

WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER

EXTREMWERTTHEORIE

SKRIPT
PROF. DR. ZAKHAR KABLUCHKO

WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER
INSTITUT FÜR MATHEMATISCHE STATISTIK

Vorwort

Dies ist ein Skript zur Vorlesung “Extremwerttheorie”, die an der Universität Ulm (im Wintersemester 2011/12 und Sommersemester 2014) und an der Universität Münster (im Sommersemester 2015) gehalten wurde. Die erste L^AT_EX-Version des Skripts wurde von Herrn Benjamin Tempel erstellt. Danach wurde das Skript von mir mehrmals korrigiert, überarbeitet und ergänzt. Ich bedanke mich bei Frau Judith Olszewski, die eine frühere Version des Skripts kritisch durchgelesen und zahlreiche Verbesserungsvorschläge gemacht hat. Ich bedanke mich außerdem bei Wolfgang König, Michael Stolz und zahlreichen weiteren Kollegen für nützliche Hinweise zu einzelnen Kapiteln des Skripts. Diese Version ist vorläufig. In Zukunft soll das Skript weiter überarbeitet und ergänzt werden.

Bei Fragen, Wünschen und Verbesserungsvorschlägen können Sie gerne eine E-Mail an
`zakhar dot kabluchko at uni-muenster dot de`
schreiben.

17. Juli 2015

Zakhar Kabluchko

Literatur

Es gibt sehr viele Bücher über Extremwerttheorie. Hier ist eine unvollständige Liste:

1. P. Embrechts, C. Klüppelberg, T. Mikosch. *Modelling Extremal Events for Insurance and Finance*.
2. S. Resnick. *Extreme Values, Regular Variation and Point Processes*.
3. L. de Haan, A. Ferreira. *Extreme Value Theory: An Introduction*.
4. S. Resnick. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*
5. M. Falk, J. Hüsler, R.-D. Reiss. *Laws of Small Numbers: Extremes and Rare Events*.
6. M. R. Leadbetter, G. Lindgren, H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes*.
7. S. Coles. *An Introduction to Statistical Modeling of Extreme Values*.
8. J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels. *Statistics of Extremes: Theory and Applications*.
9. E. J. Gumbel. *Statistics of Extremes*.
10. J. Galambos. *The Asymptotic Theory of Extreme Order Statistics*.
11. V. Nevzorov. *Records: Mathematical Theory*.
12. L. de Haan. *On regular variation and its application to the weak convergence of sample extremes*.
13. N. H. Bingham, C. M. Goldie, J. L. Teugels. *Regular Variation*.
14. D. Pfeifer. *Einführung in die Extremwertstatistik*.

Inhaltsverzeichnis

Vorwort	i
Literatur	ii
Kapitel 1. Extremwertverteilungen	1
1.1. Verteilung des Maximums	1
1.2. Definition der Extremwertverteilungen und deren Max-Anziehungsbereiche	3
1.3. Drei Beispiele von Extremwertverteilungen: Gumbel, Fréchet, Weibull	5
1.4. Satz von Fisher–Tippett–Gnedenko	9
1.5. Jenkinson–von Mises–Darstellung	11
Kapitel 2. Cauchy–Funktionalgleichung	12
2.1. Additive und multiplikative Funktionen	12
2.2. Stetige Lösungen	13
2.3. Satz von Steinhaus	14
2.4. Messbare Lösungen	16
2.5. Hamel–Funktionen	17
Kapitel 3. Satz von Fisher–Tippett–Gnedenko	18
3.1. Eindeutigkeit der Normierungskonstanten	18
3.2. Max-stabile Verteilungen	20
3.3. Charakterisierung der max-stabilen Verteilungen	21
Kapitel 4. Regulär variierende Funktionen	25
4.1. Definition der regulär variierenden Funktionen	25
4.2. Asymptotische Äquivalenz	27
4.3. Satz über die gleichmäßige Konvergenz	28
4.4. Darstellungssatz von Karamata	30
4.5. Abschätzungen für regulär variierende Funktionen	31
4.6. Charakterisierung der Grenzwertfunktionen	33
4.7. Integrale von regulär variierenden Funktionen	34
Kapitel 5. Max–Anziehungsbereiche	36
5.1. Max–Anziehungsbereich der Fréchet–Verteilung Φ_α	36
5.2. Max–Anziehungsbereich der Weibull–Verteilung Ψ_α	42
5.3. Max–Anziehungsbereich der Gumbel–Verteilung Λ	44
5.4. Beispiel: Normalverteilung	46
5.5. Von Mises–Bedingungen	50
Kapitel 6. Statistik der Extremwertverteilungen	55

6.1. Statistik der Blockmaxima: GEV-Verteilungen	55
6.2. Peaks over Threshold: Statistik der GP-Verteilungen	59
Kapitel 7. Ordnungsstatistiken	63
7.1. Allgemeine Eigenschaften der Ordnungsstatistiken	63
7.2. Extreme Ordnungsstatistiken	66
7.3. Darstellungen der Ordnungsstatistiken	70
7.4. Wiederkehrperiode und Gumbel-Überschreitungsmethode	73
Kapitel 8. Rekorde	76
8.1. Satz von Rényi	76
8.2. Anzahl der Rekorde	78
8.3. Rekordzeiten	80
8.4. Zentrale Grenzwertsätze	84
8.5. Rekordwerte	86
Kapitel 9. Poisson-Punktprozesse	89
9.1. Poisson-Verteilung	89
9.2. Beispiel zu Poisson-Prozessen	90
9.3. Definition der Zählmaße und Punktprozesse	91
9.4. Definition der Poisson-Punktprozesse	94
9.5. Superpositionssatz	95
9.6. Abbildungssatz	96
9.7. Laplace-Funktionale	98
9.8. Färbungs- und Markierungssatz für Poisson-Punktprozesse	101
Kapitel 10. Konvergenz von Punktprozessen	104
10.1. Vage Konvergenz	104
10.2. Verteilungskonvergenz von Punktprozessen	105
10.3. Bernoulli-Experimente mit kleiner Erfolgswahrscheinlichkeit	106
10.4. Konvergenz der Binomialpunktprozesse gegen die Poisson-Punktprozesse	108
10.5. Konvergenz der extremen Ordnungsstatistiken: Gumbel-Fall	109
10.6. Konvergenz der extremen Ordnungsstatistiken: Fréchet und Weibull-Fall	113
10.7. Konvergenz der Rekordzeiten gegen einen Poisson-Punktprozess	115
10.8. Konvergenz gegen den Extremwertprozess	116
10.9. Allgemeiner Extremwertprozess	118

KAPITEL 1

Extremwertverteilungen

1.1. Verteilung des Maximums

Seien X_1, X_2, \dots unabhängige und identisch verteilte (u.i.v.) Zufallsvariablen.

Definition 1.1.1. Die **Verteilungsfunktion** F von X_i ist definiert durch

$$F(t) = \mathbb{P}[X_i \leq t], \quad t \in \mathbb{R}.$$

Die **Tailfunktion** \bar{F} von X_i ist definiert durch

$$\bar{F}(t) = 1 - F(t) = \mathbb{P}[X_i > t], \quad t \in \mathbb{R}.$$

Die Verteilungsfunktion F ist monoton steigend. Die Tailfunktion \bar{F} hingegen ist monoton fallend.

Mit M_n bezeichnen wir das **Maximum** von X_1, \dots, X_n :

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

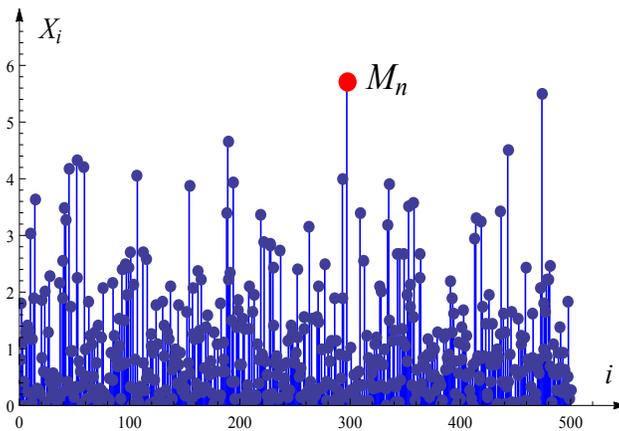


ABBILDUNG 1. Veranschaulichung von M_n .

Satz 1.1.2. Die Verteilungsfunktion von M_n ist gegeben durch

$$\mathbb{P}[M_n \leq t] = F^n(t) \quad \text{für alle } t \in \mathbb{R}.$$

BEWEIS. Aus der Definition des Maximums M_n folgt, dass

$$\mathbb{P}[M_n \leq t] = \mathbb{P}[\max\{X_1, \dots, X_n\} \leq t] = \mathbb{P}[X_1 \leq t, \dots, X_n \leq t].$$

Da X_1, \dots, X_n unabhängig und identisch verteilt sind, gilt

$$\mathbb{P}[X_1 \leq t, \dots, X_n \leq t] = \mathbb{P}[X_1 \leq t] \cdot \dots \cdot \mathbb{P}[X_n \leq t] = F^n(t).$$

Zusammen ergibt sich $\mathbb{P}[M_n \leq t] = F^n(t)$. □

Aufgabe 1.1.3. Seien X_1, \dots, X_n u.i.v. Zufallsvariablen mit Verteilungsfunktion F . Bestimmen Sie die Verteilungsfunktion des Minimums $m_n := \min\{X_1, \dots, X_n\}$.

Hinweis: Es ist einfacher, die Tailfunktion von m_n zu berechnen.

Aufgabe 1.1.4. Seien X_1, \dots, X_n u.i.v. Zufallsvariablen mit Verteilungsfunktion F . Bestimmen Sie die gemeinsame Verteilungsfunktion von m_n und M_n , d.h. berechnen Sie die Wahrscheinlichkeit $\mathbb{P}[m_n \leq x, M_n \leq y]$.

Wir werden uns für die Eigenschaften von M_n für große Werte von n interessieren. Im folgenden Satz berechnen wir den Wert, gegen den die Zufallsvariable M_n für $n \rightarrow \infty$ in Wahrscheinlichkeit konvergiert.

Definition 1.1.5. Der **rechte Endpunkt** der Verteilungsfunktion F ist definiert durch

$$x^* = \sup\{t \in \mathbb{R} \mid F(t) < 1\} = \inf\{t \in \mathbb{R} \mid F(t) = 1\} \in \mathbb{R} \cup \{+\infty\}.$$

Erreicht die Verteilungsfunktion F nie den Wert 1, so ist $x^* = +\infty$.

Der rechte Endpunkt kann endlich oder $+\infty$ sein:

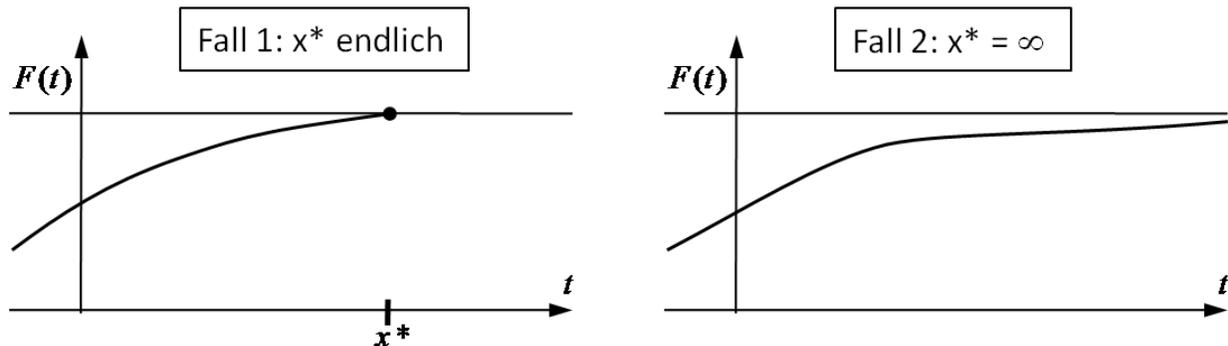


ABBILDUNG 2. Veranschaulichung von x^*

Der rechte Endpunkt x^* ist das Supremum aller Werte, die die Zufallsvariable X_1 annehmen kann, wobei aber Werte ignoriert werden, wenn ihre Wahrscheinlichkeit gleich 0 ist. Der Wert x^* heißt auch das **essentielle Supremum** von X_1 und wird mit $\text{esssup } X_1$ bezeichnet.

