

Extremwerttheorie
Clemens Grünwald

Inhaltsverzeichnis

1	Einleitung	2
2	Die Block-Maxima Methode	3
2.1	Motivation	3
2.2	Der maximale Anziehungsbereich	3
2.3	Verallgemeinerte Extremwertverteilung	4
2.4	Minima	5
2.5	Einteilung in die maximalen Anziehungsbereiche	6
2.6	Maxima von strikt stationären stochastischen Prozessen	8
2.7	Die Block-Maxima Methode	8
3	Threshold Exceedances	10
3.1	Excess-Verteilung	10
3.2	Verallgemeinerte Paretoverteilung	10
3.3	Das Theorem von Pickhands-Balkema-de Haan	11
3.4	Modellierungen	12
4	Zählprozess-Modelle	16
4.1	Überschreitungen einer Schranke bei weißem Rauschen	17
4.2	Das POT (Peaks-Over-Threshold) Modell	19
4.3	Self-Exciting Prozesse	20

1 Einleitung

Ich möchte in dieser Arbeit einen kleinen Einblick in das Gebiet der Extremwerttheorie geben und werde mich dabei im Wesentlichen an das gleichnamige Kapitel aus dem Buch "Quantitative Risk Management" von McNeil, Frey und Embrechts halten.

In weiten Teilen der Arbeit werden Folgen von unabhängigen identisch verteilten Zufallsvariablen behandelt, die für finanzielle Verluste, wie etwa betriebswirtschaftliche Verluste, Versicherungsverluste oder Verluste aus einem Kreditportfolio stehen. Später werden wir jedoch auch abhängige streng stationäre Zeitreihen, die zum Beispiel für ein Portfolio von Investitionen stehen können, betrachten.

Wie der Name der Seminararbeit bereits sagt, sind für uns besonders hohe Werte interessant. Im folgenden Kapitel werden wir die Block-Maxima Methode betrachten, deren Ziel es ist, das Maximum der Verluste aufgrund einer hohen Anzahl von Daten zu modellieren. Eine wesentliche Rolle werden hierbei die verallgemeinerten Extremwertverteilungen spielen.

Das darauffolgende Kapitel "Threshold Exceedances" beschäftigt sich mit Überschreitungen einer hohen Schranke und der Verteilung der Verluste über dieser Schranke. Auch der Tail einer Verteilung und somit die Hill-Methode werden in diesem Kapitel behandelt.

Die im letzten Kapitel vorgestellten Modelle stützen sich auf Zählprozesse. Hier spielt nicht mehr wie zuvor nur die Höhe der Überschreitung, sondern auch der Zeitpunkt und somit die Häufigkeit einer Überschreitung eine wichtige Rolle.

2 Die Block-Maxima Methode

2.1 Motivation

Wir wollen an dieser Stelle den sehr bekannten Zentralen Grenzwertungssatz betrachten, der zur Wiederholung wie folgt lautet.

Seien unabhängige identisch verteilte stochastische Größen X_1, X_2, \dots, X_n mit $Var(X_i) < \infty \forall i \in \mathbb{N}$ gegeben, so gilt

$$\lim_{n \rightarrow \infty} P \left(\frac{S_n - n\mathbb{E}(X)}{\sqrt{nVar(X)}} \leq x \right) = \Phi(x), \quad \forall x \in \mathbb{R}$$

mit $S_n = X_1 + \dots + X_n$.

Man kann erkennen, dass die Verteilung der normalisierten Summe dieser stochastischen Größen gegen die bekannte Standardnormalverteilung konvergiert.

Nun stellt sich die Frage, ob sich ähnliche Aussagen auch für normalisierte Block-Maxima $M_n = \max(X_1, \dots, X_n)$ tätigen lassen.

2.2 Der maximale Anziehungsbereich

Da M_n der n -ten Ordnungstatistik entspricht, erhält man für n unabhängige identisch verteilte Zufallsgrößen folgende Aussage über das Maximum:

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

Wir nehmen nun an die Verteilungsfunktion von M_n konvergiere bei angemessener Normalisierung gegen eine nicht-degenerierte Verteilungsfunktion $H(x)$. Das heißt, es existieren reelle Folgen (d_n) und $(c_n) > 0$, sodass folgende Aussage gilt:

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

wobei $H(x)$ eine nicht-degenerierte Verteilungsfunktion ist, das heißt $H(x)$ erfüllt

$$H(x) = \begin{cases} 0, & x < x^* \\ 1, & \text{sonst} \end{cases}, \quad x^* = \sup\{x : F(x) < 1\} \in \mathbb{R} \cup \{\infty\}$$

nicht.

Gibt es ein $H(x)$, sodass diese Gleichung erfüllt ist, so sagt man, dass F im maximalen Anziehungsbereich von H (maximum domain of attraction) liegt, und man schreibt $F \in MDA(H)$.

2.3 Verallgemeinerte Extremwertverteilung

Das **Theorem von Fisher-Tippett und Gnedenko** besagt, dass aus $F \in MDA(H)$ folgt, dass $H(x)$ eine verallgemeinerte Extremwertverteilung (generalized extreme value (GEV) distribution) ist.

