

Extremwerttheorie

Seminar aus Finanz- und
Versicherungsmathematik

WS 2008/09, Technische Universität Wien

André Surenian

Inhaltsverzeichnis

1	Extremwerttheorie	2
1.1	Maxima	2
1.1.1	Verallgemeinerte Extremwertverteilung:	2
1.1.2	Maximaler Anziehungsbereich von Extremwertverteilungen	5
1.1.3	Die Block-Maxima Methode	6
1.2	Grenzüberschreitungen (threshold exceedances)	7
1.2.1	Verallgemeinerte Paretoverteilung	7
1.2.2	Modellierung und Messung von Verteilungsenden	10
1.2.3	Der Hill Schätzer	11
1.3	Tails ausgesuchter Verteilungen	13
1.3.1	Anziehungsbereich der Fréchetverteilung	13
1.3.2	Anziehungsbereich der Gumbel Verteilung	14
1.4	Modelle für stochastische Prozesse	15
1.4.1	Das POT-Modell	17
1.4.2	Self-exciting POT Modell	19
1.4.3	Konstruktion eines self-exciting POT-Modells	20
1.5	Multivariate Maxima	21
1.5.1	Copulas für multivariate Minima	23
1.5.2	Anziehungsbereich der Copulas	24
1.6	Multivariate Grenzüberschreitungen	24
1.6.1	Modell	24
1.6.2	Praktische Vorgehensweise	25
1.6.3	Grenzcopulas	26

1 Extremwerttheorie

1.1 Maxima

Die Extremwerttheorie (extreme value theory) als Zweig der Statistik richtet ihre Aufmerksamkeit auf Werte aus einer Beobachtung(speriode), die extrem klein oder extrem groß sind. Für das Ermitteln von Parametern in statistischen Betrachtungen werden diese Werte meist als Ausreißer geringgeschätzt und wegargumentiert. Für ein Versicherungsunternehmen zum Beispiel wäre diese Vorgehensweise sehr riskant, denn mit extremen Werten hat es immer dann zu tun, wenn extrem hohe Schäden durch seltene Ereignisse (z.B. durch Flugzeugabstürze oder Naturkatastrophen) entstehen. Diese Ereignisse treten zwar sehr selten auf, richten aber einen enormen Schaden an und müssen daher in der Kalkulation miteinbezogen werden. Voraussetzung für ebendiese Kalkulation ist eine ordentliche mathematische Modellierung und genau darum geht es in der Extremwerttheorie. Seien also im folgenden $(X_i)_{i \in \mathbb{N}}$ unabhängige, identisch verteilte Zufallsvariablen (hier: vorwiegend finanzielle Verluste, etwa Schadenhöhen in der Sachversicherung oder Verluste in einem Kreditportfolio). Sei $M_1 = X_1$, $M_2 = \max\{X_1, X_2\}$, $M_n = \max\{X_1, \dots, X_n\}$ usw. Daraus folgt, dass
$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = F^n(x)$$

1.1.1 Verallgemeinerte Extremwertverteilung:

Ähnlich wie beim zentralen Grenzwertsatz für die Summe von unabhängigen und identisch verteilten Zufallsvariablen $S_n = X_1 + \dots + X_n$, der besagt, dass bei wachsenden Beobachtungen ($n \rightarrow \infty$) die standardisierte Summe $\frac{S_n - a_n}{b_n}$, mit Zentrierungs- und Normierungskonstanten $a_n = nE[X_1]$ und $b_n = \sqrt{\text{Var}[X_1]}$, in Verteilung gegen eine Standardnormalverteilung konvergiert, also dass

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - a_n}{b_n} \leq x\right) = \Phi(x), \quad x \in \mathbb{R}$$

wird versucht, Aussagen über zentrierte und normierte Maxima zu finden, d.h.: Suche Konstanten c_n, d_n sodass gilt: $c_n^{-1}(M_n - d_n) \rightarrow X$ mit Verteilungsfunktion H . Tatsächlich gilt, dass solch eine Verteilung H existiert.

Definition 1 Die Verteilung:

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

heißt **verallgemeinerte Extremwertverteilung** für $1 + \xi x > 0$. Man erhält eine dreiparametrische Familie durch die Einführung eines Skalenparameter $\sigma > 0$ und eines Lageparameters $\mu \in \mathbb{R}$, $H_{\xi, \mu, \sigma} := H_\xi\left(\frac{x - \mu}{\sigma}\right)$. Der Parameter ξ wird als Shape oder Tailparameter bezeichnet und bestimmt die Form des Verteilungsschwanzes (siehe auch weiter unten die Unterteilung der Verteilungsfunktionen in Abhängigkeit von ξ). Je größer der Parameter ξ , desto schwerer wird der Tail.

Definition 2 Der rechte Endpunkt x_F einer Verteilung ist:

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

Die verallgemeinerte Extremwertverteilung kann in Abhängigkeit von ξ wiederum in 3 Spezialfälle unterteilt werden:

1. Für $\xi = 0$ **Gumbelverteilung**: $\Lambda(x) = \exp(-e^{-x}), x \in \mathbb{R}$
2. Für $\xi > 0$ **Fréchetverteilung**: $\Phi_\alpha(x) = \exp(-x^{-\alpha}), x \geq 0, \alpha > 0$
3. Für $\xi < 0$ **Weibullverteilung**: $\Psi_\alpha(x) = \exp(-(-x)^{-\alpha}), x \leq 0, \alpha < 0$

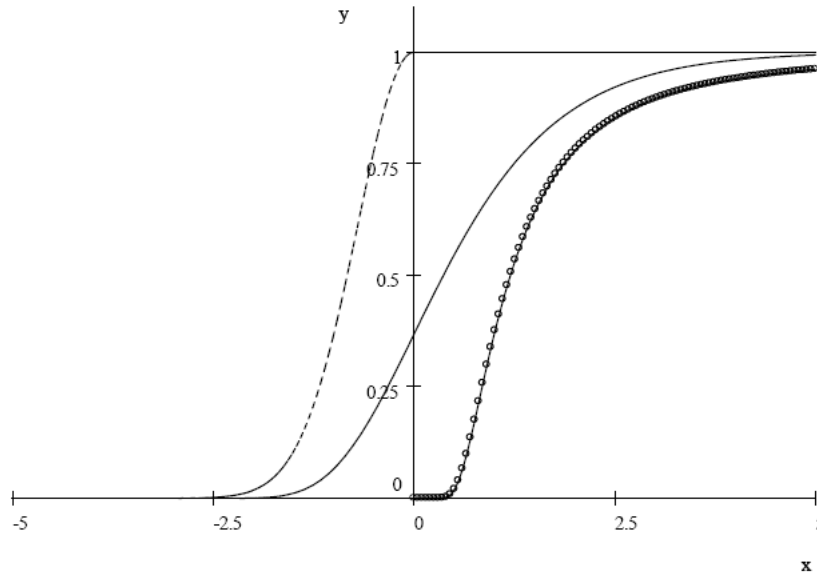


Abbildung 1: Verteilungsfunktionen der Extremwertverteilungen: Gumbelverteilung (durchgezogen), Fréchetverteilung (gepunktet, $\xi = \frac{1}{2}$), Weibullverteilung (gestrichelt $\xi = -\frac{1}{2}$)

In der Graphik kann man nun deutlich erkennen, dass die Weibullverteilung einen endlichen rechten Endpunkt x_F besitzt (in der Graphik ist der Wert $x_F = 0$, der natürlich durch den Lageparameter beliebig verschoben werden kann), wohingegen die Fréchet und die Gumbelverteilung einen unendlichen rechten Endpunkt haben, wobei die Fréchetverteilung ein schwereres Ende hat als die Gumbelverteilung. Wir halten also fest, dass die Konvergenz des Maximums M_n unabhängig und identisch verteilter Zufallsvariablen bedeutet, dass es reelle Folgen d_n und c_n gibt, wobei $c_n > 0, \forall n$, sodass:

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - d_n}{c_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x) \quad (1)$$

Bemerkung 3 Für festes x folgt, dass $\lim_{\xi \rightarrow 0} H_\xi(x) = H_0(x)$. Die verallgemeinerte Extremwertverteilung ist also stetig in ξ . Eine Tatsache, die die Handhabung der Verteilung in der Praxis erleichtert.

Definition 4 Maximum Domain of Attraction Wir sagen, die Verteilungsfunktion F liegt im maximalen Anziehungsbereich (Maximum Domain of Attraction (MDA)) von H , wenn $\lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x)$:

$$F \in MDA(H) \iff \exists c_n > 0, d_n \in \mathbb{R} : c_n^{-1}(M_n - d_n) \rightarrow H.$$

Wie sieht nun dieses $H(x)$ aus? Ist es für alle Verteilungen gleich und unter welchen Voraussetzungen konvergiert F^n gegen H ? Die große Bedeutung der verallgemeinerten Extremwertverteilung und die Antwort auf diese Fragen liefert der folgende Satz, der als zentraler Satz der Extremwerttheorie gilt.

Satz 5 (Fisher-Tippett, Gnedenko) (ohne Beweis) Ist $F \in MDA(H)$, dann folgt, dass H eine Verteilung vom Typ H_ξ ist.

Tatsächlich gilt also, dass die verallgemeinerte Extremwertverteilung die einzig in Frage kommende Grenzverteilung für zentrierte und normierte Maxima ist (modulo affiner Transformation).

Beispiel 6 (Exponentialverteilung) Ist F exponentialverteilt mit dem Parameter $\beta > 0$, so ist $F(x) = 1 - \exp(-\beta x)$, $x \geq 0$. Die Folgen werden dabei so gewählt, dass $c_n = \frac{1}{\beta}$ und $d_n = \frac{\ln n}{\beta}$, woraus wir erhalten:

$$\begin{aligned} F^n(c_n x + d_n) &= \left(1 - \frac{1}{n} \exp(-x)\right)^n, \\ \lim_{n \rightarrow \infty} F^n(c_n x + d_n) &= \exp(-e^{-x}), \end{aligned}$$

was nichts anderes bedeutet, als dass $F \in MDA(H_0)$

Bemerkung 7 (Konvergenz von Minima) Aussagen über Maxima können sofort in solche für Minima übergeführt werden, da:

$$\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n) \quad (2)$$

Man sieht leicht, dass die Konvergenz zentrierter und normierter Minima mit der Verteilungsfunktion F vorliegt, wenn $\tilde{F}(x) = 1 - F(-x)$ (die Verteilungsfunktion von $-X_1, \dots, -X_n$) im maximalen Anziehungsbereich einer Extremwertverteilung liegt. Für $M_n^* = \max(-X_1, \dots, -X_n)$ und $\tilde{F} \in MDA(H_\xi)$ folgt, dass

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n^* - d_n}{c_n} \leq x\right) = H_\xi(x)$$

und somit

$$\lim_{n \rightarrow \infty} P\left(\frac{\min(X_1, \dots, X_n) + d_n}{c_n} \leq x\right) = 1 - H_\xi(-x)$$

Die Verteilung der Minima hat also die Gestalt $1 - H_\xi(-x)$. Für symmetrische Verteilungen gilt $\tilde{F}(x) = F(x)$.