Satz 1.1.6. Für $n \rightarrow \infty$ konvergiert die Zufallsvariable M_n in Wahrscheinlichkeit gegen den Wert x^* .

BEWEIS. Wir betrachten zwei Fälle.

FALL 1. Sei zuerst x^* endlich. Für jedes $\varepsilon > 0$ gilt $F(x^* - \varepsilon) < 1$, wobei die Ungleichung strikt ist. Aus Satz 1.1.2 folgt, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n \leq x^* - \varepsilon] = \lim_{n \rightarrow \infty} F^n(x^* - \varepsilon) = 0.$$

Außerdem gilt $\mathbb{P}[M_n \geq x^* + \varepsilon] = 0$. Es folgt

$$\lim_{n \rightarrow \infty} \mathbb{P}[|M_n - x^*| \geq \varepsilon] = \lim_{n \rightarrow \infty} (\mathbb{P}[M_n \leq x^* - \varepsilon] + \mathbb{P}[M_n \geq x^* + \varepsilon]) = 0$$

Somit gilt $M_n \xrightarrow{P} x^*$.

FALL 2. Sei nun $x^* = +\infty$. Wir zeigen, dass für jedes noch so großes $A \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n < A] = 0.$$

Aus $x^* = \infty$ folgt, dass $F(A) < 1$, wobei die Ungleichung strikt ist. Mit Satz 1.1.2 folgt

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n < A] = \lim_{n \rightarrow \infty} F^n(A) = 0.$$

Somit gilt $M_n \xrightarrow{P} +\infty$. □

Aufgabe 1.1.7. Zeigen Sie, dass M_n sogar fast sicher gegen x^* konvergiert.

1.2. Definition der Extremwertverteilungen und deren Max–Anziehungsbereiche

Wir werden uns für die Gestalt der Verteilungsfunktion des Maximums $M_n = \max\{X_1, \dots, X_n\}$ für große Werte von n interessieren. Zunächst einmal erinnern wir uns an zwei klassische Sätze über die Verteilung der Summe $X_1 + \dots + X_n$.

Das starke **Gesetz der großen Zahlen** besagt, dass für u.i.v. Zufallsvariablen X_1, X_2, \dots mit endlichem Erwartungswert μ

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow \infty]{f.s.} \mu.$$

Der **zentrale Grenzwertsatz** besagt, dass für u.i.v. Zufallsvariablen X_1, X_2, \dots mit endlichem Erwartungswert μ und endlicher Varianz $\sigma^2 > 0$ die Verteilungskonvergenz

$$\frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} N(0, 1)$$

gilt, wobei $N(0, 1)$ eine Standardnormalverteilung bezeichnet.

Wir wollen nun ein Analogon des zentralen Grenzwertsatzes für das Maximum M_n herleiten. Seien also X_1, X_2, \dots u.i.v. Zufallsvariablen mit Verteilungsfunktion F . Wir fragen uns, ob es Folgen von Konstanten $a_n > 0$, $b_n \in \mathbb{R}$ und eine Verteilungsfunktion G gibt, so dass für $n \rightarrow \infty$ die folgende Verteilungskonvergenz gilt:

$$(1.2.1) \quad \frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} G.$$

Es sei bemerkt, dass wir in (1.2.1) eine Normierung von M_n mit beliebigen Folgen a_n und b_n (und nicht nur mit dem Erwartungswert und der Standardabweichung) zulassen, da in vielen interessanten Fällen der Erwartungswert und die Varianz gar nicht existieren.

Die Verteilungskonvergenz in (1.2.1) bedeutet, dass

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

für alle Stetigkeitspunkte t von G . Eine äquivalente Formulierung ist diese:

$$(1.2.2) \quad \lim_{n \rightarrow \infty} F^n(a_n t + b_n) = G(t)$$

für alle Stetigkeitspunkte t von G .

Wenn (1.2.1) bzw. (1.2.2) gilt, so sagen wir, dass G eine **Extremwertverteilung** ist und dass die Verteilungsfunktion F (bzw. die Zufallsvariablen X_i) im **Max-Anziehungsbereich** von G liegt.

Bemerkung 1.2.1. Es gibt einen Spezialfall von (1.2.1) und (1.2.2), der nicht interessant ist, und den wir deshalb ausschließen möchten. Eine Zufallsvariable Z , bzw. deren Verteilungsfunktion $G(t) = \mathbb{P}[Z \leq t]$, heißt **degeneriert**, wenn es einen Wert c mit $\mathbb{P}[Z = c] = 1$, bzw.

$$(1.2.3) \quad G(t) = \begin{cases} 0, & t < c, \\ 1, & t \geq c \end{cases}$$

gibt. Für jede Verteilungsfunktion F kann man durch die “falsche” Wahl der Konstanten a_n, b_n erreichen, dass (1.2.1) bzw. (1.2.2) mit einer degenerierten Verteilungsfunktion G gilt. Man kann zum Beispiel $b_n = 0$ und a_n derart schnell steigend wählen, dass M_n/a_n gegen 0 in Verteilung konvergiert (Übungsaufgabe). Deshalb werden wir im Folgenden die degenerierten Verteilungsfunktionen G der Form (1.2.3) aus unseren Definitionen ausschließen.

Definition 1.2.2. Der **Max-Anziehungsbereich** einer nichtdegenerierten Verteilungsfunktion G besteht aus allen Verteilungsfunktionen F , für die es zwei Folgen $a_n > 0$ und $b_n \in \mathbb{R}$ gibt, so dass

$$(1.2.4) \quad \lim_{n \rightarrow \infty} F^n(a_n t + b_n) = G(t)$$

für alle Stetigkeitspunkte t von G .

Den Max-Anziehungsbereich von G werden wir mit $\text{MDA}(G)$ (**maximum domain of attraction**) bezeichnen. Wir werden im Folgenden sehen, dass es nur sehr wenige Verteilungsfunktionen G mit einem nicht-leeren Max-Anziehungsbereich gibt.

Definition 1.2.3. Eine nichtdegenerierte Verteilungsfunktion G heißt eine **Extremwertverteilung**, wenn der Max-Anziehungsbereich von G nicht leer ist.

Somit ist G eine Extremwertverteilung, wenn es eine Verteilungsfunktion F und zwei Folgen $a_n > 0$ und $b_n \in \mathbb{R}$ gibt, so dass (1.2.4) gilt.

Bemerkung 1.2.4. Später werden wir sehen, dass alle Extremwertverteilungsfunktionen stetig sind. Deshalb kann man Definition 1.2.2 vereinfachen, indem man die Einschränkung auf die Stetigkeitspunkte von t weglässt und stattdessen verlangt, dass (1.2.4) für alle $t \in \mathbb{R}$ gilt.

Bemerkung 1.2.5. Wir machen in diesem Skript keinen Unterschied zwischen einer Verteilung (die ein Wahrscheinlichkeitsmaß auf \mathbb{R} ist) und der dazugehörigen Verteilungsfunktion. Zum Beispiel bezeichnen wir oft eine Verteilungsfunktion als “Extremwertverteilung”.

1.3. Drei Beispiele von Extremwertverteilungen: Gumbel, Fréchet, Weibull

Mit der obigen Definition ist es nicht klar, ob Extremwertverteilungen überhaupt existieren. Im Folgenden werden wir drei Beispiele von Extremwertverteilungen (oder sogar Familien von Extremwertverteilungen) konstruieren. Später werden wir zeigen, dass es bis auf lineare Transformationen keine weiteren Extremwertverteilungen gibt.

Wir erinnern daran, dass wir mit X_1, X_2, \dots u.i.v. Zufallsvariablen mit Verteilungsfunktion F bezeichnen. Weiterhin, sei $M_n = \max\{X_1, \dots, X_n\}$.

Gumbel–Verteilung $\Lambda(t) = e^{-e^{-t}}$

Definition 1.3.1. Eine Zufallsvariable hat **Gumbel–Verteilung**, wenn Ihre Verteilungsfunktion die folgende Gestalt hat:

$$\Lambda(t) = e^{-e^{-t}}, \quad t \in \mathbb{R}.$$

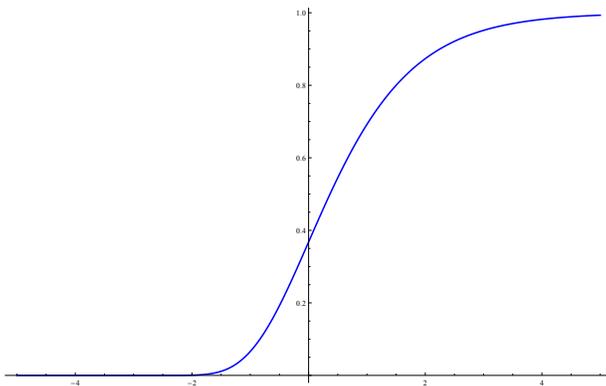


ABBILDUNG 3. Verteilungsfunktion der Gumbel–Verteilung.

Der nächste Satz zeigt, dass die Gumbel–Verteilung eine Grenzwertverteilung für Maxima von u.i.v. exponentialverteilten Zufallsvariablen ist.

Satz 1.3.2. Die Zufallsvariablen X_1, X_2, \dots seien unabhängig und exponentialverteilt mit Parameter 1, d.h.

$$F(t) = 1 - e^{-t}, \quad t \geq 0.$$

Dann konvergieren für $n \rightarrow \infty$ die Zufallsvariablen $M_n - \log n$ in Verteilung gegen Λ , d.h.

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n - \log n \leq t] = e^{-e^{-t}}, \quad t \in \mathbb{R}.$$

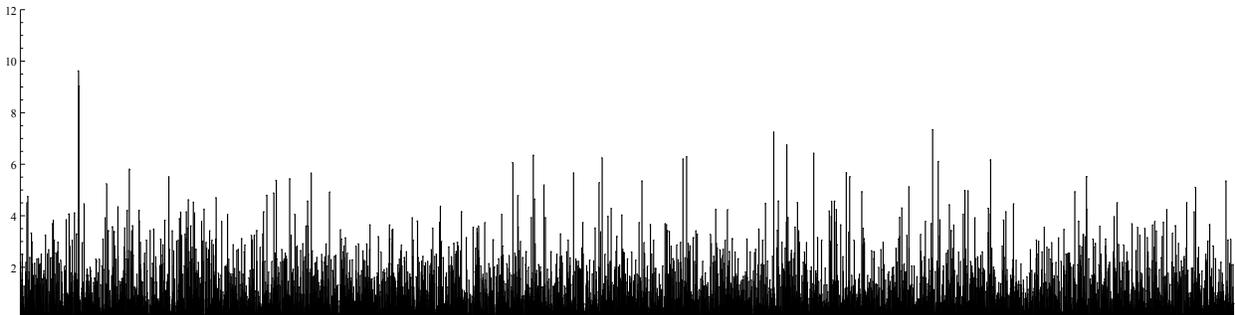


ABBILDUNG 4. Eine mit Parameter 1 exponentialverteilte Stichprobe vom Umfang $n = 5000$.