Die Verteilungsfunktion einer verallgemeinerten Extremwertverteilung ist gegeben durch:

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

für $1 + \xi x > 0$.

Man kann aus $H_{\xi, \mu, \sigma} := H_\xi(\frac{x-\mu}{\sigma})$ leicht eine drei-parametrische Familie gewinnen, wobei $\mu \in \mathbb{R}$ der Lageparameter, $\sigma > 0$ der Skalierungsparameter und ξ der Formparameter genannt wird.

Abhängig von ξ gibt es drei Typen von Extremwertverteilungen:

- Wenn $\xi > 0$, dann handelt es sich um eine **Frechetverteilung**.
- Wenn $\xi = 0$, dann handelt es sich um eine **Gumbelverteilung**.
- Wenn $\xi < 0$, dann handelt es sich um eine **Weibullverteilung**.

Hier sei zu bemerken, dass die Extremwertverteilung für ein festes x von beiden Seiten stetig bezüglich ξ ist (d.h. $\lim_{\xi \rightarrow 0} H_\xi(x) = H_0(x)$).

Die Weibullverteilung ist die einzige der drei Verteilungen, für die x_F endlich ist.

Wenn die Verteilung des normalisierten Maximums konvergiert, dann ist der Typ der Grenzverteilung eindeutig bestimmt.

Die Lage- und Skalierungsparameter der Grenzfunktion hängen jedoch von den exakten Folgen c_n und d_n ab. Man kann die Folgen immer so wählen, dass die Grenzverteilung die Stammform H_ξ besitzt.

Hierzu will ich zwei Beispiele angeben.

1. Exponentialverteilung:

$$F(x) = 1 - \exp(-\beta x), \quad \beta > 0, x \geq 0 \quad c_n = \frac{1}{\beta}, d_n = \ln\left(\frac{n}{\beta}\right)$$

$$\Rightarrow F^n(c_n x + d_n) = \left(1 - \frac{1}{n} \exp(-x)\right)^n, \quad x \geq -\ln(n)$$

$$\Rightarrow \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = \exp(-e^{-x}), \quad x \in \mathbb{R}$$

$$\Rightarrow F \in MDA(H_0)$$

2. Paretoverteilung:

$$F(x) = 1 - \left(\frac{\kappa}{\kappa + x}\right)^\alpha, \quad \alpha > 0, \kappa > 0 \quad c_n = \frac{\kappa n^{\frac{1}{\alpha}}}{\alpha}, d_n = \kappa n^{\frac{1}{\alpha}} - \kappa$$

$$\Rightarrow F^n(c_n x + d_n) = \left(1 - \frac{1}{n} \left(1 + \frac{x}{\alpha}\right)^{-\alpha}\right)^n, \quad 1 + \frac{x}{\alpha} \geq n^{-\frac{1}{\alpha}}$$

$$\Rightarrow \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = \exp\left(-\left(1 + \frac{x}{\alpha}\right)^{-\alpha}\right), \quad 1 + \frac{x}{\alpha} > 0$$

$$\Rightarrow F \in MDA(H_{\frac{1}{\alpha}})$$

2.4 Minima

Will man nun auch über die Verteilung des Minimums einer Folge unabhängiger identisch verteilter stochastischer Größen Aussagen tätigen, so greift man auf das Maximum zurück, indem man folgende Identität verwendet:

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

Wir schreiben $\tilde{F} = 1 - F(-x)$, für die Verteilungsfunktion der stochastischen Größen $-X_1, \dots, -X_n$.

Wenn nun $\tilde{F} \in MDA(H_\xi)$ erfüllt ist, folgt daraus:

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

Daraus erhält man weiters durch Umformungen:

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

woraus man erkennt, dass das normalisierte Minimum die Grenzfunktion $1 - H_\xi(-x)$ besitzt.

Im symmetrischen Fall $\tilde{F}(x) = F(x)$ erkennt man leicht, dass wenn H_ξ der Typ der Grenzverteilung fürs Maximum ist, daraus folgt, dass $1 - H_\xi(-x)$ der Typ der Grenzverteilung des Minimums ist.

Es ist nun natürlich sehr interessant, ob sich irgendwelche Aussagen tätigen lassen, ob eine Funktion im maximalen Anziehungsbereich einer Extremwertverteilung liegt.

2.5 Einteilung in die maximalen Anziehungsbereiche

Im Wesentlichen sind alle bekannten stetigen Verteilungen in $MDA(H_\xi)$ für einen gewissen Wert ξ .

Um nun erkennen zu können, um welchen Typ es sich handelt, braucht man vorerst noch die Definition der langsamen Variation.

Eine positive, Lebesgue-messbare Funktion $L(x)$ auf $(0, \infty)$ variiert langsam an ∞ , wenn gilt:

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, \quad \forall t > 0$$

Nun können wir Aussagen über die einzelnen Fälle treffen.

- Frechetverteilung ($\xi > 0$):

Die Frechet-Klasse lässt sich folgendermaßen charakterisieren:

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

wobei $L(x)$ eine Funktion ist, die langsam an ∞ variiert.

Diese Verteilungen, die im maximalen Anziehungsbereich einer Frechet-Verteilung liegen, sind die in der Extremwerttheorie am meist untersuchten.