1.1.2 Maximaler Anziehungsbereich von Extremwertverteilungen

Vorhin sprachen wir ganz Allgemein über den maximalen Anziehungsbereich und was es bedeutet, wenn eine beliebige Verteilungsfunktion F darin liegt. Es ist tatsächlich so, dass die in der Praxis am häufigsten verwendeten stetigen Verteilungen in $MDA(H)$ liegen. Im folgenden wollen wir näher untersuchen welche Eigenschaften von F zu welcher speziellen Extremwertverteilung führen.

Maximaler Anziehungsbereich der Fréchet Verteilung ($MDA(H_\xi)$):

Wie hat also eine Verteilungsfunktion F auszusehen, sodass $\lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H_\xi(x)$, $\xi > 0$, gilt? Um diese Frage beantworten zu können, brauchen wir zunächst folgende Definition.

Definition 8 Eine positive, Lebesgue-messbare Funktion $L(x)$ auf \mathbb{R}^+ heißt von **langsamer Variation**, wenn gilt

$$\lim_{x \rightarrow \infty} \frac{L(cx)}{L(x)} = 1 \quad \forall c > 0$$

Jede Funktion L , die für $x \rightarrow \infty$ gegen einen positiven Grenzwert strebt, ist von langsamer Variation. Funktionen mit logarithmisches Wachstum sind ebenfalls von langsamer Variation.

Satz 9 (Fréchet MDA, Gnedenko) (ohne Beweis) Für $\xi > 0$ gilt:

$$F \in MDA(H_\xi) \Leftrightarrow 1 - F(x) = x^{-\frac{1}{\xi}} L(x) \quad (3)$$

für eine Funktion $L(x)$ von langsamer Variation.

Solche Verteilungen, wie etwa die Paretoverteilung, die Cauchyverteilung oder die Burrverteilung gehören zu den Verteilungen mit schweren Enden und sind vor allem für Anwendungen aus der Finanzmathematik von besonderem Interesse.

Maximaler Anziehungsbereich der Gumbelverteilung ($MDA(H_0)$)

Die Charakterisierung der Verteilungen in dieser Klasse gestaltet sich etwas schwieriger als im Fréchet Fall. In $MDA(H_0)$ finden sich sowohl Verteilungen mit leichten Enden als auch solche mit mittelschweren Enden (z.B. Exponentialverteilung, Normalverteilung, Lognormalverteilung usw.). Es gilt der folgende Satz.

Satz 10 Sei $x_F \leq \infty$. Dann gilt: $F \in MDA(H_0) \iff \exists z < x_F$ und $1 - F(x) = c(x) \exp\left(-\int_z^x \frac{g(t)}{a(t)} dt\right)$, wobei $g(x)$, $c(x)$ messbar und $g(x) \rightarrow 1$, $c(x) \rightarrow c > 0$ und $\lim_{x \rightarrow x_F} a'(x) = 0$ (d.h. a soll langsamer wachsen als lineare Funktionen).

Maximaler Anziehungsbereich der Weibullverteilung ($MDA(H_\xi)$) Erinnern wir uns an die Tatsache, dass die Weibullverteilung einen endlichen rechten Endpunkt $x_F < \infty$ besitzt und deswegen keine allzu große Bedeutung im Finanz und Versicherungswesen hat. Obwohl Verluste nach oben beschränkt sind, ist es in der Praxis in der Tat angenehmer, mit Verteilungen zu arbeiten, für die $x_F \leq \infty$ gilt. Eine Charakterisierung für Verteilungen die in $MDA(\Psi_\alpha)$ liegen, gibt der folgende Satz.

Satz 11 (Weibull MDA, Gnedenko) (ohne Beweis) Für $\xi < 0$ gilt:

$$F \in MDA(H_\xi) \Leftrightarrow x_F < \infty \text{ und } 1 - F\left(x_F - \frac{1}{x}\right) = x^{\frac{1}{\xi}} L(x)$$

wobei $L(x)$ von langsamer Variation.

Wir haben bereits weiter oben festgehalten, dass die in der Praxis am häufigsten verwendeten stetigen Verteilungen in $MDA(H_\xi)$ liegen für $\xi \in \mathbb{R}$ beliebig. Nun stellt sich die Frage, wie vorhandene Daten an die verallgemeinerte Extremwertverteilung angepasst werden können, bzw. welcher spezielle Fall der Extremwertverteilung mit welchen Parametern ξ, μ und σ vorliegt. Eine Möglichkeit, dies festzustellen bietet die Block-Maxima Methode.

1.1.3 Die Block-Maxima Methode

Gehen wir davon aus, dass über eine gewisse Periode hinweg Daten beobachtet werden (z.B. Returns an Börsetagen). Für die Block-Maxima Methode benötigen wir wiederholte Beobachtungen einer gleichgroßen Periode, sodass m Blöcke zu je n Beobachtungen entstehen (z.B. 10 Jahre lang werden Windgeschwindigkeiten gemessen und es entstehen 10 Blöcke zu 365 Beobachtungen). So kann über jeden der m Blöcke das Maximum $M_{n,j}, j \in \{1, \dots, m\}$ gebildet werden. Die gesammelten Daten bilden dann die Rechnungsgrundlage für die weitere Vorgehensweise. Mit mehreren Methoden können nun diese Daten an eine verallgemeinerte Extremwertverteilung angepasst werden. Für die Maximum Likelihoodmethode ist es von Vorteil, dass die m Blöcke sehr groß sind, sodass ungeachtet der Tatsache ob die Daten abhängig voneinander sind oder nicht, die Block-Maxima als unabhängig angenommen werden können, damit die Approximation der Maxima der Blöcke durch eine verallgemeinerte Extremwertverteilung genauer wird. Über die Dichte der Extremwertverteilung

$$h_{\xi, \mu, \sigma} = \frac{1}{\sigma} e^{-(1+\xi \frac{x-\mu}{\sigma})^{-\frac{1}{\xi}}} \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}-1}$$

gelangt man zur log-likelihood Funktion

$$l(\xi, \mu, \sigma; M_{n,1}, \dots, M_{n,m}) = \sum_{i=1}^m \ln h_{\xi, \mu, \sigma}(M_{n,i}) =$$

$$-m \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^m \ln \left(1 + \xi \frac{M_{ni} - \mu}{\sigma}\right) - \sum_{i=1}^m \left(1 + \xi \frac{M_{ni} - \mu}{\sigma}\right)^{-\frac{1}{\xi}},$$

wobei $\sigma > 0$ und $1 + \xi \frac{M_{n,i} - \mu}{\sigma} > 0 \forall i$. Für $\xi > -\frac{1}{2}$ sind die Konsistenz und die asymptotische Effizienz der Maximum Likelihood Schätzer gewährleistet. Eine große Anzahl n von Beobachtungen in einem Block führen zu einer genaueren Approximation der Block-Maxima durch eine verallgemeinerte Extremwertverteilung und reduzieren die Verzerrung des Schätzers des Parameters. Eine große Anzahl m von Blöcken führt zu mehr Block-Maxima und somit einer geringeren Varianz des Schätzers.

Nun hat die Block-Maxima Methode jenen großen Nachteil, dass sie sehr verschwenderisch mit vorhandenen Daten umgeht, da schließlich m Daten aus den Blöcken aus insgesamt $n*m$ Daten verwendet werden. Als Alternative bietet sich an, nur Daten zu betrachten, die extrem in dem Sinne sind, dass sie eine vorgegebene Schranke überschreiten.

1.2 Grenzüberschreitungen (threshold exceedances)

1.2.1 Verallgemeinerte Paretoverteilung

Gerade wenn es um Extremwerte geht, sind jene Bereiche einer Dichtefunktion irrelevant, in denen sich die größte Wahrscheinlichkeit befindet. Extreme Ereignisse (\Leftrightarrow große x) besitzen sehr wenig Wahrscheinlichkeit, dürfen aber wie zu Beginn erwähnt nicht vernachlässigt werden. In diesem Abschnitt geht es darum, nur den hinteren Teil einer beliebigen Verteilung zu fokussieren und den vorderen auszublenden. Dazu definieren wir zunächst die verallgemeinerte Paretoverteilung.

Definition 12 Verallgemeinerte Paretoverteilung mit der Verteilungsfunktion:

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \frac{\xi x}{\beta})^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - \exp(-\frac{x}{\beta}), & \xi = 0, \end{cases} \quad (4)$$

wobei $\beta > 0$, und $x \geq 0$ wenn $\xi \geq 0$, und $x \in [0, -\frac{\beta}{\xi}]$ wenn $\xi < 0$. ξ ist wie bei der verallgemeinerten Extremwertverteilung der Shape- bzw. Tailparameter und β der Skalenparameter. Verallgemeinert ist diese Verteilung in dem Sinn, dass sie in Abhängigkeit von ξ verschiedene Fälle birgt. Bei $\xi > 0$ handelt es sich um eine gewöhnliche Paretoverteilung mit Parametern $\alpha = \frac{1}{\xi}$ und $\kappa = \frac{\beta}{\xi}$. Für $\xi = 0$ bekommen wir eine Exponentialverteilung und für $\xi < 0$ eine verschobene Paretoverteilung. Auch hier gilt wie bei der verallgemeinerten Extremwertverteilung, dass für festes x die Stetigkeit $G_{\xi,\beta}(x) \xrightarrow{\xi \rightarrow 0} G_{0,\beta}(x)$ gegeben ist.

Weiters gilt, dass $G_{\xi,\beta}(x) \in MDA(H_\xi) \forall \xi \in \mathbb{R}$ und dass für $\xi < 1$ der Erwartungswert gegeben ist durch

$$E(X) = \frac{\beta}{1-\xi}.$$

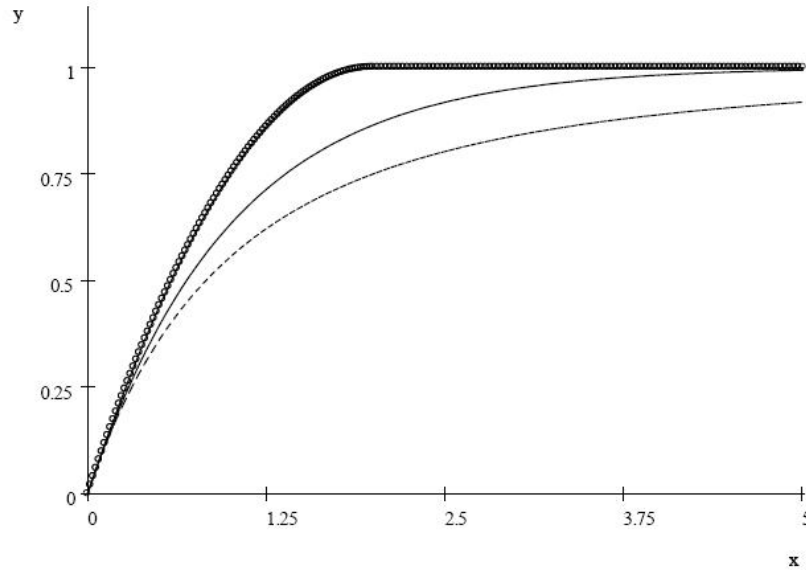


Abbildung 2: Verteilungsfunktion der verallgemeinerten Paretoverteilung mit $\xi = 0$ (durchgezeichnet), $\xi > 0$ (gestrichelt), $\xi < 0$ (gepunktet).

