_\xi(x; \mu, \sigma)$ ist. In diesem Fall gilt für die Excess-Verteilung für $\xi \in \mathbb{R}$:

$$\tilde{F}_\xi^u(x) = \tilde{F}_\xi(x; 0, \sigma + \xi(u - \mu)) \quad (17)$$

Beweis: Für eine verallgemeinerte Pareto-Verteilung $\tilde{F}_\xi(x; \mu, \sigma)$ mit $\xi \neq 0$ gilt: (vgl. (13))

$$\begin{aligned} \tilde{F}_\xi^u(x; \mu, \sigma) &= \frac{\tilde{F}_\xi(x+u; \mu, \sigma) - \tilde{F}_\xi(u; \mu, \sigma)}{1 - \tilde{F}_\xi(u; \mu, \sigma)} = 1 - \frac{\tilde{F}_\xi(x+u; \mu, \sigma)}{\tilde{F}_\xi(u; \mu, \sigma)} \\ &= 1 - \frac{\left(1 + \xi \frac{x+u-\mu}{\sigma}\right)^{-1/\xi}}{\left(1 + \xi \frac{u-\mu}{\sigma}\right)^{-1/\xi}} \\ &= 1 - \left(\frac{1 + \xi \frac{u-\mu}{\sigma} + \xi \frac{x}{\sigma}}{1 + \xi \frac{u-\mu}{\sigma}}\right)^{-1/\xi} \\ &= 1 - \left(1 + \xi \frac{x}{\sigma + \xi(u-\mu)}\right)^{-1/\xi} = \tilde{F}_\xi(x; 0, \sigma + \xi(u-\mu)) \end{aligned}$$

Dies ist jedoch wieder eine Pareto-Verteilung mit $\xi \neq 0$ und verändertem Lokalisations- und Skalierungsparameter. Völlig analog zeigt man, dass auch die verallgemeinerte Pareto-Verteilung mit $\xi = 0$ die Excess-Stability-Eigenschaft besitzt.

Für den Nachweis, dass eine Verteilungsfunktion mit der Excess-Stability-Eigenschaft stets eine verallgemeinerte Pareto-Verteilung ist, siehe z.B. RESNICK (1987).

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Aus dem letzten Satz erhält man für $\tilde{F}_0(x; 0, \sigma)$, d.h. eine gewöhnliche Exponential-Verteilung mit Parameter $\frac{1}{\sigma}$, den Zusammenhang

$$\tilde{F}_0^u(x) = \tilde{F}_0(x; 0, \sigma).$$

D.h. X und $X - u | X > u$ besitzen dieselbe Verteilung. Dies ist wieder die als **Gedächtnislosigkeit** bekannte Eigenschaft der Exponential-Verteilung (vgl. Abschnitt 4.7).

Beispiel (Mean-excess-Funktion einer verallgemeinerten Pareto-Verteilung)

Es gelte $X \sim \tilde{F}_\xi(x; \mu, \sigma)$ mit $\xi < 1$. Mit (16) und (17) erhält man dann für die Mean-excess-Funktion von X (vgl. Abschnitt 6.5):

$$e_X(u) = \mathbb{E}[X - u | X > u] = \frac{\sigma + \xi(u - \mu)}{1 - \xi} = \frac{\sigma - \xi\mu}{1 - \xi} + \frac{\xi}{1 - \xi}u \quad (18)$$

D.h. die Mean-excess-Funktion $e_X(u)$ ist linear im Threshold u mit der positiven Steigung $\frac{\xi}{1 - \xi}$. Für $\xi = 0$ erhält man daraus die konstante Mean-excess-Funktion $e_X(u) = \sigma$ der Exponential-Verteilung.

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Mit Hilfe dieses Ergebnisses erhält man für den **Expected-Shortfall** und den **Conditional-Tail-Expectation** einer verallgemeinerten Pareto-Verteilung mit $\xi < 1$ das folgende Resultat:

Folgerung (ES und CTE einer verallgemeinerten Pareto-Verteilung)

Es gelte $X \sim \tilde{F}_\xi(x; \mu, \sigma)$ mit $\xi < 1$. Dann gilt für den Expected-Shortfall und den Conditional-Tail-Expectation von X zum Sicherheitsniveau $q \in (0, 1)$:

$$\text{ES}_q(X) = \text{CTE}_q(X) = \text{VaR}_q(X) + e_X(\text{VaR}_q(X)) = \frac{\text{VaR}_q(X)}{1 - \xi} + \frac{\sigma - \xi\mu}{1 - \xi} \quad (19)$$