Sie sind besonders interessant für finanzielle Anwendungen, da sie "heavy-tailed" sind und unendliche höhere Momente besitzen (d.h. $\mathbb{E}(X^k) = \infty \quad \forall k > \frac{1}{\xi}$ für eine nicht-negative Zufallsvariable X , deren Verteilungsfunktion in $MDA(H_\xi)$ mit $\xi > 0$ liegt).

Zu ihnen gehören unter anderem die Pareto-, die Burr- und die Student-Verteilung.

- Gumbelverteilung ($\xi = 0$):

In diesem Fall ist eine Charakterisierung der Verteilungen, die in dem maximalen Anziehungsbereich einer Gumbel-Verteilung liegen, leider nicht so einfach wie im Fall der Frechetverteilung.

Man kann sagen, dass die Verteilungsfunktionen dieser Klasse im Wesentlichen Tails besitzen, die exponential fallen.

Hierzu gehören zum Beispiel die Exponentialverteilung selbst, sowie die Normal- oder die Lognormalverteilung.

Eine positive Zufallsvariable, dessen Verteilungsfunktion in $MDA(H_0)$ liegt, hat endliche Momente positiver Ordnung (d.h. $\mathbb{E}(X^k) < \infty \quad \forall k > 0$).

Wie schon erwähnt, sind die Verteilungen der Frechet-Klasse am interessantesten für Finanz-Modelle, jedoch auch die Gumbel-Klasse ist oft von Interesse, da sie einige Verteilungen beinhaltet, die um einges schwerere Enden als die Normalverteilung besitzen.

- Weibullverteilung ($\xi < 0$):

Diese Klasse ist die wohl unwichtigste aller drei im Bezug auf finanzielle Modellierungen, da die Verteilungen dieser Klasse endliche Endpunkte haben (d.h. $x_F < \infty$).

In der Praxis sind zwar alle möglichen Finanz- und Versicherungsverluste beschränkt, doch bei Modellierungen geht man lieber von unbeschränkten Verlusten aus.

Ähnlich wie im Frechet-Fall kann man auch die Verteilungen der Weibull-Klasse leicht charakterisieren:

$$F \in MDA(H_\xi) \Leftrightarrow x_F < \infty \wedge 1 - F(x_F - x^{-1}) = x^{\frac{1}{\xi}} L(x),$$

wobei $L(x)$ eine Funktion ist, die langsam an ∞ variiert.

Zu dieser Klasse gehören zum Beispiel die stetige Gleichverteilung und die Betaverteilung.

2.6 Maxima von strikt stationären stochastischen Prozessen

Bis jetzt haben wir die Maxima von unabhängigen identisch verteilten stochastischen Größen betrachtet. Wir wollen nun jedoch auch kurz die Theorie der Maxima von strikt stationären stochastischen Prozessen behandeln.

Ein stochastischer Prozess ist strikt stationär, wenn die gemeinsame Verteilung von $(Y_{t_1+k}, \dots, Y_{t_n+k})$ für alle $k \in \mathbb{Z}$ unabhängig von k ist.

Sei $(X_i)_{i \in \mathbb{Z}}$ ein strikt stationärer stochastischer Prozess mit stationärer Verteilung F und $(\tilde{X}_i)_{i \in \mathbb{N}}$ ein unabhängiger identisch verteilter Prozess mit der selben Verteilungsfunktion F , und weiters seien $M_n = \max(X_1, \dots, X_n)$ sowie $\tilde{M}_n = \max(\tilde{X}_1, \dots, \tilde{X}_n)$ wie gewohnt die Block-Maxima des ursprünglichen und des unabhängig identisch verteilten stochastischen Prozesses, dann kann für viele Prozesse gezeigt werden, dass es eine reelle Zahl $\theta \in (0, 1]$ gibt, sodass

$$\lim_{n \rightarrow \infty} P \left(\frac{\tilde{M}_n - d_n}{c_n} \leq x \right) = H_\xi(x)$$

genau dann, wenn

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

Wir nennen θ den Extremal-Index des Prozesses.

Man kann leicht nachweisen, dass $H_\xi^\theta(x)$ vom selben Verteilungstyp wie $H_\xi(x)$ ist, und sie daher den selben Formparameter ξ besitzt, sich jedoch durch den Skalierungs- oder Lageparameter unterscheiden kann.

Nicht jeder strikt stationäre stochastischer Prozess hat einen Extremal-Index, aber für diese, die für finanzielle Modelle interessant sind, existiert der Extremal-Index in der Regel.

2.7 Die Block-Maxima Methode

Nun wollen wir die bisher vorgestellten Erkenntnisse verwenden, um auf Grund von Daten die Verteilung des Maximums zu schätzen.

Seien die Daten von unabhängigen, identisch verteilten Zufallsvariablen mit der Verteilungsfunktion F , von der wir annehmen, dass sie für ein ξ im maximalen Anziehungsbereich von H_ξ liegt, so kann nun laut Theorie die Verteilung eines n -Block-Maximums M_n bei groß genug gewähltem n durch eine drei-parametrische Extremwertverteilung $H_{\xi, \mu, \sigma}$ approximiert werden.

