Matthias Löwe

Extremwerttheorie

WS 2008/09

Inhaltsverzeichnis

1 Extremwertverteilungen von iid-Beobachtungen

Rekorde sind ein Phänomen, das uns im täglichen Leben interessiert. Wir sind auf der Suche nach dem höchsten Berg, dem längsten Fluss, dem heißesten Tag. Bei Sportveranstaltungen küren wir die schnellste Frau oder den Mann, der am höchsten springt, einen Speer am weitesten wirft.

Rekorde sind aber auch von großer Wichtigkeit für praktische Belange: In den Vorlesungen über Wahrscheinlichkeitstheorie haben wir gelernt, wie sich Summen von Zufallsvariablen verhalten. Dies hilft uns, etliche praktische Problemstellungen lösen zu können. Will man aber zum Beispiel einen Deich bauen, so interessiert einen die durchschnittliche Höhe einer Flut wenig. Ebenso ist man am Extremalverhalten interessiert, wenn man die notwendige Sicherheitsreserve einer Bank oder eines Versicherungsunternehmens festlegen möchte. Auch hier hilft es einem wenig zu wissen, wie hoch die durchschnittlichen Forderungen an eine Versicherung sind – wir wollen wissen, wie groß eine solche Forderung maximal sein kann (und dies mit großer Wahrscheinlichkeit). Nun sind Wasserstände ebenso einem stochastischen Einfluss unterworfen wie die Forderungen an ein Versicherungsunternehmen. Es bleibt uns wenig anderes übrig als sie als stochastischen Prozess zu modellieren. Wir wollen daher in der Folge das Extremwertverhalten von stochastischen Prozessen untersuchen. Wir beginnen mit Folgen von iid Zufallsvariablen. Es sei also $(X_1, X_2, \ldots)$ eine Folge von iid Zufallsvariablen mit der Verteilungsfunktion F. Da wir uns für die Extrema der $(X_i)_i$ interessieren, bezeichnen wir mit

$$M_n := \max_{1 \le i \le n} X_i$$

das Maximum der ersten n der X_i und mit

$$m_n := \min_{1 \le i \le n} X_i$$

ihr Minimum. Es liegt nahe, dass wir – fragen wir nach einer Grenzwertverteilung von M_n – das M_n skalieren müssen (auch im Zentralen Grenzwertsatz benötigen wir ja z. B. eine Skalierung). Tatsächlich ist das Grenzwertverhalten ohne eine solche Skala leicht festzustellen: Mit Hilfe des Borel-Cantelli-Lemmas sieht man leicht, dass M_n $\mathbb{P}$-f.s. gegen den rechten Randpunkt der Verteilung, also gegen

$$x_R = \sup_x \{x : F(x) = 1\}$$

konvergiert. Wir fragen also, ob es Folgen (A_n) und (B_n) bzw. (a_n) und (b_n) gibt, sowie Grenzverteilungen G bzw. H gibt, so dass

$$A_n(M_n - B_n) \xrightarrow{\mathcal{D}} G$$

bzw.

$$a_n(m_n - b_n) \xrightarrow{\mathcal{D}} H$$

gilt.

Die Frage ist natürlich, wie wir überhaupt an die Verteilung der Extrema kommen. Dazu dient das folgende Lemma.

Lemma 1.1 *Für alle $n \in \mathbb{N}$ und $x \in \mathbb{R}$ gilt*

$$\begin{aligned} F^{M_n}(x) &= F^n(x) \quad \textit{und} \\ F^{m_n}(x) &= 1-(1-F(x))^n. \end{aligned}$$

Hierbei sind F^{M_n} und F^{m_n} die Verteilungsfunktionen von M_n bzw. m_n.

Beweis: Es gilt

$$\begin{aligned} F^{M_n}(x) &= \mathbb{P}(\max_{1\le i\le n} X_i \le x) \\ &= \mathbb{P}(\bigcap_{i=1}^{n}\{X_i \le x\}) \\ &= \prod_{i=1}^{n} \mathbb{P}(X_i \le x) \\ &= F^n(x). \end{aligned}$$

Hierbei haben wir die Unabhängigkeit und die identische Verteilung der X_i benutzt. Analog folgt

$$\begin{aligned} F^{m_n}(x) &= \mathbb{P}(\min X_i \le x) \\ &= 1-\mathbb{P}(\bigcap_{i=1}^{n}\{X_i > x\}) \\ &= 1-\prod_{i=1}^{n} \mathbb{P}(X_i > x) \\ &= 1-(1-F(x))^n. \end{aligned}$$

□

Die eingangs gestellte Frage lässt sich also so umformulieren, dass wir Konstanten A_n, B_n bzw. a_n, b_n und Verteilungsfunktionen G bzw. H, so dass

$$F^n\left(\frac{x}{A_n}+B_n\right) \to G(x)$$

bzw.

$$\left(1-F\left(\frac{x}{a_n}+b_n\right)\right)^n \to 1-H(x).$$

Dies soll jeweils für alle Stetigkeitspunkte x der Funktionen G bzw. H gelten. Man bezeichnet dies folgendermaßen:

Definition 1.2 *Es seien F und G Verteilungsfunktionen.*

a) F liegt im Max-Anziehungsbereich von G, falls es Konstanten $A_n > 0, B_n$ gibt, so dass

$$F^n\left(\frac{x}{A_n} + B_n\right) \underset{n\to\infty}{\longrightarrow} G(x)$$

für alle x, in denen G stetig ist (Bezeichnung für diese Menge: $C(G)$). Wir schreiben:

$$F \in \mathcal{D}(G).$$

b) F liegt im Min-Anziehungsbereich von G, falls es Konstanten $a_n > 0$ und b_n gibt mit

$$\left(1 - F\left(\frac{x}{a_n} + b_n\right)\right)^n \to 1 - G(x)$$

für alle $x \in C(G)$.

Beispiel 1.3 *a) (X_n) seien iid Exp(λ)-verteilt. Diese haben die Verteilungsfunktion*

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Man kann sich z. B. vorstellen, dass die X_i die Zerfallsdauer von radioaktiven Isotopen modellieren. Wir wählen $A_n = \lambda$ und $B_n = \frac{\log n}{\lambda}$. Dann folgt für alle $x \in \mathbb{R}$:

$$\begin{aligned}
F^{\lambda M_n - \log n}(x) &= F^{M_n}\left(\frac{x}{\lambda} + \frac{\log n}{\lambda}\right) \\
&= F^n\left(\frac{x}{\lambda} + \frac{\log n}{\lambda}\right) \\
&= (1 - (e^{-x-\log n}))^n \mathbb{1}_{(0,\infty)}\left(\frac{x}{\lambda} + \frac{\log n}{\lambda}\right) \\
&= (1 - (e^{-x-\log n}))^n \mathbb{1}_{[-\log n,\infty)}(x) \\
&= (1 - \frac{e^{-x}}{n})^n \mathbb{1}_{[-\log n,\infty)}(x) \\
&\underset{n\to\infty}{\longrightarrow} e^{-e^{-x}} =: \Lambda(x),
\end{aligned}$$

da $-\log n$ für $n \to \infty$ gegen $-\infty$ konvergiert. Man prüft leicht nach, dass $\Lambda(\cdot)$ monoton ist und dass

$$\lim_{x\to-\infty} e^{-e^{-x}} = 0 \quad \text{und} \quad \lim_{x\to\infty} e^{-e^{-x}} = 1$$

gilt. Also ist $\Lambda(\cdot)$ eine Verteilungsfunktion mit $C(\Lambda) = \mathbb{R}$.

b) Die Variablen $X_1, X_2, \ldots$ seien nun iid normalverteilt, genauer $X_1 \sim \mathcal{N}(0,1)$. Wieder suchen wir $A_n > 0$ und B_n, so dass $F^n(\frac{x}{A_n} + B_n)$ einen nicht-entarteten Limes hat. Nun ist

$$\mathbb{P}\left[A_n(M_n - B_N) \leq x\right] = \mathbb{P}\left[M_n \leq \frac{x}{A_n} + B_n\right] = \left(\Phi\left(\frac{x}{A_n} + B_n\right)\right)^n,$$

wobei Φ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist. Wenn wir

$$x_n := \frac{x}{A_n} + B_n$$

setzen, schreiben wir daher

$$\left(\Phi\left(\frac{x}{A_n} + B_n\right)\right)^n = (1 - (1 - \Phi(x_n))^n.$$

Dies konvergiert, wenn

$$(1 - \Phi(x_n)) = \frac{g(x)}{n} + \mathcal{O}(1/n)$$

und in diesem Falle ist der Limes

$$\lim_{n\to\infty} (1 - (1 - \Phi(x_n))^n = e^{-g(x)}.$$

Somit müssen wir die Folgen A_n und B_n so finden, dass

$$\frac{1}{\sqrt{2\pi}} \int_{x_n}^{\infty} e^{-y^2/2} dy = \frac{g(x)}{n}.$$

Nun lässt sich das Tail-Verhalten dieses Integrals recht gut abschätzen: Es ist

$$\frac{1}{\sqrt{2\pi}} \int_{x_n}^{\infty} e^{-y^2/2} dy \leq \int_{x_n}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{y}{x_n} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi} x_n} e^{-x_n^2/2}.$$

Ähnlich zeigt man

$$\frac{1}{\sqrt{2\pi}} \int_{x_n}^{\infty} e^{-y^2/2} dy \geq \frac{1}{x_n \sqrt{2\pi}} e^{-x_n^2/2} (1 - \frac{2}{x_n^2}).$$

Somit genügt es, x_n so zu wählen, dass

$$\frac{1}{x_n \sqrt{2\pi}} e^{-x_n^2/2} = \frac{g(x)}{n}.$$

Setzen wir x_n ein, so erhalten wir

$$\frac{g(x)}{n} = \frac{e^{-\frac{1}{2}(\frac{x}{A_n} + B_n)^2}}{\sqrt{2\pi}(\frac{x}{A_n} + B_n)} = \frac{e^{-B_n^2/2 - \frac{x^2}{2A_n^2} - \frac{B_n}{A_n} x}}{\sqrt{2\pi}(\frac{x}{A_n} + B_n)}. \tag{1.1}$$

Setzen wir $x = 0$, sehen wir, dass

$$e^{-B_n^2/2} / \sqrt{2\pi} B_n = g(0)/n$$

gilt. Wir setzen

$$B_n = \sqrt{2 \log n} + C_n$$

und bekommen dann für C_n

$$e^{-\sqrt{2 \log n C_n} - C_n^2/2} = g(0) \sqrt{2\pi} (\sqrt{2 \log n} + C_n).$$

Wir normieren g, so dass $g(0) = 1$ gilt. Es lässt sich dann zeigen, dass die führenden Terme für C_n von der Gestalt

$$C_n = -\frac{\log\log n + \log(4\pi)}{2\sqrt{2\log n}}$$

sind. Die höheren Ordnungsterme tragen nichts mehr bei. Damit uns der Faktor von x in (1.1) nicht abhaut, muss A_n approximativ so groß wie B_n sein. Wir setzen daher

$$A_n = \sqrt{2\log n}.$$

Setzt man dies in (1.1), so erhält man

$$g(x) = e^{-x} + o(\frac{1}{n}).$$

Mit unserer Eingangsüberlegung haben wir somit gezeigt, dass mit der Wahl

$$B_n := \sqrt{2\log n} - \frac{\log\log n + \log(4\pi)}{2\sqrt{2\log n}}$$

und

$$A_n := \sqrt{2\log n}$$

für alle $x \in \mathbb{R}$ gilt

$$\lim_{n\to\infty} \mathbb{P}(A_n(M_n - B_n) \le x) = e^{-e^{-x}}.$$

Bemerkung 1.4 *Wir werden den Sachverhalt aus Beispiel 1.3 b) später noch in leicht modifizierter Form verwenden. Mit den gleichen Konstantenfolgen A_n und B_n sei*

$$u_n(x) := B_n + \frac{x}{A_n}.$$

Dann gilt

$$\lim_{n\to\infty} \mathbb{P}[M_n \le u_n(x)] = e^{-e^{-x}}.$$

Wir sehen hier, dass die Extreme einer Folge von Zufallsvariablen zwar wachsen, aber relativ langsam, nämlich wie $\sqrt{\log n}$. Für Gaußsche Zufallsvariablen lässt sich mit relativ guter Präzision behaupten, dass

$$M_n \sim \sqrt{2\log n}$$

gilt. Darüber hinaus haben wir gesehen, dass für zwei verschiedene Verteilungen, die Exponential- und die Gaußverteilung, die Extreme gegen die gleiche Limesverteilung konvergieren. Die Verteilung mit der Verteilungsfunktion $\Lambda(x) = e^{-e^{-x}}$ heißt auch die Gumbelverteilung. Es erhebt sich die Frage, ob dies die einzige Limesverteilung unter geeigneten Voraussetzungen ist (ähnlich wie die Normalverteilung im Zentralen Grenzwertsatz) oder ob noch andere Verteilungen auftreten können. Hierbei muss man die gleichen Unterscheidungen wie beim Zentralen Grenzwertsatz treffen. Sind die X_i iid mit $\mathbb{E}(X_1) = \mu$ und $0 \neq \mathbb{E}((X_1 - \mu)^2) < +\infty$, so besagt der Zentrale Grenzwertsatz, dass

$$\frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n\mathbb{V}X_1}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

aber auch

$$\frac{\sum_{i=1}^n (X_i - \mu) + 2\mu}{\sqrt{2n\mathbb{V}X_1}} \xrightarrow{\mathcal{D}} \mathcal{N}(2\mu, \frac{1}{2}).$$

Natürlich würde niemand sage, dass dies verschiedene Limiten für die Summe der X_i sind. Eine entsprechende Untersuchung trifft die folgende Definition.

Definition 1.5 *Zwei Verteilungsfunktionen G und $\tilde{G}$ heißen* vom selben Typ, *falls Konstanten $c > 0$ und $d \in \mathbb{R}$ existieren mit*

$$\tilde{G}(x) = G(cx + d) \quad \forall\, x \in \mathbb{R}.$$

Ist also $F \in \mathcal{D}(G)$, dann gilt auch $F \in \mathcal{D}(\tilde{G})$.

Darüber hinaus bemerken wir noch, dass wegen

$$M_n^- := \max_{1 \le i \le n} (-X_i) = -\min_{1 \le i \le n} X_i$$

und

$$m_n^- := \min_{1 \le i \le n} (-X_i) = -\max_{1 \le i \le n} X_i$$

folgendes gilt:

- Gilt für $A_n > 0,\ B_n \in \mathbb{R}$ und stetiges G, dass
  $$\mathbb{P}^{A_n(M_n - B_n)} \xrightarrow{\mathcal{D}} G,$$
  so folgt
  $$\mathbb{P}^{A_n(m_n^- + B_n)} \xrightarrow{\mathcal{D}} H$$
  mit $H(x) := 1 - G(-x) \quad \forall\, x \in \mathbb{R}$.
- Gilt für $a_n > 0,\ b_n \in \mathbb{R}$ und stetiges H, dass
  $$\mathbb{P}^{a_n(m_n - b_n)} \xrightarrow{\mathcal{D}} H,$$
  so folgt
  $$\mathbb{P}^{a_n(M_n^- + b_n)} \xrightarrow{\mathcal{D}} G,$$
  wobei $G(x) := 1 - H(-x) \quad \forall\, x \in \mathbb{R}$. Will man dies begründen, so berechnet man
  $$\begin{aligned} \mathbb{P}(A_n(m_n^- + B_n) \le x) &= \mathbb{P}(-A_n(M_n - B_n) \le x) \\ &= 1 - \mathbb{P}(A_n(M_n - B_n) < -x) \underset{n\to\infty}{\longrightarrow} 1 - G(-x). \end{aligned}$$

Die zweite Tatsache beweist man analog. Wir können uns also auf das Studium der Maxima eines Prozesses konzentrieren (oder, wenn dies praktischer sein sollte, auf die Minima).

Wir haben in Beispiel 1.3 schon die Gumbel-Verteilung $\Lambda(x) = e^{-e^{-x}}$ als mögliche Limesverteilung kennengelernt. Wir sollen in der Folge sehen, welche Verteilungen noch mögliche Limes-Verteilungen sind. Denkt man an den CLT zurück, so ist ein wesentlicher Grund für das Auftreten der Normalverteilung dort, dass die Normalverteilung stabil ist unter Faltung, d. h. sind die $X_i \sim \mathcal{N}(0,1)$ und iid, so ist

$$\frac{\sum_{i=1}^n X_i}{\sqrt{n}} \sim \mathcal{N}(0,1).$$

Wir wollen sehen, ob Ähnliches für die Gumbelverteilung gilt:

Beispiel 1.6 *Es gelte $F^{X_i} = \Lambda(x) = e^{-e^{-x}}$. Dann gilt für $M_n - \log n$, d. h. für $A_n = 1$ und $B_n = \log n$*

$$\begin{aligned} F^{M_n - \log n}(x) &= \Lambda^n(x + \log n) \\ &= (e^{-e^{-x-\log n}})^n = (e^{-\frac{e^{-x}}{n}})^n \\ &= e^{-e^{-x}} = \Lambda(x) \quad \forall\, x \in \mathbb{R},\ \forall\, n \in \mathbb{N}. \end{aligned}$$

Λ erhält man also auch als Verteilung des Maximums von richtig skalierten Λ-verteilten iid Zufallsvariablen.

Wir nehmen dieses Beispiel zum Anlass für folgende Definition:

Definition 1.7 *Eine nicht-entartete Verteilungsfunktion G heißt genau dann* max-stabil*, wenn für geeignete Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$ gilt*

$$G^n(a_n^{-1}x + b_n) = G(x) \quad \forall\, x \in \mathbb{R} \quad \forall n \in \mathbb{N}.$$

Eine nicht-entartete Verteilung heißt max-stabil*, wenn die zugehörige Verteilungsfunktion max-stabil ist.*

Definition 1.8

- *Die Verteilung mit der Verteilungsfunktion*

$$\Phi_\alpha(x) = e^{-x^{-\alpha}} \mathbb{1}_{(0,\infty)}(x),\ \alpha > 0,$$

heißt Fréchet-Verteilung.

- *Die Verteilung mit der Verteilungsfunktion*

$$\psi_\alpha(x) = e^{-(-x)^\alpha} \mathbb{1}_{(-\infty,0)}(x) + \mathbb{1}_{[0,\infty)}(x)$$

(für $\alpha > 0$) heißt Weibull-Verteilung.

Beispiel 1.9 *(i) Es sei $F^{X_n} = \Phi_\alpha$. Für $A_n = n^{-1/\alpha}$ und $B_n = 0$ gilt dann*

$$\begin{aligned} F^{n^{-1/\alpha} M_n}(x) &= F^{M_n}(xn^{1/\alpha}) = \Phi_\alpha^n(xn^{1/\alpha}) \\ &= \left(e^{-(xn^{1/\alpha})^{-\alpha}}\right)^n \mathbb{1}_{(0,\infty)}(x_n^{1/\alpha}) \\ &= \left(e^{\frac{-x^{-\alpha}}{n}}\right)^n \mathbb{1}_{(0,\infty)}(x) \\ &= e^{-x^{-\alpha}} \mathbb{1}_{(0,\infty)}(x) \\ &= \Phi_\alpha(x) \quad \forall\, x \in \mathbb{R}, \forall\, n \in \mathbb{N}. \end{aligned}$$

Also ist die Fréchet-Verteilung max-stabil.

(ii) Es sei $F^{X_n} = \psi_\alpha$, $\alpha > 0$. Sei $A_n = n^{1/\alpha}$ und $B_n = 0$. Dann gilt:

$$\begin{aligned}
F^{M_n n^{1/\alpha}}(x) &= F^{M_n}(xn^{-1/\alpha}) = \psi_\alpha^n(xn^{-1/\alpha}) \\
&= \left(e^{-(-xn^{-1/\alpha})^\alpha}\right)^n \mathbb{1}_{(-\infty,0)}(xn^{-1/\alpha}) + \mathbb{1}_{[0,\infty)}(xn^{-1/\alpha}) \\
&= \left(e^{-((-x)^\alpha n^{-1})}\right)^n \mathbb{1}_{(-\infty,0)}(x) + \mathbb{1}_{[0,\infty)}(x) \\
&= e^{-(-x)^\alpha} \mathbb{1}_{(-\infty,0)}(x) + \mathbb{1}_{[0,\infty)}(x) \\
&= \psi_\alpha(x) \quad \forall\, x \in \mathbb{R},\ n \in \mathbb{N}.
\end{aligned}$$

Somit ist auch die Weibull-Verteilung max-stabil.

Wir wollen nun die möglichen Limesverteilungen für Maxima von iid Zufallsvariablen charakterisieren. Ein erster Schritt in diese Richtung ist die Beschreibung der max-stabilen Verteilungen. Wir beginnen mit dem folgenden Satz.

Satz 1.10 *Eine nicht-entartete Verteilungsfunktion G ist genau dann max-stabil, wenn es Folgen (F_n) von Verteilungsfunktionen und $(A_n), (B_n)$ von Konstanten mit $A_n > 0$ gibt, so dass*

$$F_n(A_{nk}^{-1}x + B_{nk}) \underset{n\to\infty}{\longrightarrow} G^{1/k}(x)$$

für alle $x \in C(G)$ und alle $k \in \mathbb{N}$ gilt.

Für den Beweis dieses Satzes benötigen wir den Satz von Khintchin, der wiederum einiger Vorbereitung bedarf.

Definition 1.11 *Sei φ eine reellwertige, monoton nicht-fallende, rechtsseitig stetige Funktion. Ferner sei*

$$\begin{aligned}
\inf(\varphi) &:= \inf\{\varphi(x), x \in \mathrm{dom}(\varphi)\} \quad \textit{und} \\
\sup(\varphi) &:= \sup\{\varphi(x), x \in \mathrm{dom}\varphi\}.
\end{aligned}$$

Dann kann die Pseudo-Inverse φ^- von φ durch

$$\begin{aligned}
&\varphi^- : (\inf(\varphi), \sup(\varphi)) \to \overline{\mathrm{dom}(\varphi)} \\
&\varphi^-(y) := \inf\{x \in \mathrm{dom}(\varphi) : \varphi(x) \geq y\}
\end{aligned}$$

definiert werden. Dabei ist $\mathrm{dom}(\varphi)$ der Definitionsbereich von φ und $\overline{\mathrm{dom}(\varphi)}$ sein Abschluss.

Wir sammeln ein paar wichtige Eigenschaften der Pseudo-Inversen.

Lemma 1.12 *φ^- sei die Pseudo-Inverse einer rechtsseitig stetigen, monoton nicht-fallenden Funktion. Dann gilt:*

(i) φ^- *ist monoton nicht-fallend und linksseitig stetig.*

(ii) Ist φ *in* $\varphi^-(y)$ *stetig, so gilt* $\varphi(\varphi^-(y)) = y$.

(iii) Ist φ^- *in* $\varphi(x) \in \mathrm{dom}(\varphi^-)$ *stetig, so gilt* $\varphi^-(\varphi(x)) = x$.

(iv) Es gilt

- $\varphi(\varphi^-(y)) \geq y \quad \forall\, y \in \mathrm{dom}(\varphi^-)$.
- $\varphi^-(\varphi(x)) \leq x \quad \forall\, x : \varphi(x) \in \mathrm{dom}(\varphi^-)$.

Beweis:

(iv) Die zweite Aussage folgt direkt aus der Definition von φ^-. Die erste Aussage folgt mit Hilfe der rechtsseitigen Stetigkeit von φ aus der Definition von φ^-.

(i) Die Monotonie von φ^- folgt direkt aus der Monotonie von φ. Es bleibt die linksseitige Stetigkeit von φ^-. Dazu sei $(y_n)_n$ monoton nicht-fallend aus $\mathrm{dom}(\varphi^-)$ und es gelte

$$\lim_{n\to\infty} y_n = y \in \mathrm{dom}(\varphi^-).$$

Dann konvergiert auch die Folge $(\varphi^-(y_n))_n$, da sie monoton nicht-fallend und durch $\varphi^-(y)$ beschränkt ist. Der Grenzwert heiße z. Zu zeigen ist, dass

$$z := \lim_{n\to\infty} \varphi^-(y_n) = \varphi^-(y)$$

gilt. Zunächst ist

$$z \leq \varphi^-(y), \quad \text{da} \quad \varphi^-(y_n) \leq \varphi^-(y)$$

für alle $n \in \mathbb{N}$ gilt. Angenommen, dass $z < \varphi^-(y)$ gilt. Dann gälte nach Definition bzw. mit Hilfe von (iv)

$$\varphi(z) < y \quad \text{und} \quad \varphi(\varphi^-(y_n)) \geq y_n \quad \forall\, n \in \mathbb{N}.$$

Daraus würde dann

$$\liminf_{n\to\infty} \varphi(\varphi^-(y_n)) \geq y > \varphi(z)$$

folgen, im Widerspruch zur Monotonie von φ, da $(\varphi^-(y_n))$ monoton nicht-fallend ist. Also gilt $z = \varphi^-(y_n)$ und φ^- ist linksseitig stetig.

(ii) Sei $y \in \mathrm{dom}(\varphi^-)$ und $\varphi^-(y)$ ein Stetigkeitspunkt von φ. $(X_n)_n$ sei monoton nicht-fallend mit

$$X_n < \varphi^-(y) \quad \forall\, n \in \mathbb{N} \quad \text{und} \quad \lim_{n\to\infty} x_n = \varphi^-(y).$$

Dann gilt nach Definition $\varphi(x_n) < y \quad \forall\, n \in \mathbb{N}$ und es folgt mit (iv)

$$\lim_{n\to\infty} \varphi(x_n) \leq y \leq \varphi(\varphi^-(y)) = \varphi(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} \varphi(x_n),$$

wobei die letzte Gleichheit gilt, da $\varphi^-(y)$ ein Stetigkeitspunkt von φ ist. Damit folgt

$$\varphi(\varphi^-(y)) = y.$$

(iii) Dies geht analog zu (ii). $\varphi(x)$ sei ein Stetigkeitspunkt von φ^- und (y_n) monoton nicht-wachsend mit

$$y_n > \varphi(x) \quad \forall\, n \in \mathbb{N} \quad \text{und} \quad \lim_{n\to\infty} y_n = \varphi(x).$$

Dann gilt nach Definition $\varphi^-(y_n) > x$ für alle $n \in \mathbb{N}$, und es folgt

$$\lim_{n\to\infty} \varphi^-(y_n) \geq x \geq \varphi^-(\varphi(x)) = \varphi^-(\lim y_n) = \lim_{n\to\infty} \varphi^-(y_n),$$

wobei die letzte Gleichung wieder wegen der Stetigkeit von φ^- in $\varphi(x)$ gilt. Also folgt

$$\varphi^-(\varphi(x)) = x.$$

□

Seinen Zweck erfüllt dieses Lemma durch Anwendung auf die Verteilungsfunktion von Zufallsgrößen. Das folgende Lemma haben wir schon beim Kolmogorov-Smirnov-Test kennengelernt.

Lemma 1.13 *(i) Sei X eine Zufallsvariable mit stetiger Verteilungsfunktion F. Dann gilt*

$$F(X) \sim \mathcal{R}(0,1)$$

(dabei ist $\mathcal{R}(a,b)$ die Rechteck-Verteilung auf dem Intervall $[a,b]$).

(ii) Sei F eine Verteilungsfunktion und $U \sim \mathcal{R}(0,1)$. Dann gilt

$$F^-(U) \sim F.$$

Beweis:

(i) Da F stetig ist, ist $F(X)$ messbar und daher eine Zufallsvariable. Außerdem ist F^- auf (0,1) definiert, da F eine Verteilungsfunktion ist. Somit sind die Mengen

$$A_x := \{y \in \mathbb{R} : F(y) \leq x\} \quad \text{und}$$
$$B_x := \{y \in \mathbb{R} : y \leq F^-(x)\}$$

für $0 < x < 1$ wohldefiniert. Zunächst zeigen wir

$$A_x = B_x \quad \forall\, x \in C(F^-). \tag{1.2}$$

"$\subseteq$": Es sei $x \in C(F^-)$ und $y \in A_x$.
1. Fall: $F(y) = x$. Dann gilt mit Lemma 1.12 (iii)

$$F^-(x) = F^-(F(y)) = y,$$

also $y \in B_x$.

2. Fall: $F(y) < x$. Angenommen, es gälte $y \geq F^-(x)$. Dann folgt mit der schwachen Monotonie von F und Lemma 1.12 (iv)

$$F(y) \geq F(F^-(x)) \geq x$$

im Widerspruch zu $F(y) < x$. Also gilt

$$y < F^-(x), \quad \text{also} \quad y \in B_x.$$

"$\supseteq$": Es sei $0 < x < 1$ und $y \in B_x$, d. h. $y \leq F^-(x)$. Mit der Monotonie und der Stetigkeit von F folgt daraus

$$F(y) \leq F(F^-(x)) = x,$$

also $y \in A_x$.

Somit ist (1.2) gezeigt. Da F^- monoton nicht-fallend ist, ist die Menge der Unstetigkeitsstellen von F^- und somit auch die Menge

$$M_{\neq} := \{x \in (0,1) : A_x \neq B_x\}$$

höchstens abzählbar. Für alle $x \in (0,1)\backslash M_{\neq}$ gilt aber:

$$\begin{aligned} \{\omega \in \Omega : F(X(\omega)) \leq x\} &= \{\omega \in \Omega : X(\omega) \in A_x\} \\ &= \{\omega \in \Omega : X(\omega) \in B_x\} \\ &= \{\omega \in \Omega : X(\omega) \leq F^-(x)\}. \end{aligned}$$

D. h. für alle $x \in (0,1)\backslash M_{\neq}$ gilt

$$\mathbb{P}(F(X) \leq x) = \mathbb{P}(X \leq F^-(x)) = F(F^-(x)) = x,$$

wobei die letzte Gleichheit aus der Stetigkeit von F folgt. Mit der Stetigkeit von F folgt die Aussage aber schon für alle $x \in \mathbb{R}$, da die höchstens abzählbar vielen Unstetigkeitsstellen von F^- isoliert liegen müssen. Also gilt $F(X) \sim \mathcal{R}(0,1)$.

(ii) Mit Lemma 1.12 (iv) und der Monotonie von F und F^- gilt für alle $0 < F(x) < 1$

$$\{\omega \in \Omega : F^-(U(\omega)) \leq x\} \subseteq \{\omega \in \Omega : U(\omega) \leq F(F^-(U(\omega)) \leq F(x)\}$$

und

$$\{\omega \in \Omega : U(\omega) \leq F(x)\} \subseteq \{\omega \in \Omega : F^-(U(\omega)) \leq F^-(F(x)) \leq x\}.$$

Daraus folgt

$$\mathbb{P}(F^-(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x)$$

für alle $0 < F(x) < 1$, wobei die letzte Gleichheit gilt, da U $\mathcal{R}(0,1)$-verteilt ist. Also besitzt $F^-(U)$ die Verteilungsfunktion F.

□

Des weiteren wird es wichtig sein, den Zusammenhang zwischen Verteilungsfunktionen und Zufallsvariablen herzustellen. Dies leistet der Satz von Skohorod.

Satz 1.14 *(Übertragungssatz von Skohorod)*
Es sei $(F_n)_n$ eine Folge von Verteilungsfunktionen und G eine Verteilungsfunktion, so dass

$$F_n(x) \underset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in C(G).$$

Dann existieren Zufallsvariablen $(X_n)_n$ und Y, so dass X_n die Verteilungsfunktion F_n hat für jedes n und Y die Verteilungsfunktion G besitzt und

$$X_n \underset{n\to\infty}{\longrightarrow} Y \quad \mathbb{P}\text{-}f.s.$$

gilt.

Beweis: Es sei U $\mathcal{R}(0,1)$-verteilt. Es wird nun gezeigt, dass die durch $X_n = F_n^-(U)$ und $Y = G^-(U)$ definierten Zufallsvariablen den gewünschten Bedingungen genügen. Aus Lemma 1.13 folgt, dass F_n für jedes n die Verteilungsfunktion von X_n ist und Y die Verteilungsfunktion G besitzt. Es bleibt die behauptete fast-sichere Konvergenz zu zeigen, also

$$\mathbb{P}(\lim X_n = Y) = 1.$$

Nun ist aber

$$\begin{aligned}\mathbb{P}(\lim_{n\to\infty} X_n = Y) &= \mathbb{P}(\lim_{n\to\infty} F_n^-(U) = G^-(U)) \\ &= \mathbb{P}^U(\lim_{n\to\infty} F_n^-(x) = G^-(x)).\end{aligned}$$

Da U auf (0,1) gleichverteilt ist und somit vom Lebesguemaß dominiert wird und jede höchstens abzählbare Menge eine $\lambda\!\!\lambda$-Nullmenge ist, genügt es zu zeigen:

$$F_n^-(x) \underset{n\to\infty}{\longrightarrow} G^-(x) \quad \text{für } \lambda\!\!\lambda\text{-fast alle } x \in (0,1).$$

Es sei daher $x_0 \in (0,1) \cap C(G^-)$ und $\varepsilon > 0$. Da G höchstens abzählbar viele Unstetigkeitsstellen besitzt, liegt $C(G)$ dicht in dom(G). Daher existieren $y, z \in C(G)$ mit

$$y < G^-(x_0) < z \quad \text{und} \quad z - y < \varepsilon. \tag{1.3}$$

Wir zeigen nun

$$G(y) < x_0 < G(z). \tag{1.4}$$

Angenommen, es gelte $G(y) \geq x_0$, dann folgt aus Lemma 1.12 (iv) und der Monotonie von G^-

$$y \geq G^-(G(y)) \geq G^-(x_0)$$

im Widerspruch zu $y < G^-(x_0)$. Umgekehrt folgt aus $G^-(x_0) < z$ mit Lemma 1.12 (iv) und der Monotonie von G

$$x_0 \leq G(G^-(x_0)) \leq G(z).$$

Angenommen, es gelte $x_0 = G(z)$. Dann wäre aber für alle $\delta > 0$ mit $x_0 + \delta < \sup(G)$

$$G^-(x_0 + \delta) = \inf\{u : G(u) \geq x_0 + \delta\} > z.$$

Da voraussetzungsgemäß $G^-(x_0) < z$ gilt, wäre das ein Widerspruch zur Stetigkeit von G^- in x_0. Damit ist (1.4) gezeigt. Da

$$F_n(x) \overset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in C(G)$$

gilt und wegen $G(y) < x_0 < G(z)$ und $y, z \in C(G)$ existiert ein $u(\varepsilon) \in \mathbb{N}$ mit

$$F_n(y) < x_0 < F_n(z) \quad \forall\, n > u(\varepsilon).$$

Daraus folgt

$$y \leq F_n^-(x_0) \leq z \quad \forall\, n > u(\varepsilon),$$

denn mit der Monotonie von F_n^- und Lemma 1.12 (iv) gilt

$$F_n^-(x_0) \leq F_n^-(F_n(z)) \leq z.$$

Für die andere Ungleichung führt man die Annahme $y > F_n^-(x_0)$ zum Widerspruch. Denn aus dieser Annahme würde mit der Monotonie von F_n und Lemma 1.12 (iv)

$$F_n(y) \geq F_n(F_n^-(x_0)) \geq x_0$$

folgen. Es gilt somit

$$y \leq F_n^-(x_0) \quad \text{und} \quad G^-(x_0) \leq z \quad \forall\, n \geq n(\varepsilon).$$

Mit (1.3) ergibt sich daher

$$|F_n^-(x_0) - G^-(x_0)| \leq z - y < \varepsilon \quad \forall\, n \geq u_\varepsilon.$$

Da G^- höchstens abzählbar viele Unstetigkeitsstellen besitzt, folgt die Behauptung. $\square$

Eine direkte Konsequenz ist

Lemma 1.15 *Es sei $(F_n)_n$ eine Folge von Verteilungsfunktionen und G eine Verteilungsfunktion, so dass*

$$F_n(x) \to G(x) \quad \forall\, x \in C(G).$$

$(\alpha_n)_n$ und (β_n) seien Folgen reeller Zahlen mit $\alpha_n > 0 \quad \forall n$,

$$\alpha_n \to \alpha > 0 \quad \textit{und} \quad \beta_n \to \beta \in \mathbb{R}.$$

Dann gilt

$$F_n(\alpha_n x + \beta_n) \underset{n\to\infty}{\longrightarrow} G(\alpha x + \beta)$$

für alle $\alpha x + \beta \in C(G)$.

Beweis: Es seien $U \sim \mathcal{R}(0,1)$, $Y := G^-(U)$ und $X_n := F_n^-(U)$ für jedes $n \in \mathbb{N}$. Wieder haben die X_n die Verteilungsfunktion F_n für jedes n und Y hat die Verteilungsfunktion G. Mit Satz 1.14 folgt

$$X_n \to Y \quad \mathbb{P}\text{-f.s.},$$

daher auch

$$\frac{X_n - \beta_n}{\alpha_n} \to \frac{Y - \beta}{\alpha} \quad \mathbb{P}\text{-f.s.}$$

Da fast-sichere Konvergenz Verteilungskonvergenz impliziert, gilt auch

$$\mathbb{P}\left(\frac{X_n - \beta_n}{\alpha_n} \le x\right) \to \mathbb{P}\left(\frac{Y - \beta}{\alpha} \le x\right)$$

für alle $\alpha x + \beta \in C(G)$, d. h.

$$F_n(\alpha_n x + \beta_n) \underset{n\to\infty}{\longrightarrow} G(\alpha x + \beta)$$

für alle $\alpha x + \beta \in C(G)$. □

Damit kann man ein wichtiges Hilfsmittel zeigen, das auch eigenständige Bedeutung besitzt.

Satz 1.16 *(Khinchin)*
Es sei G eine nicht-entartete Verteilungsfunktion und $(F_n)_n$ eine Folge von Verteilungsfunktionen. Es seien (α_n) und (β_n) Folgen reeller Zahlen mit $\alpha_n > 0$ und es gelte

$$F_n(\alpha_n x + \beta_n) \underset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in C(G).$$

Dann gilt die folgende Äquivalenz:
Für Folgen reeller Zahlen $(\tilde{\alpha}_n)$ und $(\tilde{\beta}_n)$ mit $\tilde{\alpha}_n > 0$ für alle n und eine geeignete nicht-entartete Verteilungsfunktion $\tilde{G}$ gilt genau dann

$$F_n(\tilde{\alpha}_n x + \tilde{\beta}_n) \underset{n\to\infty}{\longrightarrow} \tilde{G}(x) \quad \forall\, x \in C(G),$$

wenn Konstanten $\alpha > 0$ und $\beta \in \mathbb{R}$ existieren mit

$$\frac{\tilde{\alpha}_n}{\alpha_n} \underset{n\to\infty}{\longrightarrow} \alpha \quad \text{und} \quad \frac{\tilde{\beta}_n - \beta_n}{\alpha_n} \underset{n\to\infty}{\longrightarrow} \beta.$$

In diesem Fall gilt

$$\tilde{G}(x) = G(\alpha x + \beta),$$

d. h. G und $\tilde{G}$ sind vom selben Typ.

Neben seiner technischen Bedeutung beantwortet dieser Satz auch die Frage nach der Eindeutigkeit der Grenzverteilung normierter Maxima. Existiert für geeignet normierte Maxima $A_n(M_n - B_n)$ von iid Zufallsvariablen eine nicht-entartete Grenzverteilung, so ist diese bis auf Typ-Gleichheit auch eindeutig bestimmt. Darüber hinaus charakterisiert dieser Satz sämtliche Konstantenfolgen $A_n > 0$ und B_n, für die $F^n(A_n x + B_n)$ gegen eine nicht-entartete Verteilungsfunktion konvergiert in Abhängigkeit von einem schon gefundenen Paar (α_n), $\alpha_n > 0$ und (β_n).

Beweis: Wir setzen

$$\alpha_n^* := \frac{\tilde{\alpha}_n}{\alpha_n} \quad \text{und} \quad \beta_n^* := \frac{\tilde{\beta}_n - \beta_n}{\alpha_n} \quad \text{und} \quad F_n^*(x) = F_n(\alpha_n x + \beta_n) \quad \forall\, x \in \mathbb{R}.$$

"$\Leftarrow$": Es gebe also Konstanten $\alpha > 0$ und β mit $\alpha_n^* \to \alpha$ und $\beta_n^* \to \beta$ für $n \to \infty$. Wegen $\alpha_n, \tilde{\alpha}_n > 0 \quad \forall n$, gilt auch $\alpha_n^* > 0 \quad \forall n \in \mathbb{N}$. Nach Voraussetzung gilt

$$F_n^*(x) \to G(x) \quad \forall\, x \in C(G).$$

Damit sind die Voraussetzungen von Lemma 1.15 erfüllt. Nach diesem ergibt sich:

$$\lim_{n\to\infty} F_n(\tilde{\alpha}_n x + \tilde{\beta}_n) = \lim_{n\to\infty} F_n^*(\alpha_n^* x + \beta_n^*) = G(\alpha x + \beta) =: \tilde{G}(x) \quad \forall\, \alpha x + \beta \in C(G).$$

Somit gilt die Konvergenz auch für alle $x \in C(\tilde{G})$.

"$\Rightarrow$": Es gelte

$$F_n^*(x) \underset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in C(G)$$

und

$$F_n^*(\alpha_n^* x + \beta_n^*) \underset{n\to\infty}{\longrightarrow} \tilde{G}(x) \quad \forall\, x \in C(\tilde{G}) \tag{1.5}$$

für eine nicht-entartete Verteilungsfunktion $\tilde{G}$.

Wir zeigen zunächst: (α_n^*) und (β_n^*) sind beschränkt: Da $\tilde{G}$ eine nicht entartete Verteilungsfunktion ist, gibt es $x, y \in C(\tilde{G})$ mit $x \neq y$ und $0 < \tilde{G}(x), \tilde{G}(y) < 1$. (Zunächst ist nur die Existenz eines solchen x klar, aufgrund der rechtsseitigen Stetigkeit gibt es aber ein weiteres y mit derselben Eigenschaft.) Dann folgt aber die Beschränktheit von $(\alpha_n^* x + \beta_n^*)_n$ und $(\alpha_n^* y + \beta_n^*)$, da sonst mit (1.5)

$$\begin{aligned} F_n^*(\alpha_n^* x + \beta_n^*) &\underset{n\to\infty}{\longrightarrow} \tilde{G}(x) \in (0,1) \quad \text{und} \\ F_n^*(\alpha_n^* y + \beta_n^*) &\underset{n\to\infty}{\longrightarrow} \tilde{G}(y) \in (0,1) \end{aligned}$$

nicht gelten könnte. Dann ist aber auch die Folge

$$(\alpha_n^* x + \beta_n^* - (\alpha_n^* y + \beta_n^*)) = (\alpha_n^*(x - y))$$

beschränkt. Dann ist auch die Folge (α_n^*) beschränkt. Dann aber auch (β_n^*), da $(\alpha_n^* x + \beta_n^*)$ beschränkt ist.

Nach dem Satz von Bolzano-Weierstrass gibt es also Teilfolgen $(\alpha_{n_j})_j$ und $(\beta_{n_j})_j$ von $(\alpha_n)_n$ bzw. $(\beta_n)_n$ und Konstanten $\alpha \geq 0$ und $\beta \in \mathbb{R}$ mit

$$\alpha_{n_j}^* \underset{j\to\infty}{\longrightarrow} \alpha \quad \text{und} \quad \beta_{n_j}^* \underset{j\to\infty}{\longrightarrow} \beta.$$

Mit (1.5) folgt daher

$$F_{n_j}^*(\alpha_{n_j}^* x + \beta_{n_j}^*) \underset{j\to\infty}{\longrightarrow} \tilde{G}(x) \quad \forall\, x \in C(\tilde{G}).$$

Daher kann nicht $\alpha = 0$ gelten, denn sonst wäre $\tilde{G}$ entartet oder konstant. Somit gilt $\alpha > 0$.

Es bleibt zu zeigen: $\alpha_n^* \to \alpha$ und $\beta_n^* \to \beta$. Dazu seien $\alpha' > 0$ und $\beta' \in \mathbb{R}$ weitere Konstanten und $(\alpha_{n_k}^*)$ und $(\beta_{n_k}^*)$ weitere Teilfolgen von (α_n^*) und (β_n^*) mit

$$\alpha_{n_k}^* \underset{k\to\infty}{\longrightarrow} \alpha' \quad \text{und} \quad \beta_{n_k}^* \underset{k\to\infty}{\longrightarrow} \beta'.$$

Wir wollen zeigen: $\alpha = \alpha'$ und $\beta = \beta'$. Aus Lemma 1.15 folgt

$$G(\alpha x + \beta) = \tilde{G}(x) = G(\alpha' x + \beta') \quad \forall\, x \in C(\tilde{G}).$$

Nun gilt diese Gleichheit aber schon für alle $x \in \mathbb{R}$, denn da $C(\tilde{G})$ dicht in $\mathbb{R}$ liegt, existiert für jedes x aus dem Unstetigkeitsbereich von G eine Folge $(x_n)_n$ aus $C(\tilde{G})$ mit $x_n \downarrow x$. Mit der rechtsseitigen Stetigkeit von G und $\tilde{G}$ folgt die Behauptung.

Da G nicht entartet ist, existieren $x_1 < x_2$ mit

$$0 < y_1 := G(x_1) < y_2 := G(x_2) \le 1.$$

Es werde zunächst $y_2 < 1$ angenommen. Dann gilt für $i \in \{1,2\}$

$$\begin{aligned}
\tilde{G}^-(y_i) &= \inf\{\omega : \tilde{G}(\omega) \ge y_i\} = \inf\{\omega : G(\alpha\omega + \beta) \ge y_i\} \\
&= \inf\left\{\frac{1}{\alpha}(\omega - \beta) : G(\omega) \ge y_i\right\} \\
&= \frac{1}{\alpha}(G^-(y_i) - \beta)
\end{aligned}$$

und

$$\begin{aligned}
\tilde{G}^-(y_i) &= \inf\{w : \tilde{G}(w) \ge y_i\} = \inf\{w : G(\alpha'\omega + \beta') \ge y_i\} \\
&= \inf\left\{\frac{1}{\alpha'}(w - \beta') : G(w) \ge y_i\right\} \\
&= \frac{1}{\alpha'}(G^-(y_i) - \beta').
\end{aligned}$$

Also folgt

$$\begin{aligned}
\tilde{G}^-(y_1) &= \frac{1}{\alpha}(G^-(y_1) - \beta) = \frac{1}{\alpha'}(G^-(y_1) - \beta') \\
\tilde{G}^-(y_2) &= \frac{1}{\alpha}(G^-(y_2) - \beta) = \frac{1}{\alpha'}(G^-(y_2) - \beta').
\end{aligned} \tag{1.6}$$

Durch Subtraktion folgt

$$\frac{1}{\alpha}(G^-(y_1) - G^-(y_2)) = \frac{1}{\alpha'}(G^-(y_1) - G^-(y_2)).$$

Ist der Klammerausdruck verschieden von null, so folgt $\alpha = \alpha'$ und dann wegen (1.6) auch $\beta = \beta'$. Dies ist aber der Fall, da $G^-(y_1) < G^-(y_2)$ gilt. Denn zunächst folgt mithilfe von Lemma 1.12 (iv) aus $y_1 = G(x_1)$

$$G^-(y_1) = G^-(G(x_1)) \le x_1$$

und dann folgt mithilfe der Monotonie von G und der Definition von G^- aus $G(x_1) = y_1 < y_2$

$$G^-(y_2) > x_1 \geq G^-(y_1).$$

Somit ist in diesem Fall $\alpha' = \alpha$ und $\beta' = \beta$. Nun sei $y_2 = 1$. Wegen $G(x_2) = y_2$ ist

$$\{x : G(x) \geq 1\} \neq \emptyset.$$

Also kann die Pseudo-Inverse auf $(0, 1]$ definiert werden und somit gilt auch in diesem Fall analog $\alpha' = \alpha$ und $\beta' = \beta$. Damit galt aber schon

$$\alpha_n^* \to \alpha \quad \text{und} \quad \beta_n^* \to \beta.$$

Dies zeigt den Satz. $\square$

Unter Ausnutzung des Satzes von Khinchin lässt sich nun auch Satz 1.10 zeigen.