Definition 13 (Excess Verteilung - Excess distribution over threshold u) Sei X eine Zufallsvariable mit der Verteilungsfunktion F . Die Excess Verteilung über die Schranke u ist wie folgt definiert:

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(x + u) - F(u)}{1 - F(u)} \quad (5)$$

für $x \in [0, x_F - u)$, wobei $x_F \leq \infty$ der rechte Endpunkt von F ist.

Definition 14 (Mittlere Überschussfunktion - mean excess function) Die mittlere Überschussfunktion einer Zufallsvariable $X \sim F$ ist, falls $E[X] < \infty$ durch

$$e(u) = E[X - u | X > u] \quad (6)$$

gegeben.

In beiden Fällen wird also darauf bedingt, dass ein gewisser, vorgegebener Wert überschritten wurde. Wie ist also "der hintere Bereich" $\{X \leq x + u | X > u\}$, der ja gerade in der Extremwerttheorie im Mittelpunkt des Interesses steht, einer Verteilung verteilt? Nähere Auskunft über diese Fragestellung gibt folgender Satz.

Satz 15 (Pickands-Balkema-de Haan) (ohne Beweis) Es existiert eine positive, messbare Funktion $\beta(u)$, sodass gilt:

$$\lim_{u \rightarrow x_F} \sup_{x \in [0, x_F - u]} |F_u(x) - G_{\xi, \beta(u)}(x)| = 0 \iff F \in MDA(H_\xi), \xi \in \mathbb{R}.$$

Es ist darauf hinzuweisen, dass dieser Satz für jede Verteilungsfunktion gilt, die im maximalen Anziehungsbereich der **verallgemeinerten Extremwertverteilung** liegt ($\xi \in \mathbb{R}$ beliebig!). Je höher wir die Schranke u , beziehungsweise je näher zum rechten Endpunkt x_F wir die Schranke der zugrundeliegenden Verteilung ansetzen, umso ähnlicher wird ihre Excessverteilung der verallgemeinerten Paretoverteilung, wobei der Tailparameter ξ von H_ξ und $G_{\xi, \beta(u)}$ übereinstimmen. Eine Annahme, die sich auf den vorigen Satz bezieht, jedoch in der Praxis im Allgemeinen nicht zutrifft, die aber viele Rechnungen vereinfacht bzw. erst möglich macht ist die folgende Annahme.

Annahme 1 dass, wenn F eine Verteilungsfunktion, sodass $F \in MDA(H_\xi)$, $\xi \in \mathbb{R}$ gilt, dass dann bei vorgegebener Schranke u folgt, dass $F_u(x) = G_{\xi, \beta}(x)$ für $\beta > 0$ und $x \in [0, x_F - u)$.

Diese Annahme verwenden wir nun, um zu untersuchen, wie bei beobachteten Daten festgestellt werden kann, welcher verallgemeinerten Paretoverteilung diese gehorchen. Gehen wir also von n unabhängigen Beobachtungen (Verluste) $X_1, \dots, X_n \sim F$ aus. Naturgemäß werden nicht alle Beobachtungen die vorgegebene Schranke u überschritten haben. Wir modellieren die Anzahl der Beobachtungen, die u überschritten haben, mit einer Zufallsvariable N_u , sodass wir im weiteren Verlauf nur jene Daten beobachten, für die $X_i \geq u$ zutrifft. Unsere neuen, "gefilterten" Daten seien nun X'_1, \dots, X'_{N_u} . Die Höhe des überschrittenen Verlustes (excess loss) ergibt sich durch $Y_i = X'_i - u$. Für die Schätzung der Parameter bietet sich, wie schon bei der Block-Maxima Methode, der Maximum Likelihood Schätzer über die log-likelihood Funktion an. Es sei also

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

die Dichte der verallgemeinerten Paretoverteilung. Die log-likelihood Funktion berechnet sich demnach zu

$$\begin{aligned} \ln L(\xi, \beta; Y_1, \dots, Y_{N_u}) &= \sum_{j=1}^{N_u} \ln g_{\xi, \beta}(Y_j) \\ &= -N_u \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{j=1}^{N_u} \ln \left(1 + \xi \frac{Y_j}{\beta}\right) \end{aligned}$$

wobei $\beta > 0$ und $1 + \xi \frac{Y_j}{\beta} > 0 \forall j$. Die Lösungen der Maximum Likelihood Gleichungen liefern uns das Modell einer verallgemeinerten Paretoverteilung des Types $G_{\xi, \beta}$ für die Excess Funktion F_u .

Für manche Anwendungen im Risikomanagement ist es auch von Interesse, das Verhalten der Überschreitungen zu untersuchen, die höher als die vorgegebene Schranke u sind. Über die Erkenntnisse die wir bereits über die Excess Verteilung gesammelt haben, lassen sich gewünschte Resultate herleiten, wie zum Beispiel das nachfolgende Lemma.

Lemma 17 Unter der **Annahme 1** folgt, dass $F_v(x) = G_{\xi, \beta + \xi(v-u)}(x)$ für jede Schranke $v \geq u$.

Beweis. Wir benutzen die Definition der Excess Verteilung:

$$\begin{aligned}
\bar{F}_v(x) &= \frac{\bar{F}(v+x)}{\bar{F}(v)} \\
&= \frac{\bar{F}(u+(x+v-u))}{\bar{F}(u)} \frac{\bar{F}(u)}{\bar{F}(u+(v-u))} \\
&= \frac{\bar{F}_u(x+v-u)}{\bar{F}_u(v-u)} \\
&= \frac{\bar{G}_{\xi, \beta}(x+v-u)}{\bar{G}_{\xi, \beta}(v-u)} \\
&= \bar{G}_{\xi, \beta + \xi(v-u)}(x).
\end{aligned}$$

□

Das heißt also, dass die Excess Verteilung eine verallgemeinerte Paretoverteilung mit demselben Tailparameter ξ bleibt, jedoch der Skalenparameter in v linear wächst. Unter der Voraussetzung $\beta < 1$, folgt für die mittlere Überschussfunktion

$$e(v) = \frac{\beta + \xi(v-u)}{1-\xi} = \frac{\xi v}{1-\xi} + \frac{\beta - \xi u}{1-\xi},$$

wobei $v \in [u, \infty)$ für $\xi \in [0, 1)$, und $v \in [u, u - \frac{\beta}{\xi}]$ für $\xi < 0$.

1.2.2 Modellierung und Messung von Verteilungsenden

Wir widmen uns in diesem Abschnitt der Schätzung von Verteilungsenden, um plausible Berechnungen und Messungen zu ermöglichen, die in der Finanzmathematik (*Value at Risk, Expected Short Fall*) von erheblicher Bedeutung sind. Grundlage unserer Berechnungen ist wieder die **Annahme 1**, die uns erlaubt, die Excess Verteilung $F_u(x)$ einer Verteilung, die im maximalen Anziehungsbereich einer verallgemeinerten Extremwertverteilung liegt, durch eine verallgemeinerte Paretoverteilung zu approximieren. Zunächst halten wir folgende Beziehung fest:

Beziehung 18 Bei vorgegebener Schranke u gilt für $x \geq u$ die Wahrscheinlichkeit des Verteilungsendes:

$$\begin{aligned}
\bar{F}(x) &= P(X > u)P(X > x|X > u) \\
&= \bar{F}(u)P(X - u > x - u|X > u) \\
&= \bar{F}(u)\bar{F}_u(x - u) \\
&= \bar{F}(u)(1 + \xi \frac{x-u}{\beta})^{-\frac{1}{\xi}}
\end{aligned}$$

wenn $F(u)$ bekannt ist. Diese Beziehung hilft uns auch, über

$F(x) = -\bar{F}(u)(1 + \xi \frac{x-u}{\beta})^{-\frac{1}{\xi}} + 1$ durch Berechnen der Umkehrfunktion auf die

Quantile zu kommen, die ja in der Praxis als *Value at Risk* verwendet werden. Für $\alpha \geq F(u)$ erhalten wir

$$VaR_\alpha = q_\alpha(F) = u + \frac{\beta}{\xi} \left(\left(\frac{1-\alpha}{\bar{F}(u)} \right)^{\frac{1}{\xi}} - 1 \right). \quad (7)$$

Für den *Expected Short Fall* erhalten wir

$$ES_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 q_x(F) dx = \frac{VaR_\alpha}{1-\xi} + \frac{\beta - \xi u}{1-\xi} \quad (8)$$

Der *Expected Short Fall* kann andererseits auch durch $ES_\alpha = VaR_\alpha + e(VaR)$ ermittelt werden, wobei $e(VaR)$ die mittlere Überschussfunktion darstellt. Richten wir noch einmal den Blick auf das Verhältnis dieser Risikomaße, so stellen wir fest, dass dieses, je höher wir α ansetzen, allein vom Tailparameter ξ abhängt. Es ist:

$$\lim_{\alpha \rightarrow 1} \frac{ES_\alpha}{VaR_\alpha} = \begin{cases} (1-\xi)^{-1}, & \xi \geq 0 \\ 1, & \xi < 0 \end{cases}$$

Wie ist nun die vorgehensweise in der Praxis? Zunächst werden die vorhandenen Daten an einer Grenze u , wie vorhin schon beschrieben, abgeschnitten und unter der **Annahme 1** an eine verallgemeinerte Paretoverteilung angepasst. In den Formeln für *VaR* und *Expected Shortfall* werden die Parameter β und ξ durch ihre Schätzer $\hat{\beta}$ und $\hat{\xi}$ ersetzt. Der (empirische) Schätzer für $\bar{F}(u)$ ergibt sich aus $\frac{N_u}{n}$, wobei hier auf eine große Menge von Daten gehofft werden muss. Es ist der Schätzer für die Wahrscheinlichkeit des Endes der Verteilung für $x \geq u$:

$$\hat{\bar{F}}(x) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x-u}{\hat{\beta}} \right)^{-\frac{1}{\hat{\xi}}} \quad (9)$$

1.2.3 Der Hill Schätzer

Wir haben im vorigen Abschnitt gesehen, dass durch die Anpassung an eine verallgemeinerte Paretoverteilung Aussagen über Wahrscheinlichkeiten von Enden von Verteilungen möglich wurden. Der Hill-Schätzer stellt eine weitere Methode zur Modellierung von Verteilungsenden dar. Hier geht es dabei ausschließlich um Verteilungen mit schweren Enden (heavy tailed distributions), also solche Verteilungen, die in $MDA(H_\xi)$ für $\xi > 0$ liegen und deren Ende als

$$\bar{F}(x) = L(x)x^{-\alpha},$$

wobei $L(x)$ von einer Funktion von langsamer Variation ist und $\alpha > 0$, geschrieben werden kann.