BEWEIS. Sei $t \in \mathbb{R}$ beliebig. Mit Satz 1.1.2 gilt

$$\mathbb{P}[M_n - \log n \leq t] = \mathbb{P}[M_n \leq t + \log n] = F^n(t + \log n).$$

Die Zufallsvariablen X_i sind exponentialverteilt und $t + \log n > 0$ für n hinreichend groß. Es folgt, dass bei einem hinreichend großen n ,

$$\mathbb{P}[M_n - \log n \leq t] = (1 - e^{-(t+\log n)})^n = \left(1 - \frac{e^{-t}}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{-e^{-t}}.$$

Somit gilt $M_n - \log n \xrightarrow[n \rightarrow \infty]{d} \Lambda$. □

Bemerkung 1.3.3. Gemäß Satz 1.3.2 liegt die Exponentialverteilung im Max–Anziehungsbereich der Gumbel–Verteilung:

$$\text{Exp}(1) \in \text{MDA}(\Lambda).$$

Die Gumbel–Verteilung ist somit eine Extremwertverteilung. Man kann Satz 1.3.2 wie folgt interpretieren: Für großes n nimmt das Maximum M_n Werte an, die sich von dem Wert $\log n$ um eine approximativ Gumbel–verteilte “Fluktuation” unterscheiden.

$$\text{Fréchet–Verteilung } \Phi_\alpha(t) = e^{-t^{-\alpha}}, \quad t > 0$$

Definition 1.3.4. Eine Zufallsvariable heißt Fréchet-verteilt mit Parameter $\alpha > 0$, wenn ihre Verteilungsfunktion die folgende Gestalt hat:

$$\Phi_\alpha(t) = \begin{cases} e^{-t^{-\alpha}}, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

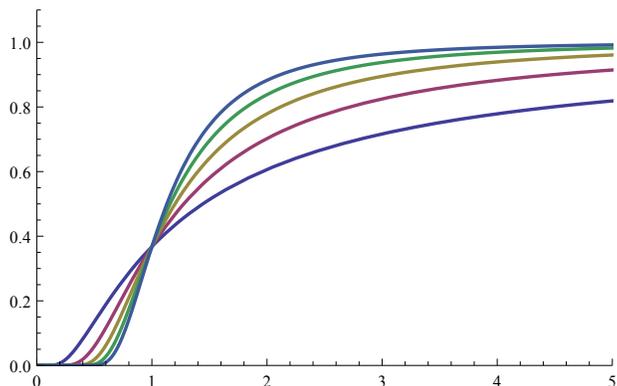


ABBILDUNG 5. Verteilungsfunktionen der Fréchet-Verteilungen mit $\alpha = 1.0, 1.5, 2.0, 2.5, 3.0$.

Der nächste Satz zeigt, dass Fréchet-Verteilung eine Extremwertverteilung ist.

Satz 1.3.5. Die Zufallsvariablen X_1, X_2, \dots seien Pareto-verteilt mit Parameter $\alpha > 0$, d.h.

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

Dann konvergieren für $n \rightarrow \infty$ die Zufallsvariablen $n^{-\frac{1}{\alpha}} M_n$ in Verteilung gegen Φ_α , d.h. es gilt

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n}{n^{1/\alpha}} \leq t \right] = \begin{cases} e^{-t^{-\alpha}}, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

BEWEIS. Sei $t > 0$ beliebig. Mit Satz 1.1.2 erhalten wir, dass

$$\mathbb{P} \left[\frac{M_n}{n^{1/\alpha}} \leq t \right] = \mathbb{P}[M_n \leq tn^{1/\alpha}] = F^n(tn^{1/\alpha}).$$

Da die Zufallsvariablen X_i Pareto-verteilt sind und $tn^{1/\alpha} > 1$ für hinreichend großes n , ergibt sich, dass

$$\mathbb{P} \left[\frac{M_n}{n^{1/\alpha}} \leq t \right] = \left(1 - \frac{1}{(tn^{1/\alpha})^\alpha} \right)^n = \left(1 - \frac{1}{t^\alpha n} \right)^n \xrightarrow{n \rightarrow \infty} e^{-t^{-\alpha}}.$$

Für $t \leq 0$ gilt $\mathbb{P}[n^{-\frac{1}{\alpha}} M_n \leq t] = 0$. Daraus folgt die Behauptung. \square

Bemerkung 1.3.6. Man kann Satz 1.3.5 wie folgt interpretieren: Für großes n nimmt das Maximum M_n sehr große Werte auf der Skala $n^{1/\alpha}$ an. Reskaliert man M_n mit dem Faktor $n^{-1/\alpha}$, so erhält man approximativ Fréchet-verteilte Werte.

Weibull-Verteilung $\Psi_\alpha(t) = e^{-(-t)^\alpha}, t < 0$

Definition 1.3.7. Eine Zufallsvariable heißt **Weibull-verteilt** mit Parameter $\alpha > 0$, wenn ihre Verteilungsfunktion die folgende Form hat:

$$\Psi_\alpha(t) = \begin{cases} e^{-(-t)^\alpha}, & t \leq 0, \\ 1, & t \geq 0. \end{cases}$$

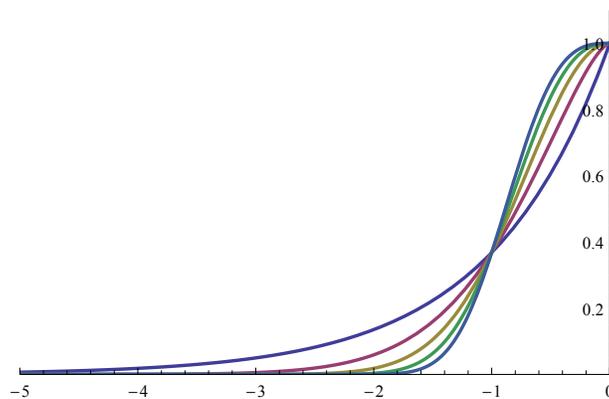


ABBILDUNG 6. Verteilungsfunktionen der Weibull-Verteilungen mit $\alpha = 1.0, 1.5, 2.0, 2.5, 3.0$.

Der nächste Satz zeigt, dass die Weibull-Verteilung eine Extremwertverteilung ist.

Satz 1.3.8. Seien die Zufallsvariablen X_1, X_2, \dots unabhängig mit der Verteilungsfunktion

$$F(t) = \begin{cases} 0, & t \leq -1, \\ 1 - (-t)^\alpha, & t \in [-1, 0], \\ 1, & t \geq 0, \end{cases}$$

wobei $\alpha > 0$ ein Parameter ist. Dann konvergieren für $n \rightarrow \infty$ die Zufallsvariablen $n^{1/\alpha}M_n$ in Verteilung gegen Ψ_α , d.h. es gilt

$$(1.3.1) \quad \lim_{n \rightarrow \infty} \mathbb{P}[n^{1/\alpha}M_n \leq t] = \begin{cases} e^{-(-t)^\alpha}, & t \leq 0, \\ 1, & t \geq 0. \end{cases}$$

BEWEIS. Sei $t \leq 0$ beliebig. Mit Satz 1.1.2 erhalten wir, dass

$$\mathbb{P}[n^{1/\alpha}M_n \leq t] = \mathbb{P}[M_n \leq tn^{-1/\alpha}] = F^n(tn^{-1/\alpha}).$$

Für n hinreichend groß ist $tn^{-1/\alpha} \in [-1, 0]$. Aus der Formel für die Verteilungsfunktion F folgt, dass

$$\mathbb{P}[n^{1/\alpha}M_n \leq t] = (1 - (-tn^{-1/\alpha})^\alpha)^n = \left(1 - \frac{(-t)^\alpha}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{-(-t)^\alpha}.$$

Für $t \geq 0$ gilt $\mathbb{P}[n^{1/\alpha}M_n \leq t] = 1$, denn $M_n \leq 0$ f.s. Daraus folgt die Behauptung. \square

Bemerkung 1.3.9. Man kann Satz 1.3.8 wie folgt interpretieren: Für großes n nähert sich das Maximum M_n dem Wert 0 von unten an. Dabei nimmt M_n sehr kleine negative Werte auf der Skala $n^{-1/\alpha}$ an. Reskaliert man M_n mit dem Faktor $n^{1/\alpha}$, so erhält man approximativ Weibull-verteilte Fluktuationen.

Bemerkung 1.3.10. Eine Fréchet-verteilte Zufallsvariable nimmt nur positive Werte an. Eine Weibull-verteilte Zufallsvariable nimmt nur negative Werte an. Eine Gumbel-verteilte Zufallsvariable kann sowohl positive als auch negative Werte annehmen.

1.4. Satz von Fisher–Tippett–Gnedenko

Wir haben folgende Extremwertverteilungen konstruiert: Die Gumbel-Verteilung Λ , die Fréchet-Verteilung Φ_α (wobei $\alpha > 0$) und die Weibull-Verteilung Ψ_α (wobei $\alpha > 0$). Weitere Beispiele von Extremwertverteilungen können konstruiert werden, indem wir auf die oben genannten Verteilungen lineare Transformationen anwenden.

Definition 1.4.1. Zwei Zufallsvariablen Z_1 und Z_2 sind **vom gleichen Typ**, wenn es $c > 0$ und $d \in \mathbb{R}$ gibt mit

$$Z_1 \stackrel{d}{=} cZ_2 + d.$$

Notation: $Z_1 \bowtie Z_2$.

Bezeichnen wir mit F_1 und F_2 die Verteilungsfunktionen von Z_1 und Z_2 , so kann man die obige Bedingung wie folgt formulieren:

$$F_1(t) = \mathbb{P}[Z_1 \leq t] = \mathbb{P}[cZ_2 + d \leq t] = \mathbb{P}\left[Z_2 \leq \frac{t-d}{c}\right] = F_2\left(\frac{t-d}{c}\right).$$

Definition 1.4.2. Zwei Verteilungsfunktionen F_1 und F_2 sind **vom gleichen Typ**, wenn es $c > 0$ und $d \in \mathbb{R}$ gibt, so dass für alle $t \in \mathbb{R}$

$$F_1(t) = F_2\left(\frac{t-d}{c}\right).$$

Notation: $F_1 \bowtie F_2$.

Beispiel 1.4.3. Die Gleichverteilung auf dem Intervall $[0, 1]$ ist vom gleichen Typ wie die Gleichverteilung auf einem beliebigen Intervall $[a, b]$. Die Normalverteilung mit beliebigen Parametern ist vom gleichen Typ wie die Standardnormalverteilung.

Aufgabe 1.4.4. Zeigen Sie, dass \bowtie eine Äquivalenzrelation ist, d. h.

- (1) $F \bowtie F$.
- (2) $F \bowtie G \Rightarrow G \bowtie F$.
- (3) $F \bowtie G, G \bowtie H \Rightarrow F \bowtie H$.

Proposition 1.4.5. Hat eine Zufallsvariable Z (mit Verteilungsfunktion $G(t)$) eine Extremwertverteilung, so hat für beliebige $c > 0$ und $d \in \mathbb{R}$ auch die Zufallsvariable $cZ + d$ (mit Verteilungsfunktion $G\left(\frac{t-d}{c}\right)$) eine Extremwertverteilung.

BEWEIS. Die Voraussetzung, dass die Zufallsvariable Z einer Extremwertverteilung gehorcht, bedeutet, dass es unabhängige identisch verteilte Zufallsvariablen X_1, X_2, \dots und Folgen $a_n > 0, b_n \in \mathbb{R}$ gibt, so dass

$$\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} Z.$$

Daraus folgt, dass

$$\frac{\max\{cX_1, \dots, cX_n\} - (cb_n - da_n)}{a_n} \xrightarrow[n \rightarrow \infty]{d} cZ + d.$$

Somit hat die Zufallsvariable $cZ + d$ ebenfalls eine Extremwertverteilung. □

Aufgabe 1.4.6. Zeigen Sie, dass die Max-Anziehungsbereiche der Verteilungsfunktionen $G(t)$ und $G\left(\frac{t-d}{c}\right)$ gleich sind.

Aus Proposition 1.4.5 folgt, dass für beliebige $\mu \in \mathbb{R}$ und $\sigma > 0$, die folgenden Verteilungen Extremwertverteilungen sind:

Verteilungen vom Gumbel-Typ:

$$(1.4.1) \quad \Lambda\left(\frac{t - \mu}{\sigma}\right) = \exp\left\{-e^{-\frac{t - \mu}{\sigma}}\right\}, \quad t \in \mathbb{R}.$$

Verteilungen vom Fréchet-Typ (mit Parameter $\alpha > 0$):

$$(1.4.2) \quad \Phi_\alpha\left(\frac{t - \mu}{\sigma}\right) = \begin{cases} \exp\left\{-\left(\frac{t - \mu}{\sigma}\right)^{-\alpha}\right\}, & \text{falls } t > \mu, \\ 0, & \text{sonst.} \end{cases}$$

Verteilungen vom Weibull-Typ (mit Parameter $\alpha > 0$):

$$(1.4.3) \quad \Psi_\alpha\left(\frac{t - \mu}{\sigma}\right) = \begin{cases} \exp\left\{-\left(-\frac{t - \mu}{\sigma}\right)^\alpha\right\}, & \text{falls } t < \mu, \\ 1, & \text{sonst.} \end{cases}$$

Ein zentraler Satz der Extremwerttheorie besagt, dass es keine weiteren Extremwertverteilungen gibt:

Satz 1.4.7 (Fisher–Tippett (1928), Gnedenko (1943)). Jede Extremwertverteilung gehört zu einer der drei Familien (1.4.1), (1.4.2), (1.4.3).