Beweis von Satz 1.10: "$\Rightarrow$": G sei max-stabil. Dann existieren nach Definition der max-Stabilität für $F_n := G^n$, $n \in \mathbb{N}$, Folgen reeller Zahlen $(a_n)_n$ und (b_n) mit $a_n > 0$ und

$$F_n(a_{nk}x + b_{nk}) = G^{nk}(a_{nk}x + b_{nk}))^{1/k} = (G(x))^{1/k}$$

für alle $k \in \mathbb{N}$, für alle $x \in \mathbb{R}$ und für alle $n \in \mathbb{N}$. Somit gilt insbesondere

$$F_n(a_{nk}x + b_{nk}) \underset{n\to\infty}{\longrightarrow} G^{1/k}(x)$$

für alle $x \in C(G)$, für alle $k \in \mathbb{N}$.

"$\Leftarrow$": Mit G ist auch $G^{1/k}$ für alle k eine nicht-entartete Verteilungsfunktion. Nach Voraussetzung gilt

$$F_n(a_{nk}x + b_{nk}) \underset{n\to\infty}{\longrightarrow} G^{1/k}(x)$$

für alle $k \in \mathbb{N}$, für $x \in C(G)$. Es sei nun $k \in \mathbb{N}$ und

$$\tilde{\alpha}_n := a_{nk},\ \tilde{\beta}_n := b_{nk},\ \tilde{G} = G^{1/k}.$$

Damit ergibt sich aus den Voraussetzungen

$$\begin{aligned} &F_n(\tilde{\alpha}_n x + \tilde{\beta}_n) \to \tilde{G}(x) \quad \forall\, x \in C(\tilde{G}) = C(G) \\ &F_n(a_n x + b_n) \to G(x) \quad \forall x \in C(G), \end{aligned}$$

wobei die zweite Zeile aus $k = 1$ folgt. Damit sind die Voraussetzungen des Satzes von Khinchin erfüllt. Daher folgt für geeignete Konstanten $\alpha_k > 0$ und $\beta_k \in \mathbb{R}$

$$\tilde{G}(x) = G^{1/k}(x) = G(\alpha_k x + \beta_k)$$

für alle $x \in \mathbb{R}$. D. h. für alle $x \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt

$$G(x) = G^n(\alpha_n x + \beta_n).$$

Also ist G max-stabil. $\square$

Wir kümmern uns nun darum, die nicht-entarteten Extremwertverteilungen genauer zu beschreiben. Der folgende Satz zeigt, dass wir mit der Gumbel-, der Fréchet- und der Weibull-Verteilung schon alle möglichen Typen von Extremwertverteilungen kennengelernt haben.

Satz 1.17 *(Fisher-Tippett-Theorem)*
Eine nicht-entartete Verteilung G ist genau dann eine Extremwertverteilung, wenn G zum Typ der Weibull-, Gumbel- oder Fréchet-Verteilung gehört.

Dieser Satz wird in zwei großen Schritten gezeigt. Zunächst leitet man her, dass die nicht-entarteten Verteilungen gerade die max-stabilen Verteilungen sind, dann beweist man, dass diese gerade mit den im Satz erwähnten Verteilungen übereinstimmen.

Wir beginnen mit dem ersten Schritt.

Satz 1.18 *Für eine nicht-entartete Verteilungsfunktion G sind die folgenden Aussagen äquivalent:*

(i) $\mathcal{D}(G) \neq \emptyset$.

(ii) G *ist max-stabil.*

Beweis: (ii) $\Rightarrow$ (i): ist offensichtlich, denn wir wissen definitionsgemäß, dass stets $G \in \mathcal{D}(G)$ gilt, wenn G eine max-stabile Verteilung ist.
(i) $\Rightarrow$ (ii): Es sei $F \in \mathcal{D}(G)$, also gelte

$$F^n(a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in C(G)$$

für geeignete Folgen (a_n), $a_n > 0$ für alle n und $(b_n), b \in \mathbb{R}$. Dann gilt aber schon

$$F^{nk}(a_{nk} x + b_{n_k}) \underset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in C(G) \quad \forall\, k \in \mathbb{N}$$

und somit

$$F^{n}(a_{nk} x + b_{nk}) \underset{n\to\infty}{\longrightarrow} G^{1/k}(x) \quad \forall x \in C(G) \quad \forall\, k \in \mathbb{N}.$$

Mit Satz 1.10 folgt, dass G max-stabil ist. $\square$

Dies ist der erste und einfachere Teil des Satzes. Alle uns bekannten Extremwertverteilungen sind also zwingend max-stabil. Wir zeigen nun, dass die bekannten Extremwertverteilungen auch die einzigen sind. Wir beginnen mit

Lemma 1.19 *Sei G eine max-stabile, nicht-entartete Verteilungsfunktion. Dann existieren messbare Funktionen*

$$\begin{aligned} & a : \mathbb{R}^+ \to \mathbb{R}^+ \\ & b : \mathbb{R}^+ \to \mathbb{R}^+ \\ \textit{mit} \quad & G^s\left(a(s)x + b(s)\right) = G(x) \quad \forall\, x \in \mathbb{R},\ \forall\, s > 0. \end{aligned}$$

Beweis: Sei G max-stabil. Bezeichne mit $\lceil\ \rceil$ die obere Gauß-Klammer. Dann gilt mit der Definition der Max-Stabilität

$$G^{\lceil ns \rceil}(a_{\lceil ns \rceil}x + b_{\lceil nx \rceil}) = G(x)$$

für alle $x \in \mathbb{R}$, für alle $s > 0$ und für alle $n \in \mathbb{N}$. Da

$$\frac{n}{\lceil ns \rceil} \underset{n\to\infty}{\longrightarrow} \frac{1}{s} \quad \forall\, s > 0$$

konvergiert, ergibt sich

$$G^n(a_{\lceil ns \rceil}x + b_{\lceil ns \rceil}) = (G^{\lceil ns \rceil}(a_{\lceil ns \rceil}x + b_{\lceil ns \rceil}))^{\frac{n}{\lceil ns \rceil}} = G^{\frac{n}{\lceil ns \rceil}}(x) \to G^{1/s}(x)$$

für alle $x \in \mathbb{R}$ und alle $s > 0$. Wählen wir insbesondere $s = 1$, so ergibt sich

$$G^n(a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in \mathbb{R}.$$

Wählen wir in den letzten beiden Konvergenzen

$$\alpha_n := a_n,\ \beta_n := b_n,\ \tilde{\alpha}_n := a_{\lceil ns \rceil} \quad \text{und} \quad \tilde{\beta}_n := b_{\lceil ns \rceil},$$

so sind die Voraussetzungen des Satzes von Khinchin erfüllt. Somit existieren für jedes $s > 0$ geeignete $a(s) > 0$ und $b(s) \in \mathbb{R}$ mit

$$\frac{a_{\lceil ns \rceil}}{a_n} \underset{n\to\infty}{\longrightarrow} a(s) \quad \text{und} \quad \frac{b_{\lceil ns \rceil} - b_n}{a_n} \underset{n\to\infty}{\longrightarrow} b(s).$$

Dies ergibt Funktionen

$$\begin{aligned} a : \mathbb{R}^+ &\to \mathbb{R}^+ \\ s &\mapsto \lim_{n\to\infty} \frac{a_{\lceil ns \rceil}}{a_n} \\ \text{und} \qquad b : \mathbb{R}^+ &\to \mathbb{R} \\ s &\mapsto \lim_{n\to\infty} \frac{b_{\lceil ns \rceil}}{a_n}. \end{aligned}$$

Diese sind als Grenzwerte messbarer Funktionen messbar und besitzen die Eigenschaft

$$G^{1/s}(x) = G(a(s)x + b(s)) \quad \forall\, x \in \mathbb{R}\ \forall\, s > 0.$$

$\square$

Nun wenden wir uns dem angekündigten letzten Schritt zu:

Satz 1.20 *Für eine nicht-entartete Verteilungsfunktion G gilt die folgende Äquivalenz:*

(i) G ist max-stabil.

(ii) G gehört zum Typ der Gumbel-, Fréchet- oder Weibullverteilung.

Beweis: Der Beweis ist aufwändig. Wir beginnen mit
(ii) $\Rightarrow$ (i) (dem einfachen Teil): Wir haben bereits nachgerechnet, dass die Gumbel-, Weibull- und Fréchet-Verteilungen max-stabil sind (siehe Beispiel 1.6 und 1.9).
(i) $\Rightarrow$ (ii): Die Idee in diesem Beweisschritt besteht darin, die Funktionalgleichung aus Lemma 1.19 zu verwenden und zu zeigen, dass deren einzige Lösung Verteilungsfunktionen der gewünschten Typen sind. Wir betrachten die Funktion

$$\begin{array}{rcl} \psi : T & \rightarrow & \mathbb{R} \\ x & \mapsto & \psi(x) := -\log(-\log G(x)), \end{array}$$

wobei

$$T = \{x \in \mathbb{R} : 0 < G(x) < 1\}$$

ist.

Behauptung: ψ besitzt eine Pseudo-Inverse ψ^- auf $\mathbb{R}$.

Beweis: Da G als Verteilungsfunktion monoton nicht-fallend ist und $\log(\cdot)$ strikt monoton steigt, ist ψ monoton nicht-fallend. Außerdem ist ψ rechtsseitig stetig. Somit bleibt noch

$$\inf_{x \in T} \psi(x) = -\infty \tag{1.7}$$

und

$$\sup_{x \in T} \psi(x) = +\infty \tag{1.8}$$

nachzuweisen. (1.7) ist gezeigt, wenn wir zeigen können, dass T kein kleinstes Element enthält, das unter G Masse hat, denn dann gilt

$$\begin{array}{ll} \inf_{x \in T} G(x) = 0 & \text{und somit} \\ \inf_{x \in T} \psi(x) = -\infty. & \end{array}$$

Angenommen, es gibt ein $x_L \in \mathbb{R}$ mit

$$G(x_L) > 0 \quad \text{und} \quad \lim_{x \uparrow x_L} G(x) = 0. \tag{1.9}$$

Aus der Max-Stabilität von G folgt

$$G^2(ax + b) = G(x) \quad \forall\, x \in \mathbb{R}$$

für geeignetes $a > 0$ und $b \in \mathbb{R}$. Gibt es nun ein x_L, das (1.9) erfüllt, so gilt:

(i) für $ax_L + b < x_L : 0 = G^2(ax_L + b) = G(x_L) > 0$,

(ii) für $ax_L + b > x_L$: Wegen $x_L > \frac{x_L - b}{a}$

$$0 < G^2(x_L) = G^2(a\frac{x_L - b}{a} + b) = G(\frac{x_L - b}{a}) = 0,$$

(iii) für $ax_L + b = x_L$:

$$G^2(x_L) = G^2(ax_L + b) = G(x_L) > 0.$$

In jedem dieser Fälle ergibt sich ein Widerspruch. Dieser ist in (i) und (ii) evident. In (iii) ergäbe sich sofort, dass $G(x_L) = 1$ ist und somit, dass G entartet ist. Also ist (1.9) zum Widerspruch geführt und es gilt (1.7).

Analog zeigt man, dass T kein größtes Element enthält, das unter G Masse trägt, denn dann gilt

$$\sup_{x \in T} G(x) = 1, \quad \text{also} \quad \sup_{x \in T} \psi(x) = +\infty.$$

Angenommen, es gäbe ein Element x_R mit

$$G(x_R) = 1 \quad \text{und} \quad \lim_{x \uparrow x_R} G(x) < 1.$$

Dann folgte wie oben

(i) für $ax_R + b < x_R : 1 > G^2(ax_R + b) = G(x_R) = 1$,

(ii) für $ax_R + b > x_R : 1 = G^2(x_R) = G^2(a\frac{x_R - b}{a} + b) = G(\frac{x_R - b}{a}) < 1$,

(iii) für $ax_R + b = x_R : \lim_{x \uparrow x_R} G^2(x) = \lim_{x \uparrow x_R} G^2(ax + b) = \lim_{x \uparrow x_R} G$.

Wie oben erhalten wir einen Widerspruch, womit auch (1.8) gezeigt ist. Daher ist gezeigt, dass ψ eine Pseudo-Inverse ψ^- auf $\mathbb{R}$ besitzt.

Da G eine nicht-entartete max-stabile Verteilung ist, können wir Lemma 1.19 anwenden. Dies liefert die Existenz messbarer Funktionen $a : \mathbb{R}^+ \to \mathbb{R}^+$ und $b : \mathbb{R}^+ \to \mathbb{R}$ mit

$$G^s(a(s)x + b(s)) = G(x) \quad \forall\, x \in \mathbb{R}, \quad \forall\, s > 0.$$

Daraus ergibt sich

$$\begin{aligned} \psi(x) &= -\log(-\log G(x)) = -\log(-\log(G^s(a(s)x + b(s)))) \\ &= -\log(-\log(G(a(s)x + b(s))) \cdot s) \\ &= \psi(a(s)x + b(s)) - \log s \end{aligned} \tag{1.10}$$

für alle $x \in \mathbb{R}$ und $s > 0$. Damit ergibt sich für ψ^-:

$$\begin{aligned} \psi^-(y) &= \inf\{x : \psi(x) \geq y\} \\ &= \inf\{x : \psi(a(s)x + b(s)) - \log s \geq y\} \\ &= \inf\left\{\frac{x - b(s)}{a(s)} : \psi(x) \geq y + \log(s)\right\} \\ &= \frac{\psi^-(y + \log(s)) - b(s)}{a(s)} \end{aligned} \tag{1.11}$$

für alle $y \in \mathbb{R}$ und für alle $s > 0$, wobei die zweite Gleichung aus (1.10) folgt. Setzen wir $y = 0$, so ergibt sich

$$\psi^-(0) = \frac{\psi^-(\log(s)) - b(s)}{a(s)} \quad \forall\, s > 0. \tag{1.12}$$

Wir subtrahieren (1.12) von (1.11) und erhalten:

$$\psi^-(y + \log(s)) - \psi^-(\log s) = a(s)(\psi^-(y) - \psi^-(0))$$

für alle $y \in \mathbb{R}$ und $s > 0$. Setzen wir

$$z := \log s, \quad \tilde{a}(z) := a(e^z), \quad g(y) := \psi^-(y) - \psi^-(0).$$

Damit erhalten wir

$$g(y + z) - g(z) = g(y)\tilde{a}(z) \quad \forall\, y, z \in \mathbb{R}. \tag{1.13}$$

Wir bestimmen nun die Lösungen von (1.13). Hierfür nehmen wir zunächst an, $\tilde{a}(z)$ wäre identisch gleich 1. Dann ist

$$g(y + z) = g(y) + g(z) \quad \forall\, y, z \in \mathbb{R} \tag{1.14}$$

zu lösen. Die Lösung dieser Funktionalgleichung ist eindeutig (bis auf Konstanten), d. h., setzen wir

$$c = g(1),$$

so ist die einzige Lösung dieser Funktionalgleichung die lineare Funktion

$$g(y) = cy.$$

Aus der Monotonie von ψ^- folgt nun

$$g(1) = \psi^-(1) - \psi^-(0) \geq 0.$$

Dann muss aber schon $g(1) > 0$ gelten, denn sonst wäre $g \equiv 0$ und somit ψ^- konstant und G entartet im Widerspruch zu den Voraussetzungen. Daher können wir $\alpha := \frac{1}{c}$ definieren und es folgt

$$g(y) = \frac{y}{\alpha} \quad \text{mit} \quad \alpha > 0.$$

Mit g ist dann aber ψ^- stetig. Setzen wir

$$\beta := -\alpha\psi^-(0),$$

so erhalten wir

$$\begin{aligned} x &= \psi^-(\psi(x)) = g(\psi(x)) + \psi^-(0) \\ &= \frac{\psi(x)}{\alpha} + \psi^-(0) \\ &= \frac{\psi(x)}{\alpha} - \frac{\beta}{\alpha} \quad \forall\, x \in T. \end{aligned}$$

Hierbei haben wir für die erste Gleichheit die Stetigkeit von ψ^- ausgenutzt, während die zweite Gleichheit aus der Definition von g folgt. Damit ergibt sich

$$\psi(x) = \alpha x + \beta \quad \forall\, x \in T,$$

also insbesondere

$$G(x) = e^{-e^{-\alpha x+\beta}} \quad \forall\, x \in \mathbb{R}.$$

Also ist G vom Typ der Gumbel-Verteilung.

Nun betrachten wir den Fall, dass $\tilde{a}(z) \not\equiv 1$ gilt. Es gibt also mindestens ein $z_0 \in \mathbb{R}$ mit $\tilde{a}(z_0) \neq 1$. Aus (1.13) folgt durch Vertauschen von y und z:

$$g(y+z) - g(y) = g(z)\tilde{a}(y) \quad \forall y, z \in \mathbb{R}. \tag{1.15}$$

Subtrahieren wir (1.15) von (1.13), so ergibt sich

$$-g(z) + g(y) = g(y)\tilde{a}(z) - g(z)\tilde{a}(y) \quad \forall\, y, z \in \mathbb{R},$$

also

$$g(z)(-\tilde{a}(y) + 1) = g(y)(-\tilde{a}(z) + 1) \quad \forall\, y, z \in \mathbb{R}.$$

Insbesondere ergibt sich für $z = z_0$ und $d := \frac{g(z_0)}{1-\tilde{a}(z_0)}$

$$g(y) = \frac{g(z_0)}{1 - \tilde{a}(z_0)}(1 - \tilde{a}(y)) = d \cdot (1 - \tilde{a}(y)) \quad \forall\, y \in \mathbb{R}. \tag{1.16}$$

Dabei ist $d \neq 0$, denn sonst wäre wieder $g \equiv 0$ und G entartet. Damit folgt aus (1.15)

$$d(1 - \tilde{a}(y+z)) - d(1 - \tilde{a}(z)) = d(1 - \tilde{a}(y))\tilde{a}(z) \quad \forall\, y, z \in \mathbb{R}.$$

Dividieren wir durch d, so erhalten wir

$$\tilde{a}(y+z) = \tilde{a}(y) \cdot \tilde{a}(z) \quad \forall\, y, z.$$

Diese Gleichung aber lässt sich auf bekannte Weise lösen: Da $\tilde{a}(y) = a(e^y) > 0$ ist, dürfen wir beide Seiten logarithmieren und erhalten

$$\log(\tilde{a}(y+z)) = \log(\tilde{a}(y)) + \log(\tilde{a}(z))$$

für alle $y, z \in \mathbb{R}$. Die einzige Lösung dieser Gleichung ist

$$\log(\tilde{a}(y)) = cy \quad \forall\, y \in \mathbb{R}$$

mit $c = \log \tilde{a}(1) = \log a(e)$. Also gilt

$$\tilde{a}(y) = e^{cy} \quad \forall\, y \in \mathbb{R}.$$

Nun ist aber $c \neq 0$, denn sonst wäre $\tilde{a} \equiv 1$ im Widerspruch zur Voraussetzung. Mit $\gamma = \frac{1}{c}$ erhalten wir also, dass die einzigen Lösungen von der Form

$$\tilde{a}(y) = e^{y/\gamma} \quad \forall\, y \in \mathbb{R}. \tag{1.17}$$

Dabei ist auch $\gamma \neq 0$. Setzen wir $\beta := \psi^-(0)$, so ergibt sich

$$\psi^-(y) = g(y) + \psi^-(0) = d(1 - \tilde{a}(y)) + \beta = d(1 - e^{y/\gamma}) + \beta \tag{1.18}$$

für alle $y \in \mathbb{R}$. Hierbei folgt die Gültigkeit der zweiten Gleichung aus (1.16) und die der dritten Gleichung aus (1.17). Wegen der Monotonie von ψ^- muss $\frac{d}{\gamma} < 0$ gelten, denn für $z > y$ ist wegen dieser Monotonie

$$\varphi^-(z) - \psi^-(y) \geq 0,$$

woraus wegen (1.18)

$$d(e^{y/\gamma} - e^{z/\gamma}) \geq 0$$

folgt. Für $d > 0$ muss also $\frac{y}{\gamma} \geq \frac{z}{\gamma}$, also $\gamma < 0$, gelten. Analog folgt aus $d < 0$, dass $\gamma > 0$ gilt. Nun gilt aber

$$\begin{aligned} x &= \psi^-(\psi(x)) = \beta + d(1 - e^{\psi(x)/\gamma}) \\ &= \beta + d\left(1 - \exp\left(\frac{-\log(-\log(G(x)))}{\gamma}\right)\right) \\ &= \beta + d(1 - (-\log G(x))^{-1/\gamma}) \quad \forall\, x \in T. \end{aligned}$$

Hierbei folgt die erste Identität aus (1.18), da dann ψ^- wieder stetig ist und die dritte Gleichheit ergibt sich direkt aus der Definition von ψ. Lösen wir obige Gleichung nach $G(x)$, so folgt

$$G(x) = \exp\left(-\left(1 - \frac{x-\beta}{d}\right)^{-\gamma}\right) \quad \forall\, x \in T. \tag{1.19}$$

Sei zunächst $\gamma > 0$. Mit

$$\hat{a} := -\frac{1}{d} > 0 \quad \text{und} \quad \hat{b} := 1 + \frac{\beta}{d}$$

ergibt sich

$$G(x) = \exp(-(\hat{a}x + \hat{b})^{-\gamma}) \quad \forall\, x \quad \text{mit} \quad 0 < G(x) < 1,$$

also für alle x mit $\hat{a}x + \hat{b} > 0$. Damit gilt

$$G(x) = \exp(-(\hat{a}x + \hat{b})^{-\gamma})\mathbb{1}_{(0,\infty)}(\hat{a}x + \hat{b}) = \Phi_\gamma(\hat{a}x + \hat{b}).$$

Also ist G vom Fréchet-Typ.

Nun gelte $\gamma < 0$ und es sei $\alpha := -\gamma > 0$. Dann folgt mit

$$\hat{a} := \frac{1}{d} \quad \text{und} \quad \hat{b} := -1 - \frac{\beta}{d}$$

aus (1.19)

$$G(x) = \exp(-(-(\hat{a}x + \hat{b}))^{\alpha}) \quad \forall\, x \;\text{ mit }\; 0 < G(x) < 1,$$

also alle x mit $\hat{a}x + b < 0$. Damit gilt

$$G(x) = \exp(-(-(\hat{a}x + \hat{b}))^{\alpha})\mathbb{1}_{(-\infty,0)}(\hat{a}x + \hat{b}) + \mathbb{1}_{[0,\infty)}(\hat{a}x + \hat{b}) = \psi_\alpha(\hat{a}x + b).$$

Also ist G vom Typ der Weibull-Verteilung mit Parameter $-\gamma$. Damit folgt (ii). $\square$

Damit sind wir auch in der Lage, das Fisher-Tippett-Theorem endgültig zu beweisen.

Beweis von Satz 1.17: G sei eine nicht-entartete Verteilungsfunktion. Dann ist der Anziehungsbereich von G nach Satz 1.18 genau dann nicht-leer, wenn G max-stabil ist. Das ist nach Satz 1.20 genau dann der Fall, wenn G vom Typ der Gumbel-, Fréchet- oder Weibull-Verteilung ist. □

Wir haben somit die Klasse der nicht-entarteten Extremwertverteilungen vollständig beschrieben. Insbesondere sehen wir, dass alle in Frage kommenden Verteilungen stetig sind. Durch den bekannten Rechentrick

$$\max = -\min -$$

lassen sich so auch die Verteilungen der Minima von iid Folgen beschreiben.

Es bleibt noch die Frage zu untersuchen, ob jede Verteilungsfunktion im Anziehungsbereich einer der drei Extremwertverteilungen liegt. Diese ist zu verneinen, wie das folgende Beispiel zeigt.

Beispiel 1.21 *Es sei F eine Verteilungsfunktion mit*

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

und es gebe ein $\rho < 1$ mit

$$\lim_{x \uparrow x_R} F(x) = \rho.$$

Dann liegt F nicht im Anziehungsbereich einer nicht-entarteten Extremwertverteilung, denn angenommen, G sei eine nicht-entartete Extremwertverteilung und $F \in \mathcal{D}(G)$. Dann gäbe es Folgen $(a_n)_n$ und $(b_n)_n$ mit $a_n > 0$ für alle $n \in \mathbb{N}$ und

$$F^n(a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x) \quad \forall\, x \in \mathbb{R},$$

denn nach dem Fisher-Tippett-Theorem gilt für jede nicht-entartete Extremwertverteilung G $C(G) = \mathbb{R}$. Da aus der rechtsseitigen Stetigkeit $F(x_R) = 1$ folgt, gälte für alle $n \in \mathbb{N}$

$$F^n(a_n x + b_n) \begin{cases} = 1 & \text{für } x \geq \frac{x_R - b_n}{a_n} \\ \leq \rho^n & \text{für } x < \frac{x_R - b_n}{a_n} \end{cases}.$$

Daraus würde aber wegen $F \in \mathcal{D}(G)$ für alle $x \in \mathbb{R}$

$$F^n(a_n x + b_n) \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{[\lim_{n\to\infty} \frac{x_R - b_n}{a_n}, \infty)}(x) = G(x)$$

folgen. Damit wäre G aber entartet in Widerspruch zur Voraussetzung.

Aus dem letzten Beispiel sehen wir, dass jedes F mit endlichem rechten Endpunkt und positiver Masse in diesem nicht im Anziehungsbereich einer nicht-entarteten Extremwertfunktion liegt. Wir haben also gesehen, dass es drei verschiedene Typen von Extremwertverteilungen gibt und dass es Verteilungen gibt, die nicht im Anziehungsbereich von einer von diesen liegen. Umgekehrt stellt sich auch die Frage, ob wir im Fall der Konvergenz auch beschreiben können, welchen Typ von Limesverteilung wir erwarten sollen, d. h. ob wir die Anziehungsbereiche der Extremwertverteilungen charakterisieren können. Wie schon eingangs erwähnt, konvergiert das nicht-normierte M_n gegen x_R, d. h. den rechten Randpunkt der Verteilung F der X_1, konvergiert. Es liegt daher nahe, dass die Frage, gegen welche Verteilung M_n konvergiert, vom Verhalten von F nahe x_R abhängt. Dies wird auch durch das letzte Beispiel gestützt. Es stellt sich dabei heraus, dass die wichtige Frage ist, "wie schnell" F in x_R hineingeht. Dies wird mit dem Begriff der Variation beschrieben. Genauer nennen wir eine Funktion regulär variierend (in $+\infty$), wenn sie sich asymptotisch wie eine Potenzfunktion verhält.

Definition 1.22 *Eine messbare Funktion $f : (0, \infty) \to (0, \infty)$ heißt* regulär variierend *mit Index $\alpha \in \mathbb{R}$, falls*

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = x^\alpha$$

für alle $x > 0$ gilt. Die Klasse der regulär variierenden Funktionen zum Index $\alpha \in \mathbb{R}$ bezeichnen wir in der Folge mit $\mathcal{R}_\alpha^\infty$.

Da für jede Verteilungsfunktion F gilt $F(x) \to 1$, wenn $x \to \infty$, ist diese immer in ∞ langsam variierend, d. h. regulär variierend mit Index 0. Allerdings ist dies nur eine sehr grobe Beschreibung des Verhaltens von F nahe x_R. Um dies genauer zu analysieren, definieren wir die Tailfunktion zu F und analysieren deren Verhalten.

Definition 1.23 *a) F sei eine Verteilungsfunktion. Die durch*

$$\bar{F}(x) := 1 - F(x) \quad \forall\, x \in \mathbb{R}$$

definierte Funktion heißt Tailfunktion *von F.*

b) $\bar{F}$ heißt regulär variierend mit Index α *in ∞, wenn*

$$\lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^\alpha \quad \forall\, x > 0$$

gilt.

c) Wir nennen

$$x_R := \sup\{x : F(x) < 1\}$$

rechten Endpunkt *der Verteilungsfunktion x.*

Mit diesen Definitionen lassen sich die Anziehungsbereiche der Extremwertverteilungen folgendermaßen charakterisieren:

Satz 1.24 *Es sei F eine Verteilungsfunktion und γ_n sei für $n \geq 2$ definiert durch*

$$\gamma_n := F^{-}\left(1 - \frac{1}{n}\right).$$

Dann gilt:

1. *F liegt genau dann im Anziehungsbereich der Fréchet-Verteilung Φ_α, wenn $\bar{F}$ in ∞ regulär variierend mit Index $-\alpha$ und $x_R = +\infty$ ist, d. h. wenn*

$$x_R = +\infty \quad \textit{und} \quad \lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha} \quad \forall\, x > 0 \tag{1.20}$$

gilt. In diesem Fall kann man die Folgen (A_n) und (B_n) für die schwache Konvergenz von $A_n(M_n - B_n)$ als $A_n = \gamma_n^{-1}$ und $B_n = 0$ wählen.

2. *F liegt genau dann im Anziehungsbereich der Weibull-Verteilung ψ_α, wenn die zu*

$$F^*(x) := F(x_R - \frac{1}{x}), \quad x > 0 \tag{1.21}$$

gehörende Tailfunktion $1 - F^*(x) = \bar{F}^*(x)$ in ∞ regulär variierend mit Index $-\alpha$ ist und $x_R < +\infty$ gilt, d. h. falls

$$x_R < +\infty \quad \textit{und} \quad \lim_{t\to\infty} \frac{\bar{F}^*(tx)}{\bar{F}^*(t)} = x^{-\alpha} \quad \forall x \in \mathbb{R}^+$$

gilt. Die Konstantenfolgen für die Konvergenz von $A_n(M_n - B_n)$ lassen sich in diesem Fall als

$$A_n := \frac{1}{x_R - \gamma_n} \quad \textit{und} \quad B_n := x_R$$

wählen.

3. *F liegt genau dann im Anziehungsbereich der Gumbel-Verteilung, wenn eine positive, messbare Funktion g existiert, so dass*

$$\lim_{t\uparrow x_R} \frac{\bar{F}(t + xg(t))}{\bar{F}(t))} = e^{-x} \quad \forall\, x \in \mathbb{R} \tag{1.22}$$

gilt. In diesem Fall können die Folgen (A_n) und (B_n) für die schwache Konvergenz von $A_n(M_n - B_n)$ als $A_n = \frac{1}{g(\gamma_n)}$ und $B_n = \gamma_n$ gewählt werden.

Es sei zunächst darauf hingewiesen, dass für hinreichend großes n stets $A_n > 0$ gilt. Dies ist im dritten Fall evident, da g positiv ist. Für die anderen Fälle betrachte $(\gamma_n)_{n\geq 2}$. Da

$$\begin{aligned} \gamma_n &:= \inf\left\{t : F(t) \geq 1 - \frac{1}{n}\right\} \quad \forall\, n \geq 2 \\ \text{und} \qquad x_R &:= \sup\{t : F(t) < 1\} \end{aligned}$$

gilt, folgt

$$\gamma_n \to x_R \quad \text{für} \quad n \to \infty.$$

Da im Fréchet-Fall $x_R = +\infty$ ist, folgt $\gamma_n > 0$ für hinreichend großes n und somit ist auch A_n positiv. Im Weibullfall gilt natürlich stets $\gamma_n \leq x_R$. Es muss noch gezeigt werden, dass nicht $\gamma_n = x_R$ gelten kann. Wäre aber $\gamma_n = x_R$ für ein n, so wäre auch $F(\gamma_n) = F(x_R)$ und daher besäße x_R unter F Wahrscheinlichkeitsmasse. Beispiel 1.21 zeigt aber, dass dann F nicht im Anziehungsbereich einer Extremwertverteilung liegen kann.

Der Beweis ist relativ lang. Bevor wir ihn beginnen, wollen wir ein Beispiel betrachten:

Beispiel 1.25 *1. Die $\mathcal{R}(0,1)$-Verteilung liegt im Anziehungsbereich der Weibullverteilung: Sei F die Verteilungsfunktion der $\mathcal{R}(0,1)$-Verteilung. Dann ist*

$$x_R = \sup_x\{x : F(x) < 1\} = \sup\{x : x < 1\} = 1.$$

Für die durch

$$F^*(x) := F\left(1 - \frac{1}{x}\right)$$

definierte Funktion F^* gilt:

$$\begin{aligned} \lim_{t\uparrow\infty} \frac{\bar{F}^*(tx)}{\bar{F}^*(t)} &= \lim_{t\uparrow\infty} \frac{1 - F(1-\frac{1}{tx})}{1-F(1-\frac{1}{t})} \\ &= \lim_{t\to\infty} \frac{\frac{1}{tx}}{\frac{1}{t}} = \frac{1}{x} \quad \forall\, x > 0. \end{aligned}$$

Also liegt F nach Satz 1.24 im Anziehungsbereich von ψ_1. Die Konstanten für die schwache Konvergenz von $A_n(M_n - B_n)$ können als

$$A_n = \frac{1}{x_R - \gamma_n} = \frac{1}{1 - F^-(1-\frac{1}{n})} = n \quad \text{und} \quad B_n = 1$$

gewählt werden.

2. *Die Pareto-Verteilung mit Verteilungsfunktion*

$$F(x) = \begin{cases} 0 & x \leq 1 \\ 1 - x^{-\alpha} & x \geq 1 \end{cases}$$

für festes $\alpha > 0$ liegt im Anziehungsbereich der Fréchet-Verteilung Φ_α: Offenbar ist

$$\frac{1}{x_R} = \sup\{x : F(x) < 1\} = +\infty.$$

Da mit $t \uparrow \infty$ auch $tx \uparrow \infty$ für alle $x > 0$ gilt, ergibt sich für alle $x > 0$

$$\lim_{t\uparrow\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = \lim_{t\uparrow\infty} \frac{1-F(tx)}{1-F(t)} = \lim_{t\uparrow\infty} \frac{(tx)^{-\alpha}}{t^{-\alpha}} = x^{-\alpha}.$$

Somit liegt F nach Satz 1.24 im Anziehungsbereich der Fréchet-Verteilung mit Index α, Φ_α. Für die Konstanten (A_n) und (B_n) kann man folgende Wahl treffen:

$$A_n = \frac{1}{F^-(1-\frac{1}{n})} = \left(\inf\left\{t : \left(1 - t^{-\alpha}\right) \mathbb{1}_{[1,\infty)}(t) \geq 1 - \frac{1}{n}\right\}\right)^{-1} = n^{-1/\alpha}$$

und $B_n = 0$.

3. *Die Exponentialverteilung mit Parameter $\lambda > 0$ liegt im Anziehungsbereich der Gumbelverteilung Λ: Sei also F die Verteilungsfunktion der Exponentialverteilung*

$$F(x) := 1 - e^{-\lambda x} \mathbb{1}_{[0,\infty)}(x).$$

Zunächst gilt offenbar

$$x_R = \sup\{x : F(x) < 1\} = +\infty.$$

Aus der Form der Verteilungsfunktion lässt sich schon ablesen, dass $\bar{F}$ schneller als jede Potenzfunktion gegen 0 konvergiert. Dies lässt schon vermuten, dass $F \in \mathcal{D}(\Lambda)$ gilt. Tatsächlich gilt für die konstante Funktion $g \equiv \lambda^{-1}$ (die natürlich messbar ist)

$$\lim_{t \uparrow x_R} \frac{\bar{F}(t + g(t)x)}{\bar{F}(t)} = \lim_{t \uparrow \infty} \frac{\bar{F}(t + \frac{x}{\lambda})}{\bar{F}(t)} = \lim_{t \to \infty} \frac{e^{-\lambda t - x}}{e^{-\lambda t}} = e^{-x}.$$

Somit folgt aus Satz 1.24, dass $F \in \mathcal{D}(\lambda)$ gilt und dass sich für die schwache Konvergenz der Folge $A_n(M_n - B_n)$ die Konstanten

$$A_n = \frac{1}{g(\gamma_n)} = \lambda \quad \textit{und} \quad B_n = \gamma_n = F^{-}\left(1 - \frac{1}{n}\right) = \frac{\log n}{n}$$

wählen lassen.

Den sehr aufwändigen Beweis von Satz 1.24 werden wir in mehreren Schritten führen. Wir beginnen mit häufiger verwendeten Lemmata:

Lemma 1.26 *Es sei F eine Verteilungsfunktion und es gelte (1.20), (1.21) oder (1.22). Ferner sei*

$$\gamma_n := F_n^{-}(1 - \frac{1}{n}) \quad \forall\, n \geq 2.$$

Dann folgt

$$\lim_{n \to \infty} n\bar{F}(\gamma_n) = 1.$$

Beweis: Zunächst folgern wir aus Lemma 1.12 (iv)

$$n\bar{F}(\gamma_n) = n\left(1 - F\left(F^{-}\left(1 - \frac{1}{n}\right)\right)\right) \leq n\left(1 - \left(1 - \frac{1}{n}\right)\right) = 1,$$

somit folgt auch

$$\limsup_{n \to \infty} n\bar{F}(\gamma_n) \leq 1.$$

Bleibt

$$\liminf_{n \to \infty} n\bar{F}(\gamma_n) \geq 1 \tag{1.23}$$

zu zeigen. Zunächst nehmen wir an, dass (1.20) gilt. Für $x \in (0,1)$ und $n \geq 2$ sei $\delta_n(x) := \gamma_n x < \gamma_n$. γ_n ist für genügend großes n positiv, da $\gamma_n \to x_R = +\infty$ gilt. Somit gilt für diese n (und alle x) $F(\delta_n(x)) < 1 - \frac{1}{n}$. Somit auch

$$n\bar{F}(\delta_n(x)) > n\left(1 - \left(1 - \frac{1}{n}\right)\right) = 1.$$

Damit ergibt sich

$$n\bar{F}(\gamma_n) > \frac{n\bar{F}(\gamma_n)}{n\bar{F}(\delta_n(x))} = \frac{\bar{F}(\gamma_n)}{\bar{F}(\gamma_n x)} \underset{n \to \infty}{\longrightarrow} x^{\alpha}$$

für alle $x \in (0,1)$. Dabei folgt die Konvergenz aus der Tatsache, dass $\bar{F}$ regulär variierend ist mit Index $-\alpha$ und $\gamma_n \to x_R = +\infty$ konvergiert. Lassen wir $x \uparrow 1$ gehen, so folgt (1.23).

Nun gelte (1.21). $(h_n)_n$ sei eine Folge positiver reeller Zahlen mit $x_R - \frac{1}{h_n} = \gamma_n$ für alle $n \geq 2$. Insbesondere folgt $h_n \to \infty$ für $n \to \infty$. Für $x \in (0,1)$ und $n \geq 2$ gilt

$$\delta_n(x) := x_R - \frac{1}{h_n x} < x_R - \frac{1}{h_n} = \gamma_n,$$

daher folgt wie oben

$$n\bar{F}(\delta_n(x)) > 1$$

für alle $x \in (0,1)$ und $n \geq 2$. Nutzt man zusätzlich (1.21), so folgt

$$n\bar{F}(\gamma_n) > \frac{n\bar{F}(\gamma_n)}{n\bar{F}(\delta_n(x))} = \frac{\bar{F}(x_R - \frac{1}{h_n})}{\bar{F}(x_R - \frac{1}{h_n x})} = \frac{F^*(h_n)}{F^*(h_n x)} \to x^\alpha \quad \forall\, x \in (0,1).$$

Mit $x \uparrow 1$ folgt wieder (1.23).

Schließlich gelte (1.22). Für $x < 0$ und $n \geq 2$ sei

$$\delta_n(x) := \gamma_n + x g(\gamma_n).$$

Dann gilt $\delta_n(x) < \gamma_n$ für alle $x < 0$ und $n \geq 2$. Daher folgt wie oben mit $n\bar{F}(\delta_n(x)) > 1$ $x_n \to x_R$ und somit

$$n\bar{F}(\gamma_n) > \frac{n\bar{F}(\gamma_n)}{n\bar{F}(\delta_n(x))} = \frac{\bar{F}(\gamma_n)}{\bar{F}(\gamma_n + x g(\gamma_n))} \underset{n\to\infty}{\longrightarrow} e^x \quad \forall\, x < 0.$$

Schicken wir x gegen 0, so folgt (1.23). □

Lemma 1.27 *Für $0 \leq \tau \leq +\infty$, eine Verteilungsfunktion F und eine Folge reeller Zahlen (u_n) sind äquivalent:*

(i) $n(1 - F(u_n)) \underset{n\to\infty}{\longrightarrow} \tau$.

(ii) $\mathbb{P}(M_n \leq u_n) = F^n(u_n) \underset{n\to\infty}{\longrightarrow} e^{-\tau}$.

Beweis: Sei zunächst $\tau < +\infty$. Es gelte 1.27 (i). Dann folgt

$$F^n(u_n) = \left[1 - \frac{1}{n}\left(n\left(1 - F(u_n)\right)\right)\right]^n \to e^{-\tau}.$$

Nun gelte 1.27 (ii). Als erstes zeigen wir, dass dann

$$1 - F(u_n) \underset{n\to\infty}{\longrightarrow} 0 \tag{1.24}$$

konvergiert, denn sonst gäbe es eine Teilfolge (u_{n_k}) von (u_n) und ein $\beta \in (0,1)$ mit $1 - F(u_{n_k}) \geq \beta$ für alle k. Dann gälte aber

$$F^{n_k}(u_{n_k}) \leq (1-\beta)^{n_k} \underset{k\to\infty}{\longrightarrow} 0,$$

was einen Widerspruch zu $\tau < +\infty$ bildet. Durch logarithmieren folgt aus 1.27 (ii)

$$n \log(1 - (1 - F(u_n)) = n \log F(u_n) \to -\tau,$$

was wegen $\tau < +\infty$ möglich ist. Entwickeln wir den Logarithmus um 1, so erhalten wir

$$\log(1-h) = -h + O(h^2), \quad h \downarrow 0$$

und daher mit (1.24)

$$n(1 - F(n_n)) \underset{n\to\infty}{\longrightarrow} \tau.$$

Nun sei $\tau = +\infty$.

(i) $\Rightarrow$ (ii) zeigen wir mit Widerspruch. Angenommen, es gäbe eine Verteilungsfunktion F und eine Folge reeller Zahlen $(u_n)_n$, für die 1.27 (i) aber nicht 1.27 (ii) gilt. Nun ist $(F_n(u_n))_n \subseteq [0,1]$ beschränkt und wenn 1.27(ii) nicht gilt, so besitzt $(F_n(u_n))$ eine konvergente Teilfolge, deren Limes wir $e^{-\mu}$ nennen wollen ($\mu \in [0,\infty)$). Mit dem oben Gezeigten folgt dann

$$\lim_{k\to\infty} n_k(1 - F(u_{n_k})) = \mu < +\infty$$

im Widerspruch zu 1.27(i).

Schließlich betrachten wir (ii) $\Rightarrow$ (i). Wieder nehmen wir an, dass eine Verteilungsfunktion F und eine Folge $(u_n)_n$ existieren, so dass

$$\lim_{n\to\infty} F^n(u_n) = 0 \quad \text{und} \quad \liminf_{n\to\infty} n(1 - F(u_n)) = \mu < +\infty$$

gilt. Dann gibt es eine Teilfolge $(u_{n_k})_k$ mit

$$\lim_{k\to\infty} n_k(1 - F(u_{n_k})) = \mu.$$

Da $\mu < +\infty$ ist, folgt (im Widerspruch zur Voraussetzung) aus dem oben Gezeigten

$$\lim_{h\to\infty} F^{n_k}(u_{n_k}) = e^{-\mu} > 0.$$

□

Wir widmen uns nun dem Anziehungsbereich der Fréchet-Verteilungen.

Satz 1.28 *Ist die Tailfunktion $\bar{F}$ einer Verteilungsfunktion F in ∞ regulär variierend mit Index $-\alpha$ und gilt $x_R = +\infty$, so gilt $F \in \mathcal{D}(\Phi_\alpha)$. Man kann $A_n = \frac{1}{\gamma_n}$ und $B_n = 0 \quad \forall\, n \geq 2$ wählen.*

Beweis: Sei $A_n = \frac{1}{\gamma_n}$ und $a_n = \frac{1}{A_n} = \gamma_n$. Wir zeigen, dass für alle x

$$n(1 - F(a_n x)) \underset{n\to\infty}{\longrightarrow} \begin{cases} x^{-\alpha} & x > 0 \\ \infty & x \leq 0 \end{cases}, \tag{1.25}$$

denn dann folgt mit Lemma 1.27

$$\lim_{n\to\infty} F^n(a_n x) = e^{-x^{-\infty}} \mathbb{1}_{(0,\infty)}(x) = \Phi_\alpha(x)$$

für alle x, also die Behauptung. Sei zunächst $x < 0$. Dann folgt aus $x_R = +\infty$, $a_n = \gamma_n \to \infty$ $(n \to \infty)$. Da $x < 0$ ist, ergibt sich $a_n x \to -\infty$, also $1 - F(a_n x) \to 1$, d. h. (1.25) für $x < 0$. Sei $x = 0$. Dann ist $(1 - F(a_n x))$ konstant und ungleich 0. Damit folgt wieder (1.25). Schließlich sei $x > 0$. Dann gilt

$$\lim_{n\to\infty} n(1 - F(a_n x)) = \lim_{n\to\infty} n(1 - F(a_n)) \frac{(1 - F(a_n x))}{1 - F(a_n)} = x^{-\alpha} \lim_{n\to\infty} n(1 - F(a_n)),$$

wobei wir im letzten Schritt wieder die reguläre Variation von $\bar{F}$ und $\gamma_n \to \infty$ ausgenutzt haben. Da mit Lemma 1.26

$$\lim_{n\to\infty} n(1 - F(a_n)) = 1$$

gilt, folgt auch für $x > 0$ (1.25). Also gilt die Behauptung. □

Nun wollen wir die Notwendigkeit (1.19) überprüfen.