Der Hill Schätzer kann auf mehrere Arten berechnet werden. Eine Möglichkeit führt über die Mittlere Überschussfunktion von $\ln X$, wobei $X \sim F$ mit $\bar{F}(x) = L(x)x^{-\alpha}$. Im folgenden bezeichnet e^* die mittlere Überschussfunktion von $\ln X$. Mit partieller Integration erhält man:

$$\begin{aligned} e^*(\ln u) &= E(\ln X - \ln u | \ln X > \ln u) \\ &= \frac{1}{\bar{F}(u)} \int_u^\infty (\ln x - \ln u) dF(x) \\ &= \frac{1}{\bar{F}(u)} \int_u^\infty \frac{\bar{F}(x)}{x} dx \\ &= \frac{1}{\bar{F}(u)} \int_u^\infty L(x)x^{-(\alpha+1)} dx \end{aligned}$$

Für u hinreichend groß kann $L(x)$ für $x \geq u$ als Konstante vor dem Integral geschrieben werden. Für $u \rightarrow \infty$ erhalten wir

$$e^*(\ln u) \sim \frac{L(u)u^{-\alpha}\alpha^{-1}}{\bar{F}(u)} = \alpha^{-1},$$

also ist $\lim_{u \rightarrow \infty} \alpha e^*(\ln u) = 1$. Für beobachtete und absteigend (!) geordnete Werte $X_{n,n}, \dots, X_{1,n}$ folgt (für großes n und geeignetes k), dass $e^*(\ln X_{k,n}) \approx \alpha^{-1}$, wobei $X_{k,n} = u$ die vorgegebene Schranke und, dass $\hat{\alpha}^{-1} = ((k-1)^{-1} \sum_{j=1}^{k-1} \ln X_{j,n} - \ln X_{k,n})$. Durch eine kleine Änderung ergibt sich der klassische **Hill Schätzer**:

$$\hat{\alpha}_{k,n}^{(H)} = \left(\frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1}, \quad 2 \leq k \leq n \quad (10)$$

Bemerkung 19 Geeignetes k bedeutet, dass es nicht zu klein sein darf, denn dann wird die Varianz des Schätzers $\frac{\alpha^2}{k}$ zu groß. Andererseits darf k nicht allzu groß sein, denn es führt dazu, dass unsere Annahme, dass $L(x)$ für $x \geq u$ konstant ist, nicht ausreichend erfüllt ist, sodass die Verzerrung des Schätzers wächst.

Basierend auf dem Hill-Schätzer erhalten wir Schätzer für die Wahrscheinlichkeit $\bar{F}(x)$. Wir betrachten dabei ein Modell mit $\bar{F}(x) = Cx^{-\alpha}$, $x \geq u > 0$ für eine hohe vorgegebene Schranke u . Wir ersetzen die Funktion von langsamer Variation durch eine Konstante C . Für geeignetes k ergibt sich der Tail-Index α über den Hill Schätzer $\hat{\alpha}_{k,n}^{(H)}$. Für die Schranke u verwenden wir den k -ten Wert der absteigend geordneten Teststatistik $X_{k,n}$. Bleibt also die Konstante C , die es zu schätzen gilt. Da sich C als $u^\alpha \bar{F}(u)$ schreiben lässt, schreiben wir das Problem um in: $\bar{F}(x) = u^\alpha \bar{F}(u) x^{-\alpha}$ und das Schätzen von $\bar{F}(x)$ wird zu einem äquivalenten Problem. Der natürliche Schätzer für $\bar{F}(u)$ ist $\frac{k}{n}$, also erhalten wir

insgesamt einen Schätzer:

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

Invertieren wir diesen Schätzer für die Verteilungsfunktion, so erhalten wir einen Schätzer für das q -Quantil x_q mit $q \approx 1$:

$$\hat{x}_q = X_{k,n} \left(\frac{n}{k} (1 - q) \right)^{-\frac{1}{\hat{\alpha}_{k,n}^{(H)}}}$$

Sowohl der Hill-Schätzer für den Tail-Index α als auch der daraus motivierte Schätzer für das Verteilungsende sind eine Funktion der k -ten Ordnungsstatistik bei Stichprobenumfang n . Erinnern wir uns an den Schätzer (9) über die verallgemeinerte Paretoverteilung für das Verteilungsende, so war dieser eine Funktion der Anzahl der Beobachtungen, die die Schranke u überschritten haben. Diese Anzahl haben wir mit der Zufallsvariablen N_u modelliert. Beide Schätzer lassen sich durch Substitutionen zusammenfügen. Schreiben wir $\hat{\xi}^{(H)}$ statt $\frac{1}{\hat{\alpha}_{k,n}^{(H)}}$, u statt $X_{k,n}$, und k für N_u so erhalten wir:

$$\hat{\bar{F}}(x) = \frac{N_u}{n} \left(1 + \xi^{(H)} \frac{x - u}{\xi^{(H)} u} \right)^{-\frac{1}{\hat{\xi}^{(H)}}} \quad (11)$$

1.3 Tails ausgesuchter Verteilungen

1.3.1 Anziehungsbereich der Fréchetverteilung

Vorhin wurde bereits gezeigt, dass im maximalen Anziehungsbereich der Fréchetverteilung, solche Verteilungen liegen, die als $\bar{F} = x^{-\alpha} L(x)$ geschrieben werden können, wobei $\alpha > 0$ und $L(x)$ eine Funktion von langsamer Variation ist. Die zentrierten und normierten Maxima von Stichproben, die zu einer solchen Verteilung gehören, konvergieren gegen eine Fréchetverteilung mit dem Parameter $\xi = \frac{1}{\alpha}$. Die Excess-Verteilungen für vorgegebene Schranken u konvergieren für solche gegen die verallgemeinerte Paretoverteilung mit dem Parameter $\xi = \frac{1}{\alpha}$. Wir betrachten nun auch die Student t Verteilung und die Gammaverteilung und zeigen, dass diese ebenfalls in der Klasse der vorhin beschriebenen Verteilungen liegen.

Beispiel 20 (Student-t-Verteilung) Nachdem die Dichte der Student-t-Verteilung auf die Form $f_\nu(x) = x^{-(\nu+1)} L(x)$ gebracht werden kann, folgt aus dem Satz von Karamata, dass für die Wahrscheinlichkeit des Endes $\bar{F}(x) = \int_x^\infty f_\nu(y) dy$ die Funktion $L(y)$ als Konstante behandelt werden kann und somit aus dem Integranden herausgeschrieben werden kann. Wir erhalten

$$\bar{F}_\nu(x) = \int_x^\infty y^{-(\nu+1)} L(y) dy \sim \nu^{-1} x^{-\nu} L(x), \quad x \rightarrow \infty,$$

woraus wir nun schließen, dass die Verteilungsfunktion F_ν der Student-t- Verteilung den Tail-Index ν hat und dass $F_\nu \in MDA(H_{\frac{1}{\nu}})$

Beispiel 21 (inverse Gammaverteilung) Ähnlich wie bei der Student-t- Verteilung kann die Dichte

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{x}\right)^{\alpha+1} \exp\left(-\frac{\beta}{x}\right), \quad x > 0,$$

der inversen Gammaverteilung auf die Form $f_{\alpha,\beta}(x) = x^{-(\alpha+1)}L(x)$ gebracht werden. Nun folgt aus der gleichen Argumentation wie im vorigen Beispiel, dass $F_{\alpha,\beta}$ den Tail-Index α hat und somit $F_{\alpha,\beta} \in MDA(H_{\frac{1}{\alpha}})$.

1.3.2 Anziehungsbereich der Gumbel Verteilung

Die Charakterisierung der Verteilungen, die im Anziehungsbereich der Gumbel Verteilung liegen, läuft über die von-Mises Verteilungsfunktion, die eine Klasse von Verteilungen enthält. Der Anziehungsbereich besteht dann aus jenen Verteilungen, die in dieser Klasse liegen, bzw. aus Funktionen, die zu diesen *tail-äquivalent* sind. Im folgenden wird näher darauf eingegangen, jedoch vorher nochmal in Erinnerung gerufen, dass sämtliche Verteilungen in dieser Klasse einen endlichen oder unendlichen rechten Endpunkt $x_F \leq \infty$ besitzen können.

Definition 22 (von Mises Verteilungsfunktion) Sei $z < x_F$, sodass für F gilt

$$\bar{F}(x) = c \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\}, \quad z < x < x_F$$

wobei $c > 0$ und $a(t)$ eine positive, gleichmäßig stetige Funktion mit $\lim_{x \rightarrow x_F} a'(x) = 0$. Dann ist F eine von Mises Verteilungsfunktion.

Definition 23 (Tail-Äquivalenz) Zwei Verteilungsfunktionen F und G sind tail-äquivalent, wenn sie dieselben rechten Endpunkte x_F und x_G besitzen und wenn gilt:

$$\lim_{x \rightarrow x_F} \frac{\bar{F}(x)}{\bar{G}(x)} = c, \quad c \in (0, \infty)$$

Wie lässt sich nun herausfinden ob eine zugrunde liegende Verteilungsfunktion F eine von Mises Verteilungsfunktion ist? Folgende Bedingung gibt darüber Aufschluss:

Bedingung 24 Sei $z < x_F$ so, dass F zwei mal differenzierbar auf (z, x_F) ist mit der Dichte $f = F'$ und $F'' < 0$ auf (z, x_F) . F ist eine von Mises Verteilungsfunktion genau dann, wenn

$$\lim_{x \rightarrow x_F} \frac{\bar{F}(x)F''(x)}{f^2(x)} = -1.$$

Beispiel 25 (Gamma Verteilung) Sei F gammaverteilt mit der Dichte $f_{\alpha,\beta} = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x}$ für $x \geq 0$. Eine einfache Rechnung liefert, dass $F''(x) = f'(x) = -f(x)(\beta + \frac{1-\alpha}{x}) < 0$ vorausgesetzt $x > \max((\frac{\alpha-1}{\beta}), 0)$. Mit der Regel von de L'Hôpital erhalten wir, dass $\lim_{x \rightarrow \infty} \frac{\bar{F}}{f(x)} = \lim_{x \rightarrow \infty} \frac{-f(x)}{f'(x)} = \frac{1}{\beta}$. Obige Bedingung 24 ist nun nach Zusammensetzen dieser beiden Rechnungen erfüllt, und es folgt, dass die Gammaverteilung in der Klasse der von Mises Verteilungen liegt.

1.4 Modelle für stochastische Prozesse

Im folgenden Abschnitt geht es darum, die Grenzüberschreitungen aus einer etwas anderen Perspektive zu betrachten. Ging es bis jetzt darum, gute Modelle bzw. Verteilungen für die Überschreitungen zu finden und Aussagen über die Höhe zu treffen, so schauen wir uns nun an, wie häufig das Ereignis "Grenzüberschreitung" überhaupt auftritt und ob es dafür ein geeignetes mathematisches Modell gibt, an welches wir Beobachtungen anpassen können. Die Höhe der Grenzüberschreitung ist somit fürs erste eher irrelevant. Die Modellierung erfolgt über stochastische Prozesse, genauer mit einem white noise prozess.