1.5. Jenkinson–von Mises–Darstellung

Es gibt eine Darstellung (*Jenkinson–von Mises–Darstellung*), die alle drei Familien (1.4.1), (1.4.2), (1.4.3) als Spezialfälle beinhaltet. Betrachte nämlich die folgende Familie von Verteilungsfunktionen (parametrisiert durch $\gamma \in \mathbb{R}$)

$$G_\gamma(t) = \begin{cases} \exp\{-(1 + \gamma t)^{-1/\gamma}\}, & \text{falls } 1 + \gamma t > 0, \\ 0, & \text{falls } \gamma > 0 \text{ und } t \leq -1/\gamma, \\ 1, & \text{falls } \gamma < 0 \text{ und } t \geq -1/\gamma. \end{cases}$$

Folgendes lässt sich nun leicht überprüfen:

- (1) Für $\gamma > 0$ ist G_γ vom gleichen Typ wie die Fréchet–Verteilung $\Phi_{1/\gamma}(t) = e^{-t^{-1/\gamma}}$, $t > 0$.
- (2) Für $\gamma < 0$ ist G_γ vom gleichen Typ wie die Weibull–Verteilung $\Psi_{-1/\gamma}(t) = e^{-(-t)^{-1/\gamma}}$, $t < 0$.
- (3) Für $\gamma = 0$ ist $(1 + \gamma t)^{-1/\gamma}$ nicht wohldefiniert. Wir interpretieren diesen Term dann als Grenzwert für $\gamma \rightarrow 0$:

$$\lim_{\gamma \rightarrow 0} (1 + \gamma t)^{-1/\gamma} = e^{-t}.$$

Somit ist $G_0(t) = e^{-e^{-t}}$, $t \in \mathbb{R}$, die Gumbel–Verteilung.

Der Satz von Fisher–Tippett–Gnedenko lässt sich also auch wie folgt formulieren.

Satz 1.5.1 (Fisher–Tippett (1928), Gnedenko (1943)). Jede Extremwertverteilung hat die Form $G_\gamma(ct + d)$ mit passenden Parametern $\gamma \in \mathbb{R}$, $c > 0$, $d \in \mathbb{R}$.

Eine in der Form $G_\gamma(ct + d)$ dargestellte Extremwertverteilung wird in der Statistik auch *GEV–Verteilung* genannt (*General Extreme–Value Distribution*).

KAPITEL 2

Cauchy–Funktionalgleichung

Die Resultate dieses Kapitels haben zwar keinen direkten Bezug zur Extremwerttheorie, werden aber später benötigt, um den Satz von Fisher–Tippett–Gnedenko zu beweisen. Außerdem werden diese Resultate benutzt, um einige Eigenschaften von regulär variierenden Funktionen herzuleiten.

2.1. Additive und multiplikative Funktionen

Definition 2.1.1. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt **additiv**, wenn

$$(2.1.1) \quad f(x + y) = f(x) + f(y) \quad \text{für alle } x, y \in \mathbb{R}.$$

Die Gleichung (2.1.1) heißt **Cauchy–Funktionalgleichung**.

Es ist klar, dass eine Funktion der Form $f(x) = cx$ additiv ist. Gibt es weitere additive Funktionen? Wir werden sehen, dass die Antwort “nein” lautet, allerdings nur wenn man an die Funktion f zusätzliche Bedingungen wie Stetigkeit, Monotonie oder Messbarkeit stellt.

Man kann auch eine multiplikative Version der Cauchy–Funktionalgleichung formulieren.

Definition 2.1.2. Eine Funktion $g : (0, \infty) \rightarrow (0, \infty)$ heißt **multiplikativ**, wenn

$$(2.1.2) \quad g(xy) = g(x)g(y) \quad \text{für alle } x, y > 0.$$

Die Funktion $g(x) = x^c$ ist für jedes $c \in \mathbb{R}$ multiplikativ und es stellt sich die Frage, ob es weitere multiplikative Funktionen gibt.

Vom Standpunkt der Algebra ist eine additive Funktion ein Homomorphismus aus der additiven Gruppe der reellen Zahlen $(\mathbb{R}, +)$ in sich selbst. Analog ist eine multiplikative Funktion ein Homomorphismus der multiplikativen Gruppe der positiven Zahlen $(\mathbb{R}_{>0}, \cdot)$ in sich selbst. Da die Exponentialfunktion und der Logarithmus

$$\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}_{>0}, \cdot), \quad \log : (\mathbb{R}_{>0}, \cdot) \rightarrow (\mathbb{R}, +)$$

einen Isomorphismus (und sein inverses) zwischen den beiden Gruppen definieren, lässt sich aus jeder multiplikativen Funktion g eine additive Funktion

$$f(x) = \log g(e^x), \quad x \in \mathbb{R},$$

basteln. Somit können alle Fragen über multiplikative Funktionen auf entsprechende Fragen über additive Funktionen reduziert werden. In Zukunft werden wir deshalb nur additive Funktionen betrachten.

2.2. Stetige Lösungen

Satz 2.2.1 (Cauchy, 1821). Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sei additiv und überall stetig. Dann hat f die Form $f(x) = cx$ für ein $c \in \mathbb{R}$.

BEWEIS. SCHRITT 1. Mit $x = y = 0$ ergibt sich $f(0) = f(0) + f(0)$, somit ist $f(0) = 0$. Mit $y = -x$ ergibt sich $f(0) = f(x) + f(-x)$, somit ist $f(-x) = -f(x)$ für alle $x \in \mathbb{R}$.

SCHRITT 2. Per Induktion zeigt man, dass die Additivität sogar für eine beliebige Anzahl von Summanden gilt:

$$f(x_1 + \dots + x_n) = f(x_1) + \dots + f(x_n), \quad x_1, \dots, x_n \in \mathbb{R}.$$

SCHRITT 3. Mit $x_1 = \dots = x_n = 1/n$ erhält man $f(1) = nf(1/n)$, also

$$f\left(\frac{1}{n}\right) = \frac{f(1)}{n} = \frac{c}{n}$$

für alle $n \in \mathbb{N}$. Dabei haben wir $c := f(1)$ gesetzt.

SCHRITT 4. Indem wir die Additivität für m Summanden, die alle gleich $1/n$ sind, benutzen, erhalten wir, dass

$$f\left(\frac{m}{n}\right) = mf\left(\frac{1}{n} + \dots + \frac{1}{n}\right) = mf\left(\frac{1}{n}\right) = c\frac{m}{n}$$

für alle $m \in \mathbb{N}$. Somit gilt für alle rationalen Zahlen r , dass $f(r) = cr$, wobei das wegen Schritt 1 auch für negatives r richtig ist.

SCHRITT 5. Bislang haben wir nur die Additivität und keine Stetigkeit benutzt. Sei $x \in \mathbb{R}$ beliebig. Es gibt eine Folge von rationalen Zahlen r_1, r_2, \dots , die gegen x konvergiert. Für jedes n gilt $f(r_n) = cr_n$ wegen Schritt 4. Nun benutzen wir die Stetigkeit:

$$f(x) = \lim_{n \rightarrow \infty} f(r_n) = \lim_{n \rightarrow \infty} cr_n = cx.$$

Somit ist die Behauptung bewiesen. □

Satz 2.2.2 (Darboux, 1875). Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sei additiv und an mindestens einer Stelle z_0 stetig. Dann hat f die Form $f(x) = cx$ für ein $c \in \mathbb{R}$.

BEWEIS. SCHRITT 1. Wir zeigen, dass aus der Stetigkeit an der Stelle z_0 (für additive Funktion f) die Stetigkeit an der Stelle 0 folgt. Die Stetigkeit an der Stelle z_0 bedeutet, dass

$$\lim_{u \rightarrow 0} f(z_0 + u) = f(z_0).$$

Wegen der Additivität gilt aber $f(z_0 + u) = f(z_0) + f(u)$. Die obige Bedingung lässt sich also wie folgt darstellen:

$$\lim_{u \rightarrow 0} f(u) = 0,$$

Somit ist die Funktion f stetig an der Stelle 0.

SCHRITT 2. Wir zeigen, dass f überall stetig ist. Sei $x \in \mathbb{R}$ beliebig. Dann gilt

$$\lim_{u \rightarrow 0} f(x + u) = \lim_{u \rightarrow 0} (f(x) + f(u)) = f(x) + \lim_{u \rightarrow 0} f(u) = f(x),$$

wobei wir Additivität und Stetigkeit an der Stelle 0 (Schritt 1) benutzt haben. Somit ist f überall stetig.

SCHRITT 3. Nun können wir den Satz von Cauchy anwenden. □

Aufgabe 2.2.3. Eine Funktion f sei additiv und auf einem Intervall (a, b) monoton. Zeigen Sie, dass f die Form $f(x) = cx$ hat.

2.3. Satz von Steinhaus

Wir werden im Folgenden zeigen, dass sich die Eigenschaft der Stetigkeit im Satz von Cauchy durch die viel schwächere Eigenschaft der Messbarkeit ersetzen lässt. Dafür benötigen wir den Satz von Steinhaus, der in diesem Abschnitt bewiesen wird.

Definition 2.3.1. Für zwei Mengen $A, B \subset \mathbb{R}$ definieren wir die **Minkowski-Summe**

$$A + B = \{a + b : a \in A, b \in B\}$$

und die **Minkowski-Differenz**

$$A - B = \{a - b : a \in A, b \in B\}.$$

Insbesondere werden wir uns für die Selbstdifferenz $A - A$ einer Menge A interessieren. Es ist klar, dass $0 \in A - A$. Der nächste Satz besagt, dass $A - A$ erstaunlicherweise sogar ein Intervall um 0 enthält, wenn das Lebesgue-Maß von A positiv ist.

Wir bezeichnen das Lebesgue-Maß von A mit $|A|$.

Satz 2.3.2 (Steinhaus, 1920). Sei $A \subset \mathbb{R}$ eine messbare Menge mit $|A| > 0$. Dann gibt es ein $\delta > 0$ mit $(-\delta, \delta) \subset A - A$.

BEWEIS. SCHRITT 1. Ohne Einschränkung der Allgemeinheit können wir annehmen, dass $|A| < \infty$. In der Tat, mindestens eine der Mengen $A_i := A \cap [i, i + 1)$ hat ein strikt positives Lebesgue-Maß. Es reicht, den Satz für eine solche Menge A_i zu beweisen, denn $A_i - A_i \subset A - A$.

SCHRITT 2. Sei also $|A| \neq 0, \infty$. Wir zeigen, dass es ein Intervall I mit $|A \cap I| > \frac{9}{10}|I|$ gibt. Somit wird fast das gesamte Intervall I von A überdeckt. Das Lebesgue-Maß von A wird

wie folgt definiert:

$$|A| = \inf \sum_{k=1}^{\infty} |I_k|,$$

wobei das Infimum über alle Überdeckungen von A durch abzählbar viele Intervalle I_1, I_2, \dots genommen wird. Wegen $|A| \neq 0, \infty$ gibt es also eine Überdeckung mit

$$(2.3.1) \quad \sum_{k=1}^{\infty} |I_k| < \frac{10}{9}|A|.$$

Wäre nun $|A \cap I_k| \leq \frac{9}{10}|I_k|$ für jedes $k \in \mathbb{N}$, so hätten wir

$$|A| \leq \sum_{k=1}^{\infty} |A \cap I_k| \leq \frac{9}{10} \sum_{k=1}^{\infty} |I_k|,$$

was wegen $|A| \neq 0, \infty$ im Widerspruch zu (2.3.1) steht. Somit ist $|A \cap I_k| > \frac{9}{10}|I_k|$ für mindestens ein k und wir können $I = I_k$ nehmen.