Satz 1.29 *Gilt $F \in \mathcal{D}(\Phi_\alpha)$ für ein α, so ist die Tailfuntion $\bar{F}$ regulär variierend mit Index $-\alpha$ und es gilt $x_R = +\infty$.*

Wir schieben zwei Lemmata voran, die aber in dieser Vorlesung nicht bewiesen werden sollen. Ihre Beweise finden sich in dem Buch von L. de Haan "On regular variation and its applications".

Lemma 1.30 *Die Tailfunktion $\bar{F}$ einer Verteilungsfunktion F ist in ∞ regulär variierend, falls für jedes $\varepsilon > 0$ Folgen $(\lambda_n(\varepsilon))_{n\in\mathbb{N}}$ und $(a_n(\varepsilon))_{n\in\mathbb{N}}$ existieren, für die folgende Aussagen gelten:*

1. $\lim_{n\to\infty} \frac{\lambda_n(\varepsilon)}{\lambda_{n+1}(\varepsilon)} > 1 - \varepsilon$.
2. $\lim_{n\to\infty} a_n(\varepsilon) = +\infty$.
3. $\lim_{n\to\infty} \lambda_n(\varepsilon)\bar{F}(a_n(\varepsilon)x) =: \lambda(x)$ *existiert und ist für alle $x > 0$ positiv und endlich.*

Lemma 1.31 *(a_n), (b_n) seien Folgen reeller Zahlen mit*

- $a_n > 0 \quad \forall\, n$;

- $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \gamma > 1$;
- $\lim_{n\to\infty} \frac{b_{n+1}-b_n}{a_n} = 0$.

Dann gilt:

$$\lim_{n\to\infty} \frac{b_n}{a_n} = 0.$$

Nun sind alle Hilfsmittel bereit gestellt, um Satz 1.29 zu beweisen:

Beweis von Satz 1.29: Nach Voraussetzung gilt für geeignete $a_n > 0,\ b_n \in \mathbb{R}$

$$\lim_{n\to\infty} F^n(a_n x + b_n) = \Phi_\alpha(x) \quad \forall\, x \in \mathbb{R}. \tag{1.26}$$

Wir zeigen zunächst:

$$\frac{b_{\lceil ns\rceil} - b_n}{a_n} \underset{n\to\infty}{\longrightarrow} 0 \quad \forall\, s > 1 \quad \text{und} \quad \frac{a_{\lceil ns\rceil}}{a_n} \underset{n\to\infty}{\longrightarrow} s^{1/\alpha} \quad \forall\, s > 1. \tag{1.27}$$

Wegen (1.26) gilt für alle $s > 1$

$$\lim_{n\to\infty} F^{\lceil ns\rceil}(a_{\lceil ns\rceil} x + b_{\lceil ns\rceil}) = \Phi_\alpha(x) \quad \forall\, x \in \mathbb{R}.$$

Damit folgt wie im Beweis von Lemma 1.19

$$\lim_{n\to\infty} F^n(a_{\lceil ns\rceil} x + b_{\lceil ns\rceil}) = \Phi_\alpha^{1/s} \quad \forall\, x \in \mathbb{R}.$$

Der Satz von Khinchin liefert

$$\frac{a_{\lceil ns\rceil}}{a_n} \underset{n\to\infty}{\longrightarrow} A_s > 0 \quad \text{und} \quad \frac{b_{\lceil ns\rceil} - b_n}{a_n} \to B_s \in \mathbb{R}$$

und $\Phi_\alpha^{1/s}(x) = \Phi_\alpha(A_s x + B_s)$ für alle $x \in \mathbb{R}$. Es folgt $A_s = s^{1/\alpha}$ und $B_s = 0$, also (1.27). Durch

$$n(1) := \left\lceil \frac{s}{s-1} \right\rceil \quad \text{und} \quad n(k+1) := \lceil n(k) \cdot s\rceil,\ k \geq 1$$

ist eine Folge $(n(k))_k$ definiert. Für diese gilt

$$n(k) \underset{k\to\infty}{\longrightarrow} \infty \quad \text{und} \quad \frac{n(k+1)}{n(k)} \underset{k\to\infty}{\longrightarrow} s. \tag{1.28}$$

Mit dem ersten Grenzübergang und (1.27) folgt

$$\lim_{k\to\infty} \frac{a_{n(k+1)}}{a_{n(k)}} = \lim_{k\to\infty} \frac{a_{\lceil n(k)\cdot s\rceil}}{a_{n(k)}} = A_s = s^{1/\alpha} > 1$$

und

$$\lim_{k\to\infty} \frac{b_{n(k+1)} - b_{n(k)}}{a_{n(k)}} = \lim_{k\to\infty} \frac{b_{\lceil n(k)\cdot s\rceil} - b_{\lceil n(k)\rceil}}{a_{n(k)}} = B_s = 0.$$

Mit Lemma 1.31 folgt aus den letzten beiden Gleichungen

$$\frac{b_{n(k)}}{a_{n(k)}} \underset{k\to\infty}{\longrightarrow} 0. \tag{1.29}$$

Aus dem Satz von Khinchin folgt nun aber

$$\lim_{k\to\infty} F^{n(k)}(a_{n(k)}x) = \Phi_\alpha(x) \quad \forall\, x \in \mathbb{R}, \tag{1.30}$$

da für $k \to \infty$ trivialerweise $\frac{a_{n(k)}}{a_{n(k)}} = 1$ mit (1.29) auch

$$\frac{b_{n(k)} - 0}{a_{n(k)}} \to 0$$

und mit (1.26) wegen $n(k) \to \infty$ auch

$$\lim_{k\to\infty} F^{n(k)}(a_{n(k)}x + b_{n(k)}) = \Phi_\alpha(x) \tag{1.31}$$

für alle $x \in \mathbb{R}$ gilt. Nun soll gezeigt werden, dass $\bar{F}$ regulär variierend mit Index $-\alpha$ ist. Wegen (1.30) gilt mit Lemma 1.27

$$n(k)\bar{F}(a_{n(k)}x) \underset{k\to\infty}{\longrightarrow} x^{-\alpha} \quad \forall\, x > 0.$$

Bleibt noch zu zeigen, dass die Voraussetzungen von Lemma 1.30 erfüllt sind:

(i) Definitionsgemäß folgt wegen $\frac{a_{n(k)+1}}{a_{n(k)}} \to s^{1/\alpha}$, dass $a_{n(k)} \to \infty$ für $k \to \infty$ gilt.

(ii) Man erinnere sich, dass die Folge $(n(k))$ von s abhängt. Für $s \downarrow 1$ gilt $\frac{1}{s} \uparrow 1$, so dass mit dem zweiten Grenzübergang in (1.28) für alle $\varepsilon > 0$ ein $s > 1$ und somit eine Folge $(n(k))$ existiert mit

$$\lim_{k\to\infty} \frac{n(k)}{n(k+1)} = \frac{1}{s} > 1 - \varepsilon.$$

Also folgt mit (1.31) aus Lemma 1.30, dass $\bar{F}$ regulär variierend mit Index $-\alpha$ ist. Es bleibt zu zeigen, dass $x_R = +\infty$ gilt. Gälte nun $x_R < +\infty$, so gäbe es wegen $a_{n(k)} > 0$ für jedes k und $a_{n(k)} \to \infty$ für $k \to \infty$ ein $x > 0$, so dass $a_{n(k)}x > x_R$ gelten würden für alle k. Für dieses x wäre dann

$$1 = \lim_{k\to\infty} F^{n(k)}(a_{n(k)}x) = \Phi_\alpha(x),$$

was im Widerspruch zu $0 < \Phi_\alpha(x) < 1 \quad \forall\, x > 0$ stünde. Damit ist die Behauptung bewiesen. □

Wir wenden uns nun dem Anziehungsbereich der Weibull-Verteilung zu. Wieder zerfällt die Behauptung in zwei Sätze.

Satz 1.32 *Eine Verteilungsfunktion F liegt im Anziehungsbereich der Weibullverteilung ψ_α, falls $x_R < +\infty$ und die Tailfunktion $\bar{F}^*$ der durch $F^*(x) = F(x_R - \frac{1}{x})$ definierten Funktion F^* regulär variierend mit Index $-\alpha$ ist. Eine mögliche Wahl der Konstanten ist*

$$A_n^{-1} = a_n = x_R - \gamma_n \quad \textit{und} \quad B_n = b_n = x_R.$$

Beweis: Mit F ist auch F^* eine Verteilungsfunktion, x_R^* bezeichne ihren rechten Randpunkt. Mit Satz 1.28 wissen wir, dass F^* im Anziehungsbereich der Fréchet-Verteilung Φ_α liegt und dass für

$$a_n^* := \gamma_n^* := F^{*-}\left(1 - \frac{1}{n}\right) \quad \text{und} \quad b_n^* := 0$$

$$\lim_{n\to\infty} F^{*n}(a_n^* x + b_n^*) = \lim_{n\to\infty} F^{*n}(a_n^* x) = \lim_{n\to\infty} F^n\left(x_R - \frac{1}{a_n^* x}\right) = \Phi_\alpha(x) \quad \forall\, x > 0$$

gilt. Mit $a_n := \frac{1}{a_n^*}$ und $b_n := x_R$ gilt daher

$$\lim_{n\to\infty} F^n\left(b_n - \frac{a_n}{x}\right) = \Phi_\alpha(x) \quad \forall\, x > 0$$

und somit

$$\lim_{n\to\infty} F^n(a_n x + b_n) = \Phi_\alpha\left(-\frac{1}{x}\right) = \psi_\alpha(x) \quad \forall\, x < 0.$$

Da $b_n = x_R$ und $a_n > 0$ gilt, ist $F(a_n x + b_n) \equiv 1$ für $x \geq 0$. Deswegen gilt natürlich

$$\lim_{n\to\infty} F^n(a_n x + b_n) = 1 \quad \forall\, x \geq 0.$$

Somit ist gezeigt, dass $F \in \mathcal{D}(\psi_\alpha)$ gilt. Bleibt noch zu zeigen: $a_n = x_R - \gamma_n$.

$$\begin{aligned}
a_n^* = \gamma_n^* &= \inf\left\{x > 0 : F^*(x) \geq 1 - \frac{1}{n}\right\} \\
&= \inf\left\{x > 0 : F\left(x_R - \frac{1}{x}\right) \geq 1 - \frac{1}{n}\right\} \\
&= \inf\left\{x > 0 : u = x_R - \frac{1}{x} \wedge F(u) \geq 1 - \frac{1}{n}\right\} \\
&= \inf\left\{x > 0 : x = (x_R - u)^{-1} \wedge F(u) \geq 1 - \frac{1}{n}\right\} \\
&= \inf\left\{(x_R - u)^{-1} > 0 : F(u) \geq 1 - \frac{1}{n}\right\} \\
&= \left(\sup\left\{x_R - u > 0 : F(u) \geq 1 - \frac{1}{n}\right\}\right)^{-1} \\
&= \left(x_R - \inf\left\{u < x_R : F(u) \geq 1 - \frac{1}{n}\right\}\right)^{-1} \\
&= (x_R - \gamma_n)^{-1}.
\end{aligned}$$

Also gilt: $a_n = (a_n^*)^{-1} = x_R - \gamma_n$. □

Wieder bleibt noch zu zeigen, dass die Bedingungen auch notwendig sind, um im Anziehungsbereich der Weibull-Verteilung zu liegen.

Satz 1.33 *Liegt F in $\mathcal{D}(\psi_\alpha)$, so gilt $x_R < +\infty$ und die Teilfunktion $\bar{F}^*$ der durch $F^*(x) := F(x_R - \frac{1}{x})$ für alle $x > 0$ definierten Funktion F^* ist in ∞ regulär variierend mit Index $-\alpha$.*

Wir beginnen auch diesen Beweis mit einem Lemma, für dessen Beweis wir auf das Buch von de Haan verweisen.

Lemma 1.34 *(a_n) und (b_n) seien zwei Folgen in $\mathbb{R}$ mit*

(i) $a_n > 0$ *für alle n;*

(ii) $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \gamma$ *mit* $0 < \gamma < 1$ *und*

(iii) $\lim_{n\to\infty} \frac{b_{n+1}-b_n}{a_n} = 0$.

Dann existiert $b = \lim_{n\to\infty} b_n$ und ist endlich und es gilt

$$\lim_{n\to\infty} \frac{b - b_n}{a_n} = 0.$$

Beweis von Satz 1.33: Der Beweis ist dem Beweis des Fréchet-Falls sehr ähnlich. Es gebe also Folgen $(a_n)_n$ und $(b_n)_n$ mit $a_n > 0$ für alle $n \in \mathbb{N}$ und

$$\lim_{n\to\infty} F^n(a_n x + b_n) = \psi_\alpha(x) \quad \forall\, x \in \mathbb{R}.$$

Wie im Beweis von Satz 1.29 kann nun gezeigt werden:

$$\lim_{n\to\infty} \frac{a_{\lceil ns\rceil}}{a_n} = A_s = s^{-1/\alpha} \quad \text{und} \quad \lim_{n\to\infty} \frac{b_{\lceil ns\rceil} - b_n}{a_n} = B_s = 0 \tag{1.32}$$

für alle $s > 1$. Wie im Beweis von Satz 1.29 wird dieselbe Folge $(n(k))_k$ in Abhängigkeit von $s > 1$ gewählt. Mit dieser Folge gilt wegen (1.32)

$$\lim_{k\to\infty} \frac{a_{n(k+1)}}{a_{n(k)}} = \lim_{k\to\infty} \frac{a_{\lceil n(k)\cdot s\rceil}}{a_{n(k)}} = s^{-1/\alpha} \tag{1.33}$$

und

$$\lim_{k\to\infty} \frac{b_{n(k+1)} - b_{n(k)}}{a_{n(k)}} = \lim_{k\to\infty} \frac{b_{\lceil n(k)\cdot n\rceil} - b_{n(k)}}{a_{n(k)}} = 0. \tag{1.34}$$

Mit Lemma 1.34 folgt daraus

$$\lim_{k\to\infty} b_{n(k)} = b \quad \text{mit} \quad -\infty < b < \infty \quad \text{und} \quad \lim_{k\to\infty} \frac{b - b_{n(k)}}{a_{n(k)}} = 0.$$

Der Satz von Khinchin liefert wie im Beweis von Satz 1.29

$$\lim_{k\to\infty} F^{n(k)}(a_{n(k)}x + b) = \psi_\alpha(x) \quad \forall\, x \in \mathbb{R}. \tag{1.35}$$

Es soll nun gezeigt werden, dass $x_R < +\infty$ gilt. Aus $\psi_\alpha(0) = 1$ folgt direkt $F(b) = 1$, also $x_R \leq b < +\infty$. Nun zeigen wir noch, dass $\bar{F}^*$ in $+\infty$ regulär variierend mit Index $-\infty$ ist. Da $\psi_\alpha(-1) < 1$ gilt, ist $F(-a_{n(k)} + b) < 1$ für hinreichend große k. Da $(a_{n(k)})$ nach (1.33) eine Nullfolge ist, kann nicht $b > x_R$ gelten, woraus mit $x_R \leq b$ trivialerweise $x_R = b$ folgt. Damit folgt aus (1.35) mit $a^*_{n(k)} := \frac{1}{a_{n(k)}}$

$$\begin{aligned}
\lim_{k\to\infty} F^{*n(k)}(a^*_{n(k)}x) &= \lim_{k\to\infty} F^{n(k)}\left(b - \frac{1}{a^*_{n(k)}x}\right) \\
&= \lim_{k\to\infty} F^{n(k)}\left(a_{n(k)}(-\frac{1}{x}) + b\right) \\
&= \psi_\alpha\left(-\frac{1}{x}\right) \quad \text{für alle} \quad x > 0.
\end{aligned}$$

Da $\psi_\alpha(-\frac{1}{x}) = \Phi_\alpha(x)$ für alle $x > 0$ gilt, folgt aus der letzten Gleichung

$$\lim_{k\to\infty} F^{*n(k)}(a^*_{n(k)}x) = \Phi_\alpha(x) \quad \forall\, x > 0.$$

Im Beweis von Satz 1.29 wurde bereits gezeigt, dass daraus folgt, dass $\bar{F}^*$ regulär variierend mit Parameter $-\alpha$ ist. □

Wir werden nun den Gumbel-Fall diskutieren.

Satz 1.35 *Eine Verteilungsfunktion F liegt in $\mathcal{D}(\Lambda)$, falls eine positive, messbare Funktion g existiert, so dass*

$$\lim_{t\uparrow x_R} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x}$$

für alle $x \in \mathbb{R}$ gilt. Die Konstanten können dann als $a_n = g(\gamma_n)$ und $b_n = \gamma_n$ gewählt werden.

Beweis: Der Beweis beruht auf Lemma 1.27. Da $\gamma_n \uparrow x_R$ für $n \to \infty$ gilt, ergibt sich mit Lemma 1.26 und den Voraussetzungen des Satzes

$$\lim_{n\to\infty} n(1 - F(\gamma_n + xg(\gamma_n)) = \lim_{n\to\infty} \frac{n(1 - F(\gamma_n + xg(\gamma_n)))}{n(1 - F(\gamma_n))} = e^{-x} \quad \text{für alle} \quad x \in \mathbb{R}.$$

Lemma 1.27 gibt dann die Behauptung des Satzes. □

Dass die genannten Bedingungen auch notwendig sind, ist wesentlich aufwendiger zu zeigen und soll hier nur skizziert werden.

Satz 1.36 *Liegt die Verteilungsfunktion F in $\mathcal{D}(\Lambda)$, so existiert eine positive, messbare Funktion g mit*

$$\lim_{t\uparrow x_R} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x}$$

für alle $x \in \mathbb{R}$.

Als ersten Beweisschritt benötigt man eine allgemeinere Version des Satzes von Khinchin, die im wesentlichen genauso bewiesen wird wie der Satz von Khinchin selbst.

Lemma 1.37 *G sei eine nicht-entartete Verteilungsfunktion, $t_0 \in \mathbb{R} \cup \{\infty\}$ und $(F_t)_{t<t_0}$ eine Familie von monoton nicht-fallenden, rechtsseitig stetigen Funktionen mit*

$$\lim_{x\to\infty} F_t(x) = 1 \quad \textit{und} \quad \lim_{x\to-\infty} F_t(x) = c_t \geq 0$$

für alle $t < t_0$. Außerdem seien $a : \mathbb{R} \to \mathbb{R}^+$ und $b : \mathbb{R} \to \mathbb{R}$ Funktionen mit

$$F_t(a(t)x + b(t)) \to G(x) \quad (t \uparrow t_0)$$

für alle $x \in C(G)$. Dann gilt folgende Äquivalenz: Für Funktionen $\alpha : \mathbb{R} \to \mathbb{R}^+$ und $\beta : \mathbb{R} \to \mathbb{R}$ und eine geeignete nicht-entartete Verteilungsfunktion $\tilde{G}$ gilt genau dann

$$F_t(\alpha(t)x + \beta(t)) \to \tilde{G}(x) \quad (t \uparrow t_0) \quad \forall\, x \in C(G),$$

wenn es Konstanten $A > 0$ und B mit

$$\frac{\alpha(t)}{a(t)} \to A \quad \textit{und} \quad \frac{\beta(t) - b(t)}{a(t)} \to B$$

für $t \uparrow t_0$ gibt. Dann gilt

$$\tilde{G}(x) = G(Ax + B) \quad \forall\, x \in \mathbb{R}.$$

Außerdem benötigen wir die folgenden Lemmata, für deren Beweis für wieder auf das Buch von de Haan verweisen:

Lemma 1.38 *Für eine Verteilungsfunktion F, die im Anziehungsbereich der Gumbel-Verteilung liegt, gilt*

$$\lim_{t\downarrow 0} \frac{F^-(1-tx) - F^-(1-t)}{F^-(1-ty) - F^-(1-t)} = \frac{\log x}{\log y} \quad \forall\, x \in \mathbb{R}^+ \quad \forall\, y \in \mathbb{R}^+\backslash\{1\}.$$

Lemma 1.39 *Für Verteilungsfunktionen F, die im Anziehungsbereich der Gumbelverteilung liegen, gilt*

$$\lim_{s\to\infty} s(1 - F(a(s)x + b(s))) = e^{-x} \quad \forall\, x \in \mathbb{R}.$$

Hierbei sei $a : \mathbb{R}^+ \to \mathbb{R}^+$ durch

$$a(s) := F^-\left(1 - \frac{1}{se}\right) - F^-(1 - \frac{1}{s})$$

und $b : \mathbb{R}^+ \to \mathbb{R}$ durch $b(s) := F^-(1 - \frac{1}{s})$ definiert.

Nun sind wir in der Lage, Satz 1.36 zu beweisen.

Beweis von Satz 1.36: Mithilfe der vorhergehenden drei Lemmata folgern wir die Behauptung des Satzes. Wichtig ist vor allem die erweiterte Version des Satzes von Khinchin. Sei F eine Verteilungsfunktion mit $F \in \mathcal{D}(\Lambda)$. Mit Lemma 1.39 wissen wir, dass

$$\lim_{s\to\infty} s(1 - F(a(s)x + b(s))) = e^{-x}$$

mit den dort definierten Funktionen $a(\cdot)$ und $b(\cdot)$ gilt. Wir substituieren s durch die Funktion

$$\begin{aligned} s : (-\infty, x_R) &\rightarrow \mathbb{R}^+ \\ t &\mapsto s(t) := \frac{1}{1 - F(t)}. \end{aligned}$$

Da $s(t) \to \infty$ für $t \uparrow x_R$ gilt, ergibt sich

$$\lim_{t\uparrow x_R} s(t)(1 - F(a(s(t))x + b(s(t)))) = \lim_{t\uparrow x_R} \frac{1 - F(a(s(t))x + b(s(t)))}{1 - F(t)} = e^{-x}$$

für alle $x \in \mathbb{R}$. Daraus folgt direkt

$$\lim_{t\uparrow x_R} \exp\left(-\frac{1 - F(a(s(t))x + b(s(t)))}{1 - F(t)}\right) = e^{-e^{-x}} \tag{1.36}$$

für alle $x \in \mathbb{R}$. Wir definieren für alle $t < x_R$ die Funktion F_t vermöge

$$F_t(x) := \exp\left(-\frac{1 - F(x)}{1 - F(t)}\right), \quad x \in \mathbb{R}.$$

Für $t < x_R$ ist die Funktion rechtsseitig stetig, monoton nicht-fallend und es gilt

$$\lim_{x\to\infty} F_t(x) = 1 \quad \text{und} \quad \lim_{x\to-\infty} F_t(x) = \exp\left(-\frac{1}{1 - F(t)}\right) > 0.$$

(1.36) impliziert für diese Funktion

$$\lim_{t\uparrow x_R} F_t(a(s(t))x + b(s(t))) = e^{-e^{-x}} \quad \forall\, x \in \mathbb{R}.$$

Es bleibt also zu zeigen, dass

$$\frac{t - b(s(t))}{a(s(t))} \to 0 \tag{1.37}$$

gilt, denn dann folgt mit der Erweiterung des Satzes von Khinchin

$$\lim_{t\uparrow x_R} F_t(a(s(t))x + t) = e^{-e^{-x}} \quad \forall\, x \in \mathbb{R}.$$

Dies ist aber äquivalent zu

$$\lim_{t\uparrow x_R} \frac{1 - F(a(s(t))x + t)}{1 - F(t)} = e^{-x} \quad \forall\, x \in \mathbb{R},$$

also zur Behauptung mit einer Funktion g, die auf $(-\infty, x_R)$ durch $g(t) := a(s(t))$ definiert ist und auf ganz $\mathbb{R}$ messbar und positiv fortgesetzt werden kann. Also zeigen wir (1.37). Nach Definition von b und s gilt:

$$b(s(t)) = F^{-}(F(t)) \leq t < b(s(t)(1+\varepsilon)),$$

da

$$F(t) = 1 - (1 - F(t)) < 1 - \frac{1 - F(t)}{1+\varepsilon}$$

$\forall\, \varepsilon > 0$. Somit folgt

$$0 \leq \frac{t - b(s(t))}{a(s(x))} \leq \frac{b(s(t)(1+\varepsilon)) - b(s(t))}{a(s(t))}.$$

Nach Lemma 1.38 konvergiert die rechte Seite dieser Ungleichung für $t \uparrow x_R$ gegen $\log(1+\varepsilon)$, denn:

$$\begin{aligned}
\frac{b(s(t)(1+\varepsilon)) - b(s(t))}{a(s(t))} &= \frac{b((1-F(t))^{-1}(1+\varepsilon)) - b((1-F(t))^{-1})}{a((1-F(t))^{-1})} \\
&= \frac{F^{-}(1-(1-F(t))(1+\varepsilon)^{-1}) - F^{-}(1-(1-F(t)))}{F^{-}(1-(1-F(t))e^{-1}) - F^{-}(1-(1-F(t)))} \\
&\underset{\substack{t\uparrow x_R \\ (1-F(t)\downarrow 0)}}{\longrightarrow} \frac{\log((1+\varepsilon)^{-1})}{\log(e^{-1})} = \log(1+\varepsilon).
\end{aligned}$$

Also gilt (1.37) und die Behauptung des Satzes. □

2 Das Pickands-Balkema-de Haan-Theorem

In diesem Kapitel werden wir eine Charakterisierung der Verteilungen geben, die in irgendeinem der erwähnten Anziehungsbereiche liegen. Die zentrale Aussage ist das sogenannte Pickands-Balkema-de Haan (PBdH)-Theorem, das besagt, dass genau die Verteilungen im Anziehungsbereich einer nicht-entarteten Extremwertverteilung liegen, wenn die sogenannte Exzessfunktion durch eine Pareto-Verteilung angenähert werden kann. Dabei ist vor allem die Richtung interessant, bei der wir wissen, dass eine Verteilung im Anziehungsbereich einer nicht-entarteten Extremwertverteilung liegt und daraus auf die Exzess-Verteilung schließen.

Wir beginnen damit, die möglichen Extremwertverteilungen in einer Verteilungsklasse zusammenzufassen.

Definition 2.1 *Die für alle ξ und x mit $1 + \xi x > 0$ durch*

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

definierte Verteilungsfunktion H_ξ heißt verallgemeinerte Extremwertverteilung. *Die zugehörige skalierte Familie $H_{\xi;\mu;\psi}$ erhält man für $\mu \in \mathbb{R}$ und $\psi > 0$ aus der verallgemeinerten Extremwertverteilung durch Substitution von x durch $\frac{x-\mu}{\psi}$. Der Definitionsbereich besteht in diesem Fall aus allen x mit $1 + \xi\frac{x-\mu}{\psi} > 0$.*

Offensichtlich ist die verallgemeinerte Extremwertverteilung für
$\xi < 0$ vom Typ der Weibull-Verteilung $\psi_{-\xi^{-1}}$
$\xi = 0$ vom Typ der Gumbel-Verteilung Λ
$\xi > 0$ vom Typ der Fréchet-Verteilung $\hat{\Phi}_{\xi^{-1}}$.

Genauer gilt für

$$\begin{aligned} &\xi < 0 : H_\xi(x) = \psi_{-\xi^{-1}}(-(1+\xi x)) \quad \forall\, 1 + \xi x > 0 \\ &\xi = 0 : H_\xi(x) = \Lambda(x) \\ &\xi > 0 : H_\xi(x) = \Phi_{\xi^{-1}}(1+\xi x) \quad \forall\, 1 + \xi x > 0. \end{aligned}$$

Wir bereiten den angestrebten Satz durch ein erstes Resultat vor:

Satz 2.2 *Es seien $\xi \in \mathbb{R}$ und F eine Verteiungsfunktion. F liegt im Anziehungsbereich von H_ξ, wenn eine positive, messbare Funktion β existiert, so dass für alle $1 + \xi x > 0$*

$$\lim_{u \uparrow x_R} \frac{\bar{F}(u + x\beta(u))}{\bar{F}(u)} = \begin{cases} (1+\xi x)^{-1/\xi}, & \xi \neq 0 \\ e^{-x}, & \xi = 0 \end{cases} \tag{2.1}$$

gilt.

Beweis:

(i) $\xi = 0$: Das haben wir schon im Fisher-Tippett-Theorem gezeigt (genauer: Satz 1.24).

(ii) $\xi > 0$: Dann ist H_ξ vom Typ $\Phi_{\xi^{-1}}$. Nach dem Fisher-Tippett-Theorem (genauer: Satz 1.24) gilt:

$$F \in \mathcal{D}(H_\xi) \Leftrightarrow x_R = +\infty \quad \text{und} \quad \lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\xi^{-1}}$$

für alle $x > 0$. Sei zunächst $F \in \mathcal{D}(H_\xi)$, β sei eine positive, messbare Funktion mit

$$\lim_{x\to\infty} \frac{\beta(x)}{x} = \xi.$$

Aus

$$x_R = +\infty \quad \text{und} \quad \lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\xi^{-1}} \quad \forall\, x > 0$$

folgt mit der Definition von β für alle $1 + \beta x > 0$

$$\lim_{u\uparrow x_R} \frac{\bar{F}(u + x\beta(u))}{\bar{F}(u)} = \lim_{u\to\infty} \frac{\bar{F}(u(1 + x\frac{\beta(u)}{u}))}{\bar{F}(u)} = (1+\xi x)^{-\xi^{-1}}.$$

Nun gebe es ein positives, messbares β mit (2.1). Für die durch $\gamma_n := F^{-}(1 - \frac{1}{n}) = \inf\{x \in \mathbb{R} : F(x) \geq 1 - \frac{1}{n}\}$ definierte Folge $(\gamma_n)_n$ gilt $\gamma_n \uparrow x_R$ für $n \to \infty$ und nach Lemma 1.26 $\bar{F}(\gamma_n) \sim \frac{1}{n}$. Damit ergibt sich für alle $1 + \xi x > 0$

$$(1+\xi x)^{-\xi^{-1}} = \lim_{n\to\infty} \frac{\bar{F}(\gamma_n + x\beta(\gamma_n))}{\bar{F}(\gamma_n)} = \lim_{n\to\infty} n\bar{F}(\gamma_n + x\beta(\gamma_n)).$$

Mithilfe von Lemma 1.27 folgt daraus

$$\lim_{n\to\infty} F^n(\beta(\gamma_n)x + \gamma_n) = \Phi_{\xi^{-1}}(1+\xi x)$$

für alle $1 + \xi x > 0$ Da β eine positive Funktion ist, gilt also $F \in \mathcal{D}(\Phi_{\xi^{-1}}) = \mathcal{D}(H_\xi)$.

(iii) $\xi < 0$: H_ξ ist jetzt vom Typ $\psi_{-\xi^{-1}}$. Nach Satz 1.24 gilt

$$F \in \mathcal{D}(H_\xi) \Leftrightarrow x_R < \infty \quad \text{und} \quad \lim_{t\to\infty} \frac{\bar{F}^*(tx)}{\bar{F}^*(x)} = x^{\xi^{-1}} \tag{2.2}$$

für alle $x > 0$. Sei zunächst $F \in \mathcal{D}(H_\xi)$ und β eine positive messbare Funktion mit

$$t\beta^*(t) := t\beta(x_R - \frac{1}{t}) \to -\xi \quad \text{für} \quad t \to \infty.$$

Es gilt

$$\lim_{u\uparrow x_R} \bar{F}(u) = \lim_{t\to\infty} \bar{F}(x_R - \frac{1}{t}) = \lim_{t\to\infty} \bar{F}^*(t)$$

und

$$\begin{aligned}
\lim_{u \uparrow x_R} \bar{F}(u + x\beta(u)) &= \lim_{t\to\infty} \bar{F}(x_R - \frac{1}{t} + x\beta(x_R - \frac{1}{t})) \\
&= \lim_{t\to\infty} \bar{F}\left(x_R - \frac{1 - tx\beta(x_R - \frac{1}{t})}{t}\right) \\
&= \lim_{t\to\infty} \bar{F}^*\left(\frac{t}{1 - tx\beta(x_r - \frac{1}{t})}\right).
\end{aligned}$$

Mit (2.2) folgt daraus schließlich

$$\lim_{u\uparrow x_R} \frac{\bar{F}(u + x\beta(u))}{\bar{F}(u)} = \lim_{t\to\infty} \frac{\bar{F}^*(t(1 - tx\beta^*(t))^{-1})}{\bar{F}^*(t)} = (1 + \xi x)^{-\xi^{-1}}$$

für alle $1 + \xi x > 0$.

Nun gebe es eine Funktion β (positiv und messbar) mit (2.1). (γ_n) sei definiert wie unter (ii). Dann folgt analog zu dort

$$(-(-\xi x - 1))^{-\xi^{-1}} = (1 + \xi x)^{-\xi^{-1}} = \lim_{n\to\infty} \frac{\bar{F}(\gamma_n + x\beta(\gamma_n))}{\bar{F}(\gamma_n)} = \lim_{n\to\infty} n\bar{F}(\gamma_n + x\beta(\gamma_n))$$

für alle $1 + \xi x > 0$. Mit Lemma 1.27 folgt daraus

$$\lim_{n\to\infty} F^n(\beta(\gamma_n)x + \gamma_n) = e^{-(-(-\xi x-1))^{-\xi^{-1}}} = \psi_{-\xi^{-1}}(-\xi x - 1)$$

für alle $-\xi x - 1 < 0$. Da β positiv ist, gilt also $F \in \mathcal{D}(\psi_{-\xi^{-1}}) = \mathcal{D}(H_\xi)$.

□

Somit haben wir den Anziehungsbereich einer nicht-entarteten Verteilung charakterisiert. Die Grenzfunktion dieses Satzes stellen auf einem eingeschränkten Bereich Teilfunktionen von Verteilungsfunktionen dar. Die zugehörige Verteiung ist die Pareto-Verteilung.

Definition 2.3 *Es seien $\xi \in \mathbb{R}$ und $\beta > 0$. Die durch*

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

für

$$x \in \mathcal{D}(\xi, \beta) := \begin{cases} [0, \infty), & \xi \geq 0 \\ [0, -\frac{\beta}{\xi}], & \xi < 0 \end{cases}$$

definierte Verteilung heißt verallgemeinerte Pareto-Verteilung. *$G_{\xi,\beta}$ bezeichnet die zugehörigen Verteilungsfunktionen. Somit ist $G_{\xi,\beta}$ auch außerhalb von $\mathcal{D}(\xi, \beta)$ definiert und nimmt dort die Werte 0 bzw. 1 an.*

Als nächstes definieren wir die Exzessfunktion einer Verteilung.

Definition 2.4 *Sei X eine Zufallsvariable mit Verteilungsfunktion F. Für festes $u < x_R$ heißt*

$$F_u(x) := \mathbb{P}(X - u \leq x | X > u), \ x \geq 0$$

Exzess-Funktion *(genauer: Exzess-Verteilungsfunktion) von X bzw. F oberhalb des Schwellwertes u. Existiert $\mathbb{E}X$, so heißt die auf $(-\infty, x_R)$ durch*

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

definierte Funktion mittlere Exzess-Funktion *von X.*

Bemerkung: Nach Definition von $e(u)$ und mit partieller Integration folgt

$$e(u) = \frac{1}{\bar{F}(u)} \int_u^{x_R} (x-u) dF(x) = \frac{1}{\bar{F}(u)} \int_u^{x_R} \bar{F}(x).$$

Man berechnet auch

$$\bar{F}(x) = \frac{e(0)}{e(x)} e^{-\int_0^x \frac{1}{e(u)} du}$$

für alle Stetigkeitspunkte von F. Also ist ein stetiges F eindeutig durch seine Exzess-Funktion bestimmt.

Diese Verteilung ist vor allem in der Finanz- und Versicherungswirtschaft von Interesse. Anders als der dort verwendete Value at Risk nahelegt, interessiert einen dort nämlich nicht nur die Wahrscheinlichkeit eines Verlustes einer bestimmten Höhe, sondern auch die Verteilung, zumindest der Erwartungswert, falls diese Grenze überschritten wird.

Wir benötigen noch ein vorbereitendes Lemma.

Lemma 2.5 *Es sei F eine Verteilungsfunktion, (a_n) und (b_n) seien reelle Folgen, ebenso wie $(\tilde{a}_n)$ und $(\tilde{b}_n)$. Es gelte $a_n, \tilde{a}_n > 0$ für alle $n \in \mathbb{N}$. Weiter sei $J \subseteq \mathbb{R}$ ein Intervall und $S, \tilde{S}$ seien stetige, streng monoton-fallende Funktionen, deren Werte für Argumente aus J in $(0, \infty)$ liegen. Es gelte*

$$\begin{aligned} \lim_{n\to\infty} n\bar{F}(a_n x + b_n) &= S(x) \quad \forall\, x \in J \\ \text{und} \qquad \lim_{n\to\infty} n\bar{F}(\tilde{a}_n x + \tilde{b}_n) &= \tilde{S}(x) \quad \forall\, x \in J. \end{aligned}$$

Dann existieren Konstanten $A > 0$ und $B \in \mathbb{R}$ mit

$$a_n^* := \frac{\tilde{a}_n}{a_n} \underset{n\to\infty}{\longrightarrow} A, \quad b_n^* := \frac{\tilde{b}_n - b_n}{a_n} \underset{n\to\infty}{\longrightarrow} B$$

und

$$\tilde{S}(x) = S(Ax + B) \quad \forall x \in I.$$

Beweis: Der Beweis ähnelt dem des Satzes von Khinchin und soll hier nicht gegeben werden. □

Wir sind nun in der Lage, das zentrale Resultat dieses Kapitels zu formulieren:

Satz 2.6 *(Pickands-Balkema-de Haan)*

1. *X sei eine Zufallsvariable, die allgemein Pareto-verteilt mit Parametern $\xi \in \mathbb{R}$ und $\beta > 0$ ist. Dann ist $\mathbb{E}X$ genau dann endlich, wenn $\xi < 1$ gilt. In diesem Fall gilt*

   (i) $\mathbb{E}(1 + \frac{\xi}{\beta}X)^{-r} = \frac{1}{1+\xi r} \quad \forall\, \xi \neq 0,\ r > -\frac{1}{\xi}$*;*

   (ii) $\mathbb{E}(\log(1 + \frac{\xi}{\beta}X))^k = \xi^k k! \quad \forall\, k \in \mathbb{N}$*;*

   (iii) $\mathbb{E}(X(\bar{G}_{\xi,\beta}(X))^r) = \frac{\beta}{(r+1-\xi)(r+1)}$ *für alle* $r + 1 > |\xi|$*;*

   (iv) $\mathbb{E}X^r = \frac{\beta^r}{\xi^{r+1}} \frac{\Gamma(\xi^{-1}-r)}{\Gamma(\xi^{-1}+1)} r!$ *für alle* $r \in \mathbb{N},\ r < \frac{1}{\xi}$.

2. *Für alle $\xi \in \mathbb{R}$ liegt eine in x_R stetige Verteilungsfunktion F in $\mathcal{D}(H_\xi)$ genau dann, wenn*

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

   für eine geeignete positive, messbare Funktion β gilt.

3. *Für alle $x_1, x_2 \in \mathcal{D}(\xi, \beta)$ gilt*

   $$\frac{\bar{G}_{\xi,\beta}(x_1 + x_2)}{\bar{G}_{\xi,\beta}(x_1)} = \bar{G}_{\xi,\beta+\xi x_1}(x_2).$$

4. *Sei N eine $Poi(\lambda)$-verteilte Zufallsvariable, die stochastisch unabhängig von der iid-Folge $(X_n)_n$ sei, welche der verallgemeinerten Pareto-Verteilung mit Parametern ξ und $\beta > 0$ genügen möge. Es sei*

   $$M_N := \max_{1 \le i \le N} X_i.$$

   Dann gilt

   $$\mathbb{P}(M_N \le x) = H_{\xi,\mu,\psi}(x),$$

   wobei

   $$\mu = \beta\, \xi^{-1}(\lambda^\xi - 1) \quad \textit{und} \quad \psi = \beta\, \lambda^\xi$$

   gilt.

5. *X genüge der verallgemeinerten Pareto-Verteilung zu den Parametern $\xi < 1$ und β. Dann gilt für jedes $u < x_R$*

   $$e(u) = \mathbb{E}[X - u | X > u] = \frac{\beta + \xi_u}{1 - \xi}.$$

Bemerkung: Teil 2 des obigen Satzes ist das eigentliche Pickands-Bolkema-de Haan-Theorem. Es legt nahe, die Verallgemeinerte Pareto-Varteilung als Approximation der Exzessfunktion zu verwenden. Hierfür muss u groß genug sein. 5 gibt eine schöne graphische Methode, um die geeigneten u zu finden. Für eine iid Folge $X_1, \ldots, X_n$, die gemäß F verteilt ist, konstruiert man die empirische mittlere Exzessfunktion $e_u(u)$

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

5 sagt nun, dass $e(u)$ für $X \sim G_{\xi,\beta}$ linear ist in u. Außerdem schätzt $e_u(u)$ die mittlere Exzessfunktion konsistent und erwartungstreu. Also sucht man ein Gebiet, bei dem $e_u(u)$ annähernd linear ist.

Wir beginnen damit, den zweiten Teil der Aussage zu beweisen, da nur hier Lemma 2.5 benötigt wird.