Wir teilen die Daten in m Blöcke der Größe n , was besonders Sinn macht, wenn die Daten auf „natürliche Weise“ geblockt werden (z.B. das Maximum täglicher finanzieller Verluste über ein Jahr hinweg gesehen).

Wir bezeichnen das Block-Maximum des j -ten Blocks mit M_{nj} und erhalten somit unsere Daten M_{n1}, \dots, M_{nm} .

Die Verteilung kann nun aus den Daten auf verschiedene Weise gewonnen werden, wobei wir nun die Maximum-Likelihood-Methode etwas genauer betrachten wollen.

Unter Verwendung der Dichte einer Extremwertverteilung:

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

(für $\xi = 0$ ist hier der Grenzwert zu verwenden), erhält man leicht die logarithmierte Likelihoodfunktion:

$$\begin{aligned} l(\xi, \mu, \sigma; M_{n1}, \dots, M_{nm}) &= \sum_{i=1}^m \ln h_{\xi, \mu, \sigma}(M_{ni}) \\ &= -m \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^m \ln \left(1 + \xi \frac{M_{ni} - \mu}{\sigma}\right) - \sum_{i=1}^m \left(1 + \xi \frac{M_{ni} - \mu}{\sigma}\right)^{-\frac{1}{\xi}} \end{aligned}$$

Die Parameter erhält man nun durch Maximierung dieser Funktion. Man darf jedoch nicht auf die beiden durch die Definitionen bestimmten Einschränkungen $\sigma > 0$ und $1 + \xi(M_{ni} - \mu)/\sigma > 0 \forall i$ vergessen.

Hier ist noch wichtig zu erwähnen, dass die Wahl von m und somit auch indirekt die von n nicht uninteressant ist.

Umso größer n ist, desto besser ist die Approximation des Maximums durch eine Extremwertverteilung.

Jedoch umso größer m ist, desto mehr Daten erhält man für den Maximum-Likelihood-Schätzer, der asymptotisch optimal ist(konsistent).

Weiters sind hier noch zwei wichtige Kennzahlen zu erwähnen.

- Angenommen H sei die wahre Verteilungsfunktion der Verteilung des n -Block-Maximums, dann ist das Return-Level durch $r_{n,k} = q_{1-\frac{1}{k}}(H)$ gegeben.

- Angenommen H sei die wahre Verteilungsfunktion der Verteilung des n -Block-Maximums, dann ist die Return-Period des Ereignisses $\{M_n > u\}$ durch $k_{n,u} = \frac{1}{H(u)}$ gegeben.

3 Threshold Exceedances

Die in Kapitel 2 vorgestellte Block-Maxima Methode hat den Nachteil, dass man für sie ziemlich viele Daten benötigt.

In diesem Kapitel wollen wir nun Methoden vorstellen, für die man nur Daten, die über einer bestimmten festgelegten hohen Schranke liegen, benötigt.

3.1 Excess-Verteilung

Sei X eine stochastische Größe mit Verteilungsfunktion F , dann ist die Excess-Verteilung über einen Schwellwert u gegeben durch:

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

$F_u(x)$ beschreibt die Verteilung um wieviel die Schranke u übertroffen wird, unter der Bedingung, dass sie überhaupt übertroffen wird.

Weiters brauchen wir in Folge die Definition der Mean-Excess-Funktion. Diese ist für eine stochastische Größe X folgendermaßen definiert:

$$e(u) = \mathbb{E}(X - u | X > u).$$

Sie beschreibt den Mittelwert von $F_u(x)$, und ist eine von u abhängige Funktion.

3.2 Verallgemeinerte Paretoverteilung

So wie bei der Methode der Block-Maxima, gibt es auch in diesem Kapitel eine sehr wichtige Verteilungsfamilie.

Es handelt sich um die verallgemeinerte Paretoverteilung, deren Verteilungsfunktion wie folgt definiert ist:

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

für $\beta > 0$ und $x \geq 0$ wenn $\xi \geq 0$ und $0 \leq x \leq -\beta / \xi$ wenn $\xi < 0$.

Auch hier wird ξ der Formparameter sowie β der Skalierungsparameter genannt.

Es gibt wiederum abhängig von ξ drei Typen von Paretoverteilungen:

- Wenn $\xi > 0$, dann handelt es sich um eine, wie bereits im 2. Kapitel vorgestellt, gewöhnliche Paretoverteilung mit den Parametern $\alpha = \frac{1}{\xi}$ und $\kappa = \frac{\beta}{\xi}$.
- Wenn $\xi = 0$, dann handelt es sich um eine Exponentialverteilung.
- Wenn $\xi < 0$, dann handelt es sich um eine Typ II Paretoverteilung.

Die Stetigkeit bezüglich ξ ist auch für die verallgemeinerte Paretoverteilung gegeben (d.h. $\lim_{\xi \rightarrow 0} G_{\xi, \beta}(x) = G_{0, \beta}(x)$).

Weiters erhalten wir, dass $G_{\xi, \beta} \in MDA(H_\xi)$ für alle $\xi \in \mathbb{R}$ und, dass $\mathbb{E}(X) = \frac{\beta}{(1-\xi)}$ für $\xi < 1$.