SCHRITT 3. Sei $\delta = \frac{1}{2}|I|$. Wir zeigen, dass $(-\delta, \delta) \subset A - A$. Sei also $|x| < \delta$. Es reicht zu zeigen, dass $A \cap (A + x) \neq \emptyset$. Hierzu reicht es zu zeigen, dass

$$(2.3.2) \quad (A \cap I) \cap ((A \cap I) + x) \neq \emptyset.$$

Nun haben die beiden Mengen $A \cap I$ und $(A \cap I) + x$ das gleiche Lebesgue-Maß $\geq \frac{9}{10}|I|$. Auf der anderen Seite sind beide Mengen im Intervall $I \cup (I + x)$ enthalten, dessen Länge wegen $|I| = 2\delta$ und $|x| < \delta$ höchstens $3\delta = \frac{3}{2}|I|$ ist. Da $2 \cdot \frac{9}{10} > \frac{3}{2}$ ist, müssen sich die Mengen $A \cap I$ und $(A \cap I) + x$ schneiden und die Behauptung (2.3.2) ist bewiesen. \square

Aufgabe 2.3.3. Sei $A \subset \mathbb{R}$ eine messbare Menge mit $|A| > 0$. Zeigen Sie ohne Verwendung des Satzes von Steinhaus, dass es in A zwei verschiedene Punkte x, y mit rationalem Abstand $|x - y| \in \mathbb{Q}$ gibt.

Aufgabe 2.3.4. Seien $A, B \subset \mathbb{R}$ zwei messbare Mengen mit $|A| > 0, |B| > 0$. Zeigen Sie, dass die Minkowski-Summe $A + B$ ein Intervall enthält, d.h. es gibt $t \in \mathbb{R}, \delta > 0$ mit $(t - \delta, t + \delta) \subset A + B$.

Wir stellen nun einige Konsequenzen des Satzes von Steinhaus vor, die bei der Untersuchung von regulär variierenden Funktionen nützlich sein werden.

Definition 2.3.5. Eine **additive Untergruppe** von \mathbb{R} ist eine Menge $S \subset \mathbb{R}$ mit den folgenden zwei Eigenschaften:

- (1) aus $x, y \in S$ folgt, dass $x + y \in S$;
- (2) aus $x \in S$ folgt, dass $-x \in S$.

Beispiel 2.3.6. Die Menge der ganzen Zahlen \mathbb{Z} , die Menge der rationalen Zahlen \mathbb{Q} , die Menge der algebraischen Zahlen sind additive Untergruppen von \mathbb{R} .

Der nächste Satz zeigt, dass es keine “massiven” additiven Untergruppen gibt mit Ausnahme von \mathbb{R} selbst.

Satz 2.3.7. Sei S eine additive Untergruppe von \mathbb{R} , die eine messbare Menge mit positivem Maß enthält. Dann ist $S = \mathbb{R}$.

BEWEIS. Wegen der Untergruppeneigenschaft gilt $S - S \subset S$. Auf der anderen Seite gibt es nach dem Satz von Steinhaus ein $\delta > 0$ mit $(-\delta, \delta) \subset S - S$. Also ist $(-\delta, \delta) \subset S$. Wegen der Untergruppeneigenschaft gilt aber auch $(-n\delta, n\delta) \subset S$ für jedes $n \in \mathbb{N}$. Somit ist $S = \mathbb{R}$. \square

Für messbare additive Untergruppen gilt das folgende “Alles-oder-fast-nichts-Prinzip”:

Korollar 2.3.8. Eine additive Untergruppe S von \mathbb{R} sei messbar. Dann ist entweder $|S| = 0$ oder $S = \mathbb{R}$.

Beweis. Folgt aus Satz 2.3.7.

2.4. Messbare Lösungen

Der nächste Satz ist eine Verstärkung des Satzes von Cauchy.

Satz 2.4.1 (Ostrowski, 1929). Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine additive und messbare Funktion. Dann hat f die Form $f(x) = cx$ für ein $c \in \mathbb{R}$.

BEWEIS. SCHRITT 1. Wir zeigen, dass es eine Menge mit positivem Lebesgue-Maß gibt, auf der die Funktion $|f|$ beschränkt ist. Betrachte hierzu die Mengen $A_n := \{x \in \mathbb{R} : |f(x)| \leq n\}$. Die Vereinigung dieser Mengen ist \mathbb{R} , also hat mindestens eine Menge A_n strikt positives Maß. Dabei ist $|f|$ auf der Menge A_n durch n beschränkt.

SCHRITT 2. Wir zeigen, dass $|f|$ auf einer Umgebung von 0 beschränkt ist. Wir wissen, dass $|f(a)| \leq n$ für $a \in A_n$ und $|A_n| > 0$. Nach dem Satz von Steinhaus gibt es ein $\delta > 0$ mit $(-\delta, \delta) \subset A_n - A_n$. Jedes $x \in (-\delta, \delta)$ kann somit in der Form $x = a - a'$ mit $a, a' \in A_n$ dargestellt werden und

$$|f(x)| = |f(a - a')| = |f(a) - f(a')| \leq |f(a)| + |f(a')| \leq 2n.$$

Somit ist f auf dem Intervall $(-\delta, \delta)$ durch $2n$ beschränkt.

SCHRITT 3. Wir zeigen, dass $|f|$ an der Stelle 0 stetig ist. Sei $m \in \mathbb{N}$ beliebig und $|t| < \delta/m$. Somit gilt $tm \in (-\delta, \delta)$ und $|f(tm)| \leq 2n$. Auf der anderen Seite gilt wegen der Additivität von f

$$f(tm) = f(t) + \dots + f(t) = mf(t).$$

Es folgt, dass $|f(t)| < 2n/m$ für alle t mit $|t| < \delta/m$. Für $m \rightarrow \infty$ erhalten wir $\lim_{x \rightarrow 0} f(x) = 0 = f(0)$. Somit ist f stetig an der Stelle 0.

SCHRITT 4. Nun können wir den Satz von Darboux anwenden. □

Eine analoge Aussage gilt für multiplikative Funktionen.

Korollar 2.4.2. Sei $g : (0, \infty) \rightarrow (0, \infty)$ eine multiplikative und messbare Funktion. Dann hat g die Form $g(x) = x^c$ für ein $c \in \mathbb{R}$.

Beweis. Die Funktion $f(x) = \log g(e^x)$ ist additiv und messbar. Nach dem Satz von Ostrowski hat sie die Form $f(x) = cx$, woraus sich $g(x) = e^{f(\log x)} = x^c$ ergibt.

2.5. Hamel–Funktionen

Wir haben gezeigt, dass jede additive Funktion f die Form $f(x) = cx$ hat, wenn man eine zusätzliche Bedingung (wie z.B. Messbarkeit) an f stellt. Kann man die zusätzlichen Bedingungen komplett entfernen? Erstaunlicherweise ist die Antwort negativ.

Beispiel 2.5.1 (Hamel, 1905). Wir konstruieren eine additive Funktion f , die nicht die Form $f(x) = cx$ hat.

\mathbb{R} als Vektorraum über \mathbb{R} hat Dimension 1. Man kann aber \mathbb{R} auch als einen Vektorraum über dem Körper \mathbb{Q} der rationalen Zahlen betrachten. Nach einem allgemeinen Satz aus der Algebra hat jeder Vektorraum über jedem Körper eine Basis. Das heißt, es gibt eine Menge B mit der Eigenschaft, dass jedes $x \in \mathbb{R}$ eine eindeutige Darstellung

$$x = r_1 b_1 + \dots + r_n b_n$$

mit passenden $n \in \mathbb{N}$, $r_1, \dots, r_n \in \mathbb{Q} \setminus \{0\}$ und $b_1, \dots, b_n \in B$ besitzt.

Wir definieren nun eine additive Funktion indem wir ihre Werte auf der Basis B beliebig vorgeben und sie dann linear (über \mathbb{Q} !) auf ganz \mathbb{R} fortsetzen. Sei $b_0 \in B$ beliebig. Definiere $f(b_0) = 1$ und $f(b) = 0$ für alle $b \in B \setminus \{b_0\}$. Die lineare (über \mathbb{Q} !) Fortsetzung von f ist definiert durch

$$f(x) = r_1 f(b_1) + \dots + r_n f(b_n).$$

Diese Funktion ist additiv per Definition. Allerdings hat sie nicht die Form $x \mapsto cx$, denn sie ist nicht identisch gleich 0 (was die Möglichkeit $c = 0$ ausschließt) und sie hat unendlich viele Nullstellen (wohingegen die Funktion $x \mapsto cx$ nur eine Nullstelle für $c \neq 0$ hat).

Die so konstruierte Funktion f ist nicht messbar (nach dem Satz von Ostrowski), nirgends stetig (nach dem Satz von Darboux), auf keinem Intervall monoton (Aufgabe 2.2.3), unbeschränkt auf jedem Intervall und sogar auf jeder Menge mit positivem Maß (das folgt aus dem Beweis des Satzes von Ostrowski). Sie ist ein wahres Monstrum.

Satz von Fisher–Tippett–Gnedenko

In diesem Kapitel beweisen wir den Satz von Fisher–Tippett–Gnedenko, der die Extremwertverteilungen beschreibt.

Satz 3.0.2 (Satz von Fisher–Tippett (1928), Gnedenko (1943)). Eine Verteilungsfunktion G ist eine Extremwertverteilung genau dann, wenn einer der drei folgenden Fälle eintritt:

- (1) G ist vom Gumbel–Typ, d.h. $G(t) = \Lambda\left(\frac{t-\mu}{\sigma}\right)$ für $\mu \in \mathbb{R}$, $\sigma > 0$.
- (2) G ist vom Fréchet–Typ, d.h. $G(t) = \Phi_\alpha\left(\frac{t-\mu}{\sigma}\right)$ für $\alpha > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.
- (3) G ist vom Weibull–Typ, d.h. $G(t) = \Psi_\alpha\left(\frac{t-\mu}{\sigma}\right)$ für $\alpha > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.

Für den Beweis benötigen wir einige Hilfsmittel.

3.1. Eindeutigkeit der Normierungskonstanten

Seien Z_1, Z_2, \dots Zufallsvariablen und $a_n > 0$ und $b_n \in \mathbb{R}$ Folgen von Normierungskonstanten mit

$$(3.1.1) \quad \frac{Z_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} Z.$$

Wir stellen uns die Frage, wie stark wir die Konstanten a_n, b_n verändern können, ohne dass die Konvergenz in (3.1.1) zerstört wird.

Proposition 3.1.1. Es gelte (3.1.1). Seien $\tilde{a}_n > 0$ und $\tilde{b}_n \in \mathbb{R}$ zwei weitere Folgen mit

$$\lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{a_n} = 1, \quad \lim_{n \rightarrow \infty} \frac{\tilde{b}_n - b_n}{a_n} = 0.$$

Dann gilt auch

$$\frac{Z_n - \tilde{b}_n}{\tilde{a}_n} \xrightarrow[n \rightarrow \infty]{d} Z.$$

BEWEIS. Es gilt

$$(3.1.2) \quad \frac{Z_n - \tilde{b}_n}{\tilde{a}_n} = \frac{a_n}{\tilde{a}_n} \cdot \frac{Z_n - \tilde{b}_n}{a_n} = \frac{a_n}{\tilde{a}_n} \cdot \left(\frac{Z_n - b_n}{a_n} - \frac{\tilde{b}_n - b_n}{a_n} \right).$$

Nach dem Lemma von Slutsky konvergiert die rechte Seite gegen Z in Verteilung. \square

Nun stellen wir eine allgemeinere Frage: Wie stark können wir die Folgen a_n, b_n verändern, so dass die Konvergenz in (3.1.1) bestehen bleibt, allerdings eventuell mit einem anderen Grenzwert als Z ?

Proposition 3.1.2. Es gelte (3.1.1). Seien $\tilde{a}_n > 0$ und $\tilde{b}_n \in \mathbb{R}$ zwei weitere Folgen mit

$$\lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{a_n} = a \in (0, \infty), \quad \lim_{n \rightarrow \infty} \frac{\tilde{b}_n - b_n}{a_n} = b \in \mathbb{R}.$$

Dann gilt auch

$$\frac{Z_n - \tilde{b}_n}{\tilde{a}_n} \xrightarrow[n \rightarrow \infty]{d} \frac{Z - b}{a}.$$

BEWEIS. Wir können nach wie vor (3.1.2) verwenden. Nach dem Lemma von Slutsky konvergiert die rechte Seite gegen $(Z - b)/a$ in Verteilung. \square

Satz 3.1.3 (“Convergence of types theorem”, Chintschin). Seien Z_1, Z_2, \dots Zufallsvariablen und $a_n > 0, b_n \in \mathbb{R}$ sowie $\tilde{a}_n > 0, \tilde{b}_n \in \mathbb{R}$ Normierungsfolgen mit

$$\frac{Z_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} Z, \quad \frac{Z_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} \tilde{Z},$$

wobei die Zufallsvariablen Z, \tilde{Z} nicht degeneriert seien. Dann existieren die Grenzwerte

$$(3.1.3) \quad a := \lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{a_n} \in (0, \infty), \quad b := \lim_{n \rightarrow \infty} \frac{\tilde{b}_n - b_n}{a_n} \in \mathbb{R}$$

und es gilt $\tilde{Z} \stackrel{d}{=} (Z - b)/a$.