Beweis von Satz 2.6.2: Sei X eine Zufallsvariable mit Verteiungsfunktion F. Zunächst gelte $F \in \mathcal{D}(H_\xi)$. Nach Satz 2.2 ist dies äquivalent zur Existenz einer positiven messbaren Funktion β mit

$$\lim_{u \uparrow x_R} \frac{\bar{F}(u + x\beta(u))}{\bar{F}(u)} = \begin{cases} (1+\xi x)^{-1/\xi}, & \xi \neq 0 \\ e^{-x}, & \xi = 0 \end{cases}$$

für alle $1 + \xi x > 0$. Daher gilt

$$F_u(\beta(u)x) = 1 - \mathbb{P}(X < u + x\beta(u) | X > u) = 1 - \frac{\bar{F}(u + x\beta(u))}{\bar{F}(u)} \underset{u \uparrow x_R}{\longrightarrow} G_\xi(x)$$

für alle $x \geq 0$ mit $1 + \xi x \geq 0$. Dann folgt insbesondere

$$\lim_{u \uparrow x_R} |F_u(x) - G_\xi(x(\beta(u))^{-1})| = \lim_{u \uparrow x_R} |F_u(x) - G_{\xi,\beta(u)}(x)| = 0 \tag{2.3}$$

für alle $x \geq 0$ mit $1 + \xi \frac{x}{\beta(u)} > 0$. Für $\xi < 0$ gilt (2.3) zunächst für alle $x \in [0, \frac{-\beta(u)}{\xi}]$. Da F_u monoton wachsend ist, $F_u(x) \leq 1$ für jedes x gilt, die verallgemeinerte Pareto-Verteilung stetig ist und bei Parametern $\beta(u) > 0$ und ξ für alle $x > \frac{-\beta(u)}{\xi}$ den Wert 1 annimt, gilt (2.3) auch für $\xi < 0$ für alle $x \geq 0$. Aufgrund der Stetigkeit der Pareto-Verteilung gilt sogar gleichmäßige Konvergenz, also

$$\lim_{u \uparrow x_R} \sup_{0 \leq x < \infty} |F_u(x) - G_{\xi,\beta(u)}(x)| = 0.$$

Dann gilt insbesondere

$$\lim_{u \uparrow x_R} \sup_{0 < x < x_R - u} |F_u(x) - G_{\xi,\beta(u)}(x)| = 0.$$

Der Beweis der Rückrichtung ist wesentlich aufwendiger. β sei also eine positive, messbare Funktion mit

$$\lim_{u \uparrow x_R} \sup_{0 < x < x_R - u} |F_u(x) - G_{\xi,\beta(u)}(x)| = 0.$$

Da $F_u(0) = G_{\xi,\beta(u)} = 0$ gilt, können wir das sup auch über $0 \leq x < x_R - u$ ausdehnen. Da $F_u(x) = 1$ für alle $x \geq x_R - u$ gilt, F in x_R stetig ist und die verallgemeinerte Pareto-Verteilung sowohl stetig als auch monoton wachsend ist, können wir das Supremum über alle Werte $x \geq 0$ ausdehnen. Also gilt auch

$$\lim_{u \uparrow x_R} \sup_{0 \leq x < \infty} |F_u(x) - G_{\xi,\beta(u)}(x)| = 0.$$

Damit folgt insbesondere auch

$$\lim_{u \uparrow x_R} \frac{\bar{F}(u + x\beta(u))}{\bar{F}(u)} = \begin{cases} (1+\xi x)^{-1/\xi}, & \xi \neq 0 \\ e^{-x}, & \xi = 0 \end{cases} \tag{2.4}$$

für alle $x \geq 0$ mit $1 + \xi x > 0$. Allerdings wollen wir die Aussage für alle x mit $1 + \xi x > 0$ zeigen. Wir zeigen die Aussage exemplarisch für $\xi < 0$. Dann gilt (2.4) zunächst für alle $x \in [0, -\frac{1}{\xi})$. Nun seien $(\gamma_n)_n$ und $(a_n)_n$ folgendermaßen definiert:

$$\begin{aligned} \gamma_n &:= F^{-}(1 - \frac{1}{n}) = \inf\{x : F(x) \geq 1 - \frac{1}{n}\} \\ a_n &:= \beta(\gamma_n) \quad \forall\, n \in \mathbb{N}. \end{aligned}$$

Damit gilt nun

$$\lim_{n\to\infty} \frac{\bar{F}(a_n x + \gamma_n)}{\bar{F}(\gamma_n)} = (1 + \xi x)^{-1/\xi}$$

für alle $x \in [0, -\frac{1}{\xi})$. Dann kann man wie im Beweis von Satz 2.2, dass

$$\lim_{n\to\infty} n\bar{F}(a_n x + \gamma_n) = (1 + \xi x)^{-1/\xi} \tag{2.5}$$

für alle $x \in [0, -\frac{1}{\xi})$ gilt. Sei nun c minimal mit der Eigenschaft, dass (2.5) für alle $x \in (c, -\frac{1}{\xi}) =: I_c$ gilt. Dann folgt

$$\lim_{n\to\infty} 2n\bar{F}(a_n x + \gamma_n) = 2(1 + \xi x)^{-1/\xi} =: S(x), \tag{2.6}$$

$$\lim_{n\to\infty} 2n\bar{F}(a_{2n} x + \gamma_{2n}) = (1 + \xi x)^{-1/\xi} := \bar{S}(x), \tag{2.7}$$

jeweils für alle $x \in I_c$. Nun sind die Voraussetzungen von Lemma 2.5 erfüllt. Dies ergibt:

$$\frac{a_{2n}}{a_n} \underset{n\to\infty}{\longrightarrow} A > 0, \qquad \frac{\gamma_{2n} - \gamma_n}{a_n} \to B$$

und

$$\bar{S}(x) = S(Ax + B) \quad \text{für alle} \quad x \in I_c.$$

Dann ist aber

$$\lim_{n\to\infty} 2n\bar{F}(a_{2n} x + \gamma_{2n}) = \lim_{n\to\infty} 2n\bar{F}(a_n(Ax + B) + \gamma_n) = 2(1 + \xi(Ax + B))^{-1/\xi} \tag{2.8}$$

für alle $Ax + B \in I_c$. Da $\bar{S}(x) = S(Ax + B)$ gilt, folgt $A = 2^\xi$ und $B = \frac{2^\xi - 1}{\xi}$. Mit

$$2(1 + \xi(Ax + B))^{-1/\xi} = 2(\xi 2^\xi x + 2^\xi)^{-1/\xi} = (1 + \xi x)^{-1/\xi}$$

ergibt sich daher

$$\lim_{n\to\infty} 2n\bar{F}(a_{2n} x + \gamma_{2n}) = (1 + \xi x)^{-1/\xi}$$

für alle $Ax + B \in I_c$, also für alle $x \in (2^{-\xi} c - \frac{1 - 2^{-\xi}}{\xi}, -\frac{1}{\xi}) =: \bar{I}_c$. Daher gilt (2.5) für die Teilfolgen der geraden Folgenglieder schon für alle $x \in \bar{I}_c$. Aus $\frac{2n+1}{2n} \to 1$ für $n \to \infty$ folgt

$$\lim_{n\to\infty} (2n + 1)\bar{F}(a_{2n} x + \gamma_{2n}) = (1 + \xi x)^{-1/\xi} \quad \forall\, x \in \bar{I}_c.$$

Für die durch $a^*_{2n+1} = a^*_{2n} = a_{2n}$ und $\gamma^*_{2n+1} = \gamma^*_{2n} = \gamma_{2n}$ definierten Folgen (a^*_n) und (γ^*_n) folgt daher

$$\lim_{n\to\infty} n\bar{F}(a^*_n x + \gamma^*_n) = (1 + \xi x)^{-1/\xi} \quad \forall\, x \in \bar{I}_c. \tag{2.9}$$

Nun kann Lemma 2.5 auf (2.9) angewandt werden. Wie oben kann man zeigen, dass (2.5) für alle $x \in \bar{I}_c$ gilt. Da c gerade als Minimum gewählt wurde, so dass (2.5) für alle $x \in I_c$ gilt, muss $I_c = \bar{I}_c$ sein und somit $c = -\infty$, da für $c \leq 0$ und $\xi < 0$ stets

$$2^{-\xi}c - \frac{1 - 2^{-\xi}}{\xi} < c$$

gilt. Also gilt (2.5) für alle $1 + \xi x > 0$ und aus Lemma 1.27 folgt dann

$$\lim_{n\to\infty} F^n(a_n x + \gamma_n) = \psi_{-\xi^-}(-(1 + \xi x))$$

für alle $1 + \xi x > 0$ und somit $F \in \mathcal{D}(H_\xi)$. Der Beweis für $\xi = 0$ und $\xi > 0$ geht analog. □

Beweis von Satz 2.6.1: X sei gemäß der verallgemeinerten Pareto-Verteilung mit Parametern ξ und β verteilt. Wir zeigen zunächst $\mathbb{E}X < +\infty \Leftrightarrow \xi < 1$. Für $\xi = 0$ ist die verallgemeinerte Pareto-Verteilung eine Exponentialverteilung, hat also einen endlichen Erwartungswert. Für $\xi < 0$ erhält man mittels partieller Integration

$$\begin{aligned}
\mathbb{E}X &= \int_0^{-\beta/\xi} \frac{x}{\beta}\left(1 + \frac{\xi}{x}x\right)^{-\frac{\xi+1}{\xi}} dx \\
&= \left[\left(1 + \frac{\xi}{\beta}x\right)^{-\frac{1}{\xi}}\left(\frac{\beta}{\xi - 1} + \frac{1}{\xi - 1}x\right)\right]^{-\beta/\xi} \\
&= \frac{\beta}{1 - \xi} < \infty.
\end{aligned}$$

Für $\xi > 0$ gilt analog

$$\begin{aligned}
\mathbb{E}X &= \int_0^{\infty} x\frac{1}{\beta}\left(1 + \frac{\xi}{\beta}x\right)^{-\frac{\xi+1}{\xi}} dx \\
&= \left[\left(1 + \frac{\xi}{\beta}X\right)^{-1/\xi}\left(\frac{\beta}{\xi - 1} + \frac{1}{\xi - 1}x\right)\right]_0^{\infty} \\
&= \left[\frac{\beta}{\xi - 1}\left(1 + \frac{\xi}{\beta}x\right)^{-1/\xi} + \frac{1}{\xi - 1}\left(x^{-\xi} + \frac{\xi}{\beta}x^{1-\xi}\right)^{-1/\xi}\right]_0^{\infty} \\
&= \lim_{x\to\infty}\left[\frac{\beta}{\xi - 1}\left(1 + \frac{\xi}{\beta}x\right)^{-1/\xi} + \frac{1}{\xi - 1}\left(x^{-\xi}\frac{\xi}{\beta}x^{-1-\xi}\right)\right] + \frac{\beta}{1 - \xi}.
\end{aligned}$$

Daher gilt $\mathbb{E}X < +\infty$ genau dann, wenn der Limes des Klammerausdrucks endlich ist. Da aber der erste Summand in der Klammer für $x \to \infty$ gegen0 konvergiert, ist die Summe genau dann endlich, wenn der Limes des zweiten Summanden endlich ist. Da $\xi > 0$ ist, ist dies der Fall, wenn die zu potenzierende Summe fr $x \to \infty$ nicht gegen 0 konvergiert. Da $x1{-}\xi$ für $x \to \infty$ gegen 0 geht, muss $x^{1-\xi}$ gegen einen Wert verschieden von 0 konvergieren. Das ist genau dann der Fall, wenn $1 - \xi > 0$, also $\xi < 0$, gilt. Wir zeigen nun

(i) $\mathbb{E}(1+\frac{\xi}{\beta}X)^{-r} = \frac{1}{1+\xi r}$ für $\xi < 1,\ \xi \neq 0,\ r > -\frac{1}{\xi}$.
Sei zunächst $\xi < 0$. dann gilt

$$\begin{aligned}\mathbb{E}(1+\frac{\xi}{\beta}X)^{-r} &= \int_{\infty}^{-\beta/\xi}(1+\frac{\xi}{\beta}x)^{-r}(1+\frac{\xi}{\beta}x)^{\frac{-\xi+1}{\xi}}\frac{\xi}{\beta}dx \\ &= \left[-\frac{1}{\beta}\frac{\xi}{1+r\xi}(1+\frac{\xi}{\beta}x)^{\frac{-1+r\xi}{\xi}}\frac{\beta}{\xi}\right]_0^{-\beta/\xi} \\ &= \frac{1}{1+r\xi}.\end{aligned}$$

Für $1 > \xi > 0$ gilt

$$\mathbb{E}(1+\frac{\xi}{\beta}X)^{-r} = \int_0^{\infty}\frac{1}{\beta}(1+\frac{\xi}{\beta}x)^{-\frac{-\xi+1+r\xi}{\xi}}dx = \left[-\frac{1}{1+r\xi}(1+\frac{\xi}{\beta}x)^{-\frac{1+r\xi}{\xi}}\right]_0^{\infty} = \frac{1}{1+r\xi},$$

da $-\frac{1+r\xi}{\xi} < 0$ gilt.

(ii) Da für $\xi < 0$ $\mathbb{P}$-f.s. $X < -\frac{\beta}{\xi}$ gilt, ist in diesem Fall $\log(1+\frac{\xi}{\beta}X)$ $\mathbb{P}$-f.s. definiert. Für $\xi < 1$ kann man

$$\mathbb{E}(\log(1+\frac{\xi}{\beta}X))^k = \xi^k k! \quad \forall\, k \in \mathbb{N}$$

per induktion nach k zeigen: Für $k = 1$ kann man im Falle $\xi < 1$ den Induktionsanfang per partieller Integration ableiten:

$$\begin{aligned}\mathbb{E}\log(1+\frac{\xi}{\beta}X) &= \frac{1}{\beta}\int_0^{-\beta/\xi}\log(1+\frac{\xi}{\beta}x)(1+\frac{\xi}{\beta}x)^{-\frac{\xi+1}{\xi}}dx \\ &= [-\log(1+\frac{\xi}{\beta}x)(1+\frac{\xi}{\beta}x)^{-1/\xi} - \xi(1+\frac{\xi}{\beta}x)^{-1/\xi}]_0^{-\beta/\xi} \\ &= \lim_{x\uparrow\frac{-\beta}{\xi}}[-\log(1+\frac{\xi}{\beta}x)(1+\frac{\xi}{\beta}x)^{-1/\xi} - \xi(1+\frac{\xi}{\beta}x)^{-1/\xi}] + \xi \\ &= \xi.\end{aligned}$$

Dabei folgt der letzte Schritt nach der l'Hospitalschen Regel. Für $1 > \xi > 0$ folgert man analog:

$$\mathbb{E}[\log(1+\frac{\xi}{\beta}X)] = \xi.$$

Den Induktionsschritt $k \mapsto k+1$ zeigen wir nur exemplarisch für $\xi < 0$. Per partieller Integration folgt

$$\begin{aligned}\mathbb{E}[\log(1+\frac{\xi}{\beta}X)]^{k+1} &= \frac{1}{\beta}\int_0^{-\beta/\xi}(\log(1+\frac{\xi}{\beta}))^{k+1}(1+\frac{\xi}{\beta}x)^{\frac{-\xi+1}{\xi}}dx \\ &\quad +(k+1)\xi\mathbb{E}(\log(1+\frac{\xi}{\beta}x))^k.\end{aligned}$$

Per Induktionsvoraussetzung erhält man:

$$\begin{aligned}\mathbb{E}[\log(1+\frac{\xi}{\beta}X)]^{k+1} &= -\lim_{x\uparrow\beta/\xi}[(\log(1+\frac{\xi}{\beta}))^{k+1}(1+\frac{\xi}{\beta}x)^{-1/\xi}] + (k+1)!\xi^{k+1} \\ &= (k+1)!\xi^{k+1},\end{aligned}$$

wobei wir für die letzte Gleichheit wieder die Regel von l'Hospital verwendet haben. Für $\xi = 0$ ist die Behauptung trivial.

(iii) Für $\xi < 1$ ist zu zeigen:

$$\mathbb{E}[X(\bar{G}_{\xi,\beta}(X))^r] = \frac{\beta}{(r+1-\xi)(r+1)}$$

für $r + 1 > |\xi|$.

Für $\xi < 0$ erhält man mit partieller Integration

$$\begin{aligned}
\mathbb{E}[X(\bar{G}_{\xi,\beta}(X))^r] &= \frac{1}{\beta}\int_0^{-\beta/\xi} x(1+\frac{\xi}{\beta}x)^{\frac{-r+\xi+1}{\xi}} dx \\
&= [-x(1+\frac{\xi}{\beta}x)^{-\frac{r+1}{\xi}}\frac{1}{r+1} - \frac{\beta}{(r+1)(r+1-\xi)}(1+\frac{\xi}{\beta})^{\frac{-r+1-\xi}{\xi}}]_0^{-\beta/\xi} \\
&= \frac{\beta}{(r+1)(r+1-\xi)}.
\end{aligned}$$

Für $\xi > 0$ ergibt sich analog

$$\begin{aligned}
&\mathbb{E}[X(\bar{G}_{\xi,\beta}(X))^r] \\
= &\lim_{x\to\infty}\left[\underbrace{-x(1+\frac{\xi}{\beta}x)^{\frac{-r+1}{\xi}}\frac{1}{r+1}}_{T_1} - \underbrace{\frac{\beta}{(r+1)(r+1-\xi)}(1+\frac{\xi}{\beta}x)^{\frac{-r+1-\xi}{\xi}}}_{T_2}\right] \\
&+\frac{\beta}{(r+1)(r+1-\xi)} \\
= &\frac{\beta}{(\gamma+1)(r+1-\xi)}.
\end{aligned}$$

Denn für $(r+1) > \xi$ konvergiert T_2 für $x \to \infty$ gegen 0 und es gilt

$$T_1 = -(x^{-\frac{\xi}{r+1}} + \frac{\xi}{\beta}x^{\frac{r+1-\xi}{r+1}})^{-\frac{r+1}{\xi}} \cdot \frac{1}{r+1}.$$

Für $r+1 > \xi$ konvergiert dann der zweite Summand gegen ∞, der erste gegen 0 und somit die potenzierte Summe gegen 0. Für $\xi = 0$ ergibt sich wieder mit partieller Integration

$$\begin{aligned}
\mathbb{E}[X(\bar{G}_{\xi,\beta}(X))^r] &= \frac{1}{\beta}\int_0^\infty xe^{-\frac{x(r+1)}{\beta}} dx \\
&= \left[-\frac{1}{\gamma+1}xe^{-\frac{x(r+1)}{\beta}} - \frac{\beta}{(r+1)^2}e^{\frac{-x(r+1)}{\beta}}\right]_0^\infty \\
&= \frac{\beta}{(r+1)^2}.
\end{aligned}$$

(iv) Für alle $\xi < 1$ ist zu zeigen:

$$\mathbb{E}X^\gamma = \frac{\beta^r}{\xi^{r+1}}\frac{\Gamma(\xi^{-1}-r)}{\Gamma(\xi^{-1}+1)}r!$$

für $r \in \mathbb{N}$ und $\frac{1}{\xi} > r$. Zunächst berechnet man

$$\begin{aligned}
\mathbb{E}X^r &= \frac{1}{\beta}\int_0^\infty x^r(1+\frac{\xi}{\beta}x)^{-\frac{\xi+1}{\xi}}dx \\
&= \left[-x^r(1+\frac{\xi}{\beta}x)^{-1/\xi}\right]_0^\infty + r\int_0^\infty x^{r-1}(1+\frac{\xi}{\beta}x)^{-1/\xi}dx \\
&= r\int_0^\infty x^{r-1}(1+\frac{\xi}{\beta}x)^{-1/\xi}dx,
\end{aligned}$$

da

$$[-x^r(1+\frac{\xi}{\beta}x)^{-1/\xi}]_0^\infty = [-(x^{-\xi r}+\frac{\xi}{\beta}x^{-\xi r+1})^{-1/\xi}]_0^\infty$$

gilt und $-\xi r + 1$ wegen $r < \frac{1}{\xi}$ positiv ist. Ferner gilt:

$$\int_0^\infty x^{r-j}(1+\frac{\xi}{\beta}x)^{-\frac{1}{\xi}-1+j}dx = \frac{\beta(r-j)}{1-\xi_j}\int_0^\infty x^{r-(j+1)}(1+\frac{\xi}{\beta}x)^{-\frac{1}{\xi}+j}dx$$

für alle $1 \le j \le r-1$. Rekursiv berechnen wir:

$$\begin{aligned}
\mathbb{E}x^r &= r\int_0^\infty x^{r-1}(1+\frac{\xi}{\beta}x)^{-1/\xi}dx \\
&= r\frac{\beta(r-1)}{1-\xi}\frac{\beta(r-2)}{1-2\xi}\cdots\frac{\beta}{1-(r-1)\xi}\int_0^\infty(1+\frac{\xi}{\beta}x)^{-1/\xi+r}dx \\
&= \frac{\beta^r r!}{\prod_{1\le j\le r}(1-j\xi)}.
\end{aligned}$$

Wir wollen nun die rechte Seite der nachzuweisenden Gleichung in derselben Weise umformen. Da $\frac{1}{\xi} > r$ vorausgesetzt ist, gilt insbesondere $\frac{1}{\xi} - j > 0$ mit $j \in \mathbb{N}$ und $j \le r$. Daher gilt

$$\Gamma\left(\frac{1}{\xi}+1\right) = \frac{1}{\xi}\Gamma\left(\frac{1}{\xi}\right) = \ldots = \prod_{j=0}^{r}\left(\frac{1}{\xi}-j\right)\Gamma\left(\frac{1}{\xi}-r\right).$$

Damit folgt

$$\begin{aligned}
\frac{\beta^r}{\xi^{r+1}}\frac{\Gamma(\frac{1}{\xi}-r)}{\Gamma(\frac{1}{\xi}+1)}r! &= \frac{\beta^r}{\xi^{r+1}}r!\ \frac{1}{\prod_{j=0}^r(\frac{1}{\xi}-j)} \\
&= \frac{\beta^r}{\xi^{r+1}}r!\ \frac{1}{\prod_{j=0}^r(\frac{1-\xi_j}{\xi})} = \frac{\beta^r r!}{\prod_{j=1}^r(1-\xi_j)}.
\end{aligned}$$

Das zeigt die Behauptung.

□

Beweis von Satz 2.6.3: Es seien $x_1, x_2 \in \mathcal{D}(\xi,\beta)$ mit $\bar{G}_{\xi,\beta}(x_1) \neq 0$. Dann gilt $x_1 + x_2 \in \mathcal{D}(\xi,\beta)$ genau dann, wenn auch $x_2 \in \mathcal{D}(\xi, \beta + x_1\xi)$ gilt. Für $x_1 + x_2 \notin \mathcal{D}(\xi,\beta)$ sind beide

Seiten der zu zeigenden Gleichung 0. Daher sei in der Folge $x_1 + x_2 \in \mathcal{D}(\xi, \beta)$. Zunächst sei $\xi \neq 0$. Dann gilt

$$\begin{aligned}
\bar{G}_{\xi,\beta}(x_1) \cdot \bar{G}_{\xi,\beta+\xi x_1}(x_2) &= (1 + \frac{\xi}{\beta^{x_1}})^{-1/\xi} \left(1 + \frac{\xi}{\beta + \xi x_1} x_2\right)^{-1/\xi} \\
&= (1 + \frac{\xi}{\beta} x_1 + \frac{\xi}{\beta} x_2)^{-1/\xi} \\
&= (1 + \frac{\xi}{\beta}(x_1 + x_2))^{-1/\xi} \\
&= \bar{G}_{\xi,\beta}(x_1 + x_2).
\end{aligned}$$

Für $\xi = 0$ und $x_1, x_2 \in \mathcal{D}(\xi, \beta)$ ergibt sich direkt

$$\bar{G}_{0,\beta}(x_1) \cdot \bar{G}_{0,\beta}(x_2) = e^{-\frac{x_1}{\beta}} e^{-\frac{x_2}{\beta}} = e^{-\frac{x_1+x_2}{\beta}} = \bar{G}_{0,\beta}(x_1 + x_2).$$

□

Beweis von Satz 2.6.4: N sei $\mathcal{P}(\lambda)$-verteilt und stochastisch unabhängig von der Folge $(X_n)_n$. Ferner sei

$$M_N := \max_{1 \leq i \leq N} X_i.$$

Dann gilt

$$\begin{aligned}
\mathbb{P}(M_N \leq x) &= \sum_{n=0}^{\infty} \mathbb{P}(\{M_N \leq x\} \cap \{N = n\}) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(M_N \leq n) \mathbb{P}(N = n) \\
&= e^{-\lambda} \sum_{n=0}^{\infty} \frac{G_{\xi,\beta}^n(x) \lambda^n}{n!} \\
&= e^{-\lambda} \exp(\lambda G_{\xi,\beta}(x)) \\
&= \exp(-\lambda \bar{G}_{\xi,\beta}(x)),
\end{aligned}$$

wobei wir die stochastische Unabhängigkeit von N und $(X_n)_n$ sowie der $(X_n)_n$ untereinander und deren identische Verteilung ausgenutzt haben. Für $\xi \neq 0$ ergibt sich daher

$$\begin{aligned}
\mathbb{P}(M_N \leq x) &= \exp(-\lambda(1 + \frac{\xi}{\beta} x)^{-1/\xi}) \\
&= \exp(-(1 + \xi \cdot \frac{x - \xi^{-1}\beta(\lambda^\xi - 1)}{\beta \lambda^\xi})^{-1/\xi}) \\
&= H_{\xi;\mu;\psi}(x)
\end{aligned}$$

mit $\mu = \frac{1}{\xi}\beta(\lambda^\xi - 1)$ und $\psi = \beta\lambda^\xi$.

Für $\xi = 0$ gilt

$$\begin{aligned}
\mathbb{P}(M_N \leq x) &= e^{-\lambda e^{-x/\beta}} \\
&= \exp(-e^{-\frac{x - \beta \log \lambda}{\beta}}) \\
&= H_{\xi;\mu;\psi}(x)
\end{aligned}$$

mit $\psi = \beta$ und $\mu = \lim_{\xi \to 0} \beta \left(\frac{\lambda^\xi - 1}{\xi} \right) = \beta \log \lambda$. $\square$

Beweis von Satz 2.4.5: X sei eine Zufallsvariable, die gemäß der verallgemeinerten Pareto-Verteilung mit Parameter $\xi < 1$ und β verteilt ist. Für $u < x_R$ ergibt sich mittels partieller Integration:

$$
\begin{aligned}
e(u) &= \mathbb{E}[X - u | X > u = \frac{1}{\bar{G}_{\xi,\beta}(u)} \int_{\{X>u\}} (X - u) d\mathbb{P} \\
&= \frac{1}{\bar{G}_{\xi,\beta}(u)} \int_u^{x_R} (x - u) \frac{dG_{\xi,\beta}(x)}{dx} dx \\
&= \frac{1}{\bar{G}_{\xi,\beta}(u)} \left[[G_{\xi,\beta}(x)(x - u)]_u^{x_R} \int_u^{x_R} G_{\xi,\beta}(x) dx \right] \\
&= \frac{1}{\bar{G}_{\xi,\beta}(u)} \left[\int_u^{x_R} 1 dx - \int_u^{x_R} G_{\xi,\beta}(x) dx \right] \\
&= \frac{1}{\bar{G}_{\xi,\beta}(u)} \int_u^{x_R} \bar{G}_{\xi,\beta}(x) dx,
\end{aligned}
$$

wobei die vorletzte Gleichheit wegen $G_{\xi,\beta}(x_R) = 1$ folgt. Also folgt für $\xi \neq 0$

$$
\begin{aligned}
e(u) &= (1 + \frac{\xi}{\beta} u)^{1/\xi} \left[\bar{G}_{\xi,\beta}(x)(1 + \frac{\xi}{\beta} x) \frac{\beta}{\xi - 1} \right]_u^{x_R} \\
&= (1 + \frac{\xi}{\beta} u) \frac{\beta}{1 - \xi} \\
&= \frac{\beta + \xi u}{1 - \xi}.
\end{aligned}
$$

Für $\xi = 0$ ergibt sich für alle $u < x_R$ und für alle $\beta + u\xi > 0$

$$
e(u) = e^{u/\beta} \int_u^\infty e^{-x/\beta} dx = \beta.
$$

$\square$

Damit haben wir das PBdH-Theorem bewiesen.

Bemerkung 2.7 *a) Das BPdH-Theorem besagt insbesondere, dass die verallgemeinerte Pareto-Verteilung für große n eine gute Approximation der Exzess-Funktion F_n darstellt, wenn F im Anziehungsbereich einer Extremwertverteilung liegt. Dies ist für viele Verteilungen bekannt.*

b) Im PBdH-Theorem stellt die Voraussetzung, dass F in x_R stetig sein muss, für die Hinrichung keine Einschränkung dar, denn eine Verteilung, die in x_R nicht stetig ist, besitzt dort positive Masse und solche Verteilungen liegen nie im Einzugsbereich einer nicht-entarteten Verteilung.

c) *Aussage 3 im PBdH-Theorem stellt eine Abschluss-Eigenschaft der verallgemeinerten Pareto-Verteilung dar. Sie besagt, dass die Wahrscheilichkeit, dass die zugrunde liegende Zufallsvariable eine Grenze $x_1 + x_2$ überschreitet, gegeben, sie überschreitet eine Grenze x_1, wieder gemäß einer verallgemeinerten Pareto-Verteilung verteilt ist.*

d) *Ist $F \in \mathcal{D}(H)$ für eine verallgemeinerte Extremwertverteilung H, so gilt gemäß Aussage 4: Die Verteilung des Maximums einer Poisson-verteilten Anzahl von iid Exzessen von oberhalb einer großen Grenze ist eine verallgemeinerte Extremwertverteilung.*

3 Konvergenzgeschwindigkeit normalisierter Extrema unabhängiger Zufallsvariablen

Das Fisher-Tippett-Theorem gibt die möglichen Extremwertverteilungen für richtig normierte Maxima von Folgen von iid Zufallsvariablen an. Um diesen Satz auch praktisch brauchbar zu machen, müssen wir wissen, wie schnell diese Konvergenz ist. d. h. wir müssen

$$|\mathbb{P}(M_n \leq \frac{x}{A_n} + B_n) - G(x)|$$

punktweise oder gleichmäßig in x abschätzen. Als Vorüberlegung verwenden wir

Lemma 3.1 *Für $0 \leq \tau < \min(2, \sqrt{2n-1}) = c_n$ gilt*

$$\frac{\tau^2 e^{-\tau}}{2n} \leq e^{-\tau} - \left(1 - \frac{\tau}{n}\right)^n. \tag{3.1}$$

Ferner gilt für alle $\tau \geq 0$

$$e^{-\tau} - \left(1 - \frac{\tau}{n}\right)^n \leq \frac{\tau^2 e^{-\tau}}{2n-1} \leq \frac{0.6}{2n-1}.$$

Beweis: Sei $h_n(\tau) = e^\tau(1 - \frac{\tau}{n})^n$ für $0 \leq \tau < c_n$. Die ersten beiden Ungleichungen in (3.1) sind dann äquivalent zu

$$1 - \frac{\tau^2}{2n} \geq h_n(\tau) \geq 1 - \frac{\tau^2}{2n-1} \tag{3.2}$$

bzw.

$$\log\left(1 - \frac{\tau^2}{2n}\right) \geq \tau + n\log\left(1 - \frac{\tau}{n}\right) \geq \log\left(1 - \frac{\tau^2}{2n-1}\right). \tag{3.3}$$

Sei

$$g_n(\tau) = \tau + n\log(1 - \frac{\tau}{n}) - \log(1 - \frac{\tau^2}{2n})$$

für $0 \leq \tau < c_n$. Es ist

$$g_n'(\tau) = 1 - \frac{1}{1 - \frac{\tau}{n}} + \frac{\tau}{n - \frac{\tau^2}{2}} = \frac{\tau - 2}{(n-\tau)(n - \frac{\tau^2}{2})} \leq 0, \tag{3.4}$$

also ist $g_n(\tau)$ schwach monoton fallend in τ mit

$$g_n(0) = 0, \quad \text{d. h. es gilt} \quad g_n(\tau) \leq 0 \tag{3.5}$$

für alle $0 \leq \tau < c_n$.

Setzt man analog

$$G_n(\tau) = \tau + n\log(1 - \frac{\tau}{n}) - \log(1 - \frac{\tau^2}{2n-1}) \tag{3.6}$$

für $0 \leq \tau \leq c_n$, so folgt

$$G_n'(\tau) = 1 - \frac{1}{1 - \frac{\tau}{n}} + \frac{\tau}{(n - \frac{1}{2}) - \frac{\tau^2}{2}} = \frac{\tau}{2}\frac{(\tau-1)^2}{(n-\tau)(n - \frac{1}{2} - \frac{\tau^2}{2})} \geq 0 \tag{3.7}$$

mit $G_n(0) = 0$, also

$$G_n(\tau) \geq 0 \quad \text{für alle} \quad 0 \leq \tau < c_n. \tag{3.8}$$

Damit ist (3.3), also auch die linke Seite von (3.1) bewiesen. Die Funktion

$$h(\tau) = \tau^2 e^{-\tau}, \quad \tau > 0$$

ist aber maximal für $\tau = 2$ mit $h(2) = 4e^{-2} \leq 0,6$, so dass hiermit auch die rechte Seite von (3.1) folgt. □

Satz 3.2 *Es sei wieder $\gamma_n = F^{-}(1 - \frac{1}{n})$, $n \geq 2$. Wir definieren die Funktion $r_n(x)$ folgendermaßen:*

- *Für $F \in \mathcal{D}(\Lambda)$*
$$r_n(x) = n(1 - F(\gamma_n + xg(\gamma_n))) - e^{-x},\ x \in \mathbb{R}$$
- *Für $F \in (\Phi_\alpha)$*
$$r_n(x) = n(1 - F(\gamma_n x)) - x^{-\alpha} \quad (x > 0).$$
- *Für $F \in (\psi_\alpha)$*
$$r_n(x) = n(1 - F(x_R + x(x_R - \gamma_n))) - (-x)^\alpha \quad \text{für} \quad x < 0.$$

Dann gilt für hinreichend großes n

- *Für $F \in \mathcal{D}(\lambda)$*
$$\begin{aligned} |\mathbb{P}(M_n \leq g(\gamma_n)x + \gamma_n) - \Lambda(x)| &\leq \frac{0.6}{2n-1} r e^{-e^{-x}} |e^{-r_n(x)} - 1| \\ &= \mathcal{O}(\max(\frac{1}{n}, |r_n(x)|)) \quad (x \in \mathbb{R}, n \to \infty). \end{aligned} \tag{3.9}$$
- *Für $F \in \mathcal{D}(\Phi_\alpha)$ gilt:*
$$\begin{aligned} |\mathbb{P}(M_n \leq x\gamma_n) - \Phi_\alpha(x)| &\leq \frac{0.6}{2n-1} + e^{-x^{-\alpha}} |e^{-r_n(x)} - 1 \\ &= \mathcal{O}(\max(\frac{1}{n}, |r_n(x)|)) \quad (x > 0, n \to \infty). \end{aligned} \tag{3.10}$$
- *Für $F \in \mathcal{D}(\psi_\alpha)$ gilt*
$$\begin{aligned} |\mathbb{P}(M_n \leq (x_R - \gamma_n)x + x_R) - \psi_\alpha(x)| &\leq \frac{0.6}{2n-1} + e^{-(-x)^\alpha} |e^{-r_n(x)} - 1| \\ &= \mathcal{O}(\max(\frac{1}{n}, |r_n(x)|) \quad x < 0, n \to \infty. \end{aligned} \tag{3.11}$$

Beweis: Sei

$$u_n(x) = \begin{cases} \gamma_n + xg(\gamma_n), & F \in \mathcal{D}(\Lambda) \\ \gamma_n x, & F \in \mathcal{D}(\Phi_\alpha) \\ (x_R - \gamma_n)x + x_R, & F \in \mathcal{D}(\psi_\alpha) \end{cases}, \tag{3.12}$$

d. h. $u_n(x) = \frac{x}{A_n} + B_n$ für die A_n und B_n aus Kapitel 1. Es ist dann (in den entsprechenden Bereichen für x) und mit G als Λ, Φ_α oder ψ_α

$$r_n(x) = n(1 - F(u_n(x))) + \log G(x). \tag{3.13}$$

Sei nun

$$\tau_n(x) = n(1 - F(u_n(x)) = r_n(x) - \log G(x). \tag{3.14}$$

Wegen $r_n(x) \to 0$ und $-\log G(x)) > 0$ ist für hinreichend großes n $\tau_n(x) \geq 0$ mit

$$\lim_{n\to\infty} \tau_n(x) = -\log G(x),$$

so dass für diese n mit Lemma 3.1 folgt

$$0 \leq e^{-\tau_n(x)} - F^n(u_n(x)) \leq \frac{0.6}{2n-1}, \tag{3.15}$$

also

$$\begin{aligned} |F^n(u_n(x)) - G(x)| &\leq |e^{-\tau_n(x)}G(x)| + |e^{-\tau n(x)} - F_n(u_n(x))| \\ &\leq \frac{0.6}{2n-1} + G(x)|e^{-r_n(x)} - 1|. \end{aligned} \tag{3.16}$$

□

Wegen $G(x) \leq 1$ kann übrigens auch global

$$|F^n(u_n(x)) - G(x)| \leq \frac{0.6}{2n-1} + 2|r_n(x)| \tag{3.17}$$

für genügend große n abgeschätzt werden (so dass $0 \leq r_n(x) - \log G(x) < \min(2, \sqrt{2n-1})$ und z. B. $|r_n(x)| \leq 1$ gilt).

Beispiel 3.3 *(Exponentialverteilung)*
Es sei

$$F(x) = 1 - e^{-\lambda x}, \quad \textit{für} \quad x > 0 \quad \textit{und} \quad \lambda > 0.$$

Bekanntlich gilt $F \in \mathcal{D}(\Lambda)$ mit

$$g(t) = \int_t^\infty e^{-\lambda(u-t)} du = \frac{1}{\lambda} \tag{3.18}$$

und

$$r_n(x) = n(1 - F(\gamma_n + \frac{x}{\lambda})) - e^{-x} = ne^{-(\lambda\gamma_n + x)} - e^{-x} = 0 \tag{3.19}$$

für $x > -\log n$. Also folgt

$$|\mathbb{P}(M_n \leq \frac{x}{\lambda} + \frac{\log n}{\lambda}) - e^{-e^{-x}}| \leq \frac{0.6}{2n-1} \tag{3.20}$$

für $x > -\log n$. Man beachte, dass für $x > 0$ mit (3.1) auch gilt

$$\frac{e^{-2x}e^{-e^{-x}}}{2n} \leq |\mathbb{P}(M_n \leq \frac{x}{\lambda} + \frac{\log n}{n}) - e^{-e^{-x}}|,$$

wir also die Konvergenzordnung mit $\mathcal{O}(\frac{1}{n})$ exakt festgestellt haben.

Beispiel 3.4 *(Normalverteilung)*
Es sei F die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung. Wir wissen schon, dass $F \in \mathcal{D}(\Lambda)$ und dass wir

$$\gamma_n^2 = 2\log n - \log\log n - \log 4\pi$$

wählen können. Weiter wählen wir $g(t) = \frac{1}{t}$. Damit ergibt sich:

$$c_n(x) = n\frac{f(\gamma_n + \frac{x}{\gamma_n})}{\gamma_n + \frac{x}{\gamma_n}} = \frac{\gamma_n^2}{\gamma_n^2 + x}\exp\left(\frac{-x^2}{2\gamma_n} + \mathcal{O}\left(\frac{\log\log n}{\log n}\right)\right)e^{-x}. \tag{3.21}$$

Nun gilt nach Abramowitz-Stegun (1984), Formel 26.2.12

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)f(x) \leq 1 - F(x) \leq \frac{1}{x}f(x)$$

für

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad x \in \mathbb{R}.$$

Daher folgt

$$\left(1 - \frac{1}{(\gamma_n + \frac{x}{\gamma_n})^2}\right)c_n(x) - e^{-x} \leq r_n(x) \leq c_n(x) - e^{-x} \tag{3.22}$$

und damit für hinreichend große n

$$\begin{aligned} |r_n(x)| &\leq e^{-x}\left(\frac{\gamma_n^2}{\gamma_n^2 + x}\left(1 + \mathcal{O}\left(\frac{\log\log n}{\log n}\right)\right) - 1\right) \\ &\leq e^{-x}\left(\frac{|x|}{\gamma_n^2 + x} + \mathcal{O}\left(\frac{\log\log n}{\log n}\right)\right) \\ &= \mathcal{O}\left(\frac{\log\log n}{\log n}\right) \quad n \to \infty. \end{aligned} \tag{3.23}$$

Damit erhalten wir

$$|\mathbb{P}(M_n \leq \frac{x}{\gamma_n} + \gamma_n) - \Lambda(x)| = \mathcal{O}\left(\frac{\log\log n}{\log n}\right) \tag{3.24}$$

für $n \to \infty$. Diese Konvergenzordnung lässt sich noch zu $0(\frac{1}{\gamma_n^2}) = 0(\frac{1}{\log n})$ verbessern, wenn man γ_n so wählt, dass

$$\frac{\gamma_n^2}{2} + \log\gamma_n = \log n - \frac{1}{2}\log 2\pi$$

wählt, wie man aus (3.23) schließen kann. Allerdings kann man auch zeigen, dass eine weitergehende Verbesserung der Konvergenzordnung nicht erreicht werden kann, unabhängig davon, welche normalisierenden Konstanten man wählt.

Wie man sieht, kann die Konvergenzordnung für normalisierte Extrema sehr schlecht werden (man vergleiche die Ordnung von $\frac{1}{\log n}$ einfach mit der Berey-Esseen-Schranke im CLT von $0(\frac{1}{\sqrt{n}})$). Entsprechendes gilt natürlich auch für richtig skalierte Minima.

Eine andere Variante von Satz 3.2 werden wir später noch kennenlernen.

Bislang haben wir die Differenz

$$|\mathbb{P}(M_n \leq \frac{x}{A_n} + B_n) - G(x)|$$

lediglich punktweise abgeschätzt. Tatsächlich liegt aber gleichmäßige Konvergenz gegen die nicht-entartete Extremwertverteilung G vor, wie der folgende Satz zeigt.

Satz 3.5 *Sei G eine beliebige stetige Verteilungsfunktion und $(F_n)_n$ eine Folge von Verteilungsfunktionen, mit*

$$\lim_{n\to\infty} F_n(x) = G(x) \quad \forall\, x \in \mathbb{R}.$$

Dann konvergiert (F_n) auch gleichmäßig gegen G, d. h. es gilt

$$\sup_{x\in\mathbb{R}} |F_n(x) - G(x)| \to 0 \quad \textit{für} \quad n \to \infty. \tag{3.25}$$

Beweis: Da G monoton ist und $\lim_{x\to-\infty} G(x) = 0$ und $\lim_{x\to\infty} G(x) = 1$ gilt, existiert für jedes $\varepsilon > 0$ ein $M_\varepsilon > 0$ mit

$$G(x) \leq \varepsilon \quad \text{und} \quad 1 - G(x) \leq \varepsilon \quad \forall\, |x| \geq M_\varepsilon. \tag{3.26}$$

Auf $[-M_\varepsilon, M_\varepsilon]$ ist G nun aber gleichmäßig stetig, d. h. es existiert δ_ε, so dass

$$|G(x) - G(y)| \leq \varepsilon \quad \text{für} \quad |x - y| \leq \delta_\varepsilon.$$

Wähle $N_\varepsilon \in \mathbb{N}$, so dass $N_\varepsilon \geq \frac{2M_\varepsilon}{\delta_\varepsilon}$ und partitioniere $[-M_\varepsilon, M_\varepsilon]$ in Intervale der Länge N_ε, d. h. setzt man

$$x_k = -M_\varepsilon + \frac{2M_\varepsilon}{N_\varepsilon} \quad \text{für} \quad 0 \leq k \leq N_2$$

sowie

$$A_k = \begin{cases} (-\infty, -M_\varepsilon] & k = 0 \\ [x_{k-1}, x_k] & 1 \leq k \leq N_\varepsilon \\ [M_\varepsilon, \infty) & k = N_\varepsilon + 1 \end{cases},$$

so gilt

$$|G(x) - G(y)| \leq \varepsilon \quad \forall\, x, y \in A_k \quad \forall\, k = 0, 1, \ldots, N_\varepsilon + 1.$$

Für $k = 1, 2, \ldots, N$ sei nun N_k^ε so gewählt, dass

$$|F_{n_k}(x) - G(x_k)| \leq \varepsilon \quad \text{und} \quad |F_n(x_{k-1}) - G(x_{k-1})| < \varepsilon \tag{3.27}$$

für $n \geq n_k^\varepsilon$, was möglich ist, da $F_n(x) \to G(x)$ für alle $x \in \mathbb{R}$. (3.27) gilt dann für

$$n \geq n_\varepsilon = \max\{n_k^\varepsilon,\ k \leq N_\varepsilon\}$$

für alle k. Dies impliziert dann

$$|F_n(x) - G(x)| \leq 2\varepsilon$$

für $x \in A_\varepsilon$ und $k = 0, 1, \ldots, N_\varepsilon + 1$. Dies ist aber die behauptete gleichmäßige Konvergenz. $\square$

Ein Beispiel für eine solche gleichmäßige Konvergenz liefert Beispiel 3.3 (die Exponentialverteilung). Etwa folgt aus (3.20), wobei nur noch das Konvergenzverhalten für $x \leq -\log n$ zu untersuchen ist. In diesem Fall gilt aber

$$|\mathbb{P}(M_n \leq \frac{x}{\lambda} + \frac{\log n}{\lambda}) - e^{-e^{-x}}| = e^{-e^{-x}} \leq e^{-e^{\log n}} = e^{-n}, \tag{3.28}$$

so dass insgesamt folgt

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(M_n \leq \frac{x}{\lambda} + \frac{\log n}{\lambda}) - e^{-e^{-x}}| \leq \frac{0.6}{2n-1} \tag{3.29}$$

für alle $n \in \mathbb{N}$.

Beispiel 3.6 *(Gleichverteilung)*
Es sei

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}.$$

Dies bedeutet insbesondere, dass $x_R = 1$ ist. Wir haben schon festgestellt dass $F \in \mathcal{D}(\psi_\alpha)$ mit $\alpha = 1$ und $\gamma_n = 1 - \frac{1}{n}$, so dass $A_n = n$ und $B_n = 1$ und

$$r_n(x) = n(1 - F(x_R + x(x_R - \gamma_n))) + x = 0 \tag{3.30}$$

für $-n \leq x < 0$ ist. Wegen

$$\mathbb{P}(M_n \leq \frac{x}{n} + 1) = 1 = \psi(x) \quad \text{für} \quad x \geq 0$$

bleibt nur noch der Bereich $x \leq -n$ zu untersuchen. Dort gilt

$$|\mathbb{P}(M_n \leq \frac{x}{n} + 1) - \psi_1(x)| = \psi_1(x) = e^x \leq e^{-n}, \tag{3.31}$$

so dass insgesamt analog zum Beispiel der Exponentialverteilung wieder folgt

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(M_n \leq \frac{x}{n} + 1) - \psi_1(x)| \leq \frac{0.6}{2n-1} \tag{3.32}$$

für alle $n \in \mathbb{N}$.

Stellt man an die Verteilungsfunktion genügende Glattheitsvoraussetzungen, so lassen sich die Abschätzungen verbessern. Darauf soll hier allerdings nicht weiter eingegangen werden.

4 Strukturelle Eigenschaften von Extrema unabhängiger Zufallsvariablen: Rekorde

Dieses Kapitel soll sich mit Struktureigenschaften von Maxima (und Minima) befassen. Darunter wollen wir die Fragestellung verstehen, wie die Teilfolgen der monoton wachsenden Maxima (bzw. monoton falenden Minima) aussieht, bzw. zu welchen Zeitpunkten sich die Werte der aktuellen Extrema ändern.

Definition 4.1 *Sei $(X_n)_n$ eine Folge von iid Zufallsvariablen mit gemeinsamer Verteilungsfunktion F. Ferner sei $x_R = +\infty$ oder x_R ein Stetigkeitspunkt von F. Die durch*

$$\begin{aligned} U_0 &= 1,\ U_{n+1} = \inf\{k : X_k > X_{U_n}\} \\ L_0 &= 1,\ L_{n+1} = \inf\{k : X_k < X_{L_n}\} \end{aligned} \tag{4.1}$$

definierten Zufallsvariablen heißen obere *bzw.* untere Rekordzeiten *für die Maxima (Minima) von (x_n). Die durch*

$$\Delta_0 = 1, \quad \Delta_n = U_n - U_{n-1} \quad (bzw.\ L_n - L_{n-1})) \tag{4.2}$$

definierten Zufallsvariablen heißen Zwischenrekordzeiten. *Die Zufallsvariablen (X_n) bzw. $(X_{L_n}$ heißen* obere *bzw.* untere Rekorde *von (X_n).*

Zunächst überlegen wir, dass unter unseren Annahmen $U_n = +\infty$ bzw. $\Delta_n = +\infty$ nur mit Wahrscheinlichkeit 0 eintritt.

Lemma 4.2 *Unter den Voraussetzungen von Definition 4.1 gilt*

$$\mathbb{P}(U_n = \infty) = \mathbb{P}(L_n = \infty) = \mathbb{P}(\Delta_n = \infty) = 0 \tag{4.3}$$

für alle $n \geq 0$.

Beweis: Wir führen den Beweis per vollständiger Induktion. Für $n = 0$ ist nichts zu zeigen. Es gelte also (4.3) für ein $n \geq 0$. Dann folgt

$$\begin{aligned} \mathbb{P}(U_{n+1} = +\infty) = \mathbb{P}(U_n = +\infty) \quad &+ \quad \sum_{k=1}^{\infty} \mathbb{P}(\{U_n = k\}) \cap \bigcap_{j \geq k+1} \{X_j \leq X_k\} \qquad (4.4) \\ &\leq \quad \sum_{k=1}^{\infty} \mathbb{P}(\bigcap_{j \geq k+1} \{X_j \leq X_k\}) = 0, \end{aligned}$$

da

$$\mathbb{P}(\bigcap_{j=2}^{\infty} \{X_j \leq X_1\}) = 0 \tag{4.5}$$

gilt. Letzteres sieht man so: Sei $\varepsilon > 0$, dann existiert ein $x_\varepsilon < x_R$, so dass

$$0 < \mathbb{P}(X_1 > x_\varepsilon) := \delta \leq \varepsilon. \tag{4.6}$$

Es folgt für jedes $m \geq 2$:

$$\begin{aligned}\mathbb{P}(\bigcap_{j=2}^{m}\{X_j \leq X_1\}) &\leq \mathbb{P}(X_1 > x_\varepsilon) + \mathbb{P}(\bigcap_{j=2}^{m}\{X_j \leq x_\varepsilon\}) \\ &\leq \varepsilon + F^{m-1}(x_\varepsilon) \\ &\leq \varepsilon + (1-\delta)^{n-1},\end{aligned} \tag{4.7}$$

da entweder $X_1 > x_\varepsilon$ ist oder alle $X_j \leq x_\varepsilon$ sein müssen. Somit folgt

$$\mathbb{P}(\bigcap_{j=2}^{\infty}\{X_j \leq X_1\}) = \lim_{n\to\infty} \mathbb{P}(\bigcap_{j=2}^{m}\{X_j \leq X_1\}) \leq \varepsilon. \tag{4.8}$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt

$$\mathbb{P}(U_{n+1} = +\infty) = 0.$$

Die Behauptung für $\{L_n\}$ und $\{\Delta_n\}$ zeigt man analog. $\square$

Die $\{U_n\}$ bzw. $\{L_n\}$ stellen also die Zeitpunkte dar, zu denen sich der Wert des aktuellen Maximums bzw. Minimums der Folge (X_k) ändert, die $\{X_{U_n}\}$ bzw. $\{X_{L_n}\}$ sind die zugehörigen Werte der Extrema. Offensichtlich besteht der folgende Zusammenhang:

$$M_k = X_{U_n} \quad \text{für} \quad U_n \leq k < U_n$$

bzw.

$$m_k = X_{L_n} \quad \text{für} \quad L_n \leq k < L_{n+1}.$$

Wir können also aus den Folgen (U_n, X_{U_n}) bzw. (L_n, X_{L_n}) die Folgen der (M_n) bzw. (m_n) eindeutig rekonstruieren (und umgekehrt). Wir kümmern uns zunächst um die Struktur der (U_n) und (L_n). Zunächst sind diese Folgen strukturell identisch. Darüber hinaus sind die Folgen strukturell unabhängig von F, solange F nur stetig ist. Ist nämlich F z. B. streng monoton zwischen x_L $(= \sup_x\{x : F(x) = 0\})$ und x_R, so bleiben die Folgen der (U_n) und (L_n) die gleichen, wenn wir statt $(X_n)_n$ die Folge der Extrema der $(F(X_n))_n$ betrachten, welche nach vorhergehenden Überlegungen $\mathcal{R}(0,1)$-verteilt ist.