3.3 Das Theorem von Pickhands-Balkema-de Haan

Dieses Theorem stellt nun einen Zusammenhang zwischen der Excess-Verteilung und der verallgemeinerten Paretoverteilung dar, denn es lautet wie folgt.

Man kann eine positive, messbare Funktion $\beta(u)$ finden, sodass

$$\lim_{u \rightarrow x_F} \sup_{0 \leq x < x_F - u} |F_u(x) - G_{\xi, \beta(u)}(x)| = 0$$

genau dann, wenn $F \in MDA(H_\xi)$, mit $\xi \in \mathbb{R}$.

Das heißt, dass wenn die Verteilung eines normalisierten Maximums gegen eine verallgemeinerte Extremwertverteilung konvergiert, so konvergiert die Excess-Verteilung gegen eine verallgemeinerte Paretoverteilung, wenn die Schranke immer höher wird. Die verallgemeinerte Extremwertverteilung und die verallgemeinerte Paretoverteilung besitzen den selben Formparameter ξ .

Wie bereits im 2. Kapitel erwähnt, sind fast alle bekannten stetigen Verteilungen im maximalen Anziehungsbereich einer verallgemeinerten Extremwertverteilung für ein bestimmtes ξ , weshalb man sagen kann, dass es Sinn macht, die Verteilungen von Verlusten über einem hohen Grenzwert durch verallgemeinerte Paretoverteilungen zu modellieren.

3.4 Modellierungen

Wir wollen nun für diesen Abschnitt auf Grund der gerade gezeigten Konvergenz folgende Annahme treffen.

Sei F eine Verlustfunktion mit rechtem Endpunkt x_F , dann gilt $F_u(x) = G_{\xi, \beta}(x)$ für $0 \leq x < x_F - u$, $\xi \in \mathbb{R}$ und $\beta > 0$.

Man darf hier nicht vergessen, dass es sich nur um eine Annahme handelt, da in der Praxis nie exakte Gleichheit erlangt wird.

- Maximum-Likelihood-Methode:

Hat man nun n unabhängige Verlustdaten von F gegeben, so beschreibt die Zufallszahl N_u die Anzahl derer, die die Schranke u überschritten haben.

Man betrachtet jetzt nur mehr diese Daten $\tilde{X}_1, \dots, \tilde{X}_{N_u}$ und erhält daraus die Höhe des Verlustes über der Schranke $Y_i = \tilde{X}_i - u$.

Nun wollen wir wieder die Maximum-Likelihood-Methode anwenden um die Parameter der GPD zu schätzen:

Unter Verwendung der Dichte einer verallgemeinerten Paretoverteilung:

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

erhält man leicht die logarithmierte Likelihoodfunktion:

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

welche nun unter den Bedingungen $\beta > 0$ und $1 + \xi \frac{Y_i}{\beta} > 0$ maximiert werden muss.

- Überschreitung einer höheren Schranke:

Gehen wir weiter von der Annahme $F_u(x) = G_{\xi, \beta}(x)$ aus, so folgt für eine höhere Schranke $v \geq u$, dass $F_v(x) = G_{\xi, \beta + \xi(v-u)}(x)$.

Hierzu möchte ich den Beweis anführen:

$$\begin{aligned}
\bar{F}_v(x) &= 1 - \frac{F(v+x) - F(v)}{1 - F(v)} = \frac{\bar{F}(v+x)}{\bar{F}(v)} \\
&= \frac{\bar{F}(u + (v+x-u))}{\bar{F}(u)} \frac{\bar{F}(u)}{\bar{F}(u + (v-u))} = \frac{\bar{F}_u(x+v-u)}{\bar{F}_u(v-u)} \\
&= \frac{\bar{G}_{\xi, \beta}(x+v-u)}{\bar{G}_{\xi, \beta}(v-u)} = \left(\frac{1 + \xi \frac{x+v-u}{\beta}}{1 + \xi \frac{v-u}{\beta}} \right)^{-\frac{1}{\xi}} \\
&= \left(\frac{\beta + \xi(x+v-u)}{\beta + \xi(v-u)} \right)^{-\frac{1}{\xi}} = \left(\frac{\beta + \xi(v-u)}{\beta + \xi(v-u)} + \frac{\xi x}{\beta + \xi(v-u)} \right)^{-\frac{1}{\xi}} \\
&= \left(1 + \frac{\xi x}{\beta + \xi(v-u)} \right)^{-\frac{1}{\xi}} = \bar{G}_{\xi, \beta + \xi(v-u)}(x)
\end{aligned}$$

Man sieht, dass die Excess-Funktion einer höheren Schranke eine verallgemeinerte Paretoverteilung mit dem selben Formparameter ξ bleibt, jedoch der Skalierungsparameter β linear ansteigt.

Weiters erhält man daraus und aus den bereits bekannten Definitionen die Mean-Excess-Funktion der höheren Schranke für $\xi < 1$:

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

- Modellierung des Tails einer Verteilung:

Wir wollen nun auf bereits Gezeigtes zurückgreifen um $\bar{F}(x)$ für $x \geq u$ zu erhalten.