Beweis: Da X eine stetige Zufallsvariable ist, gilt $\text{ES}_q(X) = \text{CTE}_q(X)$ (vgl. Abschnitt 2.5). Zusammen mit (18) liefert dies:

$$\begin{aligned} \text{ES}_q(X) &= \text{CTE}_q(X) \\ &= \mathbb{E}[X | X > \text{VaR}_q(X)] \\ &= \text{VaR}_q(X) + \mathbb{E}[X - \text{VaR}_q(X) | X > \text{VaR}_q(X)] \\ &= \text{VaR}_q(X) + e_X(\text{VaR}_q(X)) \\ &= \text{VaR}_q(X) + \frac{\sigma + \xi(\text{VaR}_q(X) - \mu)}{1 - \xi} \\ &= \frac{\text{VaR}_q(X)}{1 - \xi} + \frac{\sigma - \xi\mu}{1 - \xi} \end{aligned}$$

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan



Analog zu F_{M_n} existiert auch für F_X^u ein Grenzwertsatz. Dieser wird nach den Statistikern J. PICKANDS, A. BALKEMA und L. DE HAAN als **Satz von Pickands-Balkema-De Haan** bezeichnet:



Satz (Pickands-Balkema-De Haan (1974,1975))

Es sei F_X eine Verteilungsfunktion und $x_F := \sup \{x \in \mathbb{R} : F_X(x) < 1\}$. Dann gilt $F_X \in \text{MDA}(F_\xi(x; \mu, \sigma))$ genau dann, wenn

$$\lim_{u \uparrow x_F} \sup_{0 \leq x < x_F - u} |F_X^u(x) - \tilde{F}_\xi(x; 0, \sigma(u))| = 0 \quad (20)$$

für eine geeignete von u abhängige Funktion $\sigma(u)$ gilt.

Beweis: Siehe z.B. EMBRECHTS ET AL. (2003), Seiten 165–166. ■

Gemäß dieses Ergebnisses ist für einen hinreichend großen Threshold u und eine Zufallsvariable X , deren Verteilungsfunktion F_X im MDA einer verallgemeinerten Extremwertverteilung mit Shape-Parameter ξ liegt, die verallgemeinerte Pareto-Verteilung mit gleichem Shape-Parameter ξ der **kanonische Kandidat zur Approximation der Excess-Verteilung F_X^u** , d.h. der Verteilungsfunktion des Excess $X - u | X > u$.

7. Extremwerttheorie

7.7 Verallgemeinerte Pareto-Verteilung und Satz von Pickands-Balkema-de Haan

Bemerkungen:

- Bei Einzelschadenhöhenverteilungen F_X , bei denen die Verteilungsfunktion des standardisierten Maximums $\frac{M_n - a_n}{b_n}$ für $n \rightarrow \infty$ gegen eine verallgemeinerte Extremwertverteilung konvergiert, strebt die Excess-Verteilung F_X^u für $u \rightarrow x_F$ gegen eine verallgemeinerte Pareto-Verteilung und zwar mit **demselben Shape-Parameter ξ** .
- Da die Excess-Verteilung F_X^u durch 0 nach unten beschränkt ist, gilt für den Lokalisationsparameter der Grenzverteilung in (20) **$\mu = 0$** .
- So gut wie alle in der Versicherungsmathematik verwendeten Einzelschadenhöhenverteilungen liegen im MDA einer verallgemeinerten Extremwert-Verteilung. Das Theorem von Pickands-Balkema-de Haan besitzt somit eine **sehr allgemeine Gültigkeit**.
- Das Theorem von Pickands-Balkema-de Haan und das Theorem von Fisher-Tippett können als **zentrale Grenzwertsätze** für M_n bzw. $X - u | X > u$ aufgefasst werden.

Abschnitt 7.8

Peaks-over-Threshold-Methode

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Gegeben seien n unabhängige und identisch-verteilte Zufallsvariablen

$$X_1, \dots, X_n \sim F_X \in \text{MDA}(F_\xi(x; \mu, \sigma)).$$

Gemäß des Theorems von Pickands-Balkema-de Haan kann die Excess-Verteilung F_X^u für große u durch die verallgemeinerte Pareto-Verteilung $\tilde{F}_\xi(x; 0, \sigma(u))$ approximiert werden.