Definition 26 White noise Prozess: Ein stochastischer Prozess $(\epsilon_t)_{t \in \mathbb{N}}$ ist ein white noise prozess, wenn für alle $t \in \mathbb{N}$ gilt:

$$E(\epsilon_t) = 0, \quad E(\epsilon_t^2) =: \sigma^2 < \infty, \quad \forall h > 0 : Var(\epsilon_{t+h}, \epsilon_t) = 0$$

Betrachten wir nun solch einen white noise Prozess $(X_t)_{t \in \mathbb{N}}$, der uns Verluste modelliert, deren Erwartungswerte in den meisten Anwendungen der Finanzmathematik gleich null gefordert werden und setzen voraus, dass die gemeinsame Verteilungsfunktion des Prozess in $MDA(H_\xi)$ liegt. Für festes x sei $u_n(x) := c_n x + d_n$ eine Folge von Schranken.

Wir gehen nun von einer Stichprobe X_1, \dots, X_n aus, wobei die Beobachtungen unabhängig und identisch verteilt sind. Die Wahrscheinlichkeit einer Grenzüberschreitung der i -ten Beobachtung ist $P(X_i > u_n(x)) = \bar{F}(u_n(x))$ bzw. 0 wenn keine Grenzüberschreitung vorliegt. Es handelt sich hierbei also um eine Bernoulli-Zufallsvariable. Für die gesamte Stichprobe ist die Anzahl der Grenzüberschreitungen $N_{u_n(x)}$ binomialverteilt $N_{u_n(x)} \sim B(n, \bar{F}(u_n(x)))$ mit dem Erwartungswert $n\bar{F}(u_n(x))$, der sich für $n \rightarrow \infty$ durch eine Poissonverteilung approximieren lässt und sich wie folgt berechnet:

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) = n(1 - F(c_n x + d_n))$$

Wegen $-\ln(1-y) \sim y$ für $y \rightarrow 0$ und $y := 1 - F(c_n x + d_n)$ folgt, dass

$$\begin{aligned} & \lim_{n \rightarrow \infty} -n \ln(1 - (1 - F(c_n x + d_n))) \\ &= \lim_{n \rightarrow \infty} -n \ln(F(c_n x + d_n)) \\ &= \lim_{n \rightarrow \infty} -\ln(F^n(c_n x + d_n)) \\ &= -\ln H_\xi(x) \end{aligned}$$

Die Anzahl der Grenzüberschreitungen $N_{u_n(x)}$ einer Stichprobe konvergiert also gegen eine Poissonverteilung mit dem Parameter $\lambda(x) = -\ln H_\xi(x)$. Tatsächlich kann auch die Ankunftszeit der Grenzüberschreitung durch einen Poissonprozess modelliert werden. Hier definieren wir zunächst einen Zählprozess.

Definition 27 Zählprozess. Gehen wir von einer Folge von stochastischen Größen Y_1, \dots, Y_n aus dem Zustandsraum Y (z.B. \mathbb{R} oder \mathbb{R}^2) aus, so ist der Zählprozess über der Menge $A \subset Y$ definiert als

$$N(A) = \sum_{i=1}^n I_{\{Y_i \in A\}}$$

Es handelt sich also um eine positive, nicht fallende Treppenfunktion, die die Anzahl der Ereignisse zählt, die sich in einem vorgegebenen Bereich abspielen. Intuitiv ist es zunächst naheliegend $N(A)$ als Anzahl der Grenzüberschreitungen zu bezeichnen und die Menge $A \subset Y$ mit $(u, \infty] \subset \mathbb{R}$ zu versehen. Betrachten wir zunächst einen wichtigen Zählprozess, nämlich den Poissonprozess.

Definition 28 Poissonprozess Der Zählprozess $N(\cdot)$ ist ein Poissonprozess auf Y mit der Mittelwertfunktion Λ , wenn:

- Für $A \subset Y$, $k \geq 0$ und $0! := 1$

$$P(N(A) = k) = \begin{cases} e^{-\Lambda(A)} \frac{\Lambda(A)^k}{k!}, & \Lambda(A) < \infty \\ 0, & \Lambda(A) = \infty \end{cases}$$

- Für $m \geq 1$ und für paarweise disjunkte Mengen $A_1, \dots, A_m \subset Y$ sind die Zufallsvariablen $N(A_1), \dots, N(A_m)$ unabhängig.

Für die Mittelwertfunktion $\Lambda(\cdot)$ des Poissonprozesses $N(\cdot)$ gilt, dass $E(N(A)) = \Lambda(A)$. Hierbei bezeichnet $\lambda(x)$ die Intensität des Poissonprozesses und ist die Ableitung der Mittelwertfunktion, sodass $\Lambda(A) = \int_A \lambda(x) dx$ gilt.

Gehen wir nun zurück zu unserem white noise Prozess $(X_i)_{i \in \mathbb{N}}$ und unserer Folge von Schranken $u_n(x) = c_n x + d_n$ für festes x . Für $n \in \mathbb{N}$ und $1 \leq i \leq n$ sei $Y_{i,n} = \frac{i}{n} I_{\{X_i > u_n(x)\}}$. $Y_{i,n}$ ist sozusagen der relative Zeitpunkt einer Grenzüberschreitung, sodass sich unser Zustandsraum zu $Y = (0, 1]$ reduziert. Der Prozess, der nun unsere Grenzüberschreitungen aus einer Stichprobe mit Umfang n fählt ist durch

$$N_n(A) = \sum_{i=1}^n I_{\{Y_{i,n} \in A\}}$$

gegeben.

Für $n \rightarrow \infty$ konvergiert dieser Zählprozess $N_n(\cdot)$ in Verteilung gegen einen Poissonprozess mit der Mittelwertfunktion $\Lambda(\cdot)$, für die gilt, dass $\Lambda(A) = (t_2 - t_1)\lambda(x)$, wobei $A = (t_1, t_2) \subset Y$ und $\lambda(x) = -\ln H_\xi(x)$. Daraus folgt auch die Konvergenz für den Erwartungswert, sodass $E(N_n(A)) = E(N(A)) = \Lambda(A) = (t_2 - t_1)\lambda(x)$. Da die Intensität nicht von der

Zeit abhängt, wird sie als Konstante $\lambda := \lambda(x)$ gekennzeichnet. Der Grenzprozess der Anzahl der Grenzüberschreitungen ist also ein homogener Poissonprozess. Dass dieser homogene Poissonprozess zur Modellierung von Grenzüberschreitungen in der Praxis nicht immer zutrifft, soll uns später noch einmal beschäftigen.

1.4.1 Das POT-Modell

Im „Peak over Threshold“ Modell geht es darum, die Theorie der vorigen Abschnitte zu vereinen. Dabei gehen wir von identisch und unabhängig verteilten Beobachtungen X_1, \dots, X_n aus und betrachten lediglich die Daten, die für unsere Betrachtungen relevant sind, also solche, die eine vorgegebene Schranke (threshold) u überschreiten. Im POT-Modell treffen wir folgende Voraussetzungen:

- Das Eintreffen des Ereignisses „Grenzüberschreitung“ gehorcht einem homogenen Poissonprozess.
- Die Höhe der Grenzüberschreitung gehorcht einer verallgemeinerten Paretoverteilung.
- Die Höhe der Grenzüberschreitung ist unabhängig vom Zeitpunkt der Grenzüberschreitung

Im vorigen Abschnitt war das Auftreten und das Zählen des Ereignisses „Grenzüberschreitung“ von Bedeutung. Im POT Modell kommt also noch eine Dimension, nämlich die Höhe der Überschreitung hinzu und wir ziehen einen zweidimensionalen Poissonprozess über einen erweiterten Definitionsbereich (t, x) heran, der uns Zeitpunkt t und Höhe x der Grenzüberschreitung modelliert.

Betrachten wir also ein POT Modell mit der Stichprobe X_1, \dots, X_n und geben uns eine hohe Schranke u vor, sodass der Zustandsraum des Zählprozess sich zu $Y = U \times V$ erweitert, wobei U den relativen Zeitpunkt aus $(0, 1]$ und V die Höhe der Überschreitung aus (u, ∞) darstellt. Für $A \subset Y$ ist der zugehörige Poissonprozess

$$N(A) = \sum_{i=1}^n I_{\{(\frac{i}{n}, X_i) \in A\}}$$

mit der Intensität auf (t, x)

$$\lambda(t, x) = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi} - 1}$$

für $1 + \xi \frac{x - \mu}{\sigma} > 0$, andernfalls $\lambda(t, x) = 0$. Diese Intensität hängt nur von x ab und ist nicht anderes als die Dichte einer verallgemeinerten Paretoverteilung und wir schreiben $\lambda(x) := \lambda(t, x)$. Für eine Teilmenge $(t_1, t_2) \times (x, \infty) = A \subset Y$ berechnet sich die Mittelwertfunktion als Integral über die Intensität zu:

$$\Lambda(A) = \int_{t_1}^{t_2} \int_x^{\infty} \lambda(y) dy dt = -(t_2 - t_1) \ln H_{\xi, \mu, \sigma}(x)$$

Für eine Schranke $x \geq u$ folgt nun, dass der eindimensionale Prozess der Werte, die x überschreiten, ein Poissonprozess mit der Intensität $\tau(x) := -\ln H_{\xi, \mu, \sigma}(x)$ ist. So lässt sich nun auch die Gegenwahrscheinlichkeit $\bar{F}_u(x)$ für Grenzüberschreitungen über die Schranke u leicht berechnen durch die Anzahl der Überschreitungen über die Schranke $x + u$ und u :

$$\bar{F}_u(x) = \frac{\tau(x+u)}{\tau(u)} = \left(1 + \frac{\xi x}{\sigma + \xi(u - \mu)}\right)^{-\frac{1}{\xi}} = \bar{G}_{\xi, \beta}(x)$$

wobei $\beta = \sigma + \xi(u - \mu)$ ein positiver Skalenparameter. Genau diese Tatsache haben wir bereits im vorigen Abschnitt bei der Modellierung von Grenzüberschreitungen gezeigt, als wir den Satz von Pickands-Balkema-de Haan zitiert haben. Dieses Modell enthält aber auch die Ergebnisse aus dem ersten Abschnitt über die verallgemeinerte Extremwertverteilung. Betrachten wir zunächst das Ereignis $\{M_n \leq x\}$ für ein $x \geq u$. Für einen Zählprozess ist dieses Ereignis nichts anderes als $P(N(A) = 0)$, da keine Grenzüberschreitung für $\{M_n \leq x\}$ vorliegt. Es folgt

$$P(N(A) = 0) = \exp(-\Lambda(A)) = H_{\xi, \mu, \sigma}(x), x \geq u,$$

was ja genau die verallgemeinerte Extremwertverteilung ist.