Unter Beibehaltung unserer Annahme erhalten wir:

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

Für ein $\alpha \geq F(u)$ erhalten wir den Value at Risk (VaR) und den Expected Shortfall (ES):

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

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

woraus sich folgende Aussage tätigen lässt:

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

In der Praxis geht man nun so vor, dass $\bar{F}(u)$ durch $\frac{N_u}{n}$ geschätzt wird und die Parameter ξ und β durch die Maximum-Likelihood-Schätzer $\hat{\xi}$ und $\hat{\beta}$ ersetzt werden.

Man erhält somit einen Schätzer für $\bar{F}(x)$:

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

Für $\alpha \geq 1 - \frac{N_u}{n}$ erhält man so auch Punktschätzer für VaR_α und ES_α .

Man muss hier natürlich annehmen, dass man genügend Daten hat, die über der Schranke liegen, damit die Schätzer einen Sinn machen.

- Die Hill-Methode:

Die GPD-Methode ist nicht die einzige Methode, um den Tail einer Verteilung zu schätzen.

Man verwendet die Hill-Methode um die Tails einer heavy-tailed Verteilung zu schätzen.

Für diese Methode nehmen wir an, dass die Verteilung im maximalen Anziehungsbereich einer Frechetverteilung liegt.

Laut Kapitel 2 heißt das nichts anderes als

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

wobei $L(x)$ langsam variiert.

Bei der Hill-Methode ist man mehr am positiven Parameter α als an ξ interessiert und versucht diesen anhand von Daten zu schätzen.

Es gibt dazu mehrere Wege, wir verwenden jedoch die Mean-Excess Funktion der logarithmierten Daten:

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

Bei groß genug gewähltem u , kann $L(x)$ für $x \geq u$ als Konstante behandelt werden, daher aus dem Integral herausgezogen werden und man erhält so aus dem Theorem von Karamata für $u \rightarrow \infty$:

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

und somit $\lim_{u \rightarrow \infty} \alpha e^*(\ln u) = 1$.

Nun brauchen wir noch einen Schätzer für die Mean-Excess-Funktion, der folgendermaßen gegeben ist:

$$e_n(u) = \frac{\sum_{i=1}^n (X_i - u) \mathbb{I}_{\{X_i > u\}}}{\sum_{i=1}^n \mathbb{I}_{\{X_i > u\}}}.$$

Ordnet man nun die Daten $X_{n,n} \leq \dots \leq X_{1,n}$, so erhält man durch $e_n^*(\ln X_{k,n}) \approx \alpha^{-1}$ für großes n und k ausreichend klein und den gerade vorgestellten Schätzer einen Schätzer für α :

$$\hat{\alpha}^{-1} = \frac{1}{k-1} \sum_{j=1}^{k-1} \ln(X_{j,n}) - \ln(X_{k,n}),$$

der durch eine kleine Umänderung zum Hill-Schätzer wird:

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

Der Hill-Schätzer ist einer der am meist untersuchten Schätzer in der Extremwerttheorie. Die asymptotischen Eigenschaften wurden intensiv beobachtet.

In der Praxis wird meist so vorgegangen, dass man den Hill-Schätzer für verschiedene Werte k aufzeichnet und somit den Hill-Plot $\{(k, \hat{\alpha}_{k,n}^{(H)} : k = 2, \dots, n)\}$ erhält.

Man versucht in diesem Plot eine stabile Region zu finden, in der die Schätzer von verschiedenen Ordnungsstatistiken recht ähnlich sind.

Man darf jedoch nicht vergessen, dass man davon ausgeht, dass die Daten von Verteilungsfunktionen mit einer regulär veränderlichen Tails stammen. Ist dies nicht der Fall, so macht der Hill-Schätzer keinen Sinn.

Nun wollen wir einen Schätzer für $\bar{F}(x)$ mittels des Hill-Schätzers erlangen.

Dazu nehmen wir an, dass $\bar{F}(x) = Cx^{-\alpha}$, $x \geq u > 0$, für eine hohe Schranke u , was bedeutet, dass die langsam variierende Funktion $L(x)$ durch eine Konstante C für großes x ersetzt wird.

α wird durch den Hill-Schätzer ersetzt und die Schranke u wird durch $X_{k,n}$ ersetzt.

Es verbleibt noch C zu schätzen, welches wir wie folgt ausdrücken können: $C = u^\alpha \bar{F}(u)$.

Ersetzt man nun noch $\bar{F}(u)$ durch $\frac{k}{n}$, so erhält man die Standardform des Hill-Tail-Schätzers:

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

4 Zählprozess-Modelle

Bis jetzt haben wir die Höhe des Verlustes über einer bestimmten Schranke betrachtet.

Nun wollen wir in diesem Kapitel die Überschreitung der Schranke als Ereignis in der Zeit betrachten und Zählprozesse nützen, um die Häufigkeit dieses Ereignisses zu modellieren.

4.1 Überschreitungen einer Schranke bei weißem Rauschen

Wir nehmen nun an, dass ein striktes weißes Rauschen $(\epsilon_t) \sim IID(\sigma^2)$ finanzielle Verluste darstellt.