Bei der **Peaks-Over Threshold-Methode (POT-Methode)** werden zur Anpassung einer verallgemeinerten Pareto-Verteilung an die Excess-Verteilung F_X^u , d.h. an die Verteilung von

$$X - u | X > u,$$

die **Überschreitungen (exceedances)** von u mit

$$Y_i := X_i - u | X_i > u \quad \text{für } i = 1, \dots, N_u \quad (21)$$

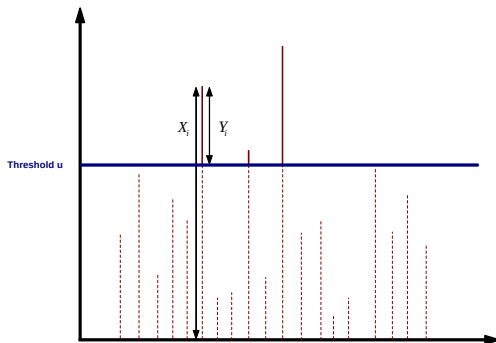
bezeichnet, wobei

$$N_u := \sum_{i=1}^n 1_{]u, \infty)}(X_i) \quad (22)$$

die (zufällige) Anzahl an Überschreitungen von u ist. Die Zufallsvariablen (21) sind ebenfalls stochastisch unabhängig und identisch-verteilt sowie für hinreichend großes u approximativ $\tilde{F}_\xi(y_i; 0, \sigma(u))$ -verteilt.

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode



Im Folgenden seien

$$y_1, \dots, y_{n_u} \quad (23)$$

die **Realisierungen** der Überschreitungen (21), wobei n_u die Realisierung von (22) ist, und $\tilde{f}_\xi(y; 0, \sigma)$ sei die Dichte der verallgemeinerten Pareto-Verteilung $\tilde{F}_\xi(y; 0, \sigma)$ (vgl. (14)). Nach einer kurzen Umformung erhält man dann für die **Log-Likelihood-funktion** der n_u Überschreitungen (23) den Ausdruck:

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

$$\ln(L(\xi, 0, \sigma; y_1, \dots, y_{n_u})) = \sum_{i=1}^{n_u} \ln(\tilde{f}_{\xi}(y_i; 0, \sigma)) \quad (24)$$
$$= \begin{cases} -n_u \ln(\sigma) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{n_u} \ln\left(1 + \xi \frac{y_i}{\sigma}\right) & \text{für } \xi \neq 0 \\ -n_u \ln(\sigma) - \sum_{i=1}^{n_u} \frac{y_i}{\sigma} & \text{für } \xi = 0 \end{cases}$$

für $y_i > 0$ mit $1 + \xi \frac{y_i}{\sigma} > 0$.

Analog zur Anpassung einer verallgemeinerten Extremwertverteilung mittels Block-Maxima-Methode (vgl. Abschnitt 7.5) erhält man durch Maximierung von (24) mittels numerischer Methoden (z.B. Newton-Raphson) die **ML-Schätzer** $\hat{\xi}, \hat{\sigma}$ für ξ und σ .

Bemerkungen:

- Bei der Maximierung von (24) müssen die Nebenbedingungen $\sigma > 0$ und $1 + \xi \frac{y_i}{\sigma} > 0$ berücksichtigt werden. Es kann jedoch gezeigt werden, dass die resultierenden ML-Schätzer für $\xi > -0,5$ **konsistent** und **asymptotisch Normal-verteilt** sind (vgl. Abschnitt 7.5).
- Mit Hilfe eines **Likelihood-Quotienten-Tests** kann wieder getestet werden, ob $\xi = 0$ gilt (vgl. Abschnitt 7.5).

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Bemerkungen (Fortsetzung):

- Bei der Wahl des Thresholds u besteht wieder ein **Zielkonflikt**:
 - Ein hoher Threshold u führt zu einer **besseren Approximation** der Excess-Verteilung F_X^u durch $\tilde{F}_\xi(y; 0, \sigma)$. Es existieren dann aber u.U. nur wenige Überschreitungen für die ML-Schätzung von ξ und σ .
 - Ein kleiner Threshold u führt zu vielen Überschreitungen und somit zu einer **geringeren Varianz** bei der ML-Schätzung von ξ und σ . Aber u.U. liefert dann $\tilde{F}_\xi(y; 0, \sigma)$ keine gute Approximation für F_X^u .