Haben wir in der Praxis nun Beobachtungen $\{\tilde{X}_j : j = 1, \dots, N_u\}$ vorliegen, welche eine vorgegebene Schranke überschritten haben, so stellt sich nun die Frage wie diese an ein POT Modell angepasst werden können. Die Likelihood Funktion eines POT Modells mit der Intensität $\lambda(t, x) = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi} - 1}$ ist

$$L(\xi, \sigma, \mu; \tilde{X}_1, \dots, \tilde{X}_{N_u}) = \exp(-\tau(u)) \prod_{j=1}^{N_u} \lambda(\tilde{X}_j)$$

Das Maximieren dieses Ausdrucks liefert die gewünschten Schätzer für ξ, σ und μ .

Eine alternativer Weg über die Loglikelihoodfunktion zu den Schätzern erhält man, wenn man das POT Modell umparametrisiert. Schreiben wir für die Intensität des eindimensionalen Poissonprozesses der Grenzüberschreitungen über die Schranke u $\tau := \tau(u) = -\ln H_{\xi, \mu, \sigma}(u)$ und $\beta = \sigma + \xi(u - \mu)$ für den Skalenparameter, so erhalten wir für die Intensität $\lambda(t, x)$

$$\lambda(x) = \lambda(t, x) = \frac{\tau}{\beta} \left(1 + \xi \frac{x - u}{\beta}\right)^{-\frac{1}{\xi} - 1},$$

wobei $\xi \in \mathbb{R}$ beliebig und $\tau, \beta > 0$. Für diese alternative Parametrisierung lässt sich die logarithmierte Likelihoodfunktion $L(\xi, \sigma, \mu; \tilde{X}_1, \dots, \tilde{X}_{N_u})$ nun schreiben als

$$\ln L(\xi, \sigma, \mu; \tilde{X}_1, \dots, \tilde{X}_{N_u}) = \ln L_1(\xi, \beta; \tilde{X}_1 - u, \dots, \tilde{X}_{N_u} - u) + \ln L_2(\tau; N_u),$$

wobei der erste Term nichts anderes ist als das vorhin schon erwähnte Schätzverfahren, um vorhandene Daten an eine verallgemeinerte Paretoverteilung anzupassen, und $\ln L_2(\tau; N_u) = -\tau + N_u \ln \tau$ die Log-Likelihoodfunktion für einen eindimensionalen homogenen Poissonprozess mit der Intensität τ .

1.4.2 Self-exciting POT Modell

Bisher gingen wir beim POT Modell von einem homogenen Poissonprozess aus. Allerdings kommt es in der Praxis häufig vor, dass das Auftreten einer Grenzüberschreitung, das Auftreten einer weiteren Grenzüberschreitung verursacht oder wahrscheinlicher macht, wie es zum Beispiel bei Erdbeben der Fall ist. Nach dem Auftreten eines Erdbebens ist es in der Zeit unmittelbar danach eher wahrscheinlich, dass die Erde erneut bebt und wir sprechen in dem Fall von einem Nachbeben. In der Finanzmathematik kann das zum Beispiel so interpretiert werden, dass ein extrem hoher Verlust einer Aktie bedeutet, dass auch einer anderen Aktie eine Grenzüberschreitung droht, z.B. kann eine ganze Branche von einer Finanzkrise betroffen sein, und es wird in einem bestimmten Zeitraum häufiger zu Grenzüberschreitungen kommen. In diesem Fall sprechen wir von einem self-exciting Prozess (*self-exciting process*).

Liegen konkrete Daten X_1, \dots, X_n und eine konkrete Schranke u vor, so gehen wir erneut davon aus, dass N_u Grenzüberschreitungen vorliegen. Erfassen wollen wir das Ganze mit der Menge $\{(i, X_i) : 1 \leq i \leq n, X_i > u\}$, und bemerken gleichzeitig, dass wir diesmal nicht vom relativen Zeitpunkt der Grenzüberschreitung ausgehen, wie wir es im POT Modell gemacht haben, sondern tatsächlich die natürliche Zeitskala der Beobachtungen übernehmen. Eine alternative Erfassung der Grenzüberschreitungen ist über die Menge $\{(T_j, \tilde{X}_j) : j = 1, \dots, N_u\}$, die diese nacheinander beschreibt.

Zunächst betrachten wir ein Modell, das lediglich den Zeitpunkt der Grenzüberschreitung liefert. Dies ist der Zählprozess $Y_i = iI_{\{X_i > u\}}$, der den Zeitpunkt (beim POT Modell war es der relative Zeitpunkt $\frac{i}{n}$) der Grenzüberschreitung liefert, falls einer auftritt, und 0 andernfalls. Der Zählprozess aller Überschreitungen ist der Prozess $N(\cdot)$ im Zustandsraum $Y = (0, n]$ der durch

$$N(A) = \sum_{i=1}^n I_{\{Y_i \in A\}}$$

für $A \subset Y$ gegeben ist und wir setzen voraus, dass dieser ein selbstständiger Zählprozess mit der *bedingten Intensität*

$$\lambda^*(t) = \tau + \psi \sum_{j: 0 < T_j < t} h(t - T_j, \tilde{X}_j - u),$$

ist, wobei $\tau > 0, \psi \geq 0$ und h eine nichtnegative Funktion ist. Jede Grenzüberschreitung (T_j, \tilde{X}_j) vor dem aktuellen Zeitpunkt t wird also in der bedingten Intensität berücksichtigt, wobei der Grad der Berücksichtigung sowohl von der abgelaufenen Zeit $(t - T_j)$ (weiter zurückliegende Grenzüberschreitungen werden

nicht so stark gewichtet werden) als auch von der Höhe ($\tilde{X}_j - u$) der Grenzüberschreitung abhängen kann.

Die Funktion h kann zum Beispiel von folgender Gestalt sein:

- $h(s, x) = \exp(\delta x - \gamma s)$, wobei $\delta, \gamma > 0$ oder
- $h(s, x) = \exp(\delta x)(s + \gamma)^{-(\rho+1)}$, wobei $\delta, \gamma, \rho > 0$.

Alle benötigten Parameter als θ zusammengefasst impliziert folgende Gestalt für die Formel der Likelihoodfunktion:

$$L(\theta; (T_j, \tilde{X}_j)) = \exp\left(-\int_0^n \lambda^*(s) ds\right) \prod_{i=1}^{N_u} \lambda^*(T_i),$$

1.4.3 Konstruktion eines self-exciting POT-Modells

Ziel dieses Abschnittes ist es, das POT Modell, das im Abschnitt 1.4.1 eingeführt wurde, nach und nach zu einem self-exciting Prozess zu verallgemeinern. Dabei gehen wir zunächst von einem selbstständigen POT Modell aus, dessen Ereignisse einer verallgemeinerten Paretoverteilung gehorchen, und *unvorhersehbar* in dem Sinne sind, dass die Ereignisse unabhängig voneinander sind. Für *vorhersehbare* Ereignisse verwenden wir erneut eine verallgemeinerte Paretoverteilung, diesmal allerdings mit einer bedingten Dichte, die sämtliche Informationen über bisherige Grenzüberschreitungen bis zum jetzigen Zeitpunkt enthält. Diese Zusatzinformationen enthält der Skalenparameter β . Und so gelangen wir zu einem Modell, welches Grenzüberschreitungen und deren Höhe in Zeiten von Krisen temporär wahrscheinlicher macht.

Diese Modelle sind also Zählprozesse $N(\cdot)$ auf dem Zustandsraum $Y = (0, n] \times (u, \infty)$, sodass $N(A) = \sum_{i=1}^n I_{\{(i, X_i) \in A\}}$ für $A \subset Y$. Weiters verwenden wir das unparametrisierte POT Modell, wo Grenzüberschreitungen einem Poissonprozess mit der Intensität τ , und die Höhe der Grenzüberschreitung einer verallgemeinerten Paretoverteilung $G_{\xi, \beta}$ gehorchen.

Modell für unvorhersehbare Ereignisse Als

$v^*(t) = \sum_{j: 0 < T_j < t} h(t - T_j, \tilde{X}_j - u)$ bezeichnen wir die self-exciting Funktion (*self-excitement function*), wobei h nichtnegativ, sodass sich die unparametrisierte Form der Intensität $\lambda(t, x) = \frac{\tau}{\beta} (1 + \xi \frac{x-u}{\beta})^{-\frac{1}{\xi}-1}$ zu

$$\lambda^*(t, x) = \frac{\tau + \psi v^*(t)}{\beta} \left(1 + \xi \frac{x-u}{\beta}\right)^{-\frac{1}{\xi}-1}$$

auf dem Zustandsraum $Y = (0, n] \times (u, \infty)$, wobei $\tau > 0$ und $\psi \geq 0$, verallgemeinern lässt. Die bedingte Intensität eines selbstständigen POT Modells wird also mit der Dichtefunktion einer verallgemeinerten Paretoverteilung kombiniert. Die

bedingte Anzahl der Grenzüberschreitungen $x \geq u$ zum Zeitpunkt t berechnet sich nun zu

$$\tau^*(t, x) = \int_x^\infty \lambda^*(t, y) dy = (\tau + \psi v^*(t)) \left(1 + \xi \frac{x - u}{\beta}\right)^{-\frac{1}{\xi}}.$$

Das Berechnen der Verteilung für die Grenzüberschreitung ergibt sich nun wieder aus dem Quotienten der Intensitäten für die Schranke $u + x$ und u

$$\frac{\tau^*(t, u + x)}{\tau^*(t, u)} = \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}} = \bar{G}_{\xi, \beta}(x)$$

und ist unabhängig von t .

Vorhandene Daten können durch das Maximieren einer Likelihoodfunktion, die die Gestalt

$$L(\theta; (T_j, \tilde{X}_j)) = \exp\left(-n\tau - \psi \int_0^n v^*(s) ds\right) \prod_{j=1}^{N_u} \lambda^*(T_j, \tilde{X}_j)$$

annimmt, an dieses Modell angepasst werden.

Modell für vorhersehbare Ereignisse Durch erneutes Verallgemeinern der bedingten Intensität $\lambda^*(t, x)$ aus dem Modell für unvorhersehbare Ereignisse erhält man

$$\lambda^*(t, x) = \frac{\tau + \psi v^*(t)}{\beta + \alpha v^*(t)} \left(1 + \xi \frac{x - u}{\beta + \alpha v^*(t)}\right)^{-\frac{1}{\xi} - 1},$$

wobei $\beta > 0$ und $\alpha \geq 0$. Die bedingte Anzahl der Grenzüberschreitungen $x \geq u$ zum Zeitpunkt t ist dieselbe wie im unvorhersehbaren Modell, allerdings ist der Parameter β durch die zeitabhängige self-exciting Funktion $\beta + \alpha v^*(t)$ ersetzt. Die Verteilung für eine Grenzüberschreitung bedingt auf vergangene Überschreitungen und darauf, dass zum Zeitpunkt t eine vorliegt, ist eine verallgemeinerte Paretoverteilung mit der Verteilungsfunktion $\bar{G}_{\xi, \beta + \alpha v^*(t)}$. Die Maximum-Likelihood Schätzfunktionen sind dieselben wie im unvorhersehbaren Modell, allerdings ist die bedingte Intensität nun eine andere.