Entsprechendes gilt (in Verteilung) auch für den allgemeinen stetigen Fall, da bei Stetigkeit von F sogenannte Bindungen, d. h. Paare $j \neq k$ mit $X_j = X_k$ nur mit Wahrscheinlichkeit 0 auftreten und damit die Werte von (X_n) und daher auch von $(F(X_n))$ fast sicher verschieden sind.

Die folgenden Resultate zeigen die Markovsche Struktur der Rekordzeiten:

Satz 4.3 *Sei (X_n) eine Folge von iid Zufallsvariablen mit stetiger Verteilungsfunktion F. Dann sind (U_n) und (L_n) homogene Markovketten mit*

$$\mathbb{P}(U_{n+1} = k | U_n = j) = \mathbb{P}(L_{n+1} = k | L_n = j) = \frac{j}{k(k-1)} \quad 1 \leq j < k \tag{4.9}$$

und

$$\mathbb{P}(U_1 = k) = \mathbb{P}(L_1 = k) = \frac{1}{k(k-1)} \tag{4.10}$$

für alle $k \geq 2$.

Beweis: Nach den Eingangsüberlegungen können wir (X_n) als iid $\mathcal{R}(0,1)$-verteilt annehmen. Wir zeigen zunächst (4.10). Für $k \geq 2$ gilt

$$\begin{aligned} \mathbb{P}(U_1 = k) &= \mathbb{P}(X_2, \dots, X_{k-1} \leq X_1 < X_k) \\ &= \mathbb{P}(\{X_2, \dots, X_{k-1}\} \in B_{k-1}(X_1) \cap \{X_1 < X_k\}). \end{aligned} \tag{4.11}$$

Hierbei ist

$$B_j(x_1) = \{(x_2, \dots, x_j) \in [0,1]^{j-1} : x_2, \dots, x_j \leq x_1\} = [0, x_1]^{j-1}.$$

Also

$$\begin{aligned} \mathbb{P}(U_1 = k) &= \int_0^1 \int_0^{x_k} \int_{B_{k-1}(x_1)} dx_2 \dots dx_k dx_1 dx_k \\ &= \int_0^1 \int_0^{x_k} x_1^{k-2} dx_1 dx_k \\ &= \int_0^1 \frac{x_k^{k-1}}{k-1} dx_k = \frac{1}{k(k-1)}. \end{aligned} \tag{4.12}$$

Sei nun $1 < k_1 < \dots < k_n$. Dann zeigt man per Induktion analog (mit $k_0 = 1$):

$$\begin{aligned} &= \mathbb{P}(\bigcap_{i=1}^{n} \{U_i = k_i\}) \\ &= \mathbb{P}(\{(X_2, \dots, X_{k_1-1}) \in B_{k_1-1}(X_1)\} \cap \bigcap_{i=1}^{n-1} \{(X_{k_{i+1}}, \dots, X_{k_{i+1}-1}) \\ &\quad \in B_{k_{i+1}-k_i-1}(X_{k_i}) \cap \{X_{k_i} < X_{k_{i+1}}\}), \\ &= \int_0^1 \int_0^{x_{k_n}} \int_{B_{k_n-k_{n-1}-1}(x_{k_n})} \dots \int_{B_{k_1-1}(x_1)} dx_2 \dots dx_{k_1-1} dx_1 \dots dx_{k_n-1} dx_{k_n} \\ &= \int_0^1 \int_0^{x_{k_n}} x_{k_{n-1}}^{k_n-k_{n-1}-1} \int_0^{x_{k_{n-1}}} x_{k_{n-2}}^{k_{n-1}-k_{n-2}} \dots \int_0^{x_{k_1}} x^{k_1-2} dx_1 dx_{k_1} \dots dx_{k_{n-1}} dx_k \\ &= \int_0^1 \int_0^{x_{k_n}} x_{k_{n-1}}^{k_n-k_{n-1}-1} \int_0^{x_{k_{n-1}}} \int_0^{x_{k_2}} \frac{x_{k_1}^{k_2-2}}{k_1-1} dx_{k-1} \dots dx_{k_n} \\ &= \dots = \int_0^1 \int_0^{x_{k_n}} \frac{x_{k_{n-1}}^{k_n-2}}{(k_1-1)(k_2-1)\dots(k_{n-1}-1)} dx_{k_{n-1}} dx_{k_n} \\ &= \prod_{i=1}^{n} \frac{1}{k_i-1} \int_0^1 x^{k_n-1} dx \\ &= \frac{1}{k_n} \prod_{i=1}^{n} \frac{1}{k_i-1} \\ &= \frac{k_{n-1}}{k_n(k_{n-1})} \frac{1}{k_n-1} \prod_{i=1}^{n-1} \frac{1}{k_i-1} \\ &= \frac{k_{n-1}}{k_n(k_n-1)} \mathbb{P}(\bigcap_{i=1}^{n-1} \{U_i = k_i\}). \end{aligned} \tag{4.13}$$

Damit ist

$$\mathbb{P}(U_n = k_n | \bigcap_{i=1}^{n-1} U_i = k_i) = \frac{k_{n-1}}{k_n(k_n - 1)}$$

nur von $U_{k_{n-1}}$ und U_{k_n} abhängig. Die (U_n) bilden daher eine (homogene) Markovkette. Die Übergangswahrscheinlichkeiten haben wir berechnet. □

Der folgende Satz liefert eine zweite anschauliche Vorstellung für die Folgen $\{U_n\}$ und $\{L_n\}$.

Satz 4.4 *Unter den Voraussetzungen von Satz 4.2 definieren wir die Folge (I_n) durch*

$$I_n = \begin{cases} 1, & \text{falls } X_n > M_{n-1} \\ 0, & \text{sonst} \end{cases} \tag{4.14}$$

(also $I_n = 1 \Leftrightarrow U_k = n$ für ein k) und $I_1 = 1$. Dann gilt: $(I_n)_n$ ist eine unabhängige Folge mit

$$\mathbb{P}(I_n = 1) = 1 - \mathbb{P}(I_n = 0) = \frac{1}{n} \tag{4.15}$$

für alle $n \in \mathbb{N}$.

Beweis: Für $1 < k_1 < \ldots < k_n$ sei

$$J_n := \{1, \ldots, k_n\} \backslash \{k_1, \ldots, k_n\}.$$

Dann ist

$$\bigcap_{i=1}^{n} \{U_i = k_i\} = \bigcap_{i=1}^{n} \{I_{k_i} = 1\} \cap \bigcap_{j \in J_n} \{I_j = 0\} \tag{4.16}$$

sowie

$$\bigcap_{i=1}^{n-1} \{U_i = k_i\} \cap \{U_n > k_n\} = \{I_{k_n} = 0\} \cap \bigcap_{i=1}^{n-1} \{I_{k_i} = 1\} \cap \bigcap_{j \in J_n} \{I_j = 0\} \tag{4.17}$$

und daher mit (4.13)

$$\begin{aligned} & \mathbb{P}(I_{k_n} = 1, I_{k_{n-1}} = \ldots = I_{k_1} = 1, I_j = 0 \; \forall j \in J_n\} && (4.18) \\ = \; & \mathbb{P}(\bigcap_{i=1}^{n} \{U_i = k_i\}) \\ = \; & \frac{1}{k_n} \prod_{i=1}^{n} \frac{1}{k_i - 1} \end{aligned}$$

sowie

$$
\begin{aligned}
& \mathbb{P}(I_{k_n}=0, I_{k_{n-1}}=\ldots=I_{k_1}=1, I_j=0\ \forall\, j\in J_n\} && (4.19)\\
= \ & \mathbb{P}(\bigcap_{i=1}^{n-1}\{U_i=k_i\}\cap\{U_n>k_n\})\\
= \ & \prod_{i=1}^{n-1}\frac{1}{k_i-1}\sum_{m>k_n}\frac{1}{m(m-1)}\\
= \ & \prod_{i=1}^{n-1}\frac{1}{k_i-1}\sum_{m>k_n}\left(\frac{1}{(m-1)}-\frac{1}{m}\right)\\
= \ & \frac{1}{k_n}\prod_{i=1}^{n-1}\frac{1}{k_i-1}.
\end{aligned}
$$

Also erhalten wir

$$
\begin{aligned}
& \mathbb{P}(I_{k_1}=\ldots=I_{k_n-1}=1, I_j=0\ \forall\, j\in J_n) && (4.20)\\
= \ & \frac{1}{k_n}(\prod_{i=1}^{n}\frac{1}{k_i-1}+\prod_{i=1}^{n-1}\frac{1}{k_i-1})=\prod_{i=1}^{n}\frac{1}{k_i-1}.
\end{aligned}
$$

Somit

$$
\begin{aligned}
& \mathbb{P}(I_{k_n}=1|I_{k_n-1}=\ldots=I_{k_1}=1, I_j=0\ \forall\, j\in J_n) && (4.21)\\
= \ & \frac{\frac{1}{k_n}\prod_{i=1}^{n}\frac{1}{k_i-1}}{\prod_{i=1}^{n}\frac{1}{k_i-1}}=\frac{1}{k_n}
\end{aligned}
$$

unabhängig von $k_1,\ldots,k_{n-1}$. Da die $k_1,\ldots,k_n$ aufsteigend aber beliebig waren, folgt die Behauptung. □

Ein entsprechender Satz gilt natürlich auch für Minima – man muss nur die I_j anders definieren.

Satz 4.4 erlaubt natürlich sofort, Aussagen über die Anzahl $S_n=\sum_{k=1}^n I_k$ von Rekorden in einer Folge von n iid Beobachtungen $X_1,\ldots,X_n$. Es ist nämlich

$$
\mathbb{E}S_n=\sum_{k=1}^{n}\frac{1}{k}=\log n+C+o(1) \tag{4.22}
$$

für $n\to\infty$ und

$$
\mathbb{V}S_n=\sum_{k=1}^{n}\frac{1}{k}-\frac{1}{k^2}=\log n+C-\frac{\pi^2}{6}+o(1). \tag{4.23}
$$

Dabei ist C die Euler-Mascheroni-Konstante. Das Gesetz der großen Zahlen und der Zentrale Grenzwertsatz implizieren sofort:

Satz 4.5 *(Rényi) Unter den Voraussetzungen von Satz 4.3 gilt*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - \log n}{\sqrt{\log n}} \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \tag{4.24}$$

für $x \in \mathbb{R}$ und

$$\lim_{n\to\infty} \frac{S_n}{\log n} = 1 \quad \mathbb{P}\text{-}f.s. \tag{4.25}$$

Beweis: Dies sind Varianten des Zentralen Grenzwertsatzes bzw. des Gesetzes der großen Zahlen, bei dem man auf die identische Verteilung verzichtet. Entsprechende Varianten finden sich beispielsweise in dem Buch von Bauer. Es ist eine Übung nachzuweisen, dass die Voraussetzungen für diese Sätze in unserer Situation erfüllt sind. □

Bemerkung 4.6 *Die durch (4.24) gegebene Normalapproximation von S_n kann asymptotisch durch eine geeignete Poissonapproximation verbesert werden. Genauer gilt*

$$\inf_{\mu\in\mathbb{R},\sigma^2>0} \sup_{x\in\mathbb{R}} |\mathbb{P}(S_n \le x) - \Phi(\sigma^2 x + \mu)| \sim \frac{a}{3\sqrt{\log n}}, \tag{4.26}$$

wenn $n \to \infty$ geht. Dabei ist wieder Φ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung und $a = 0,2784\ldots$ die positive Lösung der Gleichung

$$1 + a + \log a = 0. \tag{4.27}$$

Dagegen erhält man mit $\lambda_n = \sum_{k=1}^{n} \frac{1}{k}$:

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(S_n \le x) - \sum_{0\le k\le x} e^{-\lambda_n} \frac{\lambda_n^k}{k!}| \sim \frac{\pi^2}{12\sqrt{2\pi??}\log n} \tag{4.28}$$

für $n \to \infty$. In (4.26) kann man dabei $\mu = \mu_n$, $\sigma^2 = \sigma_n^2$ mit

$$\mu_n = \lambda_n + \frac{2-a}{3}, \quad \sigma_n^2 = \sqrt{\lambda_n - \frac{\pi^2}{6}}, \quad n \in \mathbb{N} \tag{4.29}$$

wählen ohne die Asymptotik zu verletzen, d. h. die Konstanten sind asymptotisch optimal. Man sieht also, dass die Geschwindigkeit der Normalapproximation etwa $0(\frac{1}{\sqrt{\log n}})$ ist, während man mit einer Poissonapproximation $0(\frac{1}{\log n})$ erreichen kann (was auch noch nicht furchtbar schnell ist).

Insgesamt sehen wir in Satz 4.5, dass bei iid Verteilung der Zufallsvariablen X_i die Anzahl der Rekorde nur wie $\log n$ wächst, was im Widerspruch zu der Beobachtung zu stehen scheint, dass Rekorde etwa im Sport häufig gebrochen werden. Hierauf kommen wir noch einmal zurück.

Durch kombinatorische Überlegungen lässt sich auch die exakte Verteilung von S_n berechnen (dies geht auf Arbeiten von Rényi (1966) und Knuth (1973) zurück. Es gilt nämlich

$$\mathbb{P}(S_n = k) = \frac{\left[{n \atop k}\right]}{n!} \quad (1 \le k \le n!), \tag{4.30}$$

wobei $\left[{n \atop k}\right]$ die sogenannten Stirling-Zahlen sind. Diese sind definiert durch

$$n!\binom{x}{n} = \prod_{i=0}^{n-1}(x-i) = \sum_{k=1}^{n}(-1)^{n-k}\begin{bmatrix} n \\ k \end{bmatrix} x^k \; x \in \mathbb{R}. \tag{4.31}$$

Satz 4.4 kann auch zur Lösung des sogenannten Sekretärinnen-Problems herangezogen werden. Hierbei hat ein Personalchef eine neue Sekretärin einzustellen. Unter den n Kandidatinnen will er natürlich eine möglichst gute finden, hat aber das Problem, dass er jeder Sekretärin nach dem Gespräch mit ihr sagen muss, ob sie angenommen oder abgelehnt ist. Äquivalent möchte man in einer Folge $X_1, \ldots, X_n$ von iid Zufallsvariablen mit stetiger Verteilungsfunktion F ohne Rückgriffmöglichkeit den letzten Rekord bzw. äquivalent das letzte der $I_1, \ldots, I_n$ das 1 ist finden. Es hat sich gezeigt (und lässt sich mit dem hier gewählten Zugang auch gut einsehen), dass die optimale Strategie für dieses Problem in der Wahl eines Index $c \in \{1, \ldots, n\}$ besteht, derart, dass

$$p_c = \mathbb{P}\left(\sum_{k=c}^{n} I_k = 1\right) = \sup_{1\leq d\leq n}\left(\mathbb{P}\left(\sum_{k=d}^{n} I_k = 1\right)\right) = \sup_{1\leq d\leq n} p_d \tag{4.32}$$

gilt und erst ab diesem Folgeglied eine Entscheidung zu treffen, d. h. ab diesem Index den ersten neu auftretenden Rekord zu wählen (bzw. das letzte Glied, falls ein solcher nicht existiert). p_c gibt auch die Wahrscheinlichkeit dafür an, dass der so ausgewählte Rekord der einzig verbleibende ist und somit das Maximum der Folge liefert. Wegen

$$\begin{aligned} p_d &= \mathbb{P}\left(\sum_{k=d}^{n} I_k = 1\right) = \sum_{k=d}^{n} \mathbb{P}(\{I_k = 1\} \cap \bigcap_{\substack{j=d \\ j\neq k}}^{n} \{I_j = 0\}) \qquad (4.33) \\ &= \sum_{k=d}^{n} \frac{1}{k} \prod_{\substack{j\neq k \\ j=d}}^{n} (1 - \frac{1}{j}) \\ &= \frac{1}{n} + \frac{d-1}{n} \sum_{k=d}^{n-1} \frac{1}{k} \\ &= \frac{1}{n} + \frac{(d-1)!\left[{n \atop 2}\right] - (n-1)!\left[{d \atop 2}\right]}{n!(d-2)!} \quad 1 \leq d \leq n \end{aligned}$$

lässt sich für kleine n das optimale n und die zugehörige Erfolgswahrscheinlichkeit leicht berechnen.

n	1	2	3	4	5	6
c	1	1 oder 2	2	2	3	3
p_c	1	1/2	4/9	5/112	13/330	77/180

Für große n gilt näherungsweise

$$p_d = \frac{d}{n}\log\frac{n}{d} + 0\left(\frac{\log n}{n}\right)$$

gleichmäßig in $d \in \{1, 2, \ldots, n\}$ mit

$$\max_{0<x\leq 1} x \log \frac{1}{x} = \frac{1}{e} \log e = \frac{1}{e}, \tag{4.34}$$

so dass für das optimale $c = c_n$ gilt

$$\lim_{n\to\infty} \frac{c_n}{n} = \frac{1}{e} \quad \text{und} \quad \lim_{n\to\infty} p_{c_n} = \frac{1}{e}. \tag{4.35}$$

Also $c_n = \lfloor \frac{n}{e} \rfloor$ für große n mit einer asymptotischen Erfolgswahrscheinlichkeit von $\frac{1}{e}$.

Gehen wir noch einmal zurück auf Satz 4.5. Aus (4.25) dort lässt sich direkt auf die fast sichere Konvergenz von $\log U_n$ ableiten. Da nämlich (U_n) streng monoton ist mit $\lim_{n\to\infty} U_n = +\infty$, ergibt eine Teilfolgenüberlegung

$$\lim_{n\to\infty} \frac{n+1}{\log U_n} = 1 \quad \mathbb{P}\text{-f.s.}, \tag{4.36}$$

da stets $S_{U_n} = n + 1$ $\mathbb{P}$-f.s. gilt. In anderer Schreibweise

$$\lim_{n\to\infty} \frac{1}{n} \log U_n = 1 \quad \mathbb{P}\text{-f.s.} \tag{4.37}$$

Wir wollen in der Folge untersuchen, ob diese Teilfolgenübertragung auch in (4.24) möglich wäre. Dies würde wegen

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sqrt{\log U_n} = 1 \quad \mathbb{P}\text{-f.s.}$$

nach (4.36) und der Symmetrie der Normalverteilung (formal) zu

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\log U_n - n}{\sqrt{n}} \leq x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \tag{4.38}$$

führen. Wir werden sogar ein stärkeres Resultat zeigen:

Satz 4.7 *Es existiert eine Folge von iid exponentialverteilten Zufallsvariablen (Y_n) mit Erwartungswert 1 sowie Zufallsvariablen $Z_n \geq 0$ mit $\sup \mathbb{E} Z_n \leq 1$, so dass*

$$\log U_n = Z_n + \sum_{k=1}^{n} Y_k + o(1) \quad \mathbb{P}\text{-}f.s. \tag{4.39}$$

$$\log \Delta_n = \log U_n - Y_n^* + o(1) \quad \mathbb{P}\text{-}f.s., \tag{4.40}$$

wobei

$$Y_n^* = -\log(1 - \exp(-Y_n)) \tag{4.41}$$

wieder eine Folge von iid Exp(1)-verteilten Zufallsvariablen spendiert. Ferner gilt mit $\rfloor x \lfloor = -\lfloor -x \rfloor,\ x \in \mathbb{R}$:

$$\rfloor \frac{U_{n+1}}{U_n} \lfloor = \rfloor e^{Y_{n+1}} \lfloor \quad \textit{für alle} \quad n \geq 0, \tag{4.42}$$

d. h. die Folge $\rfloor \frac{U_{n+1}}{U_n} \lfloor$ ist unabhängig und identisch verteilt wie U_1 (vgl. (4.11)). Dies alles ist natürlich nur möglich, wenn die (Y_i) auf dem gleichen Raum definiert sind wie die (X_i).

Beweis: Sei $(W_n)_n$ eine Folge unabhängiger $\mathcal{R}(0,1)$-verteilter Zufallsvariablen, die auf dem gleichen Raum definiert sind wie die (X_n) und unabhängig von diesen sind. Nach (4.10) folgt

$$\mathbb{P}(U_{n+1} > k | U_n = j) = \sum_{i=k+1}^{\infty} \frac{j}{i(i-1)} = \frac{j}{k} \tag{4.43}$$

für $1 \leq j \leq k$. Dann lässt sich zeigen, dass die durch

$$V_{n+1} = (1 - W_{n+1}) \frac{U_n}{U_{n+1}} + W_{n+1} \frac{U_n}{U_{n+1} - 1} = \frac{U_n}{U_{n+1}} \left(1 + \frac{W_{n+1}}{U_{n+1} - 1} \right) \quad \text{für alle} \quad n \geq 0 \tag{4.44}$$

gegebene Folge wieder aus unabhängigen $\mathcal{R}(0,1)$-verteilten Zufallsvariablen besteht. Dies ist die Übung. Damit ergibt aber

$$Y_n = -\log V_n \quad \text{für alle} \quad n \in \mathbb{N} \tag{4.45}$$

die gewünschte Folge. Es ist nämlich nach (4.44)

$$Y_{n+1} = \log U_{n+1} - \log U_n - \log \left(1 + \frac{W_{n+1}}{U_{n+1} - 1} \right) \tag{4.46}$$

für alle $n \geq 0$, so dass sich nach Summation wegen $\log U_0 = 0$ ergibt

$$\sum_{k=1}^{n} Y_k = \log U_n - \sum_{k=1}^{n} \log \left(1 + \frac{W_k}{U_k - 1} \right) = \log U_n - Z_n, \; n \in \mathbb{N}. \tag{4.47}$$

Hierbei ist

$$0 \leq Z_n = \sum_{k=1}^{n} \log \left(1 + \frac{W_k}{U_k - 1} \right) \leq 2 \sum_{k=1}^{n} \frac{W_k}{U_k} \leq 2 \sum_{k=1}^{n} W_k S_k \tag{4.48}$$

nach (4.46) $\mathbb{P}$-f.s., wobei

$$S_k := \exp \left(- \sum_{i=1}^{k} Y_i \right), \quad k \in \mathbb{N}. \tag{4.49}$$

Für

$$Z := \sup_n Z_n = \sum_{i=1}^{\infty} \log \left(1 + \frac{W_k}{U_k - 1} \right)$$

ergibt sich nun nach dem Satz von der monotonen Konvergenz

$$\begin{aligned} \mathbb{E}Z &= \lim_{n \to \infty} \mathbb{E}[Z_n] \\ &\leq 2 \sum_{k=1}^{\infty} \mathbb{E}[W_k] \mathbb{E}\left[\frac{1}{U_k} \right] \\ &\leq \sum_{k=1}^{\infty} \mathbb{E}(S_k) \\ &= \sum_{k=1}^{\infty} \prod_{i=1}^{k} \mathbb{E} e^{-Y_i} \\ &= \sum_{k=1}^{\infty} 2^{-k} = 1 \end{aligned} \tag{4.50}$$

wegen $\mathbb{E}e^{-Y_i} = \mathbb{E}V_i = \frac{1}{2} \quad \forall\, i \in \mathbb{N}$ (man beachte, dass nach Voraussetzung W_k und (X_n), also auch W_k und U_k unabhängig sind). Wegen $\lim_{n\to\infty} Z_n = Z$ $\mathbb{P}$-f.s. ergibt sich (4.39). Zum Beweis von (4.40) benötigen wir noch

$$\lim_{n\to\infty} n^2(1 - V_n) = \infty \quad \mathbb{P}\text{-f.s.} \tag{4.51}$$

Dies erhalten wir mittels einer Anwendung des Borel-Cantelli-Lemmas, denn es gilt für alle $c \geq 1$

$$\sum_{n=1}^{\infty} \mathbb{P}(n^2(1 - V_n) \leq c) \leq \sum_{n=1}^{\infty} \frac{c}{n^2} < \infty, \tag{4.52}$$

also $\mathbb{P}(\limsup n^2(1 - V_n) \leq c) = 0$ für alle $c \geq 0$. Aus (4.37) folgt schließlich noch

$$\lim_{n\to\infty} \frac{U_n}{n^2} = \infty \quad \mathbb{P}\text{-f.s.} \tag{4.53}$$

und damit auch

$$\lim_{n\to\infty} (1 - V_n)U_n = \infty \quad \mathbb{P}\text{-f.s.} \tag{4.54}$$

Aus (4.44) erhält man nun

$$\begin{aligned}
\frac{\Delta_{n+1}}{U_n} &= \frac{U_{n+1}}{U_n} - 1 = e^{Y_{n+1}} \left(1 + \frac{W_{n+1}}{U_{n+1} - 1}\right) - 1 \\
&= e^{Y_{n+1}}(1 - e^{-Y_{n+1}}) \left(1 + \frac{W_{n+1}}{(1 - V_{n+1})(U_{n+1} - 1)}\right) \\
&= e^{Y_{n+1}}(1 - e^{-Y_{n+1}})(1 + o(1)) \quad \mathbb{P}\text{-f.s.}
\end{aligned} \tag{4.55}$$

für $n \to \infty$ nach (4.52). Durch Logarithmieren ergibt sich

$$\log \Delta_{n+1} = \log U_n + Y_{n+1} - Y^*_{n+1} + o(1) = \log U_{n+1} - Y^*_{n+1} + o(1) \quad \mathbb{P}\text{-f.s.} \tag{4.56}$$

für $n \to \infty$ mit

$$\mathbb{P}(Y^*_n > x) = \mathbb{P}(1 - e^{-Y_1} < e^{-x}) = e^{-x}$$

für $x > 0$. Damit sind (4.39) und (4.40) bewiesen. Die Gültigkeit von (4.42) ergibt sich nun aus (4.45) vermöge

$$U_{n+1} - U_n e^{Y_{n+1}} = U_{n+1} - \frac{U_n}{V_{n+1}} = \frac{W_{n+1}U_{n+1}}{W_{n+1} + U_{n+1} - 1} \quad \text{für} \quad n \geq 0, \tag{4.57}$$

was wegen $U_{n+1} \geq 2$ zu

$$U_{n+1} = \rfloor U_n e^{Y_{n+1}} \lfloor = \rfloor \frac{U_n}{V_{n+1}} \lfloor \quad \text{für} \quad n \geq 0 \tag{4.58}$$

führt. Dies impliziert (4.43). $\square$

Eine Konsequenz aus Satz 4.7 ist:

Satz 4.8 *Es gilt*

$$\lim_{n\to\infty} \frac{1}{n} \log U_n = \lim_{n\to\infty} \frac{1}{n} \log L_n = \lim_{n\to\infty} \frac{1}{n} \log \Delta_n = 1 \quad \mathbb{P}\text{-}f.s. \tag{4.59}$$

und

$$\begin{aligned} & \lim_{n\to\infty} \mathbb{P}\left(\frac{\log U_n - n}{\sqrt{n}} \le x\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{\log L_n - n}{\sqrt{n}} \le x\right) \\ = & \lim_{n\to\infty} \mathbb{P}\left(\frac{\log \Delta_n - n}{\sqrt{n}} \le x\right) \\ = & \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \end{aligned} \tag{4.60}$$

Beweis: Mit den Bezeichnungen von Satz 4.7 gilt

$$\lim_{n\to\infty} \frac{Z_n}{n} = \lim_{n\to\infty} \frac{Z_n}{\sqrt{n}} = 0 \quad \mathbb{P}\text{-f.s.}$$

ebenso wie

$$\lim_{n\to\infty} \frac{Y_n^*}{\sqrt{n}} = 0 \quad \mathbb{P}\text{-f.s.}, \tag{4.61}$$

da wieder mit dem Borel-Cantelli-Lemma folgt

$$\sum_{n=1}^{\infty} \mathbb{P}(Y_n^* > \varepsilon\sqrt{n}) = \sum_{n=1}^{\infty} e^{-\varepsilon\sqrt{n}} < +\infty \tag{4.62}$$

für jedes $\varepsilon > 0$, also $\mathbb{P}(\limsup\{Y_n^* > \varepsilon\sqrt{n}\}) = 0$ für alle $\varepsilon > 0$. Die Aussagen des Satzes ergeben sich somit aus dem Gesetz der großen Zahlen und dem Zentralen Grenzwertsatz für die Folge (Y_n). □

Bemerkenswert an der Aussage des letzten Satzes ist, dass sich das Wachstumsverhalten von $(\log \Delta_n)$ praktisch nicht von der Folge $(\log U_n)$ (bzw. $(\log L_n)$) unterscheidet, obwohl nach (4.2) U_n mit $\sum_{k=0}^{n} \Delta_k$ identisch ist. Dies wird auch noch einmal im folgenden Satz anders beleuchtet.

Satz 4.9 *Es gilt*

$$\left.\begin{aligned} \mathbb{E} \log U_n &= n + 1 - C + o(1) \\ \mathbb{V} \log U_n &= n + 1 - \tfrac{\pi^2}{6} + o(1) \end{aligned}\right\} \quad \textit{für} \quad n \to \infty \tag{4.63}$$

$$\left.\begin{aligned} \mathbb{E} \log \Delta_n &= n - C + o(1) \\ \mathbb{V}(\log \Delta_n) &= n + \tfrac{\pi^2}{6} + o(1) \end{aligned}\right\} \quad n \to \infty. \tag{4.64}$$

Einen Beweis dieses Satzes geben wir nicht. Wir werden zunächst weitere strukturelle Eigenschaften von Rekorden bereitstellen. Wir beginnen mit

Satz 4.10 *Unter den Voraussetzungen von Definition 4.1 gilt*

a) (U_n, X_{U_n}) *und* (X_{U_n}) *sind homogene Markovketten mit den Übergangswahrscheinlichkeiten*

$$\mathbb{P}(U_{n+1} = k, X_{U_{n+1}} > y | U_n = j, X_{U_n} = x) = F^{k-j-1}(x)(1 - F(y)) \tag{4.65}$$

für $n \geq 0,\ 1 \leq j < k,\ x_L < x \leq y,$

$$\mathbb{P}(X_{U_{n+1}} > y | X_{U_n} = x) = \frac{1 - F(y)}{1 - F(x)} \tag{4.66}$$

für $n \geq 0,\ x \leq y,\ x < x_R.$

b) Die Folge (Δ_n) *ist bedingt unabhängig unter* (X_{U_n})*; es gilt*

$$\begin{aligned} & \mathbb{P}(\bigcap_{j=1}^{n} \{\Delta_j = k_j\} | X_{U_{n-1}} = x_{n-1}, \ldots, X_{U_0} = x_0) \qquad (4.67) \\ = & \prod_{j=1}^{n} \mathbb{P}(\Delta_j = k_j | X_{U_{j-1}} = x_{j-1}) \\ = & \prod_{j=1}^{n} [F^{k_j - 1}(x_{j-1})(1 - F(x_{j-1}))] \end{aligned}$$

für $n \in \mathbb{N},\ x_L < x_0 < x_1 \ldots < x_{n-1} < x_R,\ k_1, \ldots, k_n \in \mathbb{N},$ *d. h. die* Δ_j *sind bedingt geometrisch verteilt unter* $X_{U_{j-1}}$*.*

Beweis: Die (U_n) sind Stoppzeiten bzgl. der (X_n); es ist nämlich

$$\{U_1 = n\} = \{X_2, \ldots, X_{n-1} \leq X_1 < X_n\} \quad \text{für} \quad n \geq 2 \tag{4.68}$$

und

$$\{U_{k+1} = n\} = \{X_{U_k+1}, \ldots, X_{U_k+n-1} \leq X_{U_k} < X_{U_k+n}\}, \tag{4.69}$$

woraus die Markov-Verträglichkeit folgt.

Die Markov-Eigenschaft folgt nun aus allgemeinen Sätzen über Markov-Ketten, z. B. Satz A1.4 und Lemma A1.5 in Pfeifer "Einführung in die Extremwertstatistik". Man berechnet

$$\begin{aligned} & \mathbb{P}(U_{n+1} = k, X_{U_{n+1}} > y | U_n = j, X_{U_n} = x) \qquad (4.70) \\ = & \mathbb{P}(X_{j+1}, \ldots, X_{k-1} \leq x, y < X_k) = F^{k-j-1}(x)(1 - F(y)) \end{aligned}$$

für $n \geq 0,\ 1 \leq j < k,\ x_l < x \leq y$. Außerdem

$$\begin{aligned} \mathbb{P}(X_{U_{n+1}} > y | X_{U_n} = x) &= \sum_{k=j+1}^{\infty} \mathbb{P}(U_{n+1} = k, X_{U_{n+1}} > y | U_n = j, X_{U_n} = x) \quad (4.71) \\ &= \sum_{i=1}^{\infty} F^{i-1}(x)(1 - F(y)) = \frac{1 - F(y)}{1 - F(x)} \end{aligned}$$

für $x \le y, \ x < x_R$,

$$\mathbb{P}(\Delta_{n+1} = k | X_{U_n} = x) = \mathbb{P}(U_{n+1} = k + j | U_n = j, X_{U_n} = x) = F^{k-1}(x)(1 - F(x)) \quad (4.72)$$

für $x_L < x < x_R, \ k \in \mathbb{N}$. □

Satz 4.10 zeigt also, dass es sich bei $((U_n, X_{U_n}))$ um ein bedingtes Wartezeitproblem handelt.

Es lassen sich noch viele weitergehende Eigenschaften von Rekorden untersuchen. Wir wollen dies hier nicht tun und verweisen auf das Buch von Pfeifer.

Es hat auch nicht an Versuchen gefehlt, die obige Modellbildung mathematischer Rekrode auf den Bereich des Sports zu übertragen. Das in den Sätzen 4.4 und 4.5 angedeutete Phänomen der relativen Seltenheit neuer Rekorde schein im Widerspruch zu unserer Erfahrung mit Rekorden im Sport zu stehen. Man hat daher versucht, das relativ häufige Brechen alter Rekorde im Sport durch andere Phänomene zu erklären. Ein Erklärungsversuch stammt von Yang (1975), der dieses Phänomen auf das geometrische Anwachsen der Weltbevölkerung zurückführen möchte. Statt iid Zufallsvariablen betrachtet er die Folge (X_n), die durch

$$(X_n) \stackrel{\mathcal{D}}{=} (\max(Z_{n1}, \ldots, Z_{n\alpha_n}),$$

wobei die $\{Z_{nj} : 1 \le j \le \alpha_n\}$ stochastisch unabhängige, identisch verteilte Zufallsvariablen sind, die eine stetige Verteilungsfunktion F besitzen und α_n geometrisch wächst, d. h.

$$\alpha_n = [\lambda^{n-1}]$$

für eine Wachstumsrate $\lambda \ge 1$. Definiert man (U_n) und (Δ_n) wie oben, so ergibt sich

Satz 4.11 *(Yang)*
Unter den obigen Annahmen gilt

$$\lim_{n\to\infty} \mathbb{P}(\Delta_n = k) = (\lambda - 1)\lambda^{-k} = p(1-p)^{k-1}$$

mit

$$p = 1 - \frac{1}{\lambda}.$$

Da sich die Weltbevölkerung seit dem Jahr 1900 etwa alle 36 Jahre verdoppelt hat, ein olympischer Zyklus 4 Jahre beträgt, muss, wenn wir nur olympische Spiele betrachten,

$$\lambda^9 = 2 \quad \text{d. h.} \quad \lambda = 1,08$$

gelten. Somit

$$\lim_{n\to\infty} \mathbb{E}\Delta_n = \frac{1}{p} = \frac{\lambda}{\lambda - 1} = 13,5 \text{ Jahre.}$$

Tatsächlich gilt aber eher $\mathbb{E}\Delta_n \approx 1$ für den 100-Meter-Lauf, $\mathbb{E}\Delta_n \approx 2,4$ für den Marathon und $\mathbb{E}\Delta_n \approx 2,6$ für den Weitsprung.

5 Extrema stationärer Prozesse

Die natürliche Verallgemeinerung von iid Folgen von Zufallsvariablen sind stationäre Folgen, wie wir die beim Birkhoffschen Ergodensatz kennengelernt haben. Wir erinnern:

Definition 5.1 *(X_n) heißt eine stationäre Folge von Zufallsvariablen, wenn*

$$\mathbb{P}^{(X_1,\ldots,X_k)} \overset{\mathcal{D}}{=} \mathbb{P}^{(X_n,\ldots,X_{n+k})}$$

für jedes $n \in \mathbb{N}$ und $k \in \mathbb{N}$ gilt.

Es ist klar, dass Stationarität nicht die einzige Bedingung darstellen kann, um das Extremalverhalten einer zufälligen Folge zu studieren. Beispielsweise ist ja auch jede konstante Folge $X_i = X$ für alle $i \in \mathbb{N}$ stationär und deren Extremalverhalten hängt allein von dem Verhalten von X ab (für das man alles wählen kann). Die Frage soll sein, welchen Einfluss die nun möglichen Korrelationen auf das Extremalverhalten der Folge hat und eine naheliegende Vermutung ist, dass eine sehr schwache Abhängigkeit das Extremwertverhalten überhaupt nicht ändert. Die meisten Arbeiten über die Extremwerttheorie abhängiger Sequenzen beschäftigen sich mit einem solchen Ansatz. Der praktische Zweck dieser Übung liegt darin, dass es oftmals sehr mühsam sein kann, die echte Abhängigkeitsstruktur einer Folge zu ermitteln. Wenn man also ein Resultat zeigen kann, dass diese Struktur auch unwichtig ist, solange sie nur nicht zu stark ist, ist man aus dem Schneider.

Um beispielsweise konstante Folgen etc. auszuschließen, klassifiziert man die Abhängigkeitsstruktur mittels sogenannter Mischungseigenschaften.

Definition 5.2 *Eine stationäre Folge erfüllt die Mischungsbedingung $\mathcal{D}$, falls es eine fallende Folge $g(l) \downarrow 0$ gibt, so dass für alle $p, q \in \mathbb{N}$, $i < \ldots < i_p$ und $j_1 < \ldots < j_q$, so dass $j_1 - i_q > l$ ist und für alle $u \in \mathbb{R}$ gilt*

$$\begin{aligned} &|\mathbb{P}[X_{i_1} \leq u, \ldots, X_{i_p} \leq u, X_{j_1} \leq u, \ldots, X_{j_q} \leq u] \\ - \;& \mathbb{P}[X_{i_1} \leq u, \ldots, X_{i_p} \leq u] \cdot \mathbb{P}[X_{j_1} \leq u, \ldots, X_{j_q} \leq u]| \leq g(l). \end{aligned} \tag{5.1}$$

Eine etwas schwächere Bedingung richtet sich nur an ein gegebenes extremes Niveau.

Definition 5.3 *Eine stationäre Folge von Zufallsvariablen (X_i) erfüllt die Mischungsbedingung $D(u_n)$ für eine Folge u_n, falls es eine Folge $(\alpha_{n,l})$ gibt, die für ein $l_n = o(n)$,*

$$\alpha_{n,l_n} \downarrow 0 \quad n \to \infty,$$

erfüllt, so dass für alle $p, q \in \mathbb{N}$, $i_1 < \ldots < i_p$ und $j_1 < \ldots < j_q$ gibt, so dass $j_1 - i_q > l$ und für alle $n \in \mathbb{N}$

$$\begin{aligned} &|\mathbb{P}[X_{i_1} \leq u_n, \ldots, X_{i_p} \leq u_n, X_{j_1} \leq u_n, \ldots, X_{j_q} \leq u_n] \\ - \;& \mathbb{P}[X_{j_1} \leq u_n, \ldots, X_{i_p} \leq u_n] \cdot \mathbb{P}[X_{j_1} \leq u_n, \ldots, X_{j_q} \leq u_n]| \leq \alpha_{n,l}. \end{aligned} \tag{5.2}$$

Wichtig ist, dass sowohl bei D als auch $D(u_n)$ der Abfall der Differenz nur von l, dem Abstand von j_q zu i_p abhängt, aber nicht von der Anzahl der beteiligten Variablen. Das ist wichtig, weil eine zentrale Idee sein wird, einen kleinen Anteil der Variablen wegzulassen und zu hoffen, dass diese "hinreichend unabhängig" sind. Eine ähnliche Idee benutzt man, wenn man für hinreichend stark mischende Folgen eine CLT zeigen möchte. Diese Strategie wird durch das folgende Lemma vorbereitet.

Lemma 5.4 *Eine stationäre Folge von Zufallsvariablen erfülle Bedingung $D(u_n)$. Es sei $E_1, \dots, E_r$ eine endliche Familie disjunkter Teilmengen von $\{1, \dots, n\}$. Sei*

$$M(E) := \max_{i \in E} X_i.$$

Falls

$$\text{dist } (E_i, E_j) \geq k$$

für alle $1 \leq i, j \leq r$, $i \neq j$, gilt, so folgt

$$|\mathbb{P}\left(\bigcap_{i=1}^{r} M(E_i) \leq u_n\right) - \prod_{i=1}^{r} \mathbb{P}(M(E_i) \leq u_n)| \leq (r-1)\alpha_{n,k}. \tag{5.3}$$

Beweis: Wir beweisen das Lemma per Induktion. Nach Voraussetzung gilt (5.3) für $r = 2$. Nun schließen wir von $r-1$ auf r: Es ist nämlich

$$\mathbb{P}\left[\bigcap_{i=1}^{r-1}\{M(E_i) \leq u_n\} \cap \{M(E_r) \leq u_n\}\right] - \mathbb{P}\left[\bigcap_{i=1}^{r-1}\{M(E_i) \leq u_n\}\right] \mathbb{P}[\{M(E_r) \leq u_n\}]| \leq \alpha_{n,k}$$

und nach Induktionsvoraussetzung

$$|\mathbb{P}\left[\bigcap_{i=1}^{r-1}\{M(E_i) \leq u_n\}\right] - \prod_{i=1}^{r-1} \mathbb{P}[M(E_i) \leq u_n]| \leq (r-2)\alpha_{n,k}.$$

Also mit Hilfe der Dreiecksungleichung:

$$\begin{aligned}
& \left|\mathbb{P}\left[\bigcap_{i=1}^{r}\{M(E_i) \leq u_n\}\right] - \prod_{i=1}^{r} \mathbb{P}[M(E_i) \leq u_n]\right| \\
\leq & \left|\mathbb{P}\left[\bigcap_{i=1}^{r}\{M(E_i) \leq u_n\}\right] - \mathbb{P}\left[\bigcap_{i=1}^{r-1}\{M(E_i) \leq u_n\}\right] \cdot \mathbb{P}[M(E_r) \leq u_n]\right| \\
+ & \left|\mathbb{P}\left[\bigcap_{i=1}^{r-1}\{M(E_i) \leq u_n\}\right] \cdot \mathbb{P}[M(E_r) \leq u_n] - \prod_{i=1}^{r} \mathbb{P}[M(E_i) \leq u_n]\right| \\
\leq & \alpha_{n,k} + (r-2)\alpha_{n,k} = (r-1)\alpha_{n,k}.
\end{aligned}$$

□

Als erste Konsequenz erhalten wir:

Satz 5.5 *Sei $(X_n)_n$ eine stationäre Folge von Zufallsvariablen. Wenn es Folgen (A_n), $A_n > 0$ und (B_n) gibt, so dass für $M_n = \max_{1\le i\le n} X_i$ gilt,*

$$\mathbb{P}[A_n(M_n - B_n) \le x] \to G(x)$$

für alle $x \in C(G)$ für eine nicht-entartete Verteilungsfunktion G, und die Folge die Bedingung $D(A_n^{-1}x + B_n)$ für alle $x \in \mathbb{R}$ erfüllt, dann ist G vom Typ der Gumbel-, Weibull- oder Fréchet-Verteilung.