Bei den beiden folgenden **Standardansätzen** zur Bestimmung des Thresholds u wird dieser möglichst klein gewählt, aber noch groß genug, dass $\tilde{F}_\xi(y; 0, \sigma)$ eine „vernünftige Approximation“ für F_X^u liefert:

- 1) Der erste Ansatz basiert auf der **Linearität** von $e_X(u)$ für $\tilde{F}_\xi(y; 0, \sigma)$ (vgl. (18)). Hierbei wird zur Schätzung von $e_X(u)$ die **empirische Mean-excess-Funktion**

$$\hat{e}_n(u) := \frac{1}{N_u} \sum_{i=1}^{N_u} Y_i \quad \text{mit} \quad Y_i = X_i - u | X_i > u \quad \text{für } i = 1, \dots, N_u$$

verwendet und der sog. **Mean-excess-Plot** $\{u < x_F : (u, \hat{e}_n(u))\}$ erstellt. Als Threshold u wird anschließend der kleinste Wert gewählt, ab dem der Mean-excess-Plot „**annähernd linear**“ verläuft.

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Ein ansteigender Verlauf spricht dann für $\xi > 0$ (d.h. eine Pareto-Verteilung), ein horizontaler Verlauf für $\xi = 0$ (d.h. eine Exponential-Verteilung) und ein fallender Verlauf für $\xi < 0$ (d.h. eine Beta-Verteilung). Aufgrund der approximativen Normalverteiltetheit von $\hat{e}_n(u)$ können für $\hat{e}_n(u)$ **Konfidenzintervalle** berechnet werden. Da jedoch ein Mean-excess-Plot nur sehr selten perfekt linear verläuft, ist es oft nicht einfach, daraus einen geeigneten Threshold u abzulesen. Häufig werden die letzten Ausprägungen von $\hat{e}_n(u)$ aus der Betrachtung ausgeschlossen, da sie nur auf sehr wenigen Überschreitungen Y_i basieren und deshalb oft sehr volatile sind.

- 2) Der zweite Ansatz basiert auf der **Abgeschlossenheit** von $\tilde{F}_\xi(y; 0, \sigma)$ bzgl. **Thresholding** (vgl. Abschnitt 7.7). Hierbei wird für verschiedene Thresholds u jeweils eine verallgemeinerte Pareto-Verteilung angepasst und der Plot $\{u < x_F : (u, \hat{\xi})\}$ erstellt. Als Threshold u wird anschließend der kleinste Wert gewählt, ab dem der Plot „**annähernd konstant**“ verläuft, d.h. die ML-Schätzungen $\hat{\xi}$ für den Shape-Parameter ξ ungefähr konstant sind.

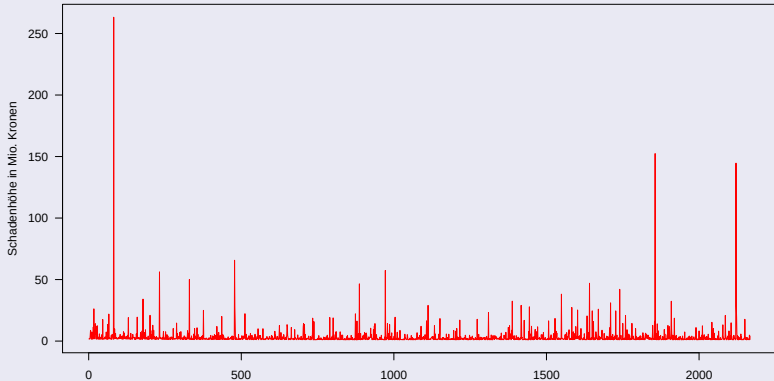
Nachdem ein geeigneter Threshold u bestimmt wurde, wird die verallgemeinerte Pareto-Verteilung an die n_u Überschreitungen Y_i mittels ML-Methode angepasst.

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Anwendung der POT-Methode)

Betrachtet werden die Daten eines dänischen Feuerversicherers im Zeitraum vom 1.1.1980 bis 31.12.1990. Es handelt sich dabei um 2167 Schäden mit einer inflationsbereinigten Schadenhöhe von jeweils mehr als 1 Mio. Dänischen Kronen.

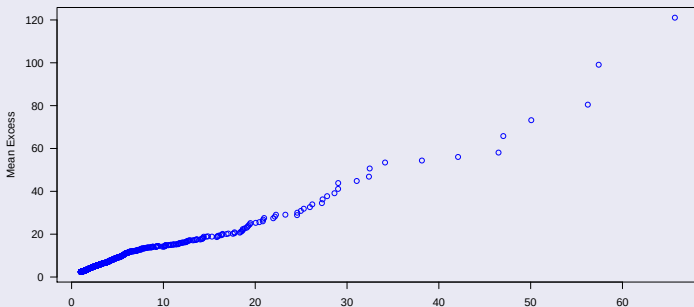


7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)

Der Mean-excess-Plot (Abbildung unten) verläuft im gesamten Bereich annähernd linear mit positiver Steigung. Bei $u = 10$ Mio. Dänischen Kronen ist jedoch ein „leichter Knick“ zu erkennen und danach ein nahezu gerades Auslaufen. Dies motiviert die Anpassung einer verallgemeinerten Pareto-Verteilung mit $\xi > 0$ und Schwellenwert $u = 10$ Mio. an die Überschreitungen $Y_i = X_i - u | X_i > u$ (d.h. zur Approximation von F_X^u). Im Falle von $u = 10$ Mio gibt es 109 Überschreitungen Y_i .