Man nennt einen stochastischen Prozess (ϵ_t) ein weißes Rauschen, wenn

- $\mathbb{E}(\epsilon_t) = 0$,
- $\mathbb{E}(\epsilon_t^2) < \infty$ und
- $\mathbb{E}(\epsilon_{t+k}\epsilon_t) = 0 \quad \forall k \neq 0$

für alle t erfüllt ist, wobei wir hier zusätzlich die Unabhängigkeit und die identische Verteilung voraussetzen.

Wir machen nun die folgenden Beobachtungen zwar für unabhängig identisch verteilte Zufallsvariablen, sie lassen sich jedoch auch auf Prozesse mit Extremal-Index $\theta = 1$ anwenden.

Im Folgenden nehmen wir an, dass die Verlustfunktion im maximalen Anziehungsbereich einer verallgemeinerten Extremwertverteilung liegt.

Wendet man auf die Gleichung

$$\lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H_\xi(x)$$

den natürlichen Logarithmus an, so erhält man

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

wodurch man weiters unter Verwendung der Beziehung $\ln(1 - y) \sim -y$ für $y \rightarrow 0$ die Gleichung

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

für ein fixes x erhält.

Weiters definieren wir uns $u_n(x) := c_n x + d_n$ und fassen diese als eine Folge von Schranken auf.

Die Anzahl von n Verlusten X_1, \dots, X_n , die die Schranke $u_n(x)$ überschreiten, ist nun binomialverteilt, $N_{u_n(x)} \sim B(n, \bar{F}(u_n(x)))$, mit Mittelwert $n\bar{F}(u_n(x))$.

Auf Grund des Zusammenhangs zwischen Binomialverteilung und Poissonverteilung folgt daraus, dass für $n \rightarrow \infty$ $N_{u_n(x)}$ gegen eine Poissonverteilung mit Mittelwert $\lambda(x) = -\ln H_\xi(x)$ konvergiert, abhängig vom fixierten x .

Für eine Folge von Zufallszahlen oder Zufallsvektoren Y_1, \dots, Y_n in einem Zustandsraum X (zum Beispiel \mathbb{R} oder \mathbb{R}^2) und für jede Teilmenge $A \subset X$ nennen wir

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

einen Zählprozess, der die Anzahl der Zufallszahlen Y_i , die in der Menge A enthalten sind, zählt.

Ein Zählprozess $N(\cdot)$ wird ein Poissonzählprozess auf X mit Mittelwertfunktion Λ genannt, wenn folgende Bedingungen erfüllt sind.

- Für $A \subset X$ und alle $k \geq 0$ gilt:

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

- Für alle $m \geq 1$ und für alle paarweise disjunkten Mengen $A_1, \dots, A_m \subset X$ sind die Zufallszahlen $N(A_1), \dots, N(A_m)$ unabhängig.

Für einen Poissonprozess $N(\cdot)$ gilt, dass $\mathbb{E}(N(A)) = \Lambda(A)$.

Weiters ist die so genannte Intensität $\lambda(x)$ nicht ohne Bedeutung, da für sie gilt, dass $\int_A \lambda(x) dx = \Lambda(A)$.

Wir wollen nun wieder das strikte weiße Rauschen $(X_i)_{i \in \mathbb{N}}$ und die dazugehörige Folge von Schranken $u_n(x) = c_n x + d_n$ für ein fixes x betrachten. Sei $Y_{i,n} = \binom{i}{n} \mathbb{I}_{\{X_i > u_n(x)\}}$ für $1 \leq i \leq n$, so erkennt man, dass $Y_{i,n}$ entweder die "normierte Zeit" oder Null zurückgibt.

Der Zählprozess $N(\cdot)$, der die Überschreitungen über die Schranke $u_n(x)$ in einem normierten Zeitraum zählt, ist für den Zustandsraum $X = (0, 1]$ durch

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

für alle $A \subset X$ gegeben.

Man erhält somit eine Folge von Zählprozessen, für die man zeigen kann, dass sie für $n \rightarrow \infty$ in der Verteilung gegen einen Poissonprozess $N(\cdot)$ konvergiert, dessen Mittelwertfunktion durch $\Lambda(A) = (t_2 - t_1)\lambda(x)$ für $A = (t_1, t_2)$ gegeben ist, wobei $\lambda(x) = -\ln H_\xi(x)$.

Es handelt sich somit um einen homogenen Poissonprozess.

Dies führt uns zu folgender Aussage:

$$\mathbb{E}(N_n(A)) \rightarrow \mathbb{E}(N(A)) = \Lambda(A) = (t_2 - t_1)\lambda(x).$$

4.2 Das POT (Peaks-Over-Threshold) Modell

Die gerade im vorherigen Abschnitt vorgestellte Theorie kombiniert mit der Theorie aus dem 3. Kapitel führt uns zu einem asymptotischen Modell für Grenzüberschreitungen von Daten aus identisch unabhängig verteilten stochastischen Größen oder aus stochastischen Prozessen mit Extremal-Index $\theta = 1$.

Das POT-Modell macht folgende Annahmen:

- Überschreitungen folgen in der Zeit einem homogenen Poissonprozess.
- Die Höhe der Überschreitungen über eine Schranke sind unabhängig sowie identisch verteilt und sie sind unabhängig von den Überschreitungszeiten.
- Die Verteilung der Höhe der Überschreitung ist eine verallgemeinerte Paretoverteilung.