Beweis: Die Hauptidee ist es zu zeigen, dass G max-stabil ist. Dazu zeigen wir, dass für jedes $k \in \mathbb{N}$ gilt:

$$\mathbb{P}[A_{n_k}(M_n - B_{n_k}) \le x] \to G^{1/x}(x) \tag{5.4}$$

für alle $x \in C(G)$. Dies aber heißt, dass wir

$$\mathbb{P}(M_{kn} \le \frac{x}{A_{nk}} + B_n k) - (\mathbb{P}[M_n \le \frac{x}{A_{nk}} - B_{nk})^k \to 0$$

zeigen müssen. Dies bringt Lemma 5.4 in Anschlag. Der naive Ansatz wäre es, die $k \cdot n$ X_i's in k Stücke der Länge n zu unterteilen. Das Problem ist, dass dann der notwendige Abstand, der in den Voraussetzungen von Lemma 5.4 gefordert wird, zwischen den Indizes nicht vorhanden ist. Der Trick ist, dass wir aus jedem Stück ein Ende der Länge fortlassen, so dass wir die k Blöcke

$$I_l := \{nl + 1, \dots, nl + (n - m)\},$$

$l = 1, \dots, k$ erhalten. Die restlichen Intervalle berechnen wir mit I_l':

$$I_l' := \{nl + n - m + 1, \dots, nl + n - 1\}.$$

Dann gilt mit

$$\begin{aligned}
& \frac{x}{A_{nk}} + B_{nk} =: u_{nk} \\
& \mathbb{P}[M_{kn} \le u_{nk}] = \mathbb{P}\left[\bigcap_{i=1}^{k} M(I_i) \le n_{nk} \bigcap_{i=1}^{k} M(I_i') \le u_{nk}\right] \qquad (5.5)\\
= \quad & \mathbb{P}\left[\bigcap_{i=1}^{k} M(I_i) \le u_{nk}\right] \\
+ \quad & \mathbb{P}\left[\bigcap_{i=1}^{k} \{M(I_i) \le u_{nk}\} \cap \bigcap_{i=1}^{k} \{M(I_i') \le u_{nk}\}\right] \\
- \quad & \mathbb{P}\left[\bigcap_{i=1}^{k} M(I_i) \le u_{nk}\right].
\end{aligned}$$

Nun gilt

$$\begin{aligned}
& \mathbb{P}\left[\bigcap_{i=1}^{k} \{M(I_i) \le u_{nk}\} \cap \bigcap_{i=1}^{k} \{M(I_i') \le u_{nk}\}\right] - \mathbb{P}\left[\bigcap_{i=1}^{k} M(I_i) \le u_{nk}\right] \qquad (5.6)\\
= \quad & \mathbb{P}\left[\bigcap_{i=1}^{k} \{M(I_i) \le u_{nk}\} \cap \bigcap_{i=1}^{k} \{M(I_i') > u_{nk}\}\right] \\
\le \quad & k \cdot \mathbb{P}[M(I_i) \le u_{nk} < M(I_i')].
\end{aligned}$$

Wir erwarten, dass dies klar wird, denn in diesem Falle wäre das Maximum der iid Zufallsvariablen über die große Menge I_l kleiner als das über die kleine Menge I_l'. Wir wollen dies mit Hilfe der Bedingung $D(u_n)$ nachweisen. Wir stellen zunächst fest, dass wir für jedes $r \in \mathbb{N}$ mit

$$r < \frac{n-2m}{2}$$

Intervalle $E_1, \ldots, E_r \leq I_1$, so dass $|E_i| = m$ und d.st $(E_i, E_j) \geq m$ sowie d ist $(E_i, I_1') \geq m$. Der Punkt ist, dass jedes der E_i so groß ist wie I_1' und wir daher auf jedem E_i' die gleiche Wahrscheinlichkeit haben wie auf I_1', die Schwelle u_{nk} zu übertreffen. Mithilfe von Lemma 5.4 bekommen wir:

$$\mathbb{P}[M(I_1) \leq u_{nk} < M(I_1')] \leq \mathbb{P}\left[\bigcap_{j=1}^{r}\{M(E_j) \leq u_{nk}\} \cap \{M(I_1') > u_{nk}\}\right] \tag{5.7}$$

$$\begin{aligned}
&= \mathbb{P}\left[\bigcap_{j=1}^{r} M(E_j) \leq u_{nk}\right] - \mathbb{P}\left[\bigcap_{j=1}^{r}\{M(E_j) \leq u_{nk}\} \cap \{M(I_1') \leq u_{nk}\}\right] \\
&\leq \mathbb{P}[M(E_1) \leq u_{nk}]^r - \mathbb{P}[M(E_1) \leq u_{nk}]^{r+1} + r\alpha_{nk,m} \\
&\leq \frac{1}{r} + r\alpha_{nk,m}.
\end{aligned}$$

Hierbei haben wir in der letzten Zeile die Ungleichung

$$0 \leq p^r(1-p) \leq \frac{1}{r} \quad \forall\, p \in [0,1]$$

verwendet. Um den ersten Term in (5.5) zu verarzten, benutzen wir abermals Lemma 5.4:

$$|\mathbb{P}\left[\bigcap_{i=1}^{k} M(I_i) \leq u_{nk}\right] - \mathbb{P}[M_1(I_1) \leq u_{nk}]^k| \leq k\alpha_{kn,m}.$$

Mit dem gleichen Argument wie in (5.2) erhalten wir

$$|\mathbb{P}[M(I_1 \cup I_1') \leq u_{nk}] - \mathbb{P}[M(I_1) \leq u_{nk}]| \leq \frac{1}{r} + r\alpha_{nk,m}.$$

Dies ergibt (nach ein wenig rechnen)

$$|\mathbb{P}[M_{kn} \leq u_{nk}] - (\mathbb{P}[M_n \leq u_{nk}])^k| \leq 2k((r+1)\alpha_{kn,m} + \frac{1}{r}).$$

Wählen wir nun

$$r \ll m \ll n \quad \text{mit} \quad r \to \infty$$

und

$$r\alpha_{kn,m} \to 0.$$

Dies ist möglich, da $\alpha_{kn,m} \to 0$ geht. □

Wenn wir nun die vollständige Übertragung der iid Sätze auf den abhängigen Fall angehen, stellt sich heraus, dass die Bedingung $D(u_n)$ nicht ausreicht. Wir definieren:

Definition 5.6 *Sei (X_n) eine stationäre Folge von Zufallsvariablen und (u_n) eine Folge in $\mathbb{R}$. Wir sagen, dass (X_n) der Bedingung $D'(u_n)$ genügt, falls*

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=1}^{[n/k]} \mathbb{P}[X_1 > u_n, X_j > u_n] = 0. \tag{5.8}$$

Satz 5.7 *Es sei (X_n) eine stationäre Folge von Zufallsvariablen und (u_n) eine Folge, so dass (X_n) die Bedingung $D(u_n)$ und $D'(u_n)$ erfüllt. Dann gilt für jedes endliche $0 \leq \tau < \infty$*

$$\lim_{n\to\infty} \mathbb{P}[M_n \leq u_n] = e^{-\tau} \tag{5.9}$$

genau dann, wenn

$$\lim_{n\to\infty} n \cdot \mathbb{P}[X_1 > u_n] = \tau \tag{5.10}$$

gilt.

Beweis: Sei $n' = \left[\frac{n}{k}\right]$. Wir beginnen mit (5.10) $\Rightarrow$ (5.9): Im Beweis von Satz 5.5 haben wir gesehen, dass bei Gültigkeit von $D(u_n)$

$$\mathbb{P}[M_n \leq u_n] \sim [\mathbb{P}(M_{n'} \leq u_n)]^k. \tag{5.11}$$

Somit folgt (5.9), wenn wir

$$\mathbb{P}[M_{n'} \leq u_n] \sim (1 - \frac{\tau}{k})$$

zeigen können. Dies folgt, da

$$\mathbb{P}[M_{n'} \leq u_n] = 1 - \mathbb{P}[M_{n'} > u_n]$$

und

$$\mathbb{P}[M_{n'} > u_n] \leq \sum_{i=1}^{n'} \mathbb{P}[X_i > u_n] = \frac{n'}{n} n\mathbb{P}[X_i > u_n] \to \frac{\tau}{k}, \quad \text{wenn} \quad n \to \infty,$$

einerseits. Umgekehrt bekommt man mit dem Einschluss-Ausschluss-Prinzip auch

$$\mathbb{P}[M_{n'} > u_n] \geq \sum_{i=1}^{n'} \mathbb{P}[X_i > u_n] - \sum_{i<j}^{n'} \mathbb{P}[X_i > u_n, X_j > u_n].$$

Wir zeigen, dass der 2. Summand vernachlässigbar ist. Hierfür benötigen wir $D'(u_n)$:

$$\sum_{i<j}^{n'} \mathbb{P}[X_i > u_n, X_j > u_n] \leq n' \sum_{j=2}^{n'} \mathbb{P}[X_1 > u_n, X_j > u_n].$$

Dies konvergiert gegen 0, wenn wir erst $n \to \infty$ und dann $k \to \infty$ schicken. Somit folgt (5.9).

Für die andere Richtung bemerken wir, dass aus (5.9) zusammen mit (5.11)

$$1 - \mathbb{P}[M_{n'} \leq u_n] \sim 1 - e^{-\tau/k}$$

folgt. Nun haben wir im vorhergehenden Schritt gerade gesehen, dass unter der Gültigkeit von $D'(u_n)$

$$1 - \mathbb{P}[M_{n'} \leq u_n] \sim n'(1 - F(u_n))$$

gilt und somit

$$n'(1 - F(u_n)) \sim k^{-1}n(1 - F(u_n)) \sim 1 - e^{-\tau/k}.$$

Schicken wir k gegen ∞, so erhalten wir

$$n(1 - F(u_n)) \to \tau.$$

□

Wir bereiten nun unsere nächsten Resultate vor:

Lemma 5.8 *Es sei (X_n) eine stationäre Folge von Zufallsvariablen und F sei Verteilungsfunktion von X_1. Weiter seien $(u_n)_n$ und (v_n) Folgen reeller Zahlen mit. Falls*

$$\lim_{n\to\infty} n(F(u_n) - F(v_n)) = 0 \tag{5.12}$$

gilt, so folgt

(i) Falls für jedes n I_n Intervalle der Länge $\nu_n = O(n)$ sind, so gilt

$$\mathbb{P}[M(I_n) \leq u_n] - \mathbb{P}(M(I_n) \leq v_n] \to 0. \tag{5.13}$$

(ii) Die Bedingungen $D(u_n)$ und $D(v_n)$ sind äquivalent.

Beweis: Wir setzen

$$F_{k_1,\ldots,k_m}(u) = \mathbb{P}\left[\bigcap_{i=1}^{m} X_{k_i} \leq u\right] \tag{5.14}$$

für $m \in \mathbb{N}$. O. B. d. A. sei $v_n \leq u_n$. Dann gilt

$$|F_{k_1,\ldots,k_m}(u_n) - F_{k_1,\ldots,k_m}(v_n)| \leq \mathbb{P}\left[\bigcup_{i=1}^{m} v_n \leq X_{k_i} \leq u_n\right] \leq m \cdot |F(u_n) - F(v_n)|.$$

Gilt also für ein $K < \infty$, $m \leq Kn$, so folgt

$$|F_{k_1,\ldots,k_m}(u_n) - F_{k_1,\ldots,k_m}(v_n)| \to 0.$$

Wählt man für $m = \nu_n$, so folgt (i) sofort.

Für (ii) nehmen wir an, dass $D(u_n)$ gilt. Setze

$$\vec{\imath} = i_1, \ldots, i_p \quad \text{und} \quad \vec{\jmath} = j_1, \ldots, j_q,$$

mit

$$i_1 < i_2 < \ldots < i_p < j_1 \ldots < j_q \quad \text{und} \quad |j_1 - i_p| > l.$$

Dann folgt

$$|F_{\vec{\imath}\vec{\jmath}}(u_n) - F_{\vec{\imath}}(u_n)F_{\vec{\jmath}}(u_n)| \le \alpha_{n,l}.$$

Aber

$$\begin{aligned} & |F_{\vec{\imath}\vec{\jmath}}(v_n) - F_{\vec{\imath}}(v_n)F_{\vec{\jmath}}(v_n)| \\ \le\ & |F_{\vec{\imath}\vec{\jmath}}(v_n) - F_{\vec{\imath}\vec{\jmath}}(u_n)| + F_{\vec{\imath}}(u_n)|F_{\vec{\jmath}}(v_n) - F_{\vec{\jmath}}(u_n)| \\ & + F_{\vec{\jmath}}(v_n)|F_{\vec{\imath}}(u_n) - F_{\vec{\imath}}(v_n)| + |F_{\vec{\imath}\vec{\jmath}}(u_n) - F_{\vec{\imath}}(u_n)F_{\vec{\jmath}}(u_n)|. \end{aligned}$$

Alle diese Terme konvergieren voraussetzungsgemäß gegen 0, wenn $n \to \infty$ und dann $l \to \infty$ konvergieren. Also folgt $D(v_n)$. □

Lemma 5.9 *Sei u_n eine Folge, so dass*

$$n(1 - F(u_n)) \to \tau$$

gilt. Sei $v_n := u_{[\frac{n}{\vartheta}]}$ für ein $\vartheta > 0$. Dann gilt:

(i) $n(1 - F(v_n)) \to \vartheta\tau$.

(ii) Falls $\vartheta \le 1$ gilt, dann impliziert $D(u_n)$ auch $D(v_n)$.

(iii) Falls $\vartheta \le 1$ gilt, dann impliziert $D'(u_n)$ auch $D'(v_n)$.

(iv) Falls für w_n gilt

$$n(1 - F(w_n)) \to \tau' \le \tau,$$

dann folgt aus $D(u_n)$ auch $D(w_n)$.

Beweis:

(i) Es gilt

$$n(1 - F(v_n)) = n(1 - F(u_{[\frac{n}{\vartheta}]})) = \frac{n}{[\frac{n}{\vartheta}]}\left[\frac{n}{\vartheta}\right](1 - F(u_{[\frac{n}{\vartheta}]}) \to \vartheta\tau.$$

(ii) Mit $\vec{\imath}$ und $\vec{\jmath}$ wie im letzten Beweis folgt

$$|F_{\vec{\imath}\vec{\jmath}}(v_n) - F_{\vec{\imath}}(v_n)F_{\vec{\jmath}}(v_n)| = |F_{\vec{\imath}\vec{\jmath}}\left(u_{[\frac{n}{\vartheta}]}\right) - F_{\vec{\imath}}\left(u_{[\frac{n}{\vartheta}]}\right)F_{\vec{\jmath}}\left(u_{[\frac{n}{\vartheta}]}\right)| \le \alpha_{[\frac{n}{\vartheta}],l}.$$

Also folgt $D(v_n)$.

(iii) Falls $\vartheta \le 1$ gilt, folgt

$$n\sum_{i=2}^{[n/k]} \mathbb{P}[X_1 > v_n, X_i > v_n] \le \frac{n}{[\frac{n}{\vartheta}]}\left[\frac{n}{\vartheta}\right]\sum_{i=2}^{[[\frac{n}{\vartheta}]/k]} \mathbb{P}[X_1 > u_{[\frac{n}{\vartheta}]}, X_i > u_{[\frac{n}{\vartheta}]}] \to 0.$$

(iv) Sei $\tau' = \vartheta\tau$. Nach (ii) gilt $D(v_n)$ und nach (i)

$$n(1 - F(v_n)) \to \vartheta\tau = \tau'.$$

Nach (ii) aus Lemma 5.8 folgt $D(w_n)$.

□

Nun folgt unmittelbar:

Satz 5.10 *Es seien (u_n) und (v_n) Folgen reeller Zahlen, so dass*

$$n(1 - F(u_n)) \to \tau \quad \textit{und} \quad n(1 - F(v_n)) \to \vartheta\tau$$

für ein $\vartheta > 0$ gilt. Es gelte $D(v_n)$ und $D'(v_n)$. Dann folgt für Intervalle I_n mit $|I_n| = [\vartheta]$

$$\lim_{n\to\infty} \mathbb{P}[M(I_n) \le u_n] = e^{-\vartheta\tau}. \tag{5.15}$$

Beweis: Der Beweis ist eine Übung. □

Da wir im Teil über iid Variablen gesehen haben, dass sich das Verhalten von M_n eindeutig durch das Verhalten von $n(1 - F(u_n))$ beschreiben lässt, besagen die letzten Sätze also, dass wir unter geeigneten Bedingungen (nämlich der Gültigkeit von $D(u_n)$ und $D'(u_n)$) die Extrema stationärer Folgen so verhalten wie die Extrema unabhängiger, identisch verteilter Folgen. Natürlich stellt sich sofort die Frage, was geschieht, wenn diese Bedingung nicht erfüllt ist. Der folgende Satz beschreibt zumindest, was wir allein aus $D(u_n)$ folgern können.

Satz 5.11 *Angenommen, für alle $\tau > 0$ gibt es $u_n(\tau)$, dergestalt, dass*

$$n(1 - F(u_n(\tau)) \to \tau$$

und dass die stationäre Folge (X_n) für alle $\tau > 0$ $D(u_n(\tau))$ genügt. Dann gibt es $0 \le \vartheta \le \vartheta' \le 1$, so dass

$$\limsup_{n\to\infty} \mathbb{P}[M_n \le u_n(\tau)]) = e^{-\vartheta\tau} \tag{5.16}$$

und

$$\liminf_{n\to\infty} \mathbb{P}[M_n \le u_n(\tau)] = e^{-\vartheta'\tau}. \tag{5.17}$$

Falls für ein τ die Folge $\mathbb{P}[M_n \le u_n(\tau)]$ konvergiert, dann gilt $\vartheta' = \vartheta$.

Beweis: Wir haben gesehen, dass unter $D(u_n)$

$$\mathbb{P}[M_n \le u_n] - (\mathbb{P}(M_{[n/k]} \le u_n))^k \to 0$$

gilt. Somit folgt aus

$$\limsup_{n\to\infty} \mathbb{P}[M_n \leq u_n(\tau)] = \kappa(\tau),$$

dass auch

$$\limsup_{n\to\infty} \mathbb{P}[M_{[n/k]} \leq u_n(\tau)] = \kappa^{1/k}(\tau).$$

Natürlich gilt auch

$$\limsup_{n\to\infty} \mathbb{P}[M_{[n/k]} \leq u_{[n/k]}(\frac{\tau}{k})] = \kappa(\frac{\tau}{k}).$$

Können wir also zeigen, dass

$$\limsup_{n\to\infty} \mathbb{P}[M_{[n/k]} \leq u_{[n/k]}(\frac{\tau}{k})] = \limsup \mathbb{P}[M_{[n/k]} \leq u_n(\tau)], \tag{5.18}$$

so folgt

$$\kappa^k(\tau/k) = \kappa(\tau) \quad \text{für alle} \quad \tau \quad \text{und} \quad k.$$

Dies wird eindeutig durch $(\kappa(\tau) = e^{-\vartheta\tau}$ für ein $\vartheta > 0$ gelöst. Um nun (5.18) zu zeigen sei o.B.d.A.

$$u_{[n/k]}(\tau/k) \geq u_n(\tau).$$

Dann gilt

$$\begin{aligned}
& |\mathbb{P}[M_{[n/k]} \leq u_{[n/k]}(\tau/k)] - \mathbb{P}[M_{[n/k]} \leq u_n(\tau)]| \\
\leq\ & \left[\frac{n}{k}\right] |F(u_{[n/k]}(\tau/k)) - F(u_n(\tau))| \\
=\ & \frac{[n/k]}{n} \left| \frac{n}{[n/k]} [n/k](1 - F(u_{[n/k]}(\tau/k)) - n(1 - F(u_n(\tau))) \right| \\
=\ & \frac{[n/k]}{n} |k(\tau/k) - \tau + o(1)) \to 0.
\end{aligned}$$

Somit folgt die Behauptung für den lim sup. Für lim inf beweist man die Behauptung analog, wobei der Wert für ϑ' möglicherweise verschieden ist von dem Wert ϑ. Natürlich stimmen ϑ und ϑ' überein, wenn für ein τ der lim sup und der lim inf der betrachteten Folge die gleichen sind. □

Definition 5.12 *Falls für eine stationäre Folge von Zufallsvariablen (X_n) und alle $\tau > 0$ eine Folge $u_n(\tau)$ existiert, so dass*

$$\begin{aligned}
u(1 - F(u_n(\tau))) &\to \tau && \textit{und} \\
\mathbb{P}(M_n \leq u_n(\tau)) &\to e^{-\vartheta\tau} && \textit{für ein} \quad 0 \leq \vartheta \leq 1
\end{aligned}$$

gilt, sagen wir, dass die Folge (X_n) den extremalen Index *ϑ hat.*

Der extremale Index kann als ein Maß für den Grad der Abhängigkeit beim Maximum gesehen werden. Man kann eine etwas andere Version von Satz 5.10 beweisen, bei dem der Aspekt eines Vergleichs mit iid Zufallsvariablen stärker zum tragen kommt. Dafür sei (X_n) eine stationäre Folge von Zufallsvariablen und F sei die Verteilungsfunktion von X_1. Mit $(\hat{X}_n)$ bezeichnen wir eine iid-Folge mit (Marginal-) Verteiungsfunktion F. Die Maxima von (X_n) bzw. $(\hat{X}_n)$ nennen wir M_n bzw. $(\hat{M}_n)$. Dann gilt

Satz 5.13 *Sei (X_n) eine stationäre Folge mit extremalem Index $\vartheta \leq 1$. Weiter sei (u_n) eine Folge und $0 \leq \rho \leq 1$. Dann gilt:*

(i) Falls $\vartheta > 0$ ist und $\mathbb{P}[\hat{M}_n \leq v_n] \to \rho$, so folgt

$$\mathbb{P}[M_n \leq v_n] \to \rho^{\vartheta}. \tag{5.19}$$

(ii) Falls $\vartheta = 0$ ist, so gilt:

a) Falls $\liminf_{n\to\infty} \mathbb{P}[\hat{M}_n \leq v_n] > 0$ ist, so gilt

$$\mathbb{P}[M_n \leq v_n] \to 1.$$

b) Falls $\limsup_{n\to\infty} \mathbb{P}[M_n \leq v_n] < 1$, so folgt

$$\mathbb{P}(\hat{M}_n \leq v_n) \underset{n\to\infty}{\longrightarrow} 0.$$

Beweis:

(i) Sei $\tau > 0$, so dass $e^{-\tau} < p$ gilt. Dann gilt

$$\mathbb{P}[\hat{M}_n \leq u_n(\tau)] \to e^{-\tau} \quad \text{und} \quad \mathbb{P}[\hat{M}_n \leq v_n] \to \rho$$

und $\rho > e^{-\tau}$. Daher gilt für hinreichend großes n $u_n(\tau) \leq vn$ und daher folgt

$$\begin{aligned} \liminf_{n\to\infty} \mathbb{P}[M_n \leq v_n] &\geq \lim_{n\to\infty} \mathbb{P}[M_n \leq u_n(\tau)] \\ &\to e^{-\vartheta\tau}. \end{aligned}$$

Da dies für alle $e^{-\tau} > \rho$ gilt, folgt

$$\liminf_{n\to\infty} \mathbb{P}[M_n \leq v_n] \geq \rho^{\vartheta}.$$

Ähnlich zeigt man

$$\limsup_{n\to\infty} \mathbb{P}[M_n \leq v_n] \leq \rho^{\vartheta},$$

was zusammen (i) ergibt.

(ii) Da $\vartheta = 0$ ist, folgt

$$\mathbb{P}[M_n \leq u_n(\tau)] \to 1 \quad \text{für alle} \quad \tau > 0.$$

Falls

$$\liminf_{n\to\infty} \mathbb{P}[\hat{M}_n \leq v_n] = \rho > 0$$

und $\rho > e^{-\tau}$, dann iglt $v_n > u_n(\tau)$ für alle hinreichend große n und daher

$$\liminf_{n\to\infty} \mathbb{P}[M_n \leq v_n] \geq \liminf_{n\to\infty} \mathbb{P}[M_n \leq u_n(\tau)] = 1.$$

Damit folgt a).

Andererseits gilt

$$\limsup_{n\to\infty} \mathbb{P}[M_n \le v_n] < 1,$$

während für alle $\tau < \infty$

$$\mathbb{P}[M_n \le u_n(\tau)] \to 1$$

gilt. Dann gilt für alle $\tau > 0$ und für schließlich alle n $v_n \le u_n(\tau)$, und somit

$$\limsup_{n\to\infty} \mathbb{P}[\hat{M}_n \le v_n] \le \lim_{n\to\infty} \mathbb{P}[\hat{M}_n \le u_n(\tau)] = e^{-\tau}.$$

Schickt man $\tau \to \infty$, so ergibt sich b).

$\square$

Aus den vorhergehenden Sätzen folgt:

- Falls der extremale Index ϑ der Folge (X_n) echt größer ist als 0, so hat $\hat{M}_n$ eine nicht-entartete Limesverteilung genau dann, wenn M_n eine solche hat. Die Limesverteilungen sind in diesem Fall dieselben und man kann die gleichen Konstantenfolgen benutzen, um M_n bzw. $\hat{M}_n$ zu normieren.
- Falls der extremale Index $\vartheta = 0$ ist, so können M_n und $\hat{M}_n$ nicht mit denselben Konstantenfolgen gegen einen nicht-entarteten Limes konvergieren.

Ein Beispiel für eine stationäre Folge (X_n) mit extremalem Index kleiner als 1 ist die autoregressive Folge ξ_n, die wir als

$$\xi_n = \frac{\xi_{n-1}}{r} + \frac{\varepsilon_n}{r} \tag{5.20}$$

definieren. Hierbei ist $r \ge 2,\ r \in \mathbb{N}$ und die ε_n sind iid-Variablen, die auf $\{0, \dots, r-1\}$ Laplace-verteilt sind und unabhängig von den $\{\xi_n\}$ sind, insbesondere sind ε_n und ξ_{n-1} unabhängig.

Falls ξ_0 gleichmäßig auf $[0, 1]$ verteilt ist, dann auch ξ_n (das ist eine Übung) für alle $n \ge 0$. Also gilt mit $u_n(\tau) = 1 - \frac{\tau}{n}$

$$n\mathbb{P}[\xi_n > u_n(\tau)] = \tau.$$

Es gilt

Satz 5.14 *Für die oben definierte Folge der (ξ_n) gilt für jedes $x \in \mathbb{R}$*

$$\mathbb{P}[M_n \le 1 - \frac{x}{n}] \to \exp(-\frac{r-1}{r}x). \tag{5.21}$$

Der Beweis beruht auf dem folgenden technischen Lemma:

Lemma 5.15 *In der Situation von Satz 5.14 sei m so, dass $1 > \frac{r^m x}{n}$. Dann gilt*

$$\mathbb{P}[M_m \leq 1 - \frac{x}{n}] = 1 - \frac{(m+1)r - m}{rn} x. \tag{5.22}$$

Beweis: Der Beweis benutzt die rekursive Definition der Zufallsvariablen ξ_n. Es gilt:

$$\begin{aligned} & \mathbb{P}[M_m \leq 1 - \frac{x}{n}] = \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}, \xi_m \leq 1 - \frac{x}{n}] \qquad (5.23) \\ = \; & \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}, \frac{\xi_{m-1}}{r} + \frac{\varepsilon_m}{r} \leq 1 - \frac{x}{n}] \\ = \; & \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}, \xi_{m-1} \leq r - \varepsilon_m - \frac{xr}{n}] \\ = \; & \sum_{\varepsilon=0}^{r-1} \frac{1}{r} \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}, \xi_{m-1} \leq r - \varepsilon - \frac{xr}{n}]. \end{aligned}$$

Nun sei $\frac{rx}{n} < 1$. Dann gilt für alle $\varepsilon \leq r - 2$

$$r - \varepsilon - \frac{rx}{n} \geq 2 - \frac{rx}{n} > 1$$

und daher, da

$$\xi_{m-1} \leq M_{m-1} \leq 1 - \frac{x}{n},$$

gilt für all diese ε trivialerweise

$$\xi_{m-1} \leq r - \varepsilon - \frac{xr}{n}.$$

Also gilt für diese x

$$\mathbb{P}[M_m \leq 1 - \frac{x}{n}] = \frac{r-1}{r} \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}] + \frac{1}{r} \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}, \xi_{m-1} \leq 1 - \frac{xr}{n}]. \tag{5.24}$$

Ebenso sehen wir, dass für $i \geq 1$ mit $\frac{r^{i+1} x}{n} < 1$ folgt

$$\begin{aligned} & \mathbb{P}[M_m \leq 1 - \frac{x}{n}, \xi_m < 1 - \frac{r^i x}{n}] \qquad (5.25) \\ = \; & \frac{r-1}{r} \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}] + \frac{1}{r} \mathbb{P}[M_{m-1} \leq 1 - \frac{x}{n}, \xi_{m-1} \leq 1 - \frac{xr^{i+1}}{n}]. \end{aligned}$$

Setzen wir also

$$A_{m,i} := \mathbb{P}[M_m \leq 1 - \frac{x}{n}, \; \xi_m < 1 - \frac{r^i x}{n}],$$

bekommen wir die Rekursion

$$A_{m,i} = \frac{r-1}{r} A_{m-1,0} + \frac{1}{r} A_{m-1,i+1}. \tag{5.26}$$

Wenn wir diese Rekursion k-mal iterien, bekommen wir für $A_{m,0}$ eine Lösung der Gestalt

$$A_{m,0} = \sum_{l=0}^{k} C_{k,l} A_{m-k,l} \tag{5.27}$$

mit Konstanten $C_{l,k}$, die wir nun berechnen wollen.

Mit (5.26) können wir die rechte Seite von (5.27) ausdrücken mittels

$$\begin{aligned} & \sum_{l=0}^{k} C_{k,l}\left[\frac{r-1}{r}A_{m-k-1,0}+\frac{1}{r}A_{m-k-1,l+1}\right] \\ = & \frac{r-1}{r}\sum_{l=0}^{k} C_{k,l}A_{m-k-1,0}+\frac{1}{r}\sum_{l=1}^{k+1} C_{k,l-1}A_{m-k-1,l} \\ = & \sum_{l=0}^{k+1} C_{k+1,l}A_{m-k-1,l}, \end{aligned} \tag{5.28}$$

wobei

$$C_{k+1,0}=\frac{r-1}{r}\sum_{l=0}^{k} C_{k,l} \tag{5.29}$$

und

$$C_{k+1,l}=\frac{1}{r}C_{k,l} \quad \text{für} \quad l \geq 1 \tag{5.30}$$

gilt. Diese Rekursion lässt sich in der Tat lösen: Setzen wir nämlich $x = 0$, so sind alle $A_{k,l} = 1$ und daher auch $\sum_{l=0}^{k} C - k, l = 1$. Also

$$C_{k,0}=\frac{r-1}{r} \quad \text{für alle} \quad k \geq 1.$$

Ebenso gilt offensichtlich $C_{0,0} = 1$. Iteriert man (5.30), so ergibt sich

$$C_{k,l}=\frac{1}{r}C_{k-l,0}=\begin{cases} r^{-l-1}(r-1), & \text{falls } k>l \\ r^{-l}, & \text{falls } k=l \end{cases}.$$

Dies setzen wir in (5.27) ein und erhalten

$$\begin{aligned} \mathbb{P}[M_m \leq 1-\frac{x}{n}] &= \sum_{l=0}^{m} C_{m,l}\mathbb{P}[M_0 \leq 1-\frac{x}{n}, \xi_0 < 1-\frac{r^l x}{n}] \\ &= \sum_{l=0}^{m} C_{m,l}\mathbb{P}[\xi_0 < 1-\frac{r^l x}{n}] \\ &= \sum_{l=0}^{m} C_{m,l}[1-\frac{r^l x}{n}] \\ &= \sum_{l=0}^{m-1}(r-1)r^{-l-1}[1-\frac{r^l x}{n}]+r^{-m}[1-r^m\frac{x}{n}] \\ &= (r-1)\frac{1-r^{-m}}{r-1}-m\frac{r-1}{r}\frac{x}{n}+r^{-m}-\frac{x}{n} \\ &= 1-\frac{r(m+1)-m}{rn}x. \end{aligned} \tag{5.31}$$

Dies ist die Behauptung des Lemmas. □

Wir sind nun in der Lage, Satz 5.14 zu beweisen.

Beweis von Satz 5.14: Wir wollen verwenden, dass für m mit $1 < \frac{r^m x}{n}$ (wie in Lemma 5.15) gilt

$$\mathbb{P}[M_n \leq 1 - \frac{x}{n}] \sim \mathbb{P}[M_m \leq 1 - \frac{x}{n}]^{n/m}. \tag{5.32}$$

Die rechte Seite konvergiert nach Lemma 5.15 gegen $\exp(-\frac{r-1}{r}x)$. Um (5.32) zu zeigen, beweisen wir, dass Bedingung $D(1-\frac{x}{n})$ erfüllt ist. Was wir tatsächlich sehen werden, ist, dass die Korrelationen der Zufallsvariablen ξ_i exponentiell schnell fallen. Konstuktionsgemäß besteht ja für $j > i$ die Zufallsvariable ξ_j aus einem großen Anteil, der unabhängig ist von ξ_i plus $r^{-(j-i)}\xi_i$. Mit den Bezeichnungen aus Lemma 5.8 und seinem Beweis ist

$$\begin{aligned} & F_{\vec{\imath}\vec{\jmath}}(u_n) - F_{\vec{\imath}}(u_n)F_{\vec{\jmath}}(u_n) \\ = \; & F_{\vec{\imath}}(u_n)(\mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n | \xi_{i_1} \leq u_n, \ldots, \xi_{i_p} \leq u_n] \\ & -\mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n]). \end{aligned} \tag{5.33}$$

Nun ist aber

$$\xi_{j_k} = \frac{1}{r}\xi_{j_k-1} + \frac{1}{r}\varepsilon_{j_k} = \ldots = r^{-l}\xi_{j_k-l} + W_{j_k}^{(l)},$$

wobei $W_{j_k}^{(l)}$ unabhängig ist von allen ξ_i mit $i \leq j_k - l$. Daher folgt

$$\begin{aligned} & \mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n | \xi_{i_1} \leq u_n, \ldots, \xi_{i_p} \leq u_n] \\ = \; & \mathbb{P}[W_{j_1}^{(j_1-i_p)} + r^{-(j_1-i_p)}\xi_{i_p} \leq u_n, \ldots, W_{j_q}^{(j_q-i_p)} + r^{-(j_q-i_p)}\xi_{i_p} \leq u_n | \xi_{i_1} \leq u_n, \ldots, \xi_{i_p} \leq u_n] \\ \leq \; & \mathbb{P}[W_{j_1}^{(j_1-i_p)} \leq u_n, \ldots, W_{j_q}^{(j_q-i_p)} \leq u_n] \end{aligned}$$

und

$$\begin{aligned} & \mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n | \xi_{i_1} \leq u_n, \ldots, \xi_{i_p} \leq u_n] \\ \geq \; & \mathbb{P}[W_{j_1}^{(j_1-i_p)} + r^{-(j_1-i_p)} \leq u_n, \ldots, W_{j_q}^{(j_q-i_p)} + r^{-(j_q-i_p)} \leq u_n]. \end{aligned}$$

Nun zeigt man ganz ähnlich

$$\mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n] \leq \mathbb{P}[W_{j_1}^{(j_1-i_p)} \leq u_n, \ldots, W_{j_q}^{(j_q-i_p)} \leq u_n]$$

und

$$\mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n] \geq \mathbb{P}[W_{j_1}^{(j_1-i_p)} + r^{-(j_1-i_p)} \leq u_n, \ldots, W_{j_q}^{(j_q-i_p)} + r^{-(j_q-i_p)} \leq u_n].$$

Daher folgern wir

$$\begin{aligned} & |\mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n | \xi_{i_1} \leq u_n, \ldots, \xi_{i_p} \leq u_n] - \mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n]| \\ \leq \; & |\mathbb{P}[W_{j_1}^{(j_1-i_p)} \leq u_n, \ldots, W_{j_q}^{(j_q-i_p)} \leq u_n] \\ & -\mathbb{P}[W_{j_1}^{(j_1-i_p)} + r^{-(j_q-i_p)} \leq u_n, \ldots, W_{j_q}^{(j_q-i_p)} + r^{-(u_q-i_p)} \leq u_n]| \\ \leq \; & \sum_{k=1}^{q} \mathbb{P}[u_n - r^{-(j_k-i_p)} \leq W_{j_k}^{(j_k-i_p)} \leq u_n]. \end{aligned}$$

Aber diese Wahrscheinlichkeiten lassen sich kontrollieren:

$$\begin{aligned}
& \mathbb{P}[u_n - r^{-(j_k-i_p)} \leq W_{j_k}^{(j_k-i_p)} \leq u_n] \\
\leq \; & \mathbb{P}[u_n - 2r^{-(j_k-i_p)} \leq \xi_{j_k} \leq u_n + r^{-(j_k-i_p)}] \\
\leq \; & 3r^{-(j_k-i_p)}.
\end{aligned}$$

Daher folgt

$$\begin{aligned}
& |\mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n | \xi_{i_1} \leq u_n, \ldots, \xi_{i_p} \leq u_n] - \mathbb{P}[\xi_{j_1} \leq u_n, \ldots, \xi_{j_q} \leq u_n] \\
\leq \; & 3 \sum_{k=1}^{q} r^{-(j_k-i_p)} \\
\leq \; & 3 \frac{r^{-l}}{r-1}.
\end{aligned}$$

Daher gilt $D(1-\frac{x}{n})$ und die Behauptung folgt. □

Bemerkung 5.16 *Im Gegensatz dazu gilt die Bedingung $D'(1-\frac{x}{n})$ nicht:*

$$\begin{aligned}
& n \sum_{i=2}^{[n/k]} \mathbb{P}[\xi_1 > u_n, \xi_j > u_n] \\
\geq \; & n \sum_{i=2}^{[n/k]} \mathbb{P}[\xi_1 > u_n, \varepsilon_j = r-1 \quad \forall\, 2 \leq j \leq i] \\
= \; & \sum_{i=2}^{[n/k]} r^{-i+1} > 0.
\end{aligned}$$

Das Auftreten eines nicht-trivialen extremalen Index stammt nämlich vor allem von der starken Korrelation benachbarter Zufallsvariablen, was die Bedingung D' genau ausschließt.

6 Extrema nicht-stationärer Zufallsvariablen

Wir wollen nun auch die Bedingung der Stationarität fallen lassen. Wenn wir uns an unsere Analyse der iid Beobachtungen erinnern, so war dort die Beziehung

$$n(1 - F(u_n)) \to \tau \Leftrightarrow \mathbb{P}[M_n \leq u_n] \to e^{-\tau} \tag{6.1}$$

zentral. Diese wird motiviert durch die Rechnung

$$\mathbb{P}[M_n \leq u_n] = F^n(u_n) = (1 - \frac{n(1 - F(u_n))}{n})^n$$

(wobei die rechte Seite gegen $e^{-\tau}$ konvergiert, wenn die linke Seite von (6.1) wahr ist). Diese Berechnungen sind natürlich nicht wahr, wenn die Bedingung der Unabhängigkeit und identische Verteilung nicht war sind. Allerdings ist (6.1) auch nicht notwendig. Wir entwickeln in der Folge ein anderes Kriterium.

Lemma 6.1 *Die Folge (A_n) erfülle für alle $s \in \mathbb{N}$*

$$A_n \leq \sum_{l=0}^{2s} \frac{(-1)^l}{l!} a_l(n) \tag{6.2}$$

und

$$A_n \geq \sum_{l=0}^{2s+1} \frac{(-1)^l}{l!} a_l(n), \tag{6.3}$$

wobei für alle $l \in \mathbb{N}$

$$\lim_{n\to\infty} a_l(n) = a^l \tag{6.4}$$

gelte. Dann folgt

$$\lim_{n\to\infty} A_n = e^{-a}. \tag{6.5}$$

Beweis: Offensichtlich folgt aus (6.2) und (6.3)

$$\limsup A_n \leq \sum_{l=0}^{2s} \frac{(-a)^l}{l!} \tag{6.6}$$

und

$$\liminf_{n\to\infty} A_n \geq \sum_{l=0}^{2s+1} \frac{(-a)^l}{l!}. \tag{6.7}$$

Die Reihen rechts in (6.6) und (6.7) konvergieren diskret gegen e^{-a}, also folgt (6.5). □

Wir verbinden dieses Lemma mit dem Maximum M_n einer nicht-stationären Folge (X_n) mittels des Einschluss-Ausschluss-Prinzips.

Satz 6.2 *Sei $(B_i)_{i\in\mathbb{N}}$ eine Folge von Ereignissen und $\mathbb{1}_{B_i}$ der Indikator eines solchen Ereignisses B_i. Dann gilt für alle $s \in \mathbb{N}$*

$$\begin{aligned} \mathbb{1}_{\bigcap_{i=1}^n B_i} &\leq \sum_{l=0}^{2s}(-1)^l \sum_{\{j_1,\ldots,j_l\}\subseteq\{1,\ldots,n\}} \mathbb{1}_{\bigcap_{r=1}^l B_{jr}^c} \\ \mathbb{1}_{\bigcap_{i=1}^n B_i} &\geq \sum_{l=0}^{2s+1}(-1)^l \sum_{\{j_1,\ldots,j_l\}\subseteq\{1,\ldots,n\}} \mathbb{1}_{\bigcap_{r=1}^l B_{jr}^c}. \end{aligned}$$

Dabei werden Terme mit $l > n$ 0 gesetzt und die Summe über $\{i_1,\ldots,i_r\}$ wird stets als die Summe über die geordneten Teilmengen $1 \leq i_1 \ldots < i_l \leq n$ verstanden.

Beweis: Der Beweis ist eine Übung. □

Korollar 6.3 *Sei (X_n) eine beliebige Folge von Zufallsvariablen, dann gilt:*

$$\begin{aligned} \mathbb{P}[M_n \leq u] &\leq \sum_{l=0}^{2s}(-1)^l \sum_{\{j_1,\ldots,j_l\}\subseteq\{1,\ldots,n\}} \mathbb{P}(X_{j_r} > 0 \quad \forall\, r = 1,\ldots,l) \quad \textit{und} \\ \mathbb{P}[M_n \leq u] &\geq \sum_{l=0}^{2s+1}(-1)^l \sum_{\{j_1,\ldots,j_l\}\subseteq\{1,\ldots,n\}} \mathbb{P}[X_{j_r} > 0 \quad \forall\, r = 1,\ldots,l]. \end{aligned}$$

Beweis: Der Beweis ist eine Übung und ergibt sich unmittelbar aus Satz 6.2. □

Durch eine Kombination von Lemma 6.1 und Korollar 6.3 bekommt man ein allgemeines Kriterium für Deiecksschemata von Zufallsvariablen.

Satz 6.4 *Sei (X_i^n), $i \leq n$, $n \in \mathbb{N}$ ein Dreiecksschema von Zufallsvariablen. Für jedes $l \in \mathbb{N}$ gelte*

$$\lim_{n\to\infty} \sum_{\{j_1,\ldots,j_l\}\subseteq\{1,\ldots,n\}} \mathbb{P}[X_{j_r}^n > u_n \quad \forall\, r = 1,\ldots,l] = \frac{\tau^l}{l!}. \tag{6.8}$$

Dann folgt

$$\lim_{n\to\infty} \mathbb{P}[M_n \leq u_n] = e^{-\tau}. \tag{6.9}$$

Beweis: Der Beweis ist nicht schwierig und wiederum eine Übung. □

Bemerkung 6.5 *(6.8) stimmt für Folgen von iid Zufallsvariablen. Der Nachweis ist eine Übung.*

Korollar 6.6 *Sei* (X_i^n) *ein Dreiecksschema von Zufallsvariablen* $(i \leq n, n \in \mathbb{N})$*, so dass für jedes* $n \in \mathbb{N}$ *die gemeinsame Verteilung von* $X_1^n, \ldots, X_n^n$ *invariant ist unter Permutationen der Indizes* $i = 1, \ldots, n$*. Falls für jedes* $l \in \mathbb{N}$

$$\lim_{n\to\infty} n^l \mathbb{P}[X_r^n > u_n \quad \forall\, r = 1, \ldots, l] = \tau^l, \tag{6.10}$$

so gilt

$$\lim_{n\to\infty} \mathbb{P}[M_n \leq u_n] = e^{-\tau}. \tag{6.11}$$

Beweis: Der Beweis ist eine einfache Übung. □

Satz 6.4 und Korollar 6.6 ergeben zusammen einfach

Satz 6.7 *Es sei* $u_n^1 > u_n^2 > \ldots > u_n^r$*, und sei* (X_i^n)*,* $n \in \mathbb{N}$*,* $i \in \{1, \ldots, n\}$ *ein Dreiecksschema von Zufallsvariablen. Für jedes* $l \in \mathbb{N}$ *und jedes* $1 \leq j \leq r$ *gelte*

$$\lim_{n\to\infty} \sum_{\{j_1,\ldots,j_l\}\subseteq\{1,\ldots,n\}} \mathbb{P}[X_{j_r}^n > u_n^s \quad \forall\, r = 1, \ldots, l] = \frac{\tau_s^l}{l!} \tag{6.12}$$

mit

$$0 < \tau_1 < \tau_2 \ldots < \tau_r < \infty,$$

dann gilt

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{P}[S_n^1 = k_1, S_n^2 - S_n^1 = k_2, \ldots, S_n^r - S_n^{r-1} = k_r] \\ = \;& \frac{\tau_1^{k_1}}{k!} \frac{(\tau_2 - \tau_1)^{k_2}}{k_2!} \cdots \frac{(\tau_r - \tau_{r-1})^{kr}}{k_r!} e^{-\tau r}. \end{aligned} \tag{6.13}$$

Hierbei ist

$$S_n(u) := \#\{i \leq n : X_i > u\} \tag{6.14}$$

die Anzahl der Überquerungen von Level u *bis zur Zeit* n *und*

$$S_n^k = S_n(n_k).$$

Beweis: Offensichtlich. □

Wir werden diese Resultate nun auf ein klassisches Problem anwenden, das sogenannte Zerlegungsproblem. Seien N Zahlen $X_1, \ldots, X_N$ gegeben. Die Aufgabe ist es, eine Unterteilung der Indizes in 2 Mengen zu finden, so dass die Summe der Zahlen aus beiden Gruppen ungefähr gleich groß ist. Diese Aufgabe kann auf vielfältige Weise übersetzt werden, etwa so, dass man eine Anzahl von Jobs auf zwei Prozessoren übertragen möchte und zwar so, dass beide beinahe die gleiche Arbeitszeit benötigen. Einige dieser Probleme haben interessante Übersetzungen in die Extremwerttheorie, insbesondere lassen sie sich als Anwendungsbeispiele für Satz 6.4 auffassen. Hierzu fassen wir die Gewichte als zufällige

Variablen auf. Des weiteren assoziieren wir mit jeder Zerlegung der Menge $\{1,\ldots,N\}$ in zwei Teilmengen Λ_1 und Λ_2, $\Lambda_1\cap\Lambda_2=\emptyset$ mit einer Abbildung

$$\sigma:\{1,\ldots,N\}\to\{-1,1\}$$

dergestalt, dass

$$\Lambda_1:=\{i:\sigma_i=1\}\qquad\text{und}\qquad\Lambda_2:=\{i:\sigma_i=-1\}$$

gilt. Dann wollen wir

$$\left|\sum_{i\in\Lambda_1}X_i-\sum_{i\in\Lambda_2}X_i\right|=\left|\sum_{i=1}^N X_i\sigma_i\right|=X_\sigma^{(N)}\tag{6.15}$$

minimieren. Offenbar gilt

$$X_\sigma^{(N)}=X_{-\sigma}^{(N)}.$$

Da uns nicht interessiert, welche der beiden Mengen wir Λ_1 und welche wir Λ_2 nennen, können wir alles modulo dieser Symmetrie betrachten. Wir spezialisieren nun den Zufallsmechanismus für die X_i und wählen diese als iid $\mathcal{N}(0,1)$-verteilt. Wir setzen

$$Y_N(\sigma):=\frac{1}{\sqrt{N}}\sum_{i=1}^N\sigma_iX_i\tag{6.16}$$

und

$$X_\sigma^{(N)}:=-|Y_N(\sigma)|.\tag{6.17}$$

Dann erhalten wir:

Satz 6.8 *Falls die X_i iid $\mathcal{N}(0,1)$-verteilt sind, so gilt*

$$\mathbb{P}[\max_{\sigma\in\{-1,+1\}^N}X_\sigma^N\le C_Nx]\to e^{-x},\tag{6.18}$$

wobei $C_N=2^{-N-1}\sqrt{2\pi}$ ist.