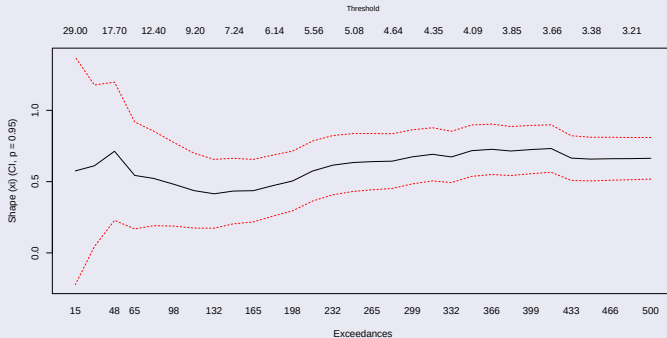


7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)

Die nächste Abbildung zeigt die ML-Schätzung $\hat{\xi}$ in Abhängigkeit vom Threshold u und der Anzahl n_u der zugehörigen Überschreitungen. Das 95%-Konfidenzband wurde mit Hilfe der geschätzten Standardfehler für $\hat{\xi}$ berechnet. Es ist zu erkennen, dass die ML-Schätzungen $\hat{\xi}$ relativ stabil sind und zum großen Teil im Intervall $(0,4;0,7)$ liegen.



7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)

Die Anpassung einer verallgemeinerten Pareto-Verteilung oberhalb des Thresholds $u = 10$ Mio. Dänischen Kronen mittels der ML-Methode liefert die folgenden Schätzungen:

Parameter	ML-Schätzung	Standardabweichung	
ξ	0,4968062	0,1362093	hoch signifikant
σ	6,9745523	1,1131016	hoch signifikant

Die angepasste verallgemeinerte Pareto-Verteilung besitzt einen sehr starken rechten Tail und wegen $\hat{\xi} \approx 0,497 < \frac{1}{2}$ fast eine unendliche Varianz.

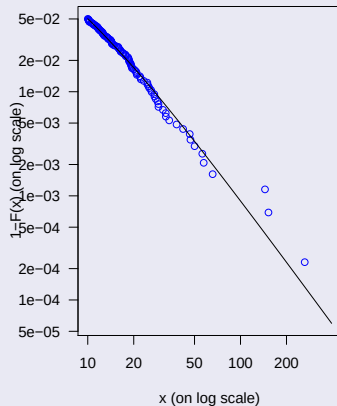
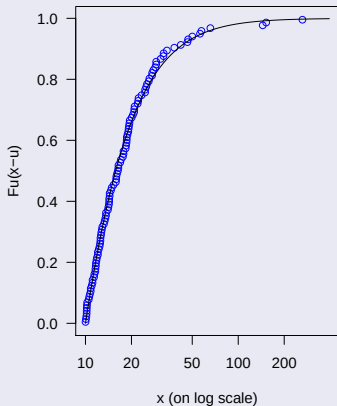
Die folgende linke Abbildung zeigt, dass die verallgemeinerte Pareto-Verteilung für die Excess-Verteilung F_X^u mit $u = 10$ Mio. eine gute Anpassung für die empirische Verteilung der $n_u = 109$ Überschreitungen $Y_i = X_i - u | X_i > u$ liefert.

Der log-log-Plot auf der rechten Seite zeigt, dass auch im rechten Tail die Anpassungsgüte über weite Strecken relativ gut ist.

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)



7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)