1.5 Multivariate Maxima

Ziel der nächsten beiden Abschnitte ist es, einen kurzen Überblick über die Theorie der multivariaten Maxima zu geben. Dabei wird häufig der bivariate Fall abgehandelt, zumal sich die mehrdimensionale Theorie nicht so leicht erfassen lässt. Das Hauptinstrument zur Charakterisierung und zur Handhabung multivariater Maxima ist die Copula, die uns die Abhängigkeitsstruktur des Zielvektors modelliert. Seien $\mathbf{X}_1, \dots, \mathbf{X}_n \in \mathbb{R}^d$ d-dimensionale Stochastische Vektoren mit einer gemeinsamen Verteilung F und den Randverteilungen F_1, \dots, F_d . Das Maximum der j -ten Komponente ($0 \leq j \leq d$) bezeichnen wir

mit $M_{n,j} = \max(X_{1,j}, \dots, X_{n,j})$. In der multivariaten Extremwerttheorie steht der d -dimensionale Vektor der Maxima $\mathbf{M}_d = (M_{n,1}, \dots, M_{n,d})'$ im Mittelpunkt des Interesses. Gesucht sind wie im eindimensionalen Fall Aussagen und Verteilungen für

$$\frac{\mathbf{M}_n - \mathbf{d}_n}{\mathbf{c}_n} = \left(\frac{M_{n,1} - d_{n,1}}{c_{n,1}}, \dots, \frac{M_{n,d} - d_{n,d}}{c_{n,d}} \right)'$$

und $n \rightarrow \infty$, wobei $\mathbf{c}_n = (c_{n,1}, \dots, c_{n,d})'$ und $\mathbf{d}_n = (d_{n,1}, \dots, d_{n,d})'$ Normierungs- und Zentrierungskonstanten sind und die $\mathbf{c}_n > 0$ erfüllen. Wir nehmen an, dass $(\mathbf{M}_n - \mathbf{d}_n)/\mathbf{c}_n$ in Verteilung gegen einen Zufallsvektor mit der gemeinsamen Verteilung H konvergiert, sodass

$$\lim_{n \rightarrow \infty} \left(\frac{\mathbf{M}_n - \mathbf{d}_n}{\mathbf{c}_n} \leq x \right) = \lim_{n \rightarrow \infty} F^n(\mathbf{c}_n x + \mathbf{d}_n) = H(x) \quad (12)$$

Definition 29 Wenn obige Beziehung (12) erfüllt wird von einer Verteilungsfunktion F und einer Grenzverteilung H so sagen wir F liegt im Maximalen Anziehungsbereich von H und schreiben $F \in MDA(H)$ und nennen H eine multivariate Extremwertverteilung.

Die Frage der Grenzverteilung wird teilweise durch die Ergebnisse der eindimensionalen Theorie beantwortet. Ist H eine nichttriviale Verteilung, so müssen die einzelnen Einträge entweder Fréchet, Gumbel oder Weibullverteilt sein. Nachdem diese auch stetig sind, können wir Sklar's Theorem zu Copulas anwenden und wir wissen, dass die Copula zu H eindeutig ist.

Satz 30 (Extremwert Copula) Ist (12) erfüllt für Verteilungsfunktionen F und H , dann gilt für die eindeutige Copula

$$C(\mathbf{u}^t) = C^t(\mathbf{u}), \quad \forall t > 0. \quad (13)$$

Jede Copula mit obiger Eigenschaft ist eine Extremwertcopula (*extreme value copula*) und kommt als Copula für eine Multivariate Extremwertverteilung in Frage. Ein wichtiges Resultat für Multivariate Extremwertverteilungen und Extremwertcopulas ist der Folgende Satz.

Satz 31 (Pickands Charakterisierung) Die Copula C ist eine Extremwertcopula, genau dann, wenn sie sich als

$$C(\mathbf{u}) = \exp \left\{ B \left(\frac{\ln u_1}{\sum_{k=1}^d \ln u_k}, \dots, \frac{\ln u_d}{\sum_{k=1}^d \ln u_k} \right) \sum_{i=1}^d \ln u_i \right\} \quad (14)$$

schreiben lässt für $B(\omega) = \int_{S_d} \max(x_1 \omega_1, \dots, x_d \omega_d) dH(\mathbf{x})$ und einer Menge $S_d = \left\{ \mathbf{x} : x_i \geq 0, i = 1, \dots, d, \sum_{i=1}^d x_i = 1 \right\}$.

Die Funktion B wird auch als *dependence function* bezeichnet und ist bis auf den bivariaten Fall schwer zu handhaben. Bezeichnen wir nämlich im bivariaten Fall $A(\omega) := B((\omega, 1 - \omega)')$ mit $\omega \in [0, 1]$, so folgt aus Pickands Charakterisierung, dass eine bivariate Copula eine Extremwertcopula ist, wenn sie folgende Form annimmt:

$$C(u_1, u_2) = \exp \left\{ (\ln u_1 + \ln u_2) A \left(\frac{\ln u_1}{\ln u_1 + \ln u_2} \right) \right\}, \quad (15)$$

wobei nun $A(\omega) = \int_0^1 \max((1-x)\omega, x(1-\omega)) dH(x)$. Es folgt, dass eine solche bivariate *dependence function* die Ungleichung

$$\max(\omega, 1-\omega) \leq A(\omega) \leq 1, \quad 0 \leq \omega \leq 1, \quad (16)$$

erfüllt und konvex ist. Umgekehrt formuliert kann jede differenzierbare, konvexe Funktion $A(\omega)$, die die obige Ungleichung erfüllt, benutzt werden, um eine Extremwertcopula zu konstruieren.

Beispiel 32 (Gumbel Copula) Die bivariate Copula, die durch

$$C_{\theta, \alpha, \beta}^{Gu}(u_1, u_2) = u_1^{1-\alpha} u_2^{1-\beta} \exp \left\{ -((-\alpha \ln u_1)^\theta + (-\beta \ln u_2)^\theta)^{\frac{1}{\theta}} \right\}$$

gegeben ist, heißt Gumbelcopula und besitzt die Eigenschaft (13). Für die *dependence function* gilt

$$A(\omega) = (1-\alpha)\omega + (1-\beta)(1-\omega) + ((\alpha\omega)^\theta + (\beta(1-\omega))^\theta)^{1/\theta}.$$

Beispiel 33 (Galambos Copula) Bei der Galambos Copula starten wir zunächst mit der *dependence function*

$$A(\omega) = 1 - ((\alpha\omega^{-\theta} + (\beta(1-\omega))^{-\theta})^{-1/\theta})$$

welche konvex ist und die Ungleichung (16) erfüllt. Somit lässt sich durch diese Funktion eine Extremwertcopula konstruieren, die die Gestalt

$$C_{\theta, \alpha, \beta}^{Gal}(u_1, u_2) = u_1 u_2 \exp \left\{ ((-\alpha \ln u_1)^{-\theta} + (-\beta \ln u_2)^{-\theta})^{-1/\theta} \right\}$$

annimmt. Das Gumbelcopula Modell wird auch als logistisches Modell bezeichnet, wohingegen Galambos Copula auch negatives logistisches Modell genannt wird. Die eben beschriebenen Copulas sind Beispiele für Extremwertcopulas.

1.5.1 Copulas für multivariate Minima

Die Struktur der Copulas für multivariate Minima kann leicht aus der für Multivariate Maxima und der Beziehung (2) hergeleitet werden. Sei F die Verteilungsfunktion der stochastischen Vektoren $\mathbf{X}_1, \dots, \mathbf{X}_n$, deren Minima und die Grenzverteilung dieser gesucht wird. Äquivalent hierzu ist das Betrachten des Maximums von $-\mathbf{X}_1, \dots, -\mathbf{X}_n$ mit der Verteilungsfunktion \tilde{F} . Liegt diese Verteilungsfunktion im Maximalen Anziehungsbereich einer Multivariaten Extremwertverteilung, so folgt, dass auch F , also die Verteilungsfunktion der Minima, eine Grenzverteilung besitzt.

Sei \mathbf{M}_n^* der Vektor der komponentenweise Maxima von $-\mathbf{X}_1, \dots, -\mathbf{X}_n$, sodass $M_{n,j}^* = \max(-X_{1,j}, \dots, -X_{n,j})$. Ist $\tilde{F} \in MDA(H)$ für nichttriviales H , dann ist

$$\lim_{n \rightarrow \infty} P \left(\frac{\mathbf{M}_n^* - \mathbf{d}_n}{\mathbf{c}_n} \leq \mathbf{x} \right) = \lim_{n \rightarrow \infty} \tilde{F}^n(\mathbf{c}_n \mathbf{x} + \mathbf{d}_n) = H(\mathbf{x}) \quad (17)$$

für geeignete vektorwertige Folgen $\mathbf{c}_n > 0$ und \mathbf{d}_n und einer multivariaten Extremwertverteilung H der Form $H(\mathbf{x}) = C(H_{\xi_1}(x_1), \dots, H_{\xi_d}(x_d))$, wobei C (13) erfüllt. Ist nun \mathbf{m}_n der der Vektor der komponentenweise Minima, so folgt, dass

$$\lim_{n \rightarrow \infty} P \left(\frac{\mathbf{m}_n + \mathbf{d}_n}{\mathbf{c}_n} \geq \mathbf{x} \right) = H(-\mathbf{x}) \quad (18)$$

und damit, dass $H(-\mathbf{x}) = C(H_{\xi_1}(-x_1), \dots, H_{\xi_d}(-x_d))$.

1.5.2 Anziehungsbereich der Copulas

Wie im eindimensionalen Fall ist es von Interesse, Eigenschaften einer Verteilungsfunktion F zu untersuchen und dadurch Rückschlüsse auf die Multivariate Extremwertverteilung und die Extremwertcopula zu ziehen. In diesem Zusammenhang ist der folgende Satz sehr hilfreich.

Satz 34 Sei $F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d))$ für stetige Randverteilungen F_1, \dots, F_d und einer Copula C . Sei weiters $H(\mathbf{x}) = C_0(H_1(x_1), \dots, H_d(x_d))$ eine multivariate Extremwertverteilung mit der Extremwertcopula C_0 . Dann ist $F \in MDA(H)$ genau dann, wenn für die einzelnen F_i gilt, dass $F_i \in MDA(H_i)$, $1 \leq i \leq d$, und dass

$$\lim_{t \rightarrow \infty} C^t(u_1^{1/t}, \dots, u_d^{1/t}) = C_0(u_1, \dots, u_d), \quad \mathbf{u} \in [0, 1]^d. \quad (19)$$

Die Copula der Grenzverteilung ergibt sich also nur durch die Copula der zugrundeliegenden Verteilung. Diese Beziehung motiviert die folgende Definition.

Definition 35 Ist (19) von einer Copula C und einer Extremwertcopula C_0 erfüllt, so sagen wir, dass die Copula C im maximalen Anziehungsbereich der Copula C_0 liegt und schreiben dafür $C \in MDA(C_0)$.