Wir beginnen den Beweis mit der folgenden

Proposition 6.9 *Sei $K_N=\frac{1}{C_N}=2^{N+1}/\sqrt{2\pi}$. Schreibe $\sum_{\sigma^1,\ldots,\sigma^l\in\{\pm1\}^N}(\cdot)$ für die Summe über alle geordneten Folgen in $\{\pm1\}^N$. Dann gilt für alle $l\in\mathbb{N}$ und alle Konstanten $c_j>0$, $j=0,\ldots,l$,*

$$\lim_{N\to\infty}\sum_{\sigma^1,\ldots,\sigma^l\in\{\pm1\}^N}\mathbb{P}[K_N|Y_N(\sigma^j)|<c_j\quad\forall\, j=1,\ldots,l]=\prod_{j=1}^l c_j/l!.\tag{6.19}$$

Tatsächlich ist mithilfe von Satz 6.4 klar, dass wir Satz 6.8 bewiesen haben, wenn wir Proposition 6.9 zeigen können. Wir skizzieren zunächst, warum Proposition 6.9 wahr sein sollte. Zunächst sind die $Y_N(\sigma)$ für jedes σ Gaußsche Zufallsvariablen (als Summe von Gaußvariablen) mit Erwartungswert 0 und Kovarianzmatrix $B_N(\sigma^1, \ldots, \sigma^l)$, deren Einträge durch

$$b_{m,n} = \operatorname{cov}(Y_N(\sigma^m), Y_N(\sigma^n)) = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^m \sigma_i^n \tag{6.20}$$

gegeben sind ($1 \leq m, n \leq l$). Die rechte Seite wird für zufällig ausgewählte Spinkonfigurationen ungefähr 0 sein (also o(1)). Die Wahrscheinlichkeit, dabei einen Wert zu erhalten, der wesentlich von 0 verschieden ist (also $\geq c > 0$ im Betrag) ist exponentiell klein in n. Für die Konfigurationen mit einer Kovarianz ungefähr 0 erhalten wir ein Verhalten wie bei unabhängigen Zufallsvariablen, die $\mathcal{N}(0,1)$-verteilt sind (denn $V(Y_N^n) = 1$) und für Gauß-Variablen legt die Kovarianz die komplette Abhängigkeitsstruktur fest. Die Wahrscheinlichkeit aus (6.19) ist dann die gleiche wie die Wahrscheinlichkeit, dass diese Gauß-Variablen in die (exponentiell kleinen) Intervalle

$$[-c_j 2^{-N-1}\sqrt{2\pi}, c_j 2^{-N-1}\sqrt{2\pi}]$$

fallen. Da die Gaußdichte bei 0 von der Größe $\frac{1}{\sqrt{2\pi}}$ ist und wir sie auf kleinen Intervallen durch Konstanten approximieren können, kommen wir dazu, dass die gesuchte Wahrscheinlichkeit durch $\prod_{j=1}^{l} 2c_j/2^{N+1}$ approximiert werden kann. Bislang ist dies allerdings nur eine heuristische Lösung, da wir sowohl alle Approximationen ohne Fehlerterme benutzt als auch den (exponentiell kleinen) Anteil von Spinpaaren mit nicht-verschwindenden $\frac{1}{N}\sum_{i=1}^{N} \sigma_i^n \sigma_j^m$ ignoriert haben. Um die Konfigurationen aus $(\{-1,+1\}^N)^l$, $\sigma^1, \ldots, \sigma^l$ zu betrachten, für die

$$b_{i,j} \to 0 \qquad \text{für mindestens ein Paar} \qquad 1 \leq i \neq j \leq l \tag{6.21}$$

für $N \to \infty$. Man könnte vermuten, dass diese Terme unproblematisch sein sollen, denn diese Terme bilden nur einen exponentiell kleinen Bruchteil aller Terme. Tatsächlich kann dies höchstens dadurch aufgewogen werden, dass die Wahrscheinlichkeiten exponentiell größer sind als die anderen. Wir werden sehen, dass dies nur im Falle degenerierter Kovarianzmatrizen geschehen kann. Ist nämlich $B_N(\sigma^1, \ldots, \sigma^l)$ nicht degeneriert, so ist die in Frage stehende Wahrscheinlichkeit (durch Approximation der exp-Funktion um 0 durch 1) von der Ordnung

$$\frac{1}{\sqrt{det\, B_N(\sigma^1, \ldots, \sigma^l)}} \prod_{j=1}^{l} \frac{2}{\sqrt{2\pi}} c_j / K_N. \tag{6.22}$$

Betrachtet man die Definition der $b_{i,j}$, so ist klar, dass $\frac{1}{\sqrt{B_N(\sigma^1, \ldots, \sigma^2)}}$ höchstens wie ein Polynom wachsen kann, also ist die in Frage stehende Wahrscheinlichkeit bis auf einen polynomiellen Faktor K_N^{-l}, wohingegen die Anzahl von Summanden in (6.19), die überhaupt hier betrachtet wird (die, wo es nicht-verschwindende Kovarianzen gibt), exponentiell kleiner ist als K_N. Somit ist der Beitrag, der von solchen $\sigma^1, \ldots, \sigma^l$ zur Gesamtsumme in (6.19) geliefert wird, exponentiell klein.

Wenn die Kovarianzmatrix, die von den $\sigma^1, \ldots, \sigma^l$ erzeugt wird, $B_N(\sigma^1, \ldots, \sigma^l)$, degeneriert ist, können die $(Y_N(\sigma^i))_{i=1}^l$ linear abhängig sein und daher auch die Wahrscheinlichkeiten in (6.19) exponentiell viel größer als K_N^{-l} sein. Wir werden trotzdem in der Lage sein zu zeigen, dass diese Terme einen verschwindenden Beitrag liefern.

Wir beginnen nun mit dem Beweis von Proposition 6.9.

Beweis von Proposition 6.9: Sei $C(\vec{\sigma})$ die $l \times N$ Matrix, die aus den Elementen σ_i^r, $i = 1, \dots, N,\ r = 1, \dots, l$, besteht. Es gilt

$$B(\vec{\sigma}) := B(\sigma^1, \dots, \sigma^r) = \frac{1}{N} C^t(\vec{\sigma}) C(\vec{\sigma}).$$

Wir unterteilen die Summe in (6.19) um:

$$\sum_{\sigma^1,\dots,\sigma^l \in \{-1,+1\}^N} [\cdot] = \sum_{\substack{\sigma^1,\dots,\sigma^l \in \{-1,+1\}^N: \\ rank\,(c(\vec{\sigma}))=l}} \mathbb{P}[\cdot] + \sum_{\substack{\sigma^1,\dots,\sigma^l \in \{\pm 1\}^N \\ rank\, c(\vec{\sigma})<l}} \mathbb{P}[\cdot], \tag{6.23}$$

wobei $[\cdot]$ natürlich für das Ereignis steht, über dessen Wahrscheinlichkeit summiert wird. Der erste Term sollte natürlich für die rechte Seite in (6.19) verantwortlich sein, während der zweite gegen 0 gehen sollte.

Lemma 6.10 *Die Matrix $C(\vec{\sigma})$ enthalte alle 2^l möglich verschiedenen Spalten und angenommen $\tilde{\sigma} \in \{\pm 1\}^N$ ist dergestalt, dass sie eine Linearkombination der Spalten der Matrix $C(\vec{\sigma})$ ist. Dann gibt es $1 \le i, j \le l$, so dass entweder $\tilde{\sigma} = \sigma^j$ oder $\tilde{\sigma} = -\sigma^j$ gilt.*

Beweis: Wir wissen also, dass $\tilde{\sigma}$ die Darstellung

$$\tilde{\sigma}_i = \sum_{\gamma=1}^{l} z_r \sigma_i^r \tag{6.24}$$

besitzt. O.B.d.A. sei $\sigma_i^1 \equiv 1$. Da alle verschiedenen Vorzeichenkombinationen in der Reihe vorkommen, gibt es ein i, so dass $\sigma_i^r = -\sigma_1^r$ für alle $r \ge 2$. Somit gilt insbesondere

$$\tilde{\sigma}_1 = z_1 + \sum_{r=2}^{l} z_r \sigma_1^r \qquad \text{und} \qquad \tilde{\sigma}_i = z_1 - \sum_{r=2}^{l} z_r \sigma_1^r. \tag{6.25}$$

Addiert man diese Gleichungen, sieht man, dass

$$z_1 \in \{-1, 0, 1\}$$

gelten muss. Ist $z_1 = 0$, so sind wir für den Fall $l = 2$ fertig. Anderenfalls können wir l um eins reduzieren und induktiv fortfahren.

Ist $z_1 \neq 0$, so ist $\tilde{\sigma}_i = \tilde{\sigma}_1 = z_1$ und wir erhalten

$$\sum_{r=2}^{l} z_i \sigma_1^r = 0.$$

In der Tat gilt ja $\sum_{r=2}^{l} z_r \sigma_j^r = 0$ für alle j, so dass $\tilde{\sigma}_j = z_j$ ist. Angenommen, es gibt ein k, so dass $\tilde{\sigma}_k = -z_1$ gilt, dann muss es auch ein k' geben mit

$$\sigma_{k'}^r = -\sigma_k^r \qquad \text{für alle} \quad r \ge 2.$$

Also erhalten wir

$$\sum_{r=2}^{l} z_r \sigma_k^r = -2z_1 \qquad \text{und} \qquad -\sum_{r=2}^{l} z_r \sigma_k^r = \tilde{\sigma}_{k'} - z_1.$$

Dann haben wir aber $2z_1 = \tilde{\sigma}'_k - z_1$ und das ist unmöglich. Also bleibt der Fall $\tilde{\sigma}_j = z_1$ für alle j. Also

$$\sum_{r=2}^{l} z_r \sigma_j^r = 0 \qquad \text{für alle} \quad j.$$

Nun betrachten wir i, so dass $\sigma_i^2 = \sigma_j^2$ und $\sigma_i^r = \sigma_j^3$ für alle $r \geq 3$. Dann folgt wie zuvor $z_2 = 0$ und daher

$$\sum_{r=3}^{l} z_r \sigma_j^l = 0 \qquad \text{für alle} \quad j.$$

Wir schließen daher, dass für $z_1 \neq 0$ gilt:

$$z_1 = \tilde{\sigma}_i \quad \text{für alle} \quad i \qquad \text{und} \qquad z_r = 0 \quad \text{für alle} \quad r \geq 2.$$

Ist $z_1 = 0$, führen wir das Argument weiter, bis wir ein k gefunden haben, so dass $z_k = \tilde{\sigma}_i$ für alle gilt und alle anderen z_r sind wieder gleich 0. Dies beweist das Lemma. □

Aus diesem Lemma können wir schon folgendes schließen: Es seien $r < l$ linear unabhängige Vektoren $\sigma^{i_1}, \ldots, \sigma^{i_r}$ unter den l Vektoren $\sigma^1, \ldots, \sigma^l$ gibt (und es gibt höchstens $(2^r - 1)^N$ solcher Vektoren). Wenn die Matrix $C(\sigma^{i_1}, \ldots, \sigma^{i_r})$ alle 2^r verschiedenen Zeiten enthält, so sind nach Lemma 6.10 alle übrigen σ^j, $j \neq i_1, \ldots, i_r$, gleich einem der $\sigma^{i_1}, \ldots, \sigma^{i_r}$. Dies ist aber unmöglich, da wir über verschiedene $\sigma^1, \ldots, \sigma^l$ summieren. Also gibt es höchstens $O((2^r - 1)^N)$ verschiedene Arten, diese Reihen zu konstruieren. Außerdem gibt es nur eine von N unabhängige Anzahl, um die $l - r$ Konfigurationen zu$C(\sigma^1, \ldots, \sigma^l)$ zu ergänzen. Wir schätzen nun die entsprechenden Wahrscheinlichkeiten ab:

Lemma 6.11 *Es gibt eine Konstante $C > 0$ unabhängig von N, so dass für alle festen $\sigma^1, \ldots, \sigma^l \in \{-1, +1\}^N$ und jeden Rang $r = rank(C(\sigma^1, \ldots, \sigma^l)) \leq l$ und alle $N > 1$*

$$\mathbb{P}[|Y(\sigma^j)| < \frac{c_j}{K_N} \quad \forall\, j = 1, \ldots, l] \leq C \cdot K_N^{-r} N^{r/2}. \tag{6.26}$$

Beweis: Wir entfernen linear abhängige Spalten und verbleiben mit r linear unabhängigen Spalten. Diese entsprechen r Konfiguarionen σ^j, $j \in A_r := \{j_1, \ldots, r_j\} \subseteq \{1, \ldots, l\}$. Wir bezeichnen mit $\bar{C}^r(\vec{\sigma})$ die $(N \times r)$-Matrix, die aus diesen Konfigurationen konstruiert wird und mit $B^r(\vec{\sigma})$ die zugehörige Kovarianzmatrix. Dann ist die Wahrscheinlichkeit

$\mathbb{P}[|Y(\sigma^j)| < c_j/K_N \, \forall\, j = 1, \ldots, l]$ nicht größer als wenn wir nur über die $j \in A_r$ gehen:

$$\begin{aligned} & \mathbb{P}[|Y(\sigma_j)| < \frac{c_j}{K_N} \quad \forall\, j = 1, \ldots, l] \qquad (6.27)\\ \leq\; & \mathbb{P}[|Y(\sigma_j)| < \frac{c_j}{K_N} \quad \forall\, j \in A_r]\\ \leq\; & \frac{1}{(2\pi)^{r/2}\sqrt{det(B^r(\vec{\sigma}))}} \int_{-c_{j_1}/K_N}^{c_j/K_N} \int_{-c_{j_r}/K_N}^{c_{j_r}/K_N} \exp(-\sum_{s,s'=1}^{r} x_{i_s}|B^r(\vec{\sigma})|^{-1}_{s,s'}) \prod_{j \in A_r} dx_j\\ \leq\; & \frac{1}{(2\pi)^{r/2}\sqrt{det(B^r(\vec{\sigma}))}} K_N^{-r} 2^r \prod_{s=1}^{r} c_{j_s}. \end{aligned}$$

Schließlich beobachtet man, dass jeder Matrixeintrag von $B^r(\vec{\sigma})$ von der Form $\frac{1}{N^z}$ für ein $z \in \mathbb{Z}$ ist. Daher ist $det(B^r(\vec{\sigma}))$ von der Form N^{-r} und der Determinante einer Matrix mit nur ganzzahligen Einträgen. Aber die Determinante einer Matrix mit ausschließlich ganzzahligen Einträgen ist eine ganze Zahl. Da wir vorausgesetzt haben, dass $B^r(\vec{\sigma})$ vollen Rang hat, ist diese Zahl nicht 0. Somit ist

$$|\, det\,(B^r(\vec{\sigma}))\,| \geq N^{-r}.$$

Setzen wir dies ein, folgt die Behauptung des Lemmas. □

Lemma 6.11 impliziert, dass jeder Term der zweiten Summe in (6.23) durch

$$C \cdot K_N^{-r} N^{r/2} \sim 2^{-Nr}$$

beschränkt werden kann. Somit ist nach den Vorüberlegungen die Summe dieser Terme von der Ordnung $0(((2^r-1)2^{-r})^N)$, was für $n \to \infty$ gegen 0 geht.

Wir betrachten nun die erste Summe auf der rechten Seite in (6.23), d. h. die Terme mit nicht-degenerierter Kovarianzmatrix. Sei $\alpha \in (0, \frac{1}{2})$ beliebig aber fest und sei $\mathcal{R}^{\alpha}_{l,N} \leq \{-1,+1\}^N$ definiert durch

$$\mathcal{R}^{\alpha}_{l,N} = \{\sigma^1, \ldots, \sigma^l \in \{-1,+1\}^N : |\sum_{i=1}^{N} \sigma_i^m \sigma_i^r| < N^{\alpha+1/2} \quad \forall\, 1 \leq m < r \leq l\}.$$

Durch eine Art große Abweichungsargument bekommt man

$$|\{\{-1,+1\}^N\}^l \backslash \mathcal{R}^{\alpha}_{l,N}| \leq l^2 2^{Nl} \exp(-N^{2\alpha}). \qquad (6.28)$$

Definitionsgemäß gilt für $(\sigma^1, \ldots, \sigma^l) \in \mathcal{R}^{\alpha}_{N,l}$, dass die Elemente $b_{k,m}$ der Kovarianzmatrix für $k \neq m$

$$|b_{k,m}| = |\frac{1}{N}\sum_{i=1}^{N} \sigma_i^k \sigma_i^m| \leq N^{\alpha-1/2}.$$

Daher gilt für alle $\sigma^1, \ldots, \sigma^l \in \mathcal{R}^{\alpha}_{l,N}$, $det\, B_N(\vec{\sigma}) = 1 + O(1)$ und insbesondere $rank(C(\vec{\sigma})) = l$.

Wegen Lemma 6.11 und (6.28) folgt

$$\sum_{\substack{\sigma^1,\ldots,\sigma^l \notin \mathcal{R}^\alpha_{l,N} \\ rank\, C(\vec{\sigma})=l}} \mathbb{P}[\cdot] \leq 2^{nL} e^{-N^{2\alpha}} C N^{3l/2} K_N^{-l} \to 0.$$

Schließlich müssen wir noch zeigen, dass

$$\sum_{\sigma^1,\ldots,\sigma^l \in \mathcal{R}^\alpha_{l,N}} \mathbb{P}[\cdot] = \prod_{j=1} \neg l c_j.$$

Dies folgt aus einer Rechnung wie unter (6.27), wobei im Falle $r = l$ die Ungleichheitszeichen durch Gleichheitszeichen ersetzt werden können. Da die Determinante in diesem Fall von der Größenordnung $1 + 0(1)$ ist, folgt die Behauptung. □

7 Extremale Prozesse

In diesem Abschnitt wollen wir nicht nur das Konvergenzverhalten der richtig skalierten Maxima $A_n(M_n - B_n)$ betrachten, sondern wir wollen allgemeiner das Konvergenzverhalten richtig skalierter stochastischer Prozesse

$$\{Y_n(t) : t > 0\} = \{A_n(M_{[nt]} - B_n) : t > 0\}$$

betrachten. Diese können wir als zufällige Funktionen auffassen. Die Ergebnisse dieses Abschnitts verhalten sich zu denen aus Kapitel 1 etwa wie der Satz von Donsker zum Zentralen Grenzwertsatz. Die resultierenden Grenzprozesse (im Sinne der schwachen Konvergenz) heißen extremale Prozesse und lassen weitere Einsichten in die Struktur von Extrema zu.

Definition 7.1 *Eine Familie $\{Z(t) : t \geq 0\}$ von Zufallsvariablen mit Werten in einem messbaren Raum $(\mathcal{X}, B)$ heißt regulärer Markov-Prozess, wenn für alle monotonen Folgen $\{t_n\} \subseteq \mathbb{R}^+$ $(Z(t_n))_n$ eine Markovkette ist. Der Markov-Prozess heißt homogen, falls für alle $(t_n)_n$ mit Gitterstruktur, d. h. $t_{n+1} - t_n = const. > 0$ für alle $n \in \mathbb{N}$, die Markovkette $(Z(t_n))$ homogen ist. Der Markovprozess $(Z_t)_{t\geq 0}$ heißt reiner Sprungprozess, wenn jede Realisierung $(Z_t(\omega))_t$ eine rechtsseitig stetige Treppenfunktion ist, d. h. wenn*

$$Z(t, \omega) = \sum_{k=1}^{\infty} Z_{k-1}(\omega) \mathbb{1}_{(T_{k-1}(\omega) \leq t < T_k(\omega)} \tag{7.1}$$

für alle $\omega \in \Omega$ gilt. Hierbei sind $(Z_k, T_k)_{k\geq 0}$ geeignete Zufallsvariable mit

$$T_0 = 0 \quad \textit{und} \quad 0 < T_1 < T_2 < \ldots \quad \mathbb{P}\text{-f.s.}$$

sind die (Z_k) paaarweise verschieden, so nennt man die (T_n) (echte) Sprungzeiten.

Definition 7.2 *Ein homogener Markov-Prozess $(E(t))_{t>1}$ heißt F-extremaler Prozess, wenn er ein Sprungprozess ist und*

$$\mathbb{P}(E(1) \leq x) = F(x) \quad \forall\, x \in \mathbb{R}, \tag{7.2}$$

$$\mathbb{P}(E(t) \leq y | E(s) = x) = F^{t-s}(y) \mathbb{1}_{x \leq y} \tag{7.3}$$

für alle $1 \leq s \leq t$ und alle $x, y \in \mathbb{R}$ gilt.

Die Bedeutung eines solchen F-extremalen Prozesses liegt neben noch nachzuweisenden Grenzwerteigenschaften auch darin, dass wir mit ihrer Hilfe genauere Informationen über die Struktur der Maxima gewinnen wollen.

Lemma 7.3 *Unter den Voraussetzungen von Definition 7.2 gilt*

$$\mathbb{P}\left(\bigcap_{i=1}^{k} E(t_i) \leq x_i\right) = F^{t_1}(x_1) \prod_{i=2}^{k} F^{t_i - t_{i-1}}(x_i) \tag{7.4}$$

für alle $1 \leq t_1 < \ldots < t_k$, $x_1 \leq \ldots \leq x_k$, $k \in \mathbb{N}$. Darüber hinaus gilt: Die Folge $E(n)$ ist verteilt wie die Folge des (M_n).

Beweis: (7.4) ergibt sich unmittelbar aus (7.2) und (7.3). So ist etwa

$$\begin{aligned}\mathbb{P}(E(t) \le x) &= \int \mathbb{P}(E(t) \le x | E(1) = z)\mathbb{P}^{E(1)}(dz) \\ &= \int_{-\infty}^{x} F^{t-1}(x)\mathbb{P}^{E(1)}(dz) \\ &= F^{t-1}(x)\mathbb{P}(E(1) \le x) \\ &= F^{t-1}(x)F(x) = F^{z}(x)\end{aligned} \tag{7.5}$$

für alle $t \ge 1$ und $x \in \mathbb{R}$. Ebenso rechnet man

$$\begin{aligned}\mathbb{P}(E(t) \le x, E(s) \le y) &= \int_{-\infty}^{y} \mathbb{P}(E(t) \le x | E(s) = z)\mathbb{P}^{E(s)}(dz) \\ &= \int_{-\infty}^{y} F^{t-s}(x)\mathbb{P}^{E(s)}(dz) \\ &= F^{t-s}(x)F^{s}(y) \quad \text{für} \quad 1 \le s \le t,\ y \le x.\end{aligned} \tag{7.6}$$

Die andere Behaptung ergibt sich, da $(E(n))_n$ eine Markovkette ist:

$$\mathbb{P}(E(n+1) \le y | E(n) = x) = F(y)\mathbb{1}_{\{y \ge x\}} = \mathbb{P}(M_{n+1} \le y | M_n = x) \tag{7.7}$$

für $n \in \mathbb{N}$ und $x, y \in \mathbb{R}$, sowie

$$\mathbb{P}(E(1) \le x) = F(x) = \mathbb{P}(M_1 \le x) \quad \forall\, x \in \mathbb{R}. \tag{7.8}$$

□

Es lässt sich sogar zeigen, dass stets ein F-extremaler Prozess auf $(\Omega, \mathcal{F}, \mathbb{P})$ existiert, der (M_n) fast sicher interpoliert. Dies werden wir aber hier nicht tun.

Das Sprungverhalten F-extremaler Prozesse wird im folgenden Satz beschrieben:

Satz 7.4 *Unter den Voraussetzungen von Lemma 7.3 seien $(T_k)_\varepsilon$ die Sprungzeigen eines F-extremalen Prozesses $(E(t))_{t \ge 1}$ für Maxima. Dann gilt*

a) $(E(T_k))$ bildet eine homogene Markovkette mit Übergangswahrscheinlichkeiten

$$\mathbb{P}(E(T_k) > y | E(T_{k-1}) = x) = \frac{-\log F(y)}{-\log F(x)} \tag{7.9}$$

$k \in \mathbb{N},\ x_L < x \le y$, wobei mit $T_0 = 1$ begonnen wird.

b) $(T_k - T_{k-1})_{1 \le k \le n}$ sind bedingt unabhängig und $E(T_k)_{0 \le k \le n}$ mit

$$\begin{aligned}&\mathbb{P}\left(\bigcap_{k=1}^{n} T_k - T_{k-1} > t_k | E(T_0) = x_0, \ldots, E(T_k) = x_k\right) \\ &= \prod_{k=1}^{n} \mathbb{P}(T_k - T_{k-1} > t_k | E(T_{k-1}) = x_{k-1}) \\ &= \prod_{k=1}^{n} F^{t_k}(x_{k-1})\end{aligned} \tag{7.10}$$

für $n \in \mathbb{N},\ t_1, \ldots, t_n > 0,\ x_0, \ldots, x_n \in \mathbb{R}$.

Beweis:

a) Die Markoveigenschaft von $(E(T_k))$ ist offensichtlich. Man berechnet

$$\begin{aligned}
& \mathbb{P}(E(T_k) \in (x,y]|E(T_{k-1}) = x) \qquad (7.11)\\
= & \lim_{h\downarrow 0} \frac{\mathbb{P}[(E(t+h) \in (x,y]|E(x) = x]}{\mathbb{P}(E(t+h) \in (x,\infty)|E(t) = x)} \\
= & \lim_{h\downarrow 0} \frac{\mathbb{P}(E(1+h) \in (x,y]|E(1) = x)}{\mathbb{P}(E(1+h) \in (x,\infty)|E(0) = x)} \\
= & \lim_{h\downarrow 0} \frac{F^h(y) - F^h(x)}{1 - F^h(x)} \\
= & \lim_{h\downarrow 0} \left[1 - \frac{1 - F^h(y)}{1 - F^h(x)}\right] \\
= & 1 - \frac{\lim_{h\downarrow 0} \frac{1-F^h(y)}{h}}{\lim_{h\downarrow 0} \frac{1-F^h(x)}{h}} \\
= & 1 - \frac{-\log F(y)}{-\log F(x)} \quad \forall\, k \in \mathbb{N},\ t \geq 1,\ x_n < x \leq y,
\end{aligned}$$

woraus sich (7.9) ergibt.

b) Dies folgt aus dem folgenden Satz, da mit den dortigen Bezeichnungen gilt

$$\begin{aligned}
\lambda(x) &= \lim_{h\downarrow 0} \frac{1}{h}\mathbb{P}(E(t+h) \neq x|E(t) = x) \\
&= \lim_{h\downarrow 0} \frac{1}{h}\mathbb{P}(E(1+h) > x|E(1) = x) \\
&= \lim_{h\downarrow 0} \left[\frac{1 - F^h(x)}{h}\right] = -\log F(x)
\end{aligned}$$

für $t \geq 1,\ x > x_L$. Es ist dann

$$\mathbb{P}(T_k - T_{k-1} > t|E(T_{k-1}) = x) = e^{-\lambda(x)t} = F^t(x) \quad \text{für} \quad t > 0,\ x > x_L, \qquad (7.12)$$

woraus (7.10) folgt.

□

Im Beweis haben wir dabei den folgenden Satz verwendet (Teil d für Teil a und Teil a, b und c) für Teil b):

Satz 7.5 *Ein homogener Markovscher Sprungprozess* $(Z_t)_{t\geq 0}$ *mit (echten) Sprungzeiten* $(T_k)_k$ *besitzt genau die folgende Struktur:*

a) $T_k - T_{k-1}$ *und* $Z(T_k)$ *sind bedingt unabhängig unter* $Z/T_{k-1})$ *für alle* $k \in \mathbb{N}$.

b) *Es existiert eine messbare Abbildung* λ,

$$\lambda(\cdot) : (\mathcal{X}, B) \to ((0,\infty), B^1)$$

derart, dass

$$\mathbb{P}(T_k - T_{k-1} > t | Z(T_{k-1}) = z) = e^{-\lambda(z)t} \tag{7.13}$$

für $t \geq 0,\ t \in \mathcal{X},\ k \in \mathbb{N}$, *d. h.* $T_k - T_{k-1}$ *ist bedingt exponentialverteilt mit einem vom letzten "Zustand"* $Z(T_{k-1}) = z$ *abhängigen Parameter* $\lambda(z)$ *und es gilt*

$$\lambda(z) = \lim_{n \downarrow 0} \frac{1}{n} \mathbb{P}(Z(t+h) \neq z | Z(t) = z)$$

unabhängig von $t \geq 0$.

c) $(T_k - T_{k-1} : 1 \leq k \leq n)$ *sind bedingt unabhängig unter* $(Z(T_k))_{0 \leq k \leq n}$ *mit der in (7.13) gegebenen bedingten Verteilung, d. h.*

$$\begin{aligned} & \mathbb{P}\left(\bigcap_{k=1}^{n} T_k - T_{k-1} > t_k | Z(T_0) = z_0, \ldots, Z(T_k) = z_k\right) \quad (7.14) \\ = & \prod_{k=1}^{n} \mathbb{P}(T_k - T_{k-1} > t_k | Z(T_{k-1}) = z_{k-1}) \\ = & \exp\left(\sum_{k=1}^{n} -\lambda(z_{k-1}) t_k\right) \end{aligned}$$

für alle $n \in \mathbb{N},\ t_1, \ldots, t_n > 0,\ z_0, \ldots, z_n \in \mathcal{X}$).

d) *Die Folge* $(Z(T_k))$ *bildet eine homogene Markovkette mit*

$$\mathbb{P}(T(T_k) \in B | Z(T_{k-1}) = z) = \lim_{n \downarrow 0} \frac{\mathbb{P}(Z(t+h) \ni B | Z(t) = z)}{\mathbb{P}(Z(t+h) \neq z | Z(t) = z)}$$

für alle $z \in \mathcal{X},\ z \notin B,\ B \in \mathcal{B}$ *unabhängig von* $t \geq 0$.

Beweis: Der Beweis findet sich im Buch von Breiman "Probability", Abschnitte 15.5 und 15.6. □

Anschaulich sind die Sprungzeiten des F-extremalen Prozesses so etwas wie die Rekordzeiten $(U_n)_n$ und $(X_{u_n})_n$ entspricht dann der Folge der $(E(T_k))_k$. Wir wollen nun studieren, wie sich die extremalen Prozesse im (richtig skalierten) Limes verhalten.

Lemma 7.6 *Es sei* G *vom Typ der Fréchet-, Weibull- oder Gumbelverteilung. Es sei* F *eine Verteilungsfunktion mit* $F \in \mathcal{D}(G)$ *und*

$$\mathbb{P}(A_n(M_n - B_n)) \to G(x)$$

für alle x *und geeignete* $(A_n)_n$ *und* $(B_n)_n$, $A_n > 0$. *Dann konvergieren sämtliche endlichdimensionalen Randverteilungen des Prozesses* $(Y_n(t))_k$ *gegen diejenigen eines* G*-extremalen Prozesses für Maxima (hierbei sei wieder*

$$Y_n(t) := A_n(M_{[nt]} - B_n) \tag{7.15}$$

und der Limes ist der Limes $n \to \infty$).

Beweis: Sei $k \in \mathbb{N}$ und $0 = t_0,\ 1 \leq t_1 < t_2 \ldots < t_k,\ x_1 \leq \ldots \leq x_k$. Nach Lemma 7.3 gilt

$$\begin{aligned} & \mathbb{P}\left(\bigcap_{i=1}^k Y_n(t_i) \leq x_i\right) = \mathbb{P}\left(\bigcap_{i=1}^k M_{[nt_i]} \leq \frac{x_i}{A_n} + B_n\right) \\ = \quad & \prod_{i=n}^k F^{[nt_i]-[nt_{i-1}]}\left(\frac{x_i}{A_n} + B_n\right) \\ = \quad & \prod_{i=1}^k (F^n(\frac{x_i}{A_n} + B_n))^{\frac{[nt_i]-[nt_{i-1}]}{n}} \\ \underset{n\to\infty}{\longrightarrow} \quad & \prod_{i=1}^k G^{t_i - t_{i-1}}(x_i). \end{aligned} \tag{7.16}$$

□

Lemma 7.6 ist ein erster Indikator, dass auch die Prozesse $(Y_n(t))_t$ in einem geeigneten Funktionenraum konvergieren. Um dies genauer zu studieren, müssen wir allerdings die geeigneten Prozesse und Techniken bereitstellen. Die interessanten Prozesse entstehen dabei, wenn wir auf die Punkte nahe am Extremum schauen. Diese haben die Struktur eines Punktprozesses. Punktprozesse sind gemacht, um die probabilistische Struktur von zufälligen Puntkmengen in $\mathbb{R}^d$ zu studieren. Eine technisch bequeme Art, solche Punktwolken zu studieren, ist es, den Punkten eine Einheitsmasse zuzuerkennen und das zugehörige Punktmaß zu studieren. Sei zunächst $x \in \mathbb{R}^d$ ein Punkt. Wir betrachten das zugehörige Diracmaß

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

Definition 7.7 *Ein Punktmaß ist nun ein Maß μ auf $\mathbb{R}$, so dass es eine abzählbare Punktmenge $(x_n)_n \subseteq \mathbb{R}^d$ gibt und wir*

$$\mu = \sum_{i=1}^{\infty} \delta_{x_i} \tag{7.17}$$

schreiben können. Darüber hinaus fordern wir, dass μ lokal endlich ist, d. h.

$$\mu(K) < +\infty \quad \text{für alle} \quad K \subseteq \mathbb{R}^d \quad \text{kompakt.}$$

Wir haben ausdrücklich dabei nicht ausgeschlossen, dass einige der x_i gleich sind.

Die Menge

$$S_\mu = \{x \in \mathbb{R}^d : \mu(x) \neq 0\}$$

heißt der Träger eines Punktmaßes μ. Ein Punktmaß, so dass $\mu(x) \leq 1$ für alle $x \in \mathbb{R}^d$ gilt, heißt einfach, hier sind offenbar die (x_i) verschieden.

Definition 7.8 *Wir bezeichnen mit $\mathcal{M}_P(\mathbb{R}^d)$ die Menge aller Punktmaße auf $\mathbb{R}^d$. Wir statten diese Menge mit der σ-Algebra $\mathcal{S}_P(\mathbb{R}^d)$ aus. Dies ist einerseits die kleinste σ-Algebra, die alle Mengen der Form*

$$\{\mu \in \mathcal{M}_P(\mathbb{R}^d) : \mu(F) \in B)$$

für $F \in \mathcal{B}(\mathbb{R}^d)$ und $B \in \mathcal{B}([0,\infty)]$ enthält. Andererseits ist es auch die kleinste σ-Algebra, die von

$$\mu \mapsto \mu(F), \ F \in \mathcal{B}(\mathbb{R}^d)$$

erzeugt wird.

Definition 7.9 *Ein Punktprozess N ist eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit Werten in $\mathcal{M}_p(\mathbb{R}^d)$ (messbar bzgl. $\mathcal{S}_P(\mathbb{R}^d)$).*

Dies ist in der Tat weniger verrückt alls es aussieht.

Satz 7.10 *N ist ein Punkt-Prozess genau dann, wenn*

$$N(\cdot, F) : \omega \mapsto N(\omega, F)$$

eine messbare Abbildung von $(\Omega, \mathcal{F})$ nach $(\mathbb{R}\mathcal{B}([0,\infty)))$ ist für jede Borel-Menge F, d. h. falls $N(F)$ eine reellwertige Zufallsvariable ist.

Beweis: Die Bedingung ist notwendig, denn falls

$$\omega \mapsto N(\omega, \cdot)$$

messbar ist mit Werten in $\mathcal{M}_P(\mathbb{R}^d)$ und

$$\mu \mapsto \mu(F)$$

messbar ist von $\mathcal{M}'_P(\mathbb{R}^d)$ nach $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, so ist auch die Komposition dieser Abbildungen messbar. Nun beweisen wir, dass die Bedingung auch notwendig ist. Sei

$$\mathcal{G} := \{A \in \mathcal{S}_P(\mathbb{R}^d) : N^{-1}A \in \mathcal{F}\}.$$

Als Urbild von σ-Algebra-Elementen ist $\mathcal{G}$ eine σ-Algebra und ist messbar als Abbildung

$$N : (\Omega, \mathcal{F}) \to (M_P(\mathbb{R}^d), \mathcal{G}).$$

Aber $\mathcal{G}$ enthält alle Mengen der Form

$$\{\mu \in M_P(\mathbb{R}^d) : \mu(F) \in \mathcal{B}\},$$

$B \in \mathcal{B}([0,\infty))$, da

$$N^{-1}\{\mu \in M_p(\mathbb{R}^d) : \mu(F) \in \mathcal{B}\} = \{\omega : N(\omega, F) \in \mathcal{B}\}$$

gilt, da $N(\cdot, F)$ messbar ist. Daher gilt $\mathcal{G} \supseteq \mathcal{S}_P(\mathbb{R}^d) \in \mathcal{F}$ und N als Abbildung von der kleinen σ-Algebra ist zwangsläufig messbar. □

Da wir unsere Extremwerte als Punktprozesse auffassen wollen und untersuchen wollen, was für $N \to \infty$ geschieht, müssen wir Konvergenzkriterien für Punktprozesse herleiten.

Definition 7.11 *Dazu sei $\mathcal{B}$ die Borelsche σ-Algebra über einem metrischen Raum E. Dann heißt $\mathcal{T} \subseteq \mathcal{B}$ ein π-*System*, falls $\mathcal{T}$ unter endlichen Durchschnitten abgeschlossen ist. $\mathcal{G} \subseteq \mathcal{B}$ heißt λ-*System *oder eine σ-*additive Klasse*, falls*

(i) $E \in \mathcal{G}$;

(ii) $A, B \in \mathcal{G}$ *und* $A \supseteq B \Rightarrow A \backslash B \in \mathcal{G}$.

(iii) *Ist* $(A_n)_n$ *eine Folge in* $\mathcal{G}$ *mit* $A_n \subseteq A_{n+1} \Rightarrow \lim_{n\to\infty} A_n \in \mathcal{G}$.

Die folgende Beobachtung heißt Satz von Dynkin, den Beweis findet man im Buch von Durrett.

Satz 7.12 *(Dynkin): Falls $\mathcal{T}$ ein π-System ist und $\mathcal{G}$ ein λ-System ist, dann folgt aus $\mathcal{G} \supseteq \mathcal{T}$, dass $\mathcal{G} \subseteq \sigma(\mathcal{T})$.*

Dies sieht zunächst wie eine eher esoterische Beobachtung aus. Der Satz hat aber wichtige praktische Konsequenzen. Die wichtigste ist die, dass zwei Wahrscheinlichkeitsmaße, die auf einem π-System übereinstimmen, der die zugrunde liegende σ-Algebra erzeugt, auf der σ-Algebra übereinstimmen, da das System, auf dem zwei Wahrscheinlichkeitsmaße übereinstimmen, immer ein λ-System ist.

In unserer Anwendung können wir die Kriterien, die N als Punktprozess definieren, abschwächen. Insbesondere können wir die F, für die die $N(\cdot, F)$ messbar sein müssen, als Rechtecke wählen.

Satz 7.13 *Sei $\mathcal{T}$ eine Teilmenge relativ kompakter Mengen in $\mathcal{B}$, die den folgenden Bedingungen genügen:*

(i) $\mathcal{T}$ *ist ein* π*-System;*

(ii) $\sigma(\mathcal{T}) = \mathcal{B}$.

(iii) *Entweder es gibt eine Folge* $E_n \in \mathcal{T}$ *mit* $E_n \uparrow E$ *oder es gibt eine Zerlegung* $(E_n)_n$ *von* E *mit* $\bigcup E_n = E,\ E_n \in \mathcal{T}$.

Dann ist N ein Punktprozess auf $(\Omega, \mathcal{F})$ in $(E, \mathcal{B})$ genau dann, wenn die Abbildungen

$$N(\cdot, I) : \omega \to N(\omega, I)$$

für alle $I \in \mathcal{T}$ messbar sind.

Beweis: Übung. □

Übung 7.14 *Ist das System aller endlichen Vereinigungen halboffener Rechtecke in* $\mathbb{R}^d$ *ein* π*-System?*

Korollar 7.15 $\mathcal{T}$ *erfülle die Voraussetzung von Satz 7.13. Ferner sei*

$$\mathcal{G} := \{\{\mu : \mu(I_j) = n_j,\ 1 \leq j \leq k\},\ k \in \mathbb{N},\ I_j \in \mathcal{T},\ n_j > 0\}.$$

Dann gilt

$$\sigma(\mathcal{G}) = \mathcal{S}_P(\mathbb{R}^d)$$

und $\mathcal{G}$ *ist ein* π*-System.*

Beweis: Übung. □

Nun zeigen wir noch, dass die Verteilung $\mathbb{P}^N$ eines Punktprozesses N durch die Verteilungen geeigneter $N(F_n),\ F_n \in \mathcal{B}(\mathbb{R}^d)$ bestimmt ist.

Satz 7.16 *Es sei* N *ein Punkt-Prozess auf* $(\mathbb{R}^d, \mathcal{B}^d)$ *und* $\mathcal{T}$ *sei wie in Satz 7.13. Wir definieren die Massefunktionen*

$$\mathbb{P}_{I_1,\ldots,I_k}(n_1, \ldots, n_k) := \mathbb{P}[N(I_j) = n_j,\ \forall\, 1 \leq j \leq k]. \tag{7.18}$$

Dann ist $\mathbb{P}^N$ *eindeutig bestimmt durch die Familie*

$$\{\mathbb{P}_{I_1,\ldots,I_k},\ k \in \mathbb{N},\ I_j \in \mathcal{T}\}.$$

Beweis: Übung □

Schließlich müssen wir noch ein paar Definitionen einführen.

Definition 7.17 *Sind* N_1 *und* N_2 *zwei Punktprozesse (mit Werten in demselben Raum* (E, B)*), so sagen wir, dass sie* unabhängig *sind, falls für alle* $F_j \in B$ *und* $G_i \in \mathcal{B}$ *die Vektoren*

$$(N_1(F_j),\ 1 \leq j \leq k) \quad \textit{und} \quad (N_2(G_j),\ 1 \leq j \leq l)$$

unabhängige Zufallsvektoren sind.

Definition 7.18 *Das* Intensitätsmaß λ *eines Punktprozesses* N *ist definiert als*

$$\lambda(F) := \mathbb{E}N(F) = \int_{\mathcal{M}_P(\mathbb{R}^d)} \mu(F)\mathbb{P}^N(d\mu)$$

für $F \in B$.

Schließlich definieren wir noch für messbare Funktionen $f : \mathbb{R}^d \to \mathbb{R}$

$$N(\omega, f) = \int_{\mathbb{R}^d} f(x) N(\omega, dx).$$

Somit ist $N(\cdot, f)$ eine Zufallsvariable und

$$\mathbb{E}N(f) = \lambda(f) = \int_{\mathbb{R}^d} f(x)\lambda(dx).$$

Wir wollen uns nun einem der wichtigsten technischen Hilfsmittel bei der Betrachtung von Punktprozessen (und nicht nur da) zuwenden, den Laplace-Funktionalen.

Definition 7.19 *Sei Q ein Wahrscheinlichkeitsmaß auf $(\mathcal{M}_P, \mathcal{S}_P)$. Die* Laplace-Transformierte *ist eine Abbildung*

$$\begin{aligned} \psi : \mathcal{B}(\mathbb{R}^d, \mathbb{R}^+) &\to \mathbb{R}^+ \\ f &\mapsto \psi(f) = \int_{\mathcal{M}_P} \exp\left(- \int_{\mathbb{R}^d} f(x)\mu(dx) \right) Q(d\mu). \end{aligned}$$

Ist N ein Punktprozess, ist daher sein Laplacefunktional *definiert als*

$$\begin{aligned} \psi_N(f) : &= Ee^{-N(f)} \\ &= \int e^{-N(\omega,f)} \mathbb{P}(d\omega) \\ &= \int_{\mathcal{M}_P} \exp\left(- \int_{\mathbb{R}^d} f(x)\mu(dx) \right) P^N(d\mu). \end{aligned}$$

Dabei ist $\mathcal{B}(\mathbb{R}^d, \mathbb{R})$ der Raum der Borel-messbaren Funktionen von $\mathbb{R}^d$ nach $\mathbb{R}$.

Satz 7.20 *Ein Punktprozess N ist durch sein Laplace-Funktional ψ_N eindeutig bestimmt.*

Beweis: Für $k \geq 1,\ F_1, \ldots, F_k \in \mathcal{B}^d$ und $c_1, \ldots, c_k \geq 0$ sei

$$f := \sum_{i=1}^{k} c_i \mathbb{1}_{F_i}(x).$$

Dann ist

$$N(\omega, f) = \sum_{i=1}^{k} c_i N(\omega, F_i)$$

und

$$\psi_N(f) = \mathbb{E}\ \exp(-\sum_{i=1}^{k} c_i N(F_i)).$$

Dies ist nichts anderes als die Laplace-Transformierte des Vektors $(N(F_i),\ 1 \leq i \leq k)$. Diese bestimmt eindeutig seine Verteilung. Daher folgt die Behauptung aus Satz 7.17. □

Die wichtigsten Punktprozesse für unsere Zwecke sind die Poissonschen Punktprozesse.

Definition 7.21 *Es sei λ ein lokal endliches, positives Maß auf $\mathbb{R}^d$. Dann heißt ein Punktprozess N ein Poissonscher Punktprozess mit Intensitätsmaß λ (PPP(λ) abkürzend), falls*

(i) Für jedes $F \in \mathcal{B}^d$ und $k \in \mathbb{N}$

$$\mathbb{P}(N(F) = k) = \begin{cases} \frac{e^{-\lambda(F)}\lambda(F)^k}{k!}, & \text{falls } \lambda(F) < +\infty \\ 0, & \text{falls } \lambda(F) = +\infty \end{cases}$$

(ii) Falls $F, G \in \mathcal{B}^d$ disjunkte Teilmengen sind, so sind $N(F)$ und $N(G)$ unabhängig.

Der nächste Satz zeigt die Existenz Poissonscher Punktprozesse zu jedem gegebenen Intensitätsmaß. Der Beweis konstruiert solche Prozese explizit.

Satz 7.22 *(i) PPP(λ) existiert und seine Verteilung ist durch die Definition eindeutig bestimmt.*

(ii) Das Laplace-Funktional des PPP(λ) ist für $f \geq 0$ durch

$$\psi_N(f) = \exp(-\int_{\mathbb{R}^d} (1 - e^{-f(x)})\lambda(dx)) \tag{7.19}$$

gegeben.