R-Code:

```
1  ###
2  # Anpassung einer verallgemeinerten Pareto-Verteilung an die Schadenhöhen eines Feuerversicherers
3  # mittels POT-Methode
4  ###
5
6  library(evir)
7  data(danish)
8  # Plot der Daten
9  par(mfrow=c(1,1),cex=1.5,mar=c(2,5,1,1))
10 plot(danish, type="l", ylab="Schadenhöhe in Mio. Kronen", col="red",
11       xlab="Schäden im Zeitraum 01.01.1980 - 31.12.1990", las=1)
12
13 # Mean-of-Excess-Plot
14 par(mfrow=c(1,1),cex=1.5,mar=c(2,5,1,1))
15 meplot(danish, col="blue",las=1)
16
17 # Plot der ML-Schätzung für den Shape-Parameters xi für verschiedene Thresholds u
18 par(mfrow=c(1,1),cex=1.5,mar=c(2,5,3,3))
19 shape(danish, models=30, start=15, end=500, ci=0.95, revers=FALSE, las=1)
20
21 # Anpassung einer verallgemeinerten Pareto-Verteilung über dem Threshold u=10 Mio.
22 fit<-gpd(danish, threshold=10)
23 fit
24 par(mfrow=c(1,2),cex=1.5,mar=c(4,5,3,3))
25 plot(fit, col="blue",lty=2, las=1)
```

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Für die Excess-Verteilung einer Zufallsvariablen X folgt für $u < x_F$:

$$\begin{aligned} F_X^u(x-u) &= \mathbb{P}(X-u \leq x-u | X > u) \\ &= \mathbb{P}(X \leq x | X > u) \\ &= \frac{\mathbb{P}(u < X \leq x)}{\mathbb{P}(X > u)} \\ &= \frac{F_X(x) - F_X(u)}{1 - F_X(u)} \\ &= \frac{-(1 - F_X(x))}{\bar{F}_X(u)} + 1 \end{aligned}$$

D.h. es gilt

$$F_X(x) = 1 - \bar{F}_X(u) \bar{F}_X^u(x-u). \quad (25)$$

Im Folgenden wird nun angenommen, dass der Threshold u hinreichend groß ist, so dass aufgrund des Satzes von Pickands-Balkema-De Haan (siehe Abschnitt 7.7) angenommen werden kann, dass es sich (näherungsweise) bei der Excess-Verteilung F_X^u um eine $\tilde{F}_\xi(x; 0, \sigma)$ -Verteilung bzw. bei $F_X^u(x-u)$ um eine $\tilde{F}_\xi(x; u, \sigma)$ -Verteilung handelt.

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Dann folgt

$$\begin{aligned} q &:= F_X(x) \\ &= \begin{cases} 1 - \bar{F}_X(u) \left(1 + \xi \frac{x-u}{\sigma}\right)^{-1/\xi} & \text{für } \xi \neq 0 \\ 1 - \bar{F}_X(u) \exp\left(-\frac{x-u}{\sigma}\right) & \text{für } \xi = 0 \end{cases}. \end{aligned} \quad (26)$$

Durch Auflösen dieser Gleichung nach x erhält man für den **Value-at-Risk** von X zum Sicherheitsniveau $q \in (0, 1)$:

$$\begin{aligned} \text{VaR}_q(X) &= F_X^{-1}(q) \\ &= \begin{cases} u + \frac{\sigma}{\xi} \left(\left(\frac{1-q}{\bar{F}_X(u)} \right)^{-\xi} - 1 \right) & \text{für } \xi \neq 0 \\ u - \sigma \ln \left(\frac{1-q}{\bar{F}_X(u)} \right) & \text{für } \xi = 0 \end{cases} \end{aligned} \quad (27)$$

Gilt zusätzlich $\xi < 1$, dann resultiert daraus für den **Expected-Shortfall** von X (vgl. (19))

$$\text{ES}_q(X) = \text{VaR}_q(X) + e_X(\text{VaR}_q(X)) = \frac{\text{VaR}_q(X)}{1-\xi} + \frac{\sigma - \xi u}{1-\xi}. \quad (28)$$

Dabei wird jedoch angenommen, dass das Sicherheitsniveau $q \in (0, 1)$ hinreichend groß ist, so dass die verallgemeinerte Pareto-Verteilung eine gute Approximation für die Verteilung von $X - \text{VaR}_q(X) | X > \text{VaR}_q(X)$ ist.

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Werden nun in (27) und (28) die Parameter ξ und σ durch ihre ML-Schätzungen $\hat{\xi}$ bzw. $\hat{\sigma}$ sowie $\bar{F}_X(u)$ durch den Schätzer

$$\hat{\bar{F}}_X(u) := \frac{1}{n} \sum_{i=1}^n 1_{]u, \infty)}(X_i) = \frac{N_u}{n} \quad (29)$$

(vgl. (22)) ersetzt, dann erhält man folgende **Schätzer für den Value-at-Risk und den Expected-Shortfall**

$$\widehat{\text{VaR}}_q(X) = \begin{cases} u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1-q) \right)^{-\hat{\xi}} - 1 \right) & \text{für } \hat{\xi} \neq 0 \\ u - \hat{\sigma} \ln \left(\frac{n}{N_u} (1-q) \right) & \text{für } \hat{\xi} = 0 \end{cases} \quad (30)$$

bzw.