BEWEIS. Weggelassen. Referenz: P. Billingsley, *Probability and measure*, Seite 193, Thm. 4.2. \square

Bemerkung 3.1.4. Bezeichnen wir mit G_1, G_2, \dots bzw. G, \tilde{G} die Verteilungsfunktionen von Z_1, Z_2, \dots bzw. Z, \tilde{Z} , so können wir den obigen Satz auch wie folgt formulieren: Aus

$$\lim_{n \rightarrow \infty} G_n(a_n t + b_n) = G(t), \quad \lim_{n \rightarrow \infty} G_n(\tilde{a}_n t + \tilde{b}_n) = \tilde{G}(t),$$

mit nicht degenerierten Grenzwertverteilungen G und \tilde{G} folgt, dass (3.1.3) gilt und dass $\tilde{G}(t) = G(at + b)$.

Lemma 3.1.5. Sei F eine nicht degenerierte Verteilungsfunktion und $c > 0, d \in \mathbb{R}$ Konstanten mit $F(ct + d) = F(t)$ für alle $t \in \mathbb{R}$. Dann ist $c = 1, d = 0$.

BEWEIS. Weggelassen. Referenz: P. Billingsley, ab Seite 193. \square

3.2. Max-stabile Verteilungen

Definition 3.2.1. Eine nicht degenerierte Verteilungsfunktion G heißt **max-stabil**, falls es für alle $n \in \mathbb{N}$ Konstanten $c_n > 0$ und $d_n \in \mathbb{R}$ gibt mit

$$G^n(c_n t + d_n) = G(t).$$

Mit anderen Worten, für jedes $n \in \mathbb{N}$ ist G^n vom gleichen Typ wie G .

Sind X_1, X_2, \dots u.i.v. Zufallsvariablen mit einer max-stabilen Verteilung, dann gilt für alle $n \in \mathbb{N}$

$$\frac{\max\{X_1, \dots, X_n\} - d_n}{c_n} \stackrel{d}{=} X_1.$$

Das heißt, das Maximum von n u.i.v. Zufallsvariablen mit einer max-stabilen Verteilung hat bis auf eine affine Transformation die gleiche Verteilung wie eine einzige Zufallsvariable.

Beispiel 3.2.2. Die Gumbel-Verteilungsfunktion $\Lambda(t) = e^{-e^{-t}}$ ist max-stabil, denn

$$\Lambda^n(t + \log n) = e^{-ne^{-(t+\log n)}} = e^{-e^{-t}} = \Lambda(t).$$

Analog lässt sich zeigen, dass Fréchet-Verteilung Φ_α und Weibull-Verteilung Ψ_α max-stabil sind.

Die Klasse der max-stabilen Verteilungen stimmt mit der Klasse der Extremwertverteilungen überein:

Satz 3.2.3. Eine Verteilungsfunktion G ist max-stabil genau dann, wenn G eine Extremwertverteilung ist.

BEWEIS. SCHRITT 1. “ \Rightarrow ” Sei G max-stabil. Dann gibt es $c_n > 0, d_n \in \mathbb{R}$, so dass $G^n(c_n t + d_n) = G(t)$. Es gilt also für alle $t \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} G^n(c_n t + d_n) = G(t),$$

weshalb G eine Extremwertverteilung ist. Wir haben übrigens gezeigt, dass die Verteilungsfunktion G in ihrem eigenen Max-Anziehungsbereich liegt.

SCHRITT 2. “ \Leftarrow ” Sei G eine Extremwertverteilung. Dann gibt es eine Verteilungsfunktion F und $a_n > 0, b_n \in \mathbb{R}$, so dass

$$(3.2.1) \quad \lim_{n \rightarrow \infty} F^n(a_n t + b_n) = G(t)$$

für alle Stetigkeitspunkte von G . Damit gilt für alle $k \in \mathbb{N}$

$$\lim_{n \rightarrow \infty} F^{nk}(a_{nk} t + b_{nk}) = G(t).$$

Indem wir die k -te Wurzel ziehen, erhalten wir

$$(3.2.2) \quad \lim_{n \rightarrow \infty} F^n(a_{nk} t + b_{nk}) = G^{1/k}(t).$$

Wir wenden nun den Chintschin–Satz 3.1.3 auf (3.2.1) und (3.2.2) an. Es folgt, dass G und $G^{1/k}$ vom gleichen Typ sind, d.h. es gibt $c_k > 0$ und $d_k \in \mathbb{R}$ mit

$$G^{1/k}(t) = G(c_k t + d_k) \quad \text{bzw.} \quad G(t) = G^k(c_k t + d_k).$$

Das bedeutet aber, dass G max-stabil ist. □

Von nun an besteht unser Ziel darin, die max-stabilen Verteilungen zu beschreiben. Die nächste Proposition zeigt, dass die Eigenschaft $G^n(c_n t + d_n) = G(t)$ sogar auf nicht-ganzzahlige Werte von n erweitert werden kann.

Proposition 3.2.4. Sei G eine max-stabile Verteilungsfunktion. Dann gibt es messbare Funktionen $c : (0, \infty) \rightarrow (0, \infty)$ und $d : (0, \infty) \rightarrow \mathbb{R}$, so dass für alle $s > 0$ (nicht notwendigerweise ganzzahlig) gilt:

$$(3.2.3) \quad G^s(c(s)t + d(s)) = G(t).$$

BEWEIS. Wir bezeichnen mit $[t]$ die Gaußklammer einer reellen Zahl t :

$$[t] = \max\{n \in \mathbb{Z} \mid n \leq t\}.$$

Sei G eine max-stabile Verteilungsfunktion. Dann gibt es $c_n > 0, d_n \in \mathbb{R}$, so dass für alle $n \in \mathbb{N}$

$$(3.2.4) \quad G^n(c_n t + d_n) = G(t).$$

Für beliebiges $s > 0$ folgt daraus, dass

$$G^{[ns]}(c_{[ns]}t + d_{[ns]}) = G(t).$$

Daraus ergibt sich, dass für alle $t \in \mathbb{R}$

$$(3.2.5) \quad G^n(c_{[ns]}t + d_{[ns]}) = (G^{[ns]}(c_{[ns]}t + d_{[ns]}))^{1/[ns]} = G^{1/[ns]}(t) \xrightarrow{n \rightarrow \infty} G^{1/s}(t).$$

Gleichzeitig gilt aber wegen (3.2.4) auch

$$(3.2.6) \quad G^n(c_n t + d_n) \xrightarrow{n \rightarrow \infty} G(t).$$

Mit Satz 3.1.3 folgt aus (3.2.5) und (3.2.6), dass die folgenden Grenzwerte existieren:

$$(3.2.7) \quad c(s) := \lim_{n \rightarrow \infty} \frac{c_{[ns]}}{c_n} \in (0, \infty), \quad d(s) := \lim_{n \rightarrow \infty} \frac{d_{[ns]} - d_n}{c_n} \in \mathbb{R},$$

und dass $G^{1/s}(t) = G(c(s)t + d(s))$. Insgesamt folgt also $G(t) = G^s(c(s)t + d(s))$.

Außerdem folgt aus der Darstellung (3.2.7), dass die Funktionen c und d als punktweise Grenzwerte von Folgen messbarer Funktionen, selber messbar sind. □

3.3. Charakterisierung der max-stabilen Verteilungen

Wegen Satz 3.2.3 können wir den Satz von Fisher–Tippett–Gnedenko nun wie folgt formulieren:

Satz 3.3.1. Jede max-stabile Verteilungsfunktion G ist vom gleichen Typ wie eine der folgenden Verteilungen: Gumbel Λ , Fréchet Φ_α mit $\alpha > 0$ oder Weibull Ψ_α mit $\alpha > 0$.

BEWEIS. Sei G eine max-stabile Verteilungsfunktion. Laut Proposition 3.2.4 gibt es messbare Funktionen $c(s) > 0$, $d(s) \in \mathbb{R}$ mit

$$(3.3.1) \quad G^s(c(s)t + d(s)) = G(t) \text{ für alle } s > 0, t \in \mathbb{R}.$$

Im Folgenden werden wir diese Funktionalgleichung lösen. Zuerst werden wir G eliminieren und die Funktionen c und d bestimmen.

SCHRITT 1. Für $s > 0$ betrachte die affine Transformation $\varphi_s : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\varphi_s(t) = c(s)t + d(s), \quad t \in \mathbb{R}.$$

Es gilt mit dieser Notation für alle $s > 0$ und $t \in \mathbb{R}$, dass

$$(3.3.2) \quad G(\varphi_s(t)) = G^{1/s}(t).$$

Wir zeigen, dass die Abbildung $s \mapsto \varphi_s$ ein Homomorphismus aus der multiplikativen Gruppe der positiven reellen Zahlen $(\mathbb{R}_{>0}, \cdot)$ in die Gruppe der affinen Transformationen von \mathbb{R} ist, d.h. für alle $s_1, s_2 > 0$, $t \in \mathbb{R}$

$$(3.3.3) \quad \varphi_{s_1 s_2}(t) = \varphi_{s_2}(\varphi_{s_1}(t)).$$

Indem wir in (3.3.2) mehrmals benutzen, erhalten wir, dass

$$G(\varphi_{s_1 s_2}(t)) = G^{1/(s_1 s_2)}(t) = (G^{1/s_1}(t))^{1/s_2} = (G(\varphi_{s_1}(t)))^{1/s_2} = G(\varphi_{s_2}(\varphi_{s_1}(t))).$$

Aus Lemma 3.1.5 folgt (3.3.3).

SCHRITT 2. Mit $\varphi_s(t) = c(s)t + d(s)$ erhalten wir

$$\begin{aligned} \varphi_{s_1 s_2}(t) &= c(s_1 s_2)t + d(s_1 s_2), \\ \varphi_{s_2}(\varphi_{s_1}(t)) &= c(s_2)(c(s_1)t + d(s_1)) + d(s_2). \end{aligned}$$

Somit führt (3.3.3) zum folgenden System von Funktionalgleichungen: Für alle $s_1, s_2 > 0$ und alle $t \in \mathbb{R}$,

$$\begin{cases} c(s_1 s_2) = c(s_1)c(s_2), \\ d(s_1 s_2) = c(s_2)d(s_1) + d(s_2). \end{cases}$$

Die erste Gleichung ist eine Cauchy-Funktionalgleichung in multiplikativer Form. Da c eine messbare Funktion ist, gibt es ein $\rho \in \mathbb{R}$ mit

$$c(s) = s^\rho, \quad s > 0.$$

SCHRITT 3. Wir betrachten zunächst den Fall $\rho = 0$, also $c(s) = 1$. Die Funktionalgleichung für d sieht folgendermaßen aus: Für alle $s_1, s_2 > 0$

$$d(s_1 s_2) = d(s_1) + d(s_2).$$

Betrachte die Funktion $h(x) = d(e^x)$, $x \in \mathbb{R}$. Für diese gilt

$$h(x_1 + x_2) = d(e^{x_1 + x_2}) = d(e^{x_1} e^{x_2}) = d(e^{x_1}) + d(e^{x_2}) = h(x_1) + h(x_2).$$

Somit ist h additiv. Da h messbar ist, folgt aus dem Satz von Ostrowski, dass $h(x) = ax$ für ein $a \in \mathbb{R}$ und somit $d(s) = a \log s$. Im Fall $\rho = 0$ haben wir also die folgende Lösung erhalten:

$$c(s) = 1, \quad d(s) = a \log s.$$

Die Gleichung (3.3.1) für die Verteilungsfunktion G vereinfacht sich also zu

$$G(t + a \log s) = G^{1/s}(t).$$

Der Fall $a = 0$ ist ausgeschlossen. In diesem Fall wäre nämlich $G(t) = G^{1/s}(t)$ für alle $s > 0$, was bedeuten würde, dass $G(t)$ nur die Werte 0 oder 1 annimmt. Widerspruch, denn G ist nicht degeneriert. Sei also $a \neq 0$. Mit $t = 0$ und $y = a \log s \in \mathbb{R}$ erhalten wir

$$G(y) = G^{1/s}(0) = (G(0))^{e^{-y/a}} = \exp \left\{ (\log G(0)) \cdot e^{-\frac{y}{a}} \right\} = \exp \left\{ -e^{-\frac{y}{a} + \log(-\log G(0))} \right\}.$$

Das gilt wegen $a \neq 0$ für alle $y \in \mathbb{R}$. Es sei bemerkt, dass $\log G(0) < 0$ und dass die Fälle $G(0) = 1$ oder 0 ausgeschlossen sind, denn sonst wäre G identisch 0 oder 1, Widerspruch. Somit ist G vom Gumbel-Typ.