Beweis: Da wir wissen, dass ein Laplace-Funktional einen Punktprozess eindeutig bestimmt, werden wir, um zu zeigen, dass die Bedingungen an PPP(λ) diesen eindeutig bestimmen, nachweisen, dass diese die Form (7.19) des Laplace-Funktionals eindeutig bestimmen. Sei also N ein PPP(λ). Sei $f = c\mathbb{1}_F$. Dann gilt

$$\begin{aligned} \psi_N(f) &= \mathbb{E}\exp(-N(f)) \\ &= \mathbb{E}\exp(-cN(F)) \\ &= \sum_{k=0}^{\infty} e^{-ck} e^{-\lambda(F)} \frac{(\lambda(F))^k}{k!} \\ &= e^{(e^{-c}-1)\lambda(F)} \\ &= \exp(-\int (1 - e^{-f(x)})\lambda(dx)), \end{aligned}$$

also (7.19). Sind nun die F_i disjunkte Teilmengen des $\mathbb{R}^d$ (Borel-messbar) und $c_i \in \mathbb{R}$, $i = 1, \ldots, k$, und schließlich

$$f := \sum_{i=1}^{k} c_i \mathbb{1}_{F_i}.$$

Dann folgt direkt

$$\begin{aligned} \psi_N(f) &= \mathbb{E}_k(\exp(-\sum_{i=1}^{k} c_i N(f_i))) \\ &= \prod_{i=1}^{k} \mathbb{E}\exp(-c_i N(F_i)) \end{aligned}$$

aufgrund der Unabhängigkeitsannahme (ii) in Definition 7.21. Hieraus folgt wie oben die Form (7.19) des Laplace-Funktionals. Für allgemeines f wählen wir nun wie üblich eine Folge von Treppenfunktionen (f_n) mit $f_n \uparrow f$. Der Satz von der monotonen Konvergenz impliziert dann auch $N(f) \geq 0$ und somit $e^{-N(f)} \leq 1$. Wir bekommen daher aus dem Satz über dominierte Konvergenz

$$\psi_N(f_n) = Ee^{-N(f_n)} \to \mathbb{E}e^{-N(f)} = \psi_N(f).$$

Da aber auch

$$1 - e^{-f_n(x)} \uparrow 1 - e^{-f(x)}$$

gilt, erhalten wir abermals mit monotoner Konvergenz:

$$\psi_N(f_n) = \exp(\int (1 - e^{-f_n(x)})\lambda(dx)) \uparrow \exp(\int (1 - e^{-f(x)})\lambda(dx)).$$

Andererseits folgt aus einer Form (7.19) des Laplace-Funktionals sofort, dass die Bedingungen aus Definition 7.21 erfüllt sind, indem man geeignete f einsetzt (das ist eine Übung).

Schließlich wollen wir noch einen PPP(λ) konstruieren. Zunächst beginnen wir mit dem Fall eines endlichen Referenzmaßes λ, also eines λ mit $\lambda(\mathbb{R}^d) < +\infty$. Nun konstruieren wir

a) Eine Poissonsche Zufallsgröße τ mit Erwartungswert $\lambda(\mathbb{R}^d)$.

b) Eine Familie von iid Zufallsvariablen mit gemeinsamer Verteilung $\lambda = \frac{\lambda(\cdot)}{\lambda(\mathbb{R}^d)}$, die unabhängig ist von τ.

Dann definieren wir

$$N^* := \sum_{i=1}^{\tau} \delta_{x_i}.$$

Man rechnet schnell nach, dass N^* ein PPP(λ) ist. Für den Fall $\lambda(\mathbb{R}^d) = +\infty$ erinnern wir, dass λ lokal-endlich sein sollte. Daher existiert eine Folge disjunkter, messbarer Mengen (F_n), so dass $\mathbb{R}^d = \bigcup_{n=1}^{\infty} F_n$ und $\lambda(F_n) < +\infty$. Nun konstruieren wir nach obigen Schema einen PPP(λ_n) auf F_n (wobei $\lambda_n = \lambda(F_n)$) und nennen diesen N_n. Setzt man

$$N = \sum_{n=1}^{\infty} N_n,$$

so ist N ein PPP(λ). □

Um die Konvergenz der Extremalprozese zu diskutieren, müssen wir beschreiben, was wir mit der Konvergenz von Punktprozessen meinen. Da es sich bei einem Punktprozess um eine Wahrscheinlichkeitsverteilung auf der Menge der Punktmaße auf $\mathbb{R}^d$ handelt, ist es natürlich, eine Form der schwachen Konvergenz als geeigneten Konvergenzbegriff zu vermuten. Wir würden dann sagen, dass eine Folge von Punktprozessen $(N_n)_n$ gegen einen Punktprozess N konvergiert, falls für alle stetigen beschränkten Funktionen f auf der Menge der Punktprozesse

$$\mathbb{E}\, f(N_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}\, f(N)$$

gilt. Um dieser Gleichung eine Bedeutung zu geben, müssen wir allerdings wissen, was eine stetige Funktion auf der Menge der Punktprozesse ist, d. h. wir müssen diese Menge topologisieren. Wir wählen hier die Topologie der vagen Konvergenz. Dazu statten wir $\mathbb{R}^d$ mit der üblichen Euklidischen Norm aus und erinnen uns, dass $\mathbb{R}^d$ damit ein vollständiger, separabler, metrischer Raum ist.

Definition 7.23

- *$C_0(\mathbb{R}^d)$ sei die Menge aller stetigen Funktionen mit kompaktem Träger auf $\mathbb{R}^d$.*
- *$C_0^+(\mathbb{R}^d)$ sei die Menge aller Funktionen $f \in C_0(\mathbb{R}^d)$ mit $f \geq 0$.*
- *$\mathcal{M}^+(\mathbb{R}^d)$ sei die Menge aller lokal endlichen, positiven Maße auf $\mathbb{R}^d$.*
- *$\mathcal{S}^+(\mathbb{R}^d)$ sei die σ-Algebra auf $\mathcal{M}^+(\mathbb{R}^d)$, die durch die Abbildungen*

  $$m \mapsto m(f),\ f \in C_0^+(\mathbb{R}^d)$$

  erzeugt wird.

Definition 7.24 *Wir sagen, dass eine Folge (μ_n) in $\mathcal{M}^+(\mathbb{R}^d)$ gegen ein $\mu \in \mathcal{M}^+(\mathbb{R}^d)$* vag *konvergiert, falls für alle $f \in C_0^+(\mathbb{R}^d)$ gilt*

$$\mu_n(f) \to \mu(f)$$

für $n \to \infty$.

Bemerkung 7.25 *Eine Subbasis dieser Topologie der vagen Konvergenz wird durch alle Mengen der Form*

$$B_{f_1,\ldots,f_k},(\mu) := \{\nu \in \mathcal{M}^+(\mathbb{R}^d) : |\nu(f_i) - \mu(f_i)| < \varepsilon \quad \forall\, i = 1,\ldots,k\}$$

gegeben.

Bemerkung 7.26 *Gegeben die vage Topologie auf $\mathcal{M}^+(\mathbb{R}^d)$, können wir natürlich ihre Borelsche σ-Algbra $\mathcal{B}(\mathcal{M}^+(\mathbb{R}^d))$ betrachten. Es stellt sich heraus, dass*

$$\mathcal{B}(\mathcal{M}^+(\mathbb{R}^d)) = S^+(\mathbb{R}^d)$$

gilt.

Es gelten die folgenden Eigenschaften für vage Konvergenz (Details und Beweise findet man z. B. im Maßtheoriebuch von Parthasaraty).

Satz 7.27 *Sei $(\mu_n)_n$ eine Folge in $\mathcal{M}^+(\mathbb{R}^d)$ und $\mu \in \mathcal{M}^+(\mathbb{R}^d)$. Dann sind äquivalent:*

(i) μ_n *konvergiert vage gegen* μ *(in Zeichen:* $\mu_n \xrightarrow{v} \mu$*).*

(ii) $\mu_n(B) \to \mu(B)$ *für alle relativ kompakten* $B \subseteq \mathbb{R}^d$ *mit* $\mu(\partial B) = 0$.

(iii) $\limsup_{n\to\infty} \mu_n(K) \leq \mu(K)$ *und* $\liminf \mu_n(G) \geq \mu(G)$ *für alle kompakten Mengen* K *und alle relativ kompakten offenen Mengen* G.

Der nächste Satz stellt einen Zusammenhang zwischen vag konvergenten Folgen von Punktmaßen und der Konvergenz der zugehörigen Punktfolgen her.

Satz 7.28 *Sei μ_n, $\mu \in \mathcal{M}^P(\mathbb{R}^d)$ für alle $n \in \mathbb{N}$. Es gelte*

$$\mu_n \xrightarrow{v} \mu.$$

Es sei K kompakt mit $\mu(\partial K) = 0$. Dann gibt es eine Nummerung der Punkte von μ_n, so dass für alle n groß genug (größer als ein $n(K)$)

$$\mu_n(\cdot \cap K) = \sum_{i=1}^{p} \delta_{x_i^{(n)}} \quad \text{und} \quad \mu(\cdot \cap K) = \sum_{i=1}^{p} \delta_{x_i}$$

gilt und

$$(x_1^{(n)}, \dots, x_p^{(n)}) \to (x_1, \dots, x_p).$$

Außerdem gilt:

Satz 7.29 *$\mathcal{M}_p(\mathbb{R}^d) \subseteq \mathcal{M}^+(\mathbb{R}^d)$ ist vag abgeschlossen.*

Schließlich benötigen wir noch die Metrisierbarkeit der vagen Topologie:

Satz 7.30 *Die Topologie der vagen Konvergenz ist metrisierbar. Mit dieser Metrik wird $\mathcal{M}^+(\mathbb{R}^d)$ zu einem vollständigen, separablen metrischen Raum.*

Beweis: Die Idee wird sein, eine abzählbare Menge von Funktionen $(k_n)_n$ aus C_0^+ zu finden, die die vage Konvergenz determinieren, d. h. so dass

$$\mu_n \xrightarrow{v} \mu \Leftrightarrow \mu_n(k_i) \to \mu(k_i) \quad \forall\, i = 1, 2, \dots$$

gilt. Sei dazu $(G_i)_i$ eine Familie relativ kompakter Mengen, so dass diese Familie abgeschlossen unter endlichen Vereinigungen und Durchschnitten ist und das $\mathcal{B}^d$ erzeugt. Der Satz von Urisohn sagt nun, dass wir die Indikatoren der G_i durch stetige Funktionen approximieren können, da G relativ kompakt ist, stammen diese Mengen aus $C_0^+(\mathbb{R}^d)$. Also gibt es $(f_{i,n})_n \subseteq C_0^+(\mathbb{R}^d)$ und $(g_{i,n})_n \subseteq C_0^+(\mathbb{R}^d)$, so dass

$$f_{i,n} \uparrow \mathbb{1}_{G_i} \quad \text{und} \quad g_{i,n} \downarrow \mathbb{1}_{G_i},$$

wobei bei beiden Limiten $n \to \infty$ genommen wird.

Unsere Menge (k_n) besteht nun aus allen $(k_{i,n})_{i,n}$ und $(g_{i,n})_{i,n}$. Nun ist $\mu \in \mathcal{M}^+(\mathbb{R}^d)$ durch seine Werte auf den $(k_n)_n$ festgelegt, da die (G_i) die vage Topologie erzeugen. Genauer gilt einerseits

$$\begin{aligned} & \mu(f_{i,n}) \uparrow \mu(G_i) \quad \text{für} \quad n \to \infty \\ \text{und} \qquad & \mu(g_{i,n}) \downarrow \mu(G_i) \quad \text{für} \quad n \to \infty, \end{aligned}$$

andererseits sind die (G_i) ein π-System, das $\mathcal{B}(\mathbb{R}^d)$ erzeugt, also determinieren die $\mu(G_i)$ das Maß μ. Somit

$$\mu_n \xrightarrow{v} \mu \Longleftrightarrow \mu_n(h_i) \to c_i = \mu(h_i)\ \forall\, i.$$

Wir setzen als Metrik:

$$d(\mu, \nu) := \sum_{i=1}^{\infty} 2^{-i}(1 - e^{-|\mu(h_i) - \nu(h_i)|}).$$

Es ist einfach zu sehen, dass

$$d(\mu_n, \mu) \to 0 \Rightarrow \mu_n(h_i) \to \mu(h_i)\ \forall\, i.$$

Andererseits impliziert

$$\mu_n(h_i) \to \mu(h_i) \quad \forall\, i = 1, 2, \ldots$$

das folgende: Wähle N so, dass

$$\sum_{i=1}^{N} 2^{-i} \geq 1 - \varepsilon$$

für $\varepsilon > 0$. Außerdem gibt es dann ein n_0, so dass

$$1 - \exp(|\mu_n(h_i) - \mu(h_i)|) \leq \varepsilon \quad \text{für alle} \quad n \geq n_\varepsilon$$

und alle $i = 1, \ldots, N$. Also

$$d(\mu_n, \mu) \leq \sum_{i=1}^{N} 2^{-i}(1 - e^{-|\mu_n(h_i) - \mu(h_i)|}) + \varepsilon \leq \varepsilon + \varepsilon = 2\varepsilon.$$

Also konvergiert auch $d(\mu_n, \mu) \to 0$. Man rechnet auch nach, dass mit dieser Metrik $(\mathcal{M}^+(\mathbb{R}^d), d)$ vollständig und separabel ist. Dies ist eine Übung. □

Da wir nun $\mathcal{M}^+(\mathbb{R}^d)$ topologisiert und sogar metrisiert haben, können wir nun die schwache Konvergenz von Maßen auf $\mathcal{M}^+(\mathbb{R}^d)$ erklären.

Ein hilfreiches Werkzeug hierbei ist die Verallgemeinerung des Satzes von Skohorod, den wir schon aus Kapitel 1 auf $\mathbb{R}$ kennen.

Satz 7.31 *Sei (X_n) eine Folge von Zufallsvariablen auf einem vollständigen, metrischen Raum. Dann konvergiert X_n schwach gegen eine Zufallsvariable X_0, dann und nur dann, wenn es eine Familie (X_n^*) von Zufallsvariablen auf $([0,1], B_{[0,1]}, \lambda\!\!\lambda'_{[0,1]})$ gibt, so dass* *(i)* $X_n \stackrel{\mathcal{D}}{=} X_n^* \quad \forall\, n.$

(ii) $X_n^* \to X_0^* \quad \lambda\!\!\lambda'$*-f.s.*

Beweis: Siehe Billingsley "Weak Convergence of Probability Measure". □

Ein sehr praktisches Resultat, um schwache Konvergenz zu überprüfen, ist Kallenbergs Satz.

Satz 7.32 *Es sei ξ ein einfacher Punktprozess auf einem metrischen Raum E und $\mathcal{T}$ ein π-System von relativ-kompakten offenen Mengen. Weiter gelte für $I \in \mathcal{T}$*

$$\mathbb{P}[\xi(\partial I) = 0] = 1.$$

Wenn $(\xi_n)_n$ eine Folge von Punktprozessen auf E ist und für alle $I \in \mathcal{T}$

$$\lim_{n\to\infty} \mathbb{P}[\xi_n(I) = 0] = \mathbb{P}(\xi(I) = 0)$$

und

$$\lim_{n\to\infty} \mathbb{E}\,\xi_n(I) = \mathbb{E}\,\xi(I) < \infty \tag{7.20}$$

gilt, dann folgt

$$\xi_n \xrightarrow{\omega} \xi.$$

Bemerkung 7.33 *Ist $E = \mathbb{R}^d$, so kann man das System der halboffenen Rechtecke als π-System wählen.*

Beweis von Satz 7.32. Die Idee geht auf den Fall $E = \mathbb{R}^d$ zurück. Hier ist ein Punktprozess eine (geordnete) Folge von Punkten. Kennt man die Gebiete, in die keine Punkte fallen, so kennt man die Abstände zwischen den Punkten und somit den Punktprozess. Wir schreiben zu diesem Zweck ein Punktmaß μ als

$$\mu = \sum_{y\in S} c_y \delta_y,$$

wobei S der Träger von μ ist, also alle Punkte, die nicht mit 0 gemessen werden und $c_y \in \mathbb{N}$. Mit μ assoziieren wir das einfache Punktmaß

$$T^*\mu = \mu^* = \sum_{y\in S} \delta_y.$$

Die Abbildung T^* ist offensichtlich messbar und falls ξ_1 und ξ_2 Punktmaße sind, so dass für alle $I \in \mathcal{T}$ gilt

$$\mathbb{P}[\xi_1(I) = 0] = \mathbb{P}[\xi_2(I) = 0],$$

so folgt

$$\xi_1^* \stackrel{\mathcal{D}}{=} \xi_2^*. \tag{7.21}$$

Ein kleiner Beweis hierfür geht folgendermaßen: Sei

$$C = \{\{\mu \in \mathcal{M}^p(E) : \mu(I) = 0\}, I \in \mathcal{T}\}.$$

C ist ein π-System (Übung). Da nach Annahme die Verteilungen $\mathbb{P}_i$ der Punktprozesse ξ_i auf diesem π-System übereinstimmen, stimmen sie auch auf $\sigma(C)$ überein. Wir zeigen nun, dass

$$T^* : (\mathcal{M}^p, \sigma(C)) \to (\mathcal{M}_p, \mathcal{S}_P)$$

messbar ist. Dies ist wahr, falls für jedes $I \in \mathcal{T}$ die Abbildung

$$T_1^* : \mu \to \mu^*(I)$$

messbar von $(\mathcal{M}_p, \sigma(C))$ nach $\{0, 1, 2, \ldots\}$ ist. Nun führen wir eine Familie endlicher Überdeckungen der (relativ kompakten) Menge I ein und nennen diese $((A_{n,j})_j)_{n=1}^{k_n}$ und fordern, dass dann $(A_{n,j}) < \frac{1}{n}$ für alle j und alle n gilt. Darüber hinaus wählen wir die $(A_{n,j})_{j,n}$ so, dass

$$A_{n+1,j} \subseteq A_{n,i} \quad \text{für jedes } j \text{ und ein } i.$$

Dann gilt:

$$T^*\mu = \mu^*(I) = \lim_{n\to\infty} \sum_{j=1}^{k_n} \mu(A_{n,j}) \wedge 1,$$

denn wenn n groß genut ist, enthält kein $A_{n,j}$ mehr als einen Punkt von μ. Sei

$$T_2^*\mu := (\mu(A_{n,j} \wedge 1).$$

Offenbar gilt:

$$(T_2^*)^{-1}(\{0\}) = \{\mu : \mu(A_{n,j}) = 0\} \in \sigma(C).$$

Also ist T_2^* messbar, somit auch T_1^* als monotoner Limes einer endlichen Summe messbarer Abbildungen. Nun folgt aber

$$\mathbb{P}[\xi_1^* \in B] = \mathbb{P}[T^*\xi_1 \in B] = \mathbb{P}[\xi_1 \in (T^*)^{-1}(B)] = \mathbb{P}_1[(T^*)^{-1}(B)].$$

Da wir aber nun gezeigt haben, dass $(T^*)^{-1}(B) \in \sigma(C)$ gilt, folgt gemäß unserer Annahmen

$$\mathbb{P}_1[(T^*)^{-1}(B)] = \mathbb{P}_2[(T^*)^{-1}(B)],$$

somit ist dies dasselbe wie $\mathbb{P}[\xi_2^* \in B]$, was (7.21) zeigt.

Nun impliziert Bedingung (7.20) schon die Straffheit der Folge $(\xi_n)_n$ (Übung). Somit hat nach dem Satz von Prohorov jede Teilfolge $(\xi_{n'})_{n'}$ eine schwach konvergente Teilfolge $(\xi_{n''})_{n''}$. Deren Limes sei η. Da $\mathcal{M}_p$ kompakt ist, ist η ein Punktprozess. Wir nehmen zunächst an, dass

a) η ist einfach;

b) für alle relativ kompakten A gilt

$$\mathbb{P}[\xi(\partial A) = 0] = 1 \Rightarrow \mathbb{P}[\eta(\partial A) = 0] = 1. \tag{7.22}$$

Dann ist die Abbildung $\mu \mapsto \mu(I)$ η-f.s. stetig und daher gilt, falls $\xi_{n''} \to \eta$, auch

$$\mathbb{P}[\xi_{n''}(I) = 0] \to \mathbb{P}(\eta(I) = 0).$$

Somit folgt aus den vorhergehenden Beobachtungen und der Tatsache, dass η und ξ einfach sind, $\xi = \eta$. Also müssen wir noch zeigen, dass η einfach ist und (7.22) gilt. Für (7.22) zeigen wir, dass für jede kompakte Menge K gilt

$$\mathbb{P}[\eta(K) = 0] \geq \mathbb{P}[\xi(K) = 0]. \tag{7.23}$$

Dies wenden wir dann auf $\bar{A}$ an und bekommen somit auch Abschätzungen für ∂A. Für ein solches kompaktes K gibt es Funktionen $f_j \in C_0^+(\mathbb{R}^d)$ und kompakte Mengen K_j (im wesentlichen wieder aufgrund von Urysohns Satz), so dass

$$\mathbb{1}_K \leq f_j \leq \mathbb{1}_{K_j}$$

und

$$\mathbb{1}_{K_j} \downarrow \mathbb{1}_K$$

gilt. Daher folgt

$$\mathbb{P}[\eta(K) = 0] \geq \mathbb{P}[\eta(f_j) = 0] = \mathbb{P}[\eta(f_j) \leq 0].$$

Nun konvergiert $\xi_{n''}(f_j)$ gegen $\eta(f_j)$ und somit ergibt sich mit Hilfe des Lemmas von Fatou

$$\mathbb{P}[\eta(f_j) \leq 0] \geq \limsup_{n' \to \infty} \mathbb{P}[\xi_{n'}(I_j) \leq 0] = \mathbb{P}[\xi(K_j) \leq 0],$$

also (7.23).

Schließlich zeigen wir noch, dass η auch einfach ist. Dazu sei $I \in \mathcal{T}$ und wir wollen zeigen, dass die Wahrscheinlichkeit, dass η Doppelpunkte in I hat, gleich Null ist. Nun hat I Doppelpunkte, wenn $\eta(I) \geq \eta^*(I)$ und

$$\mathbb{P}[\eta(I) > \eta^*(I)] = \mathbb{P}[\eta(I) - \eta^*(I) \geq \frac{1}{2}] \leq 2 \cdot [\mathbb{E}\,\eta(I) - \mathbb{E}\,\eta^*(I)], \tag{7.24}$$

wobei die Ungleichung aus der Markov-Ungleichung mit $g(x) = x$ folgt. Nun konvergiert aber die Folge der Intermitätsmaße $\mathbb{E}\,\xi_n(I)$ gegen $\mathbb{E}\,\xi(I)$ und ξ ist ein einfacher Punktprozess. Somit ist

$$\mathbb{E}\,\eta(I) = \mathbb{E}\,\xi(I) = \mathbb{E}\,\xi^*(I) = \mathbb{E}\,\eta^*(I),$$

d. h. die rechte Seite ist 0. □

Bemerkung 7.34 *Wenn wir die beiden Bedingungen von Satz 7.32 betrachten, so ist die wesentliche davon die Konvergenz der sogenannten "Vermeidungsfunktionen"*

$$\mathbb{P}[\xi(I_n) = 0] \to \mathbb{P}[\xi(I) = 0].$$

Die Konvergenz der Intensitätsmaße

$$\mathbb{E}\,\xi_n(I) \to \mathbb{E}(\xi(I))$$

haben wir nur für die Straffheit und das letzte Argument gebraucht. Diese Bedingung ist hinreichend, aber nicht notwendig. Man kann sie durch eine Straffheitsbedingung ersetzen, etwa: Für alle $I \in \mathcal{T}$ und alle $\varepsilon > 0$ gibt es ein $\mathcal{R} \in \mathbb{N}$, so dass

$$\mathbb{P}[\xi_n(I) > \mathcal{R}] \leq \varepsilon \quad \textit{für alle} \quad n \in \mathbb{N}, \tag{7.25}$$

falls man zusätzlich einen Weg findet, um zu zeigen, dass die Limespunktprozesse nur einfache Punktprozesse sein können. Offenbar folgt aus der Markov-Ungleichung natürlich aus (7.20) auch (7.25), aber nicht umgekehrt und tatsächlich lassen sich Beispiele finden, in denen (7.25) gilt, (7.20) aber nicht.

Wir werden uns nun den extremalen Punktprozessen erneut zuwenden. Dabei wollen wir verschiedene Aspekte beachten:

(i) Wie schon im ersten Teil die Verteilung des größten Wertes des Prozesses. Falls $u_n(x)$ die Skalenfunktion ist, so dass

$$\mathbb{P}[M_n \leq u_n(x)] \xrightarrow{\omega} G(x)$$

gilt, wäre ein natürliches Objekt der Punktprozess

$$N_n := \sum_{i=1}^{n} \delta_{u_n^{-1}(x_i)}. \tag{7.26}$$

Schickt man $n \to \infty$, so werden aufgrund der extremalen Skala die meisten Punkte gegen $-\infty$ verschwinden, man kann aber hoffen, dass N_n als Punktprozess konvergiert.

(ii) Die räumliche Struktur der extremalen Werte: Ähnlich wie in Kapitel 2 können wir einen Level u_n festlegen und uns fragen, wie die Werte i verteilt sind, für die x_i diesen Level u_n überschreitet. Um diese Überschreitungen als Punktprozesse zu schreiben, betten wir die $1 \leq i \leq n$ in das halboffene Einheitsintervall $(0, 1]$ ein. Dies funktioniert über die Abbildung

$$i \mapsto \frac{i}{n}.$$

Wir betrachten dann den Punktprozess der Überschreitungen:

$$N_n := \sum_{i=1}^{n} \delta_{i/n} \mathbb{1}_{X_i > u_n}. \tag{7.27}$$

(iii) Man kann (i) und (ii) kombinieren und bekommt den folgenden Punktprozess auf $\mathbb{R} \times (0, 1]$

$$N_n := \sum_{i=1}^{n} \delta_{(u_n^{-1}(x_i), i/n)}. \tag{7.28}$$

Wir beginnen unster Studium mit dem Punktprozess der Überschreitungen aus (7.27), weil dieser am einfachsten ist. Wir zeigen den folgenden Satz. Wir brauchen hierfür nur stationäre Zufallsvariablen.

Satz 7.35 *Es sei $(X_n)_n$ eine stationäre Folge von Zufallsvariablen und sei F ihre (marginale) Verteilungsfunktion. Dann gilt*

(i) Es sei $\tau > 0$ und es mögen die Bedingungen $D(u_n)$ und $D'(u_n)$ gelten, wobei wir $u_n := u_n(\tau)$ so wählen, dass

$$n(1 - F(u_n(\tau)) \to \tau.$$

Sei N_n definiert wie in (7.27). Dann konvergiert (N_n) schwach gegen den Punktprozess N auf $[0, 1]$. N hat das Intensitätsmaß τdx.

(ii) Falls die Bedingungen auf (i) für alle $\tau > 0$ gelten, so konvergiert der Punktprozess

$$\tilde{N}_n := \sum_{i=-\infty}^{\infty} \delta_{i/n} \mathbb{1}_{x_i > u_n(\tau)}$$

schwach gegen den Punktprozess $\tilde{N}$ auf $\mathbb{R}$ mit Intensitätsmaß τdx (hierzu beachte man, dass sich eine stationäre Folge natürlich mit Indizes auf $\mathbb{Z}$ schreiben lässt).

Beweis: Natürlich wollen wir Kallenbergs Satz verwenden. Zunächst überprüfen wir, dass die Intensitätsmaße gegen den richtigen Limes konvergieren:

$$\mathbb{E}N_n((c, d]) = \sum_{i=1}^{n} \mathbb{P}[X_i > u_n(\tau)] \mathbb{1}_{i/nin(c,d)} = n(d - c)(1 - F(u_n(\tau))) \to \tau(d - c),$$

somit erhalten wir asymptotisch das richtige Intensitätsmaß. Als nächstes zeigen wir, dass

$$\mathbb{P}[N_n(I) = 0] \to e^{-\tau|I|}$$

für $n \to \infty$ und jede endliche disjunkte Vereinigung von Intervallen I. Zu diesem Zweck beginnen wir mit dem Fall, dass I nur ein Intervall ist. Dann gilt:

$$\mathbb{P}[N_n(I) = 0] = \mathbb{P}[X_i \leq u_n \quad \forall \frac{i}{n} \in I] \underset{n\to\infty}{\longrightarrow} e^{-\tau|I|}.$$

Hierbei haben wir für die Konvergenz Satz 5.10 verwendet. Für endliche, disjunkte Vereinigungen gilt dann dasselbe Resultat, wenn man ausnutzt, dass die Distanz der Intervalle, wenn man sie wieder auf $\mathbb{N}$ abbildet, wie n skaliert, so dass wir Lemma 5.4 verwenden können, um für $I = \bigcup_k^N I_k$ zu zeigen

$$\mathbb{P}(N_n(I) = 0) \sim \prod_{k=1}^{N} \mathbb{P}[N_n(I_k) = 0]$$

(dieser Schritt ist natürlich für iid Variablen klar). Für (i) sind alle Intervalle kürzer als 1, also genügt es, die Bedingungen $D(u_n)$ und $D'(u_n)$ für $u_n = u_n(\tau)$ und das gegebene

τ zu kennen. In (ii) benötigen wir allerdings auch größere Intervalle, so dass wir die entsprechenden Bedingungen für alle $\tau > 0$ benötigen. $\square$

Wir wenden uns nun dem Prozess der Extrema (7.26) zu. Hier zeigen wir

Satz 7.36 *Sei $(X_n)_n$ eine stationäre Folge von Zufallsvariablen mit (marginaler) Verteilungsfunktion F. Außerdem mögen die Bedingungen $D(u_n)$ und $D'(u_n)$ für alle $u_n = u_n(\tau)$ mit*

$$n(1 - F(u_n(\tau))) \to \tau$$

gelten. Dann konvergiert der Punktprozess

$$\mathcal{P}_n := \sum_{i=1}^{n} \delta_{u_n^{-1}(x_i)}$$

schwach gegen den Punktprozess $\mathcal{P}$ auf $\mathbb{R}^+$, dessen Intensitätsmaß das Lebesguemaß ist.

Beweis: Wieder benutzen wir den Satz von Kallenberg. Wir beginnen mit der Konvergenz der Intensitätsmaße. Für jedes Intervall $(a, b] \leq \mathbb{R}^+$ gilt

$$\mathbb{E}\mathcal{P}_n[(a,b]) = n\mathbb{P}[n_n^{-1}(X_1) \in (a,b]] = n(F(u_n(a)) - F((u_n(b))) \to b - a.$$

Als nächsten müssen wir uns um die Wahrscheinlichkeit des Ereignisses $\{\mathcal{P}_n((a,b]) = 0\}$ kümmern. Es gilt

$$\mathbb{P}[\mathcal{P}((a,b]) = 0] = \sum_{k=0}^{\infty} \mathbb{P}[\#\{i : X_i > u_n(a)\} = k \wedge \{i : X_i \leq u_n(b)\} = n - k].$$

Nun zerlegen wir das Intervall $(1, \ldots, n)$ in disjunkte Teilintervalle I_l und I'_l, $l = 1, \ldots, r$ wie im Beweis von Satz 5.5, um die Terme, die in der Summe auftreten, abzuschätzen. Denn gilt

$$\mathbb{P}[\exists l : M(I'_l) \geq u_n(b)] \leq rm(1 - F(u_n(b))) \leq \frac{mrb}{n} \to 0.$$

Daher folgt

$$\begin{aligned} & \mathbb{P}[\#\{i : X_i > u_n(a)\} = k \wedge \#\{i : X_i \leq u_n(b)\} = n - k] \\ \leq\ & \mathbb{P}[\#\{i \in \bigcup_l I_l : X_i > u_n(a)\} = k \wedge \{i \in \bigcup_l I_l : X_i \leq u_n(b)\} = n - k] \\ & +\mathbb{P}[\exists l : M(I'_l) \geq u_n(b)] \end{aligned}$$

und

$$\begin{aligned} & \mathbb{P}[\#\{i : X_i > u_n(a)\} = k \wedge \#\{i : X_i \leq u_n(b)\} = n - k] \\ \geq\ & \mathbb{P}[\#\{i \in \bigcup_l I_l : X_i > u_n(a)\} = k \wedge \#\{i \in \bigcup_l I_l : X_i \leq u_n(b)\} = n - k] \\ & -\mathbb{P}[\exists l : M(I'_l) \geq u_n(b)]. \end{aligned}$$

Außerdem erhalten wir wegen Bedingung $D'(u_n)$

$$\begin{aligned} & \mathbb{P}[\exists l : \#\{i \in I_l : X_i \geq u_n(a)\} \geq 2] \\ \leq\ & r \sum_{1\neq j\in I_1} \mathbb{P}[X_1 > u_n(a), X_j > u_n(b)] \\ \leq\ & n \sum_{1\neq j\in I_1} \mathbb{P}[X_1 > u_n(b), X_j > u_n(b)] \downarrow 0. \end{aligned}$$

Wir können uns daher auf das Ereignis konzentrieren, dass keiner der Intervalle I_l mehr als eine Überschreitung des Niveaus $u_n(a)$ enthält. Daher gibt es nur eine Art, wie k der X_i das Niveau $u_n(a)$ überschreiten, während alle unterhalb von $u_n(b)$ bleiben: In genau k der Intervalle I_l muss das Maximum $u_n(a)$ überschreiten, während es in den anderen unterhalb von $u_n(b)$ bleiben muss. Aufgrund der Stationarität ergibt sich

$$\begin{aligned} & \mathbb{P}[\#\{i : X_i > u_n(a)\} = k \wedge \{i : X_i \leq u_n(b)\} = n-k] \\ = & \binom{r}{k} (\mathbb{P}[M(I_1) > u_n(a)])^k (\mathbb{P}[M(I_1) \leq u_n(b)])^{n-k} + 0(1). \end{aligned}$$

Schließlich gilt auch

$$\mathbb{P}[M(I_1) > u_n(a)] \sim \frac{n}{r}(1 - F(u_n(a))) \sim \frac{a}{r}$$

und

$$\mathbb{P}[M(I_1) \leq u_n(b)] \sim 1 - \frac{n}{r}(1 - F(u_n(b))) \sim 1 - \frac{b}{r}.$$

Somit ergibt sich

$$\binom{r}{k} (\mathbb{P}[M(I_1) > u_n(a)])^k (\mathbb{P}[M(F_1) \leq u_n(b)])^{n-k} \sim \frac{1}{k!} a^k e^{-b}.$$

Setzt man dies ein und summiert über k, so erhält man

$$\mathbb{P}[\mathcal{P}_n((a,b]) = 0] = \sum_{k=0}^{\infty} \frac{1}{k!} a^k e^{-b} (1 + 0(1)) = e^{a-b}(1 + 0(1)),$$

wobei $0(1)$ gegen 0 geht, wenn $n \to \infty$, $r \to \infty$, $n \to \infty \frac{r}{n} \to 0$ und $\frac{mr}{n} \to 0$ geht. Dies zeigt mit dem Satz von Kallenberg die Behauptung. $\square$

Satz 7.35 besagt, dass die Anzahl Überquerungen eines gegebenen (extremalen) Niveaus in disjunkte Intervalle asymptotisch unabhängig sind. Wir wollen diesen Satz weiter verschärfen, indem wir verschiedene Level für verschiedene Intervalle zulassen und insofern als dass tatsächlich Extremwertprozesse über disjunkten Intervallen unabhängig sind. Hierzu sei $r \in \mathbb{R}$ fest. Wir betrachten für $k = 1, \ldots, r$ Folgen u_n^k, so dass

$$n(1 - F(u_n^k)) \to \tau_k$$

gilt. Hierbei sei

$$u_n^1 \geq u_2^n \geq \ldots \geq u_n^r.$$

Wir führen eine Verschärfung der (bekannten) Bedingung $D(u_n)$ ein und nennen diese $D_r(\vec{u}_n)$.

Definition 7.37 *Eine stationäre Folge* (X_i) *erfüllt Bedingung* $D_r(\vec{u}_n)$ *für eine Folge* $\vec{u}_n$ *mit*

$$u_n^1 \geq \ldots \geq u_n^r$$

genau dann, wenn

$$|F_{\vec{\imath}\vec{\jmath}}(\vec{v},\, \vec{w}) - F_{\vec{\imath}}(\vec{v})F_{\vec{\jmath}}(\vec{w})| \leq \alpha_{n,l},$$

wobei $\vec{\imath}$ *und* $\vec{\jmath}$ *gewählt sind wie in der Definition von* $D(u_n)$ *und*

$$\vec{v} = (v_1, \ldots, v_p),\ \vec{w} = (w_1, \ldots, w_q),$$

wobei die v_i, w_j *gewählt sind als*

$$v_1, v_2, \ldots, w_1, \ldots, w_q \in \{u_n^1, \ldots, u_n^r\},$$

$\alpha_{n,l}$ *ist so wie in der Definition von* $D(u_n)$.

Wir bekommen sofort

Lemma 7.38 *Es gelte die Bedingung* $D_r(\vec{u}_n)$. *Für* $E_1, \ldots, E_s \leq \{1, \ldots, n\}$ *mit*

$$dist(E_j, E_j) \geq l \quad \forall\, j \neq j'$$

gilt

$$|\mathbb{P}[M(E_j) \leq u_{n,j}\, \forall\, j = 1, \ldots, s] - \prod_{j=1}^{s} \mathbb{P}[M(E_j) \leq u_{n,j}]| \leq \alpha_{n,l}(s-1),$$

falls für alle u_{n_j} *gilt* $u_{n_j} \in \{u_n^1, \ldots, u_n^r\}$.

Der Beweis ist eine Übung.

Satz 7.39 *Es seien* $J_1, \ldots, J_s \leq \{1, \ldots, n\}$, *so dass* $|J_j| \sim \Theta_j n \quad \forall\, j = 1, \ldots, s$. *Es gelte* $D_r(\vec{u}_n)$ *und die* u_{n_j} *seien gewählt wie im vorhergehenden Lemma. Dann gilt:*

(i) $\mathbb{P}[M(J_j) \leq u_{n_j} \quad \forall\, j = 1, \ldots, s] - \prod_{j=1}^{s} \mathbb{P}[M(J_j) \leq u_{n_j}] \to 0$ *mit* $n \to \infty$.

(ii) Sind die $J_1, \ldots, J_s$ *für festes* m *disjunkte Teilmengen von* $\{1, \ldots, nm\}$ *mit* $|J_J| \sim n\Theta_j$ *und* $\Sigma\Theta_j \leq m$. *Gilt dann* $D_r(\vec{u}_n)$ *für*

$$\vec{u}_n = (u_n(\tau_1 m), \ldots, u_n(\tau_s m)),$$

wobei $u_n(\tau)$ *so ist, dass*

$$n(1 - F(u_n(\tau)) \to \tau$$

gilt, dann gilt auch (i).

Beweis: Auch diesen Beweis lassen wir als Übung. □

Wenn wir zusätzlich D' fordern, erhalten wir

Korollar 7.40 *(i) Es gelten die Bedingungen aus Satz 7.39 und die Bedingung $D'(u_{n,k})$ für alle $k = 1, \ldots, r$. Gilt dann*

$$\sum_{k=1}^{r} \Theta_k \leq 1,$$

so folgt

$$\mathbb{P}[M(J_k) \leq u_{n,k} \quad \forall\, k = 1, \ldots, r] \to \exp(-\sum_{k=1}^{r} \Theta_k \tau_k').$$

Hierbei ist τ_k' bestimmt durch

$$\tau_k' := \lim_{n\to\infty} n(1 - F(u_{n,k})).$$

(ii) Gleiches gilt, falls die Bedingungen aus Satz 7.39 Teil (ii) gelten und zusätzlich Bedingung $D'(v_n)$ mit $v_n = U_n(\Theta_k \tau_{k'})$ für $k = 1, \ldots, r$.

Beweis: Übung □

Wir sind nun in der Lage, auch die Überschreitung mehrerer Level durch den Extremwertprozess zu analysieren. Dazu sei die Folge $(u_n^{(k)})$ wie oben gewählt mit

$$u_n^{(1)} \geq \ldots \geq u_n^{(r)}.$$

Wir definieren den folgenden Punktprozess auf $\mathbb{R}^2$

$$N_n = \sum_{i=1}^{n} \sum_{k=1}^{r} \mathbb{1}_{\{X_i > u_n^{(k)}\}} \delta_{(l_k, i/n)}.$$

Hierbei sind die (l_k) von der Form

$$l_1 < \ldots < l_r,$$

z. B. $l_k = l^{-\tau k}$ mit

$$l_K = \lim_{n\to\infty} n(1 - F(u_n^{(k)}))$$

(letzteres ist aber nicht notwendig).

Wir wollen als erstes den möglichen Limesprozess der N_n beschreiben. Dazu sei $\mathcal{P}_r$ ein Poissonscher Punktprozess mit Intensitätsmaß τ_r auf $\mathbb{R}$. Seine Atome seien (x_r), also

$$\mathcal{P}_r = \sum_j \delta_{x_j}.$$

Ferner seien (β_j) iid Zufallsvariablen, die unabhängig von $\mathcal{P}_r$ seien. Jedes β_j nehme Werte in $\{1, \ldots, r\}$ an. Es gelte

$$\mathbb{P}[\beta_j = s] = \begin{cases} (\tau_{r-s+1} - \tau_{r-s})/\tau_r, & s = 1, \ldots, r-1 \\ \frac{\tau_1}{\tau_r}, & s = r \end{cases}.$$

Wir setzen

$$N := \sum_{j=1}^{\infty} \sum_{k=1}^{\beta_j} \delta_{(l_k, x_j)}.$$

Wenn wir die Einschränkung dieses Prozesses auf die Geraden $x = l_k$ einen Poissonschen Punktprozess mit Intensität τ_k. Die obige Konstruktion konstruiert r solcher Prozesse, wobei jeder eine Ausd?? des vorhergehenden ist. Daher erscheint es natürlich, dass N der Limes der N_n ist. In der Tat wisen wir ja schon, dass die Marginalverteilungen von N_n auf den Geraden $x = l_k$ gegen die Marginalverteilungen von N konvergieren. Außerdem sind die aufeinanderfolgenden Punktprozesse auf den Geraden bei N Ausde?? voneinander. Es gilt

Satz 7.41 *(i) Es gelte $D_r(\vec{u}_n)$ und $'(u_n^{(k)})$ für alle $1 \leq k \leq r$. Dann konvergiert der Punktprozess N_n schwach in Verteilung gegen N.*

(ii) Gilt zusätzlich für $0 \leq \tau < \infty$ und $u_n(\tau)$

$$n(1 - F(u_n(\tau))) \to \tau$$

und gilt $D_r(\vec{u}_n)$ für

$$u_n = (u_n(m\tau_1), \ldots, u_n(m\tau_r))$$

für alle $m \geq 1$ und gilt $D'(u_n(\tau))$ für alle $\tau > 0$, dann konvergiert N_n gegen N (als Punktporzess auf $\mathbb{R}^+ \times \mathbb{R}$).

Beweis: Wieder verwenden wir den Satz von Kallenberg. Nachzuweisen , dass die Intensitätsmaße konvergieren, dass also

$$\mathbb{E}\, N_n(B) \underset{n\to\infty}{\longrightarrow} \mathbb{E}\, N(B)$$

für alle B gilt, ist eine einfache Übung.

Wir zeigen nun, dass auch

$$\mathbb{P}[N_n(B) = 0] \to \mathbb{P}[N(B) = 0]$$

für $n \to \infty$ und alle B, die sich als disjunkte Vereinigung halboffener Rechtecke schreiben lassen, also etwa

$$B = \bigcup_k (c_k, d_k] \times (\gamma_k, \delta_K] =: \bigcup_k F_k.$$

Durch Umschreiben können wir B auch in der Form

$$B = \bigcup_j (c_j, d_j] \times E_j =: \bigcup_j F_j$$

schreiben, wobei E_j endliche Vereinigungen halboffener Intervalle sind. Nun können wir die Struktur der prozesse N_n ausnutzen. Es gilt $N_n(F_j) = 0$ genau dann, wenn die niedrigste Gerade $x = l_k$, F_j schneidet keine Punkte hat. Sei dies die Gerade mit der Nummer m_j. Dann gilt

$$\{N_n(F_j) = 0\} = \{\mathcal{P}_{m_j}((c_j, d_j]) = 0\}.$$

Dies ist aber das Ereignis

$$\{M([c_j n, d_j n]) \leq u_n, m_j\}.$$

Aus Korollar 7.40 erhalten wir somit

$$\mathbb{P}[N_n(B) = 0] \to \exp(-\sum_{j=1}^{s}(d_j - c_j)\tau_{m_j}).$$

Die rechte Seite ist aber $\mathbb{P}[N(B) = 0]$, wie man leicht nachrechnet. Das zeigt den Satz. □

Schließlich wenden wir uns der Charakterisierung des gesamten Extremprozesses wie in (7.28 zu. Wieder sei $u_n(\tau)$ so, dass

$$n(1 - F(u_n(\tau))) \to \tau.$$

Wir setzen

$$\mathcal{N}_n := \sum_{i=1}^{\infty} \delta_{(i/n, u_n^{-1}(X_i))}.$$

Dies ist ein Punktprozess auf $(0, 1) \times \mathbb{R}^+$. Es gilt:

Satz 7.42 *Für $u_n(\tau)$ wie oben gelte für jedes $\tau > 0$ $D'(u_n(\tau))$ und für jedes $r \in \mathbb{N}$ und jedes*

$$\vec{u}_n = (u_n(\tau_1), \ldots, u_n(\tau_r))$$

gelte $D_r(\vec{u}_n)$. Dann konvergiert der Punktprozess $\mathcal{N}_n$ gegen den Punktprozess $\mathcal{N}$ auf $\mathbb{R}^+ \times \mathbb{R}^+$ mit Intensitätsmaß λ^2.

Beweis: Wieder verwenden wir den Satz von Kallenberg. Ist $B = (c, d] \times (\gamma, \delta]$, dann folgt

$$\begin{aligned}
\mathbb{E}\, N(B) &= ([nd] - [nc])\mathbb{P}[\gamma < u_n^{-1}(X_1) \leq \delta] \\
&\sim n(d - c)\mathbb{P}[u_n(\gamma) < X_1 \leq u_n(\delta)] \\
&= n(d - c)(F(u_n(c)) - F(u_n(d))) \\
&\to (d - c)(\delta - \gamma) \\
&= \mathbb{E}\mathcal{N}(B).
\end{aligned}$$

Um auch die Konvergenz der Vermeidungsfunktionen zu zeigen, drücken wir die (für festes B) auftretenden Wahrscheinlichkeiten mithilfe der Prozesse N_n aus Satz 7.41 aus. Ist z. B. B ein Rechteck, so ist B ohne Punkte durch $\mathcal{N}_n$ belegt, genau dann, wenn die Anzahl der Überschreitungen des Niveaus $u_n(\gamma)$ und $u_n(\delta)$ in $(c, d]$ die gleichen sind. Dies lässt sich mithilfe des Prozesses N_n für die Niveaus $u_n(\gamma)$ und $u_n(\delta)$ ausdrücken. N_n aber konvergiert schwach, also konvergieren die zugehörigen Wahrscheinlichkeiten gegen die, die mithilfe des Prozesses N ausgedrückt werden. Diese Wahrscheinlichkeiten aber lassen sich leicht berechnen: Jede Anzahl von Punkten in dem unteren Prozess ist erlaubt, vorausgesetzt, dass die entsprechenden β_j gleich 2 sind. Dies ergibt

$$N(B) = \sum_{l=0}^{\infty} e^{-(d-c)\delta} \frac{[(d-c)\delta]^l}{l!} \left(\frac{\gamma}{\delta}\right)^l = e^{-(d-c)(\delta-\gamma)}.$$

Dies ist die Behauptung. □