$$\widehat{\text{ES}}_q(X) = \frac{\widehat{\text{VaR}}_q(X)}{1 - \hat{\xi}} + \frac{\hat{\sigma} - \hat{\xi} u}{1 - \hat{\xi}}. \quad (31)$$

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Wiederkehrwert und Wiederkehrperiode

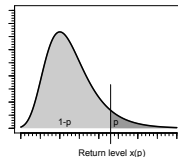
Mit Hilfe der angepassten Verteilung

$$\hat{F}_X(x) = \begin{cases} 1 - \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x-u}{\hat{\sigma}} \right)^{-1/\hat{\xi}} & \text{für } \hat{\xi} \neq 0 \\ 1 - \frac{N_u}{n} \exp \left(-\frac{x-u}{\hat{\sigma}} \right) & \text{für } \hat{\xi} = 0 \end{cases}$$

(vgl. (26) und (29)) kann

- a) für eine vorgegebene Überschreitungswahrscheinlichkeit $p \in (0, 1)$ der zugehörige **Wiederkehrwert (return level)** $\hat{x}(p)$ bzw.
- b) für einen vorgegebenen Wiederkehrwert $\hat{x}(p)$ die zugehörige **Wiederkehrperiode (return period)** T

ermittelt werden. Dies erfolgt durch Auflösen der Gleichung



$$p = \mathbb{P}(X > \hat{x}(p)) \approx 1 - \hat{F}_X(\hat{x}(p)) = \begin{cases} \frac{N_u}{n} \left(1 + \hat{\xi} \frac{\hat{x}(p)-u}{\hat{\sigma}} \right)^{-1/\hat{\xi}} & \text{für } \hat{\xi} \neq 0 \\ \frac{N_u}{n} \exp \left(-\frac{\hat{x}(p)-u}{\hat{\sigma}} \right) & \text{für } \hat{\xi} = 0 \end{cases}$$

nach $\hat{x}(p)$. Dies liefert dann:

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

$$\hat{x}(p) = \begin{cases} u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{N_u}{pn} \right)^{\hat{\xi}} - 1 \right) & \text{für } \hat{\xi} \neq 0 \\ u + \hat{\sigma} \ln \left(\frac{N_u}{pn} \right) & \text{für } \hat{\xi} = 0 \end{cases}$$

Beim Wiederkehrwert $\hat{x}(p)$ handelt es sich um eine ML-Schätzung für den Wert, der mit der Wahrscheinlichkeit p überschritten wird, d.h. der durchschnittlich alle

$$T = \frac{1}{p}$$

Zeiteinheiten übertroffen wird.

Mit Hilfe der [Delta-Methode](#) kann wieder eine Schätzung für die Varianz von $\hat{x}(p)$ ermittelt werden (vgl. Abschnitt 7.5).

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Schätzungen für MeF, VaR und ES)

Betrachtet wird wieder die Situation aus dem letzten Beispiel (Schadenhöhen eines dänischen Feuerversicherers zwischen 1.1.1980 und 31.12.1990). Mit (18) erhält man dann für die erwartete Überschreitung $Y = X - u | X > u$ mit Threshold $u = 10$ Mio. die Schätzung

$$\widehat{e}_X(u) = \widehat{\mathbb{E}}[X - u | X > u] = \frac{\widehat{\sigma} + \widehat{\xi}u}{1 - \widehat{\xi}} = 23,73 \text{ Mio.}$$

Die folgende Tabelle zeigt für verschiedene Sicherheitsniveaus $q \in (0, 1)$ die mittels (30) und (31) berechneten Schätzungen für $\text{VaR}_q(X)$ bzw. $\text{ES}_q(X)$:

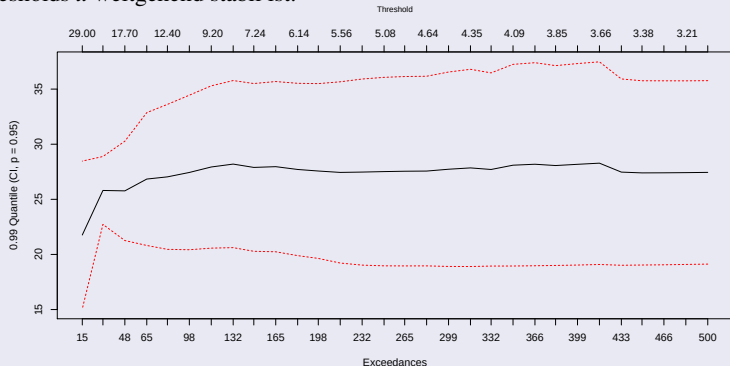
q	$\widehat{\text{VaR}}_q(X)$	$\widehat{\text{ES}}_q(X)$
0,9000	5,94	15,79
0,9500	10,04	23,94
0,9900	27,28	58,21
0,9950	40,16	83,80
0,9990	94,29	191,37
0,9999	304,62	609,37