SCHRITT 4. Nun betrachten wir den Fall $\rho \neq 0$. Es gilt $c(s) = s^\rho$ und die Gleichung für d nimmt die folgende Form an: Für alle $s_1, s_2 > 0$

$$d(s_1 s_2) = s_2^\rho d(s_1) + d(s_2).$$

Indem wir s_1 und s_2 vertauschen, erhalten wir

$$d(s_2 s_1) = s_1^\rho d(s_2) + d(s_1).$$

Somit ergibt sich für alle $s_1, s_2 > 0$

$$s_2^\rho d(s_1) + d(s_2) = s_1^\rho d(s_2) + d(s_1).$$

Sei nun $s_2 = 2$. Aus $\rho \neq 0$ folgt, dass $2^\rho - 1 \neq 0$ und wir erhalten, dass für alle $s_1 > 0$

$$d(s_1) = \frac{d(2)}{2^\rho - 1} (s_1^\rho - 1) = (s_1^\rho - 1)\mu,$$

wobei $\mu = \frac{d(2)}{2^\rho - 1} \in \mathbb{R}$. Wir haben die Funktionen c und d bestimmt:

$$c(s) = s^\rho, \quad d(s) = (s^\rho - 1)\mu,$$

wobei $\rho \in \mathbb{R}$ und $\mu \in \mathbb{R}$ zwei Parameter sind.

Die Gleichung (3.3.1) für die Verteilungsfunktion G nimmt somit die folgende Gestalt an: Für alle $s > 0, t \in \mathbb{R}$,

$$G^s(s^\rho t + (s^\rho - 1)\mu) = G(t).$$

Betrachte die Verteilungsfunktion $H(u) = G(u - \mu)$, $u \in \mathbb{R}$. Die Gleichung für H sieht wie folgt aus: Für alle $s > 0, t \in \mathbb{R}$,

$$H^s(s^\rho t) = H(t).$$

Mit $t = 0$ erhalten wir $H^s(0) = H(0)$ für alle $s > 0$, somit ist $H(0) = 0$ oder $H(0) = 1$.

FALL 1. Sei $H(0) = 0$. Da H eine Verteilungsfunktion ist, gilt $H(y) = 0$ für alle $y \leq 0$. Sei nun $y > 0$. Mit $t = 1$ und $s = y^{1/\rho} > 0$ erhalten wir

$$H(y) = H^{1/s}(1) = H^{y^{-1/\rho}}(1) = \exp \left\{ (\log H(1)) \cdot y^{-1/\rho} \right\}, \quad y > 0.$$

Es sei bemerkt, dass $\log H(1) < 0$. Es folgt, dass H und somit auch G vom gleichen Typ wie die Fréchet-Verteilung $\Phi_{1/\rho}$ ist.

FALL 2. Sei $H(0) = 1$. Da H eine Verteilungsfunktion ist, gilt $H(y) = 1$ für alle $y \geq 0$. Sei nun $y < 0$. Mit $t = -1$ und $s = (-y)^{1/\rho} > 0$ erhalten wir

$$H(y) = H^{1/s}(-1) = H^{(-y)^{-1/\rho}}(-1) = \exp \{ (\log H(-1)) \cdot (-y)^{-1/\rho} \}, \quad y < 0.$$

Es sei bemerkt, dass $\log H(-1) < 0$. Es folgt, dass H und somit auch G vom gleichen Typ wie die Weibull-Verteilung $\Psi_{-1/\rho}$ ist. □

Regulär variierende Funktionen

Unser nächstes Ziel ist es, die Max-Anziehungsbereiche der Extremwertverteilungen zu beschreiben. Dies wird im nächsten Kapitel geschehen. Wir haben bereits gesehen, dass die Verteilung mit der Tailfunktion $\bar{F}(x) = x^{-\alpha}$, $x \geq 1$, im Max-Anziehungsbereich der Fréchet-Verteilung Φ_α liegt. Wir werden später sehen, dass der Max-Anziehungsbereich von Φ_α aus allen Verteilungen besteht, deren Tailfunktionen sich im gewissen Sinne “wie $x^{-\alpha}$ ” verhalten. Um die exakte Bedingung zu formulieren, brauchen wir den Begriff der regulären Variation, den wir in diesem Kapitel einführen. Wir werden hier nur auf einige Aspekte der regulären Variation eingehen, für eine umfassende Darstellung dieses Gebiets verweisen wir auf das Buch von N. H. Bingham, C. M. Goldie, J. L. Teugels “*Regular Variation*”.

4.1. Definition der regulär variierenden Funktionen

Definition 4.1.1 (Karamata, 1930). Eine Funktion $L : (A, \infty) \rightarrow (0, \infty)$ ist **langsam variierend** in $+\infty$, wenn

$$\lim_{x \rightarrow +\infty} \frac{L(\lambda x)}{L(x)} = 1 \text{ für jedes } \lambda > 0.$$

Bemerkung 4.1.2. Der genaue Wert von A spielt keine Rolle, denn in der obigen Definition geht es nur um das Verhalten der Funktion L für $x \rightarrow +\infty$. Es hat keinen Einfluss auf die Eigenschaft der langsamen Variation, wenn man die Funktion L auf einem endlichen Intervall beliebig verändert.

Beispiel 4.1.3. Eine Funktion L , für die der Grenzwert $c := \lim_{x \rightarrow +\infty} L(x)$ in $(0, \infty)$ existiert, ist langsam variierend, denn

$$\lim_{x \rightarrow +\infty} \frac{L(\lambda x)}{L(x)} = \frac{c}{c} = 1.$$

Ist aber der Grenzwert c gleich $+\infty$ oder 0 , so muss die Funktion nicht langsam variierend sein. Zum Beispiel sind die Funktionen $f_1(x) = x$ und $f_2(x) = 1/x$ nicht langsam variierend.

Beispiel 4.1.4. Die Funktion $L(x) = c(\log x)^\beta$, mit $c > 0$, $\beta \in \mathbb{R}$ ist langsam variierend, denn für jedes $\lambda > 0$ gilt

$$\frac{L(\lambda x)}{L(x)} = \frac{c(\log(\lambda x))^\beta}{c(\log x)^\beta} = \left(\frac{\log x + \log \lambda}{\log x} \right)^\beta \rightarrow 1$$

für $x \rightarrow +\infty$.

Beispiel 4.1.5. Die Funktionen

$$\log x, \quad \log \log x, \quad \log \log \log x, \dots$$

sind langsam variierend. Funktionen der Form

$$c(\log x)^{\beta_1}(\log \log x)^{\beta_2} \dots (\log \dots \log x)^{\beta_k}$$

mit $c > 0$, $\beta_1, \dots, \beta_k \in \mathbb{R}$ sind ebenfalls langsam variierend.

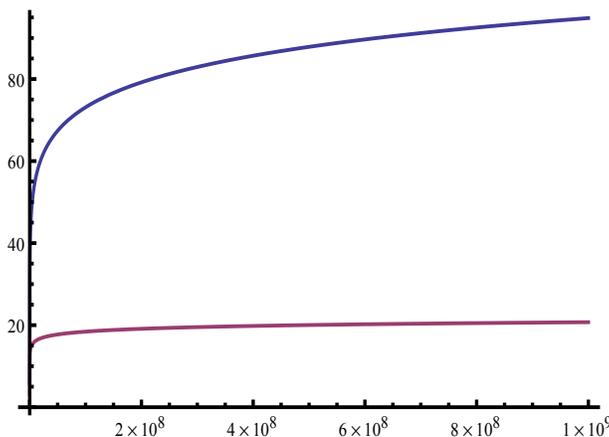


ABBILDUNG 1. Zwei langsam variierende Funktionen: $\exp\{\sqrt[3]{\log x}\}$ (blau) und $\log x$ (rot).

Beispiel 4.1.6. Für $\alpha \neq 0$ ist die Funktion $f(x) = x^\alpha$ nicht langsam variierend. Wir werden später sehen, dass für jedes $\varepsilon > 0$ eine langsam variierende Funktion nicht schneller als x^ε steigen darf und nicht schneller als $x^{-\varepsilon}$ gegen 0 gehen darf.

Beispiel 4.1.7. Die Funktionen $L(x) = e^{(\log x)^\beta}$ sind für alle $|\beta| < 1$ langsam variierend (Übungsaufgabe). Für $\beta = \pm 1$ ergeben sich die Funktionen $L(x) = x$ bzw. $L(x) = 1/x$, die nicht langsam variierend sind. Allerdings ist die Funktion

$$L_1(x) = e^{\log x / \log \log x}$$

langsam variierend (Übungsaufgabe).

Beispiel 4.1.8. Nicht jede beschränkte Funktion $L : (0, \infty) \rightarrow (0, \infty)$ ist langsam variierend. Beispiel: $L(x) = 2 + \sin x$ ist nicht langsam variierend (Übung).

Beispiel 4.1.9. Es gibt eine langsam variierende Funktion L , die zwischen 0 und $+\infty$ oszilliert in dem Sinne, dass $\limsup_{x \rightarrow \infty} L(x) = +\infty$ und gleichzeitig $\liminf_{x \rightarrow \infty} L(x) = 0$. Ein Beispiel ist gegeben durch (Übungsaufgabe)

$$(4.1.1) \quad L(x) = \exp \left\{ \sqrt{\log x} \cdot \cos \left(\sqrt[3]{\log x} \right) \right\}, \quad x \geq 1.$$

Diese Funktion oszilliert unendlich oft zwischen der “unteren Grenze” $e^{-\sqrt{\log x}}$ (die im Fall $\cos(\sqrt[3]{\log x}) = -1$ erreicht wird) und der “oberen Grenze” $e^{\sqrt{\log x}}$ (die im Fall $\cos(\sqrt[3]{\log x}) = 1$ erreicht wird). Allerdings sind die Oszillationen sehr langsam: Zum Beispiel erreicht L die obere Grenze $e^{\sqrt{\log x}}$ an den Stellen $e^{8\pi^3 k^3}$, $k = 0, 1, 2, \dots$

Aufgabe 4.1.10. Seien L_1 und L_2 langsam variierende Funktionen. Zeigen Sie, dass $L_1 L_2$ und $L_1 + L_2$ ebenfalls langsam variierend sind.

Definition 4.1.11 (Karamata, 1930). Eine messbare Funktion $R : (A, \infty) \rightarrow (0, \infty)$ heißt **regulär variierend** in $+\infty$ mit Index $\alpha \in \mathbb{R}$, falls

$$\lim_{x \rightarrow +\infty} \frac{R(\lambda x)}{R(x)} = \lambda^\alpha \text{ für jedes } \lambda > 0.$$

Bezeichnung: $f \in \text{RV}_\alpha$.

Bemerkung 4.1.12. Eine Funktion ist langsam variierend genau dann, wenn sie regulär variierend mit Index $\alpha = 0$ ist.

Beispiel 4.1.13. Die Funktion $R(x) = cx^\alpha$, wobei $c > 0$, ist regulär variierend mit Index α , denn

$$\frac{R(\lambda x)}{R(x)} = \frac{c(\lambda x)^\alpha}{cx^\alpha} = \lambda^\alpha \text{ für alle } \lambda > 0.$$

Beispiel 4.1.14. Die Funktion $f(x) = e^{[\log x]}$ (wobei $[\cdot]$ die Gauß-Klammer ist) ist nicht regulär variierend obwohl die sehr ähnliche Funktion $R(x) = e^{\log x} = x$ regulär variierend mit Index $\alpha = 1$ ist.

Aufgabe 4.1.15. Seien $f \in \text{RV}_\alpha$ und $g \in \text{RV}_\beta$. Zeigen Sie, dass $fg \in \text{RV}_{\alpha+\beta}$.