1.6 Multivariate Grenzüberschreitungen

1.6.1 Modell

Wie im eindimensionalen Fall ist es natürlich auch von Interesse, den hinteren Teil einer multivariaten Verteilung zu charakterisieren und Wahrscheinlichkeiten zuzuordnen, selbst wenn diese mehrdimensional ist. Dabei stellt sich zunächst die Frage, wie eine vorgegebene Schranke definiert wird und wie über diese hinaus die Excess Verteilung aussieht. Starten wir also mit stochastischen Vektoren $\mathbf{X}_1, \dots, \mathbf{X}_n$, die eine gemeinsame Verteilung $F = C(F_1(x_1), \dots, F_d(x_d))$ besitzen

für irgendeine Copula C und den Randverteilungen F_1, \dots, F_d . Weiters setzen wir voraus, dass F im Maximalen Anziehungsbereich einer multivariaten Extremwertverteilung liegt. Führen wir diesen mehrdimensionalen Fall auf den eindimensionalen zurück, so bedeutet das, dass für einen Vektor von Schranken $\mathbf{u} = (u_1, \dots, u_d)'$ und darüberliegenden Werten $x_j \geq u_j$ komponentenweise durch die Verteilung

$$\tilde{F}_j(x_j) = 1 - \lambda_j \left(1 + \xi_j \frac{x_j - u_j}{\beta_j} \right)^{-1/\xi_j} \quad (20)$$

gegeben ist (also wie im eindimensionalen Fall die verallgemeinerte Paretoverteilung), wobei $\lambda_j = \bar{F}_j(u_j)$. Für den Vektor $\mathbf{x} \geq \mathbf{u}$ motiviert dieses Resultat die Approximation $F(\mathbf{x}) \approx C(\tilde{F}_1(x_1), \dots, \tilde{F}_d(x_d))$. Hier ist allerdings zu bemerken, dass C ebenfalls unbekannt und daher approximiert werden muss. Der Satz über den Anziehungsbereich der Copulas legt folgende Abschätzung nahe: Für einen Vektor $v \in (0, 1)^d$ und hinreichend großes t ist $C(v^{1/t}) = C_0^{1/t}(v)$. Schreiben wir $\omega = v^{1/t}$, erhalten wir über Bedingung (13), dass

$$C(\omega) \approx C_0^{1/t}(\omega^t) = C_0(\omega). \quad (21)$$

Diese Approximation wird genauer, je größer ω ist, da $v^{1/t} \rightarrow 1$ für $t \rightarrow \infty$.

Setzen wir nun voraus, dass wir für große Werte die Copula C durch ihre Extremwertcopula C_0 substituieren können, erhalten wir das allgemeine Modell für die Verteilung

$$\tilde{F}(\mathbf{x}) = C_0(\tilde{F}_1(x_1), \dots, \tilde{F}_d(x_d)). \quad (22)$$

1.6.2 Praktische Vorgehensweise

Liegen Beobachtungen $\mathbf{X}_1, \dots, \mathbf{X}_n$ einer Verteilung vor, die die obige Approximation gestattet, und haben wir einen Vektor von Schranken $\mathbf{u} = (u_1, \dots, u_d)'$, so ist es naheliegend, dass nur einige wenige Beobachtungen über der gemeinsamen Schranke ($\mathbf{x} \geq \mathbf{u}$) liegen. Vielmehr wird es der Fall sein, dass eine Beobachtung ihre Schranke überschreitet, die anderen jedoch nicht, die wiederum von einer anderen Beobachtung überschritten werden und so weiter. Die erforderlichen Parameter $\xi_j, \beta_j, \lambda_j$ für $j = 1 \dots d$ und der Copula Parameter (oder Vektor) θ erhalten wir über den Maximum Likelihood Schätzer über gefilterte Daten.

Gehen wir davon aus, dass m_i Überschreitungen eines Datenvektors $\mathbf{X}_i = (X_{1,i}, \dots, X_{d,i})'$ über den Vektor der Schranken $\mathbf{u} = (u_1, \dots, u_d)'$ vorliegen. Die restlichen $d - m_i$ Komponenten, die die Schranke nicht überschritten haben, werden insofern gefiltert, als dass an Stelle der Daten die Schranke u_i gesetzt wird. Für einen Datenvektor \mathbf{X}_i ist die Likelihoodfunktion durch

$$L_i = L_i(\xi, \beta, \lambda, \theta; \mathbf{X}_i) = \frac{\partial^{m_i} \tilde{F}(x_1, \dots, x_d)}{\partial x_{j_1} \dots \partial x_{j_{m_i}}} \Bigg|_{\max(\mathbf{X}_i, \mathbf{u})}$$

gegeben, wobei die Indizes j_1, \dots, j_{m_i} zu den Komponenten des Vektors gehören, die ihre Schranke überschritten haben.

Im bivariaten Modell mit der vorhin zitierten Gumbelcopula ergibt sich die Likelihoodfunktion zu

$$L_i = \begin{cases} C_\theta^{Gu}(1 - \lambda_1, 1 - \lambda_2), & X_{i,1} \leq u_1, \quad X_{i,2} \leq u_2, \\ C_{\theta,1}^{Gu}(\tilde{F}_1(X_{i,1}), 1 - \lambda_2)\tilde{f}_1(X_{i,1}), & X_{i,1} > u_1, \quad X_{i,2} \leq u_2, \\ C_{\theta,2}^{Gu}(1 - \lambda_1, \tilde{F}_2(X_{i,2}))\tilde{f}_2(X_{i,2}), & X_{i,1} \leq u_1, \quad X_{i,2} > u_2, \\ c_{\theta,2}^{Gu}(\tilde{F}_1(X_{i,1}), \tilde{F}_2(X_{i,2}))\tilde{f}_1(X_{i,1})\tilde{f}_2(X_{i,2}) & X_{i,1} > u_1, \quad X_{i,2} > u_2, \end{cases} \quad (23)$$

wobei \tilde{f}_j die univariate Dichte der Verteilungen \tilde{F}_j aus (20) und $c_\theta^{Gu}(u_1, u_2)$ die Dichte der Gumbelcopula und $C_{\theta,j}^{Gu} := (\partial/\partial u_i)C_\theta^{Gu}(u_1, u_2)$ die bedingte Verteilung der Copula ist.

1.6.3 Grenzcopulas

Welche Copula erhalten wir, bedingt darauf, dass von einem Vektor manche Komponenten ein Grenzüberschreitung vorweisen und andere nicht? Im univariaten Fall haben wir gesehen, dass das Ende einer Verteilung (eben auf das Überschreiten einer vorgegebenen Schranke bedingt) durch eine verallgemeinerte Paretoverteilung approximierbar ist. Gibt es nun auch analog zum univariaten Fall Grenzcopulas und sind diese eindeutig und unter welchen Voraussetzungen existieren sie?

Untere Grenzcopulas Gehen wir aus von einem Zufallsvektor (X_1, X_2) mit stetigen Randverteilungen F_1 und F_2 und einer symmetrische Copula C ($C(u_1, u_2) = C(u_2, u_1)$). Wir betrachten die Verteilung von (X_1, X_2) unter der Bedingung, dass das v -Quantil nicht überschritten wurde, also das Ereignis $A_v = \{X_1 \leq F_1^{-1}(v), X_2 \leq F_2^{-1}(v)\}$, $0 < v \leq 1$. Für $C(v, v) \neq 0$ ist die Wahrscheinlichkeit für X_1 unter dem x_1 Quantil und X_2 unter dem x_2 Quantil durch

$$P(X_1 \leq F_1^{-1}(x_1), X_2 \leq F_2^{-1}(x_2) | A_v) = \frac{C(x_1, x_2)}{C(v, v)}, \quad x_1, x_2 \in [0, v]$$

gegeben. Als Funktion von x_1 und x_2 aufgefasst definiert die obige Darstellung eine bivariate Verteilungsfunktion auf $[0, v]^2$ und aus dem Satz von Sklar folgt, dass

$$\frac{C(x_1, x_2)}{C(v, v)} = C_v^0(F_{(v)}(x_1), F_{(v)}(x_2)), \quad x_1, x_2 \in [0, v], \quad (24)$$

für eine eindeutige Copula C_v^0 und der stetigen Randverteilung

$$F_{(v)}(x) = P(X_1 \leq F_1^{-1}(x) | A_v) = \frac{C(x, v)}{C(v, v)}, \quad x \in [0, v], \quad (25)$$

die auch an den eindimensionalen Fall der Excessverteilung erinnert. Diese eindeutige Copula wird als

$$C_v^0(u_1, u_2) = \frac{C(F_{(v)}^{-1}(u_1), F_{(v)}^{-1}(u_2))}{C(v, v)} \quad (26)$$

geschrieben und wird als untere Grenzcopula bezeichnet (*lower threshold Copula C at level v*).

Obere Grenzcopulas Gehen wir erneut von einem Zufallsvektor (X_1, X_2) aus und betrachten nun das Ereignis $\bar{A}_v = \{X_1 > F_1^{\leftarrow}(v), X_2 > F_2^{\leftarrow}(v)\}$, $0 < v \leq 1$, so erhalten wir für die bedingte Wahrscheinlichkeit

$$P(X_1 > F_1^{\leftarrow}(x_1), X_2 > F_2^{\leftarrow}(x_2) | \bar{A}_v) = \frac{\bar{C}(x_1, x_2)}{\bar{C}(v, v)}, \quad x_1, x_2 \in [v, 1],$$

wobei $\bar{C}(x_1, x_2)/\bar{C}(v, v)$ eine Verteilungsfunktion der Gestalt $1 - F(x)$ auf $[v, 1]^2$ ist und als

$$\frac{\bar{C}(x_1, x_2)}{\bar{C}(v, v)} = \hat{C}_v^1(\bar{G}_{(v)}(x_1), \bar{G}_{(v)}(x_2)), \quad x_1, x_2 \in [v, 1], \quad (27)$$

zu schreiben, wobei \hat{C}_v^1 die *survival copula* einer Copula C_v^1 ist und für die Randverteilungsfunktion

$$\bar{G}_{(v)}(x) = P(X_1 > F_1^{\leftarrow}(x) | \bar{A}_v) = \frac{\bar{C}(x, v)}{\bar{C}(v, v)}, \quad x \in [v, 1], \quad (28)$$

gilt. Die Copula C_v^1 ist die obere Grenzcopula (*upper threshold copula at level v*). In der Praxis ist das Untersuchen oberer Grenzcopula äquivalent zum untersuchen der unteren Grenzcopula und umgekehrt und somit ausreichend, wie folgendes Lemma die Theorie der Grenzcopulas zusammenfasst

Lemma 36 Die *survival copula* \hat{C}_v^1 der oberen Grenzcopula C_v^1 von C bei v ist die untere Grenzcopula C_v^0 der Copula \hat{C} bei $1 - v$.

Literatur

- [1] Alexander J. McNeil, Rüdiger Frey, Paul Embrechts - Quantitative Risk Management: Concepts, Techniques and Tools. - Princeton, NJ: Princeton Univ. Press, 2005
- [2] Rolf D. Reiss, Michael Thomas - Statistical Analysis of Extreme Values: From Insurance, Finance, Hydrology and Other Fields. - Basel: Birkhäuser, 1997
- [3] Wikipedia, die freie Enzyklopädie; Portal: Mathematik