7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)

Die nächste Abbildung zeigt die Schätzung $\widehat{\text{VaR}}_{0,99}(X)$ in Abhängigkeit vom Threshold u und das zugehörige 95%-Konfidenzband (berechnet mittels der approximativen Normalverteilung der ML-Schätzungen und der Delta-Methode, vgl. Abschnitt 7.5). Es ist zu erkennen, dass die Schätzung $\widehat{\text{VaR}}_{0,99}(X)$ bzgl. einer Veränderung des Thresholds u weitgehend stabil ist.



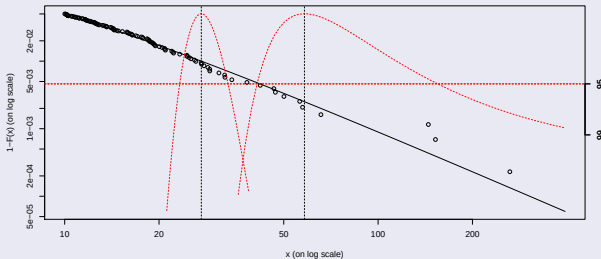
7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)

Die folgende Abbildung zeigt den log-log-Plot sowie die Schätzungen $\widehat{\text{VaR}}_{0,99}(X)$ und $\widehat{\text{ES}}_{0,99}(X)$ (vertikale Linien). Die 95%-Konfidenzbänder für diese beiden Schätzungen ergeben sich jeweils durch die beiden Schnittpunkte der horizontalen Linie mit der entsprechenden Kurve für $\widehat{\text{VaR}}_{0,99}(X)$ bzw. $\widehat{\text{ES}}_{0,99}(X)$.

	Untere Schranke	Schätzung	Obere Schranke
$\widehat{\text{VaR}}_{0,99}(X)$	23,36	27,28	33,16
$\widehat{\text{ES}}_{0,99}(X)$	41,21	58,21	154,89



7. Extremwerttheorie

7.8 Peaks-over-Threshold-Methode

Beispiel (Fortsetzung)

R-Code:

```
1  ###
2  # Schätzung des VaR und des ES für die angepasste verallgemeinerte Pareto-Verteilung
3  ###
4
5  library(evir)
6  data(danish)
7  # Anpassung einer verallgemeinerten Pareto-Verteilung über dem Threshold u=10 Mio.
8  fit<-gpd(danish, threshold=10)
9
10 # Bestimmung des VaR und ES für verschiedene SN
11 q<-c(0.90,0.95,0.99,0.995,0.999,0.9999)
12 riskmeasures(fit,q)
13
14 # Ausgabe des 0.95-Konfidenzbands für den VaR zum SN 0.99
15 par(mfrow=c(1,1),cex=1.3,mar=c(5,5,5,5))
16 quant(danish, p=0.99, models=30, start=15, end=500, ci=0.95, revers=FALSE)
17
18 # Log-Log-Plot und Konfidenzbänder für den VaR und ES
19 par(mfrow=c(1,1),cex=1.3,mar=c(5,5,5,5))
20 tail<-tailplot(fit)
21 gpd.q(tail, pp=0.99, ci.p=0.95)
22 gpd.sfall(tail, pp=0.99, ci.p=0.95)
```

Abschnitt 7.9

Literatur

7. Extremwerttheorie

7.9 Literatur



EMBRECHTS, P., KLÜPPELBERG, C., MIKOSCH, T. (2003). *Modelling Extremal Events: for Insurance and Finance*. Springer.



BALKEMA, A. A., DE HAAN, L. (1974). *Residual lifetime at great age*. The Annals of Probability, 5/2, p. 792–804.



FISHER, R. A., TIPPETT, L. H. C. (1928). *Limiting forms of the frequency distribution of the largest or smallest member of a sample*. Mathematical Proceedings of the Cambridge Philosophical Society, 24/2, p. 180–190.



LEADBETTER, M. R., LINDGREN, G., ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer.



PICKANDS, J. (1975). *Statistical inference using extreme order statistics*. The Annals of Statistics, 3/1, p. 119–131.



RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point processes*. Springer.