Es gibt mehrere alternative Wege das Modell zu beschreiben, wie zum Beispiel durch den zwei-dimensionalen Poissonprozess, bei dem die Punkte (t, x) für Zeitpunkt und Höhe der Grenzüberschreitung stehen.

Wir nehmen an X_1, \dots, X_n seien Zufallsvariablen, die einen Verlust beschreiben und u sei eine festgesetzte hohe Schranke.

Weiters nehmen wir an, dass der auf dem Zustandsraum $X = (0, 1] \times (u, \infty)$ definierte Zählprozess $N(A) = \sum_{i=1}^n \mathbb{I}_{\{(\frac{i}{n}, X_i) \in A\}}$ ein Poissonprozess ist, dessen Intensität durch

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

für $(1 + \xi \frac{x - \mu}{\sigma}) > 0$, und $\lambda(t, x) = 0$ sonst, gegeben ist.

Man beachte, dass die Intensität nur von x aber nicht von t abhängig ist und, dass der zwei-dimensionale Poissonprozess nicht homogen ist.

Wir vereinfachen daher die Schreibweise der Intensität zu $\lambda(x) := \lambda(t, x)$ und erhalten somit für die Menge $A = (t_1, t_2) \times (x, \infty) \subset X$ die Mittelwertfunktion:

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

Es folgt daraus, dass für ein $x \geq u$ der implizierte ein-dimensionale Prozess von Überschreitungen der Schranke x ein homogener Poissonprozess mit Intensität $\tau(x) := -\ln H_{\xi, \mu, \sigma}(x)$ ist.

Die Wahrscheinlichkeit des Tails einer Excess-Funktion $\bar{F}_u(x)$ kann als Quotient der Intensitäten von $u+x$ und x angesehen werden, was uns zu bereits Bekanntem führt:

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

mit dem Skalierungsparameter $\beta = \sigma + \xi(u - \mu)$.

Es sei auch anzumerken, dass dieses Modell auch das aus dem 2. Kapitel bekannte Modell für Maxima impliziert.

Man betrachte das Ereignis $\{M_n \leq x\}$ für ein $x \geq u$, das bedeutet, dass in der gesamten Zeit kein Wert über x eintritt und somit der Zählprozess gleich Null für die Menge $A = (0, 1] \times (x, \infty)$ ist.

Wendet man dies nun an, so sieht man den Zusammenhang zu Kapitel 2:

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

für $x \geq u$.

4.3 Self-Exciting Prozesse

Bisher hatten wir homogene Poissonprozesse um die Eintrittszeiten von Levelüberschreitungen zu modellieren, jetzt wollen wir self-exciting Zählprozesse betrachten.

In diesen Modellen kann eine kürzlich eingetretene Serie von Überschreitungen das Risiko einer neuerlichen Überschreitung erhöhen.

Die Hauptanwendung dieser Modelle findet traditionell bei der Modellierung von Erdbeben und deren Nachbeben statt, sie scheinen jedoch auch zur Modellierung von Einbrüchen auf den Finanzmärkten nützlich zu sein.

Deshalb möchte ich noch über dieses Modell einen kleinen Überblick geben.

Wie üblich haben wir die Daten X_1, \dots, X_n und eine Schranke u gegeben, wobei die Zufallsvariable N_u die Anzahl der Überschreitungen beschreibt. Von jetzt an wollen wir die Zeit der Überschreitung in der natürlichen Skalierung angeben und schreiben die Daten $\{(i, X_i) : 1 \leq i \leq n, X_i > u\}$ in $\{(T_j, \tilde{X}_j) : j = 1, \dots, N_u\}$ um.

Die Zufallszahl $Y_i = i \mathbb{I}_{\{X_i > u\}}$ gibt den Zeitpunkt der Überschreitung zurück, sofern sie überhaupt stattfindet.

Der Zählprozess der Überschreitungen $N(\cdot)$ mit dem Zustandsraum $X = (0, n]$ ist durch $N(A) = \sum_{i=1}^n \mathbb{I}_{\{Y_i \in A\}}$ für ein $A \subset X$ gegeben.

Wir nehmen an, der Zählprozess $N(\cdot)$ sei ein self-exciting Prozess mit bedingter Intensität

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

wobei $\tau > 0$, $\psi \geq 0$ und h eine positive Funktion ist.

Jede vorherige Überschreitung (T_j, \tilde{X}_j) spielt in die bedingte Dichte mit ein, einerseits spielt die Zeit wie lange die Überschreitung her ist $(t - T_j)$, andererseits die Höhe der Überschreitung $(\tilde{X}_j - u)$ hier mit ein.

Nun möchte ich noch zwei Beispiele für die positive Funktion h geben:

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

Literatur

- [1] ALEXANDER J. MCNEIL, RÜDIGER FREY, UND PAUL EMBRECHTS:
Quantitative Risk Management, Princeton Series in Finance, Kapitel 7
- [2] W.GURKER: *Skriptum zur Vorlesung Angewandte Statistik*, S 2007