Viele Beispiele von regulär variierenden Funktionen können konstruiert werden, indem man eine langsam variierende Funktion $L(x)$ mit x^α multipliziert. Dadurch wird die Klasse der regulär variierenden Funktionen ausgeschöpft.

Satz 4.1.16. Sei R eine regulär variierende Funktion mit Index α . Dann gibt es eine langsam variierende Funktion L , sodass

$$R(x) = x^\alpha L(x).$$

BEWEIS. Setze $L(x) = \frac{R(x)}{x^\alpha}$. Dann muss man nur noch zeigen, dass $L(x)$ langsam variierend ist:

$$\frac{L(\lambda x)}{L(x)} = \frac{R(\lambda x)/(\lambda x)^\alpha}{R(x)/x^\alpha} = \lambda^{-\alpha} \frac{R(\lambda x)}{R(x)} \rightarrow \lambda^{-\alpha} \cdot \lambda^\alpha = 1, \quad x \rightarrow +\infty,$$

da R nach Voraussetzung regulär variierend ist. Deshalb ist L langsam variierend. \square

Im Rest dieses Kapitels werden wir einige Eigenschaften der langsam und regulär variierenden Funktionen herleiten. Die wichtigste Anwendung dieser Eigenschaften in diesem Skript ist die Beschreibung der Max-Anziehungsbereiche.

4.2. Asymptotische Äquivalenz

Definition 4.2.1. Zwei Funktionen $f : (A, \infty) \rightarrow (0, \infty)$ und $g : (A, \infty) \rightarrow (0, \infty)$ heißen **asymptotisch äquivalent** (in $+\infty$), falls

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = 1.$$

Bezeichnung: $f(x) \sim g(x)$ für $x \rightarrow +\infty$.

Beispiel 4.2.2. $x \sim x + 3$ und $x^2 + x \sim x^2 + 2x$ für $x \rightarrow +\infty$.

Aufgabe 4.2.3. Seien $f_1, f_2, g_1, g_2 : (A, \infty) \rightarrow (0, \infty)$ Funktionen mit $f_1 \sim f_2$ und $g_1 \sim g_2$. Zeigen Sie, dass

- (1) $f_1 g_1 \sim f_2 g_2$.
- (2) $f_1 + g_1 \sim f_2 + g_2$.

Die Klasse der regulär variierenden Funktionen mit Index α ist abgeschlossen bezüglich der asymptotischen Äquivalenz:

Aufgabe 4.2.4. Es gelte $f \sim g$ und $f \in \text{RV}_\alpha$. Zeigen Sie, dass dann auch $g \in \text{RV}_\alpha$ ist.

4.3. Satz über die gleichmäßige Konvergenz

Satz 4.3.1 (Über die gleichmäßige Konvergenz, Karamata, 1930). Sei L eine langsam variierende Funktion. Dann gilt für alle $0 < a < b < \infty$

$$\lim_{x \rightarrow \infty} \sup_{\lambda \in [a, b]} \left| \frac{L(\lambda x)}{L(x)} - 1 \right| = 0.$$

BEWEIS. SCHRITT 1. Zunächst einmal werden wir aus der “multiplikativen” Notation in die “additive” Notation wechseln. Betrachte die messbare Funktion

$$h(x) = \log L(e^x), \quad x > \log A.$$

Die Definition der langsamen Variation für L übersetzt sich wie folgt: Für jedes $u \in \mathbb{R}$ gilt

$$(4.3.1) \quad \lim_{x \rightarrow \infty} (h(x+u) - h(x)) = 0.$$

Wir werden zeigen, dass die Konvergenz in (4.3.1) gleichmäßig in $u \in I := [\log a, \log b]$ ist.

SCHRITT 2. Angenommen, die Konvergenz ist nicht gleichmäßig in $u \in I$. Dann gibt es ein $\varepsilon > 0$ und eine Folge $x_1, x_2, \dots \rightarrow \infty$ mit

$$\sup_{u \in I} |h(x_n + u) - h(x_n)| > 2\varepsilon \text{ für } n = 1, 2, \dots$$

Es gibt also eine Folge $u_1, u_2, \dots \in I$ mit

$$(4.3.2) \quad |h(x_n + u_n) - h(x_n)| > 2\varepsilon \text{ für } n = 1, 2, \dots$$

Ohne Einschränkung der Allgemeinheit können wir annehmen, dass u_n gegen einen Grenzwert u konvergiert (andernfalls betrachte eine konvergente Teilfolge) und dass $|u_n - u| \leq 1$ für alle $n \in \mathbb{N}$ (andernfalls entferne endlich viele Terme in den Folgen).

SCHRITT 3. Laut (4.3.1) gilt für jedes $y \in \mathbb{R}$, dass

$$(4.3.3) \quad |h(x_n + y) - h(x_n)| < \varepsilon \text{ für } n \text{ hinreichend groß.}$$

Aus (4.3.2) und (4.3.3) ergibt sich, dass

$$|h(x_n + u_n) - h(x_n + y)| > \varepsilon \text{ für } n \text{ hinreichend groß.}$$

Definiere messbare Mengen I_1, I_2, \dots durch

$$I_k = \{y \in [-1, 1] : |h(x_n + u_n) - h(x_n + y)| > \varepsilon \text{ für alle } n \geq k\}.$$

Es gilt dann $\cup_{k=1}^{\infty} I_k = [-1, 1]$, somit existiert ein k so dass I_k ein strikt positives Lebesgue-Maß hat.

SCHRITT 4. Betrachte die Mengen $Z_n = u_n - I_k = \{u_n - y : y \in I_k\}$ und

$$Z = \{\omega \in [-1, 1] : \omega \in Z_n \text{ für unendlich viele } n\} = \bigcap_{j=1}^{\infty} \bigcup_{n=j}^{\infty} Z_n.$$

Es gilt $|Z_n| = |I_k| > 0$ für alle $n \in \mathbb{N}$. Da die Mengen Z_n im Intervall $[u - 2, u + 2]$ enthalten sind, dürfen wir die Stetigkeit des Lebesgue-Maßes verwenden:

$$|Z| = \lim_{j \rightarrow \infty} \left| \bigcup_{n=j}^{\infty} Z_n \right| \geq \lim_{j \rightarrow \infty} |Z_j| = |I_k| > 0.$$

SCHRITT 5. Somit ist Z nicht leer. Sei $\omega \in Z$, dann gilt für unendlich viele n , dass $\omega \in u_n - I_k$ und somit

$$|h(x_n + u_n) - h(x_n + u_n - \omega)| > \varepsilon.$$

Das ist aber ein Widerspruch zu (4.3.1), denn $x_n + u_n \rightarrow \infty$. □

Für regulär variierende Funktionen nimmt der Satz über die gleichmäßige Konvergenz die folgende Form an:

Aufgabe 4.3.2. Zeigen Sie: Ist R eine mit Index α regulär variierende Funktion, so gilt für alle $0 < a < b < \infty$

$$\lim_{x \rightarrow \infty} \sup_{\lambda \in [a, b]} \left| \frac{R(\lambda x)}{R(x)} - \lambda^\alpha \right| = 0.$$

Hinweis: $R(x) = L(x)x^\alpha$, wobei L langsam variierend ist.

Aufgabe 4.3.3. Seien $f \in \text{RV}_\alpha$ und $g \in \text{RV}_\beta$, und es gelte außerdem $\lim_{x \rightarrow +\infty} g(x) = +\infty$. Zeigen Sie, dass die Komposition $f(g(x))$ ebenfalls regulär variierend mit Index $\alpha\beta$ ist.

Satz 4.3.4 (Über die lokale Beschränktheit). Sei L eine langsam variierende Funktion. Dann gibt es ein B , so dass L auf jedem Intervall der Form $[B, B + c]$ beschränkt ist.

BEWEIS. Nach dem Satz über die gleichmäßige Konvergenz gibt es ein $B > 0$, so dass für alle $x \geq B$

$$\sup_{\lambda \in [1,2]} \left| \frac{L(\lambda x)}{L(x)} - 1 \right| < 1.$$

Somit gilt für alle $x \geq B$

$$\sup_{\lambda \in [1,2]} \frac{L(\lambda x)}{L(x)} < 2.$$

Mit $x = 2^k B$, $k = 0, 1, \dots$, folgt, dass

$$\sup_{z \in [2^k B, 2^{k+1} B]} L(z) \leq 2L(2^k B).$$

Per Induktion erhält man, dass die Funktion L auf dem Intervall $[B, 2^k B]$ durch $2^k L(B)$ beschränkt ist. \square

Aufgabe 4.3.5. Zeigen Sie: Ist R eine regulär variierende Funktion, so gibt es ein B , so dass R auf jedem Intervall der Form $[B, B + c]$ beschränkt ist.

4.4. Darstellungssatz von Karamata

Satz 4.4.1 (Darstellungssatz für langsam variierende Funktionen, Karamata, 1930). Eine Funktion $L : (A, \infty) \rightarrow (0, \infty)$ ist genau dann langsam variierend, wenn es eine Darstellung der Form

$$(4.4.1) \quad L(x) = c(x) \exp \left\{ \int_a^x \frac{\varepsilon(u)}{u} du \right\}, \quad x \geq a,$$

gibt, wobei $a > A$ und $c(x) > 0$, $\varepsilon(x)$ messbare Funktionen sind mit

$$(4.4.2) \quad c := \lim_{x \rightarrow \infty} c(x) \in (0, \infty), \quad \lim_{x \rightarrow \infty} \varepsilon(x) = 0.$$

Aufgabe 4.4.2. Finden Sie eine solche Darstellung explizit für die Funktionen $L_1(x) = \log x$, $L_2(x) = \log \log x$, $L_3(x) = e^{(\log x)^\beta}$, $|\beta| < 1$.

BEWEIS. SCHRITT 1. Es ist eine Übungsaufgabe, zu zeigen, dass jede in der Form (4.4.1), (4.4.2) dargestellte Funktion L langsam variierend ist.

SCHRITT 2. Es sei L eine langsam variierende Funktion. Wir zeigen, dass es eine Darstellung der Form (4.4.1) gibt. Zunächst einmal werden wir aus der “multiplikativen” in die “additive” Notation wechseln. Betrachte dazu die Funktion

$$h(x) := \log L(e^x), \quad x > \log A.$$

Die Bedingung der langsamen Variation von L nimmt die folgende Gestalt an: Für alle $u \in \mathbb{R}$ gilt

$$(4.4.3) \quad \lim_{x \rightarrow +\infty} (h(x+u) - h(x)) = 0.$$

Wir werden zeigen, dass es für h folgende Darstellung gibt:

$$(4.4.4) \quad h(x) = d(x) + \int_B^x e(t) dt,$$

wobei $B \in \mathbb{R}$ und $d(x), e(x)$ messbare Funktionen sind mit

$$(4.4.5) \quad d := \lim_{x \rightarrow \infty} d(x) \in \mathbb{R}, \quad \lim_{x \rightarrow \infty} e(x) = 0.$$

Daraus ergibt sich für L die Darstellung

$$L(x) = e^{h(\log x)} = e^{d(x)} \exp \left\{ \int_B^{\log x} e(t) dt \right\} = e^{d(x)} \exp \left\{ \int_B^x \frac{e(\log u)}{u} du \right\}$$

wobei wir den Ansatz $t = \log u$ gemacht haben. Das würde die Behauptung beweisen, denn

$$c(x) := e^{d(x)} \rightarrow e^d \in (0, \infty), \quad \varepsilon(x) := e(\log x) \rightarrow 0$$

für $x \rightarrow \infty$, was die Gültigkeit von (4.4.2) zeigt.

SCHRITT 3. Wir beweisen die Existenz der Darstellung (4.4.4). Wir würden sehr gerne $e(t) = h'(t)$ setzen, das geht allerdings nicht, denn h muss nicht differenzierbar sein. Wir setzen deshalb $e(t) = h(t+1) - h(t)$. Es sei bemerkt, dass $\lim_{x \rightarrow \infty} e(x) = 0$ nach (4.4.3). Für $x > B$ gilt

$$\int_B^x e(t) dt = \int_B^x (h(t+1) - h(t)) dt = \int_x^{x+1} h(t) dt - \int_B^{B+1} h(t) dt,$$

wobei B so groß sei, dass alle beteiligten Funktionen auf $[B, x]$ beschränkt (und somit integrierbar) seien. Somit

$$h(x) = \int_B^x e(t) dt + \int_0^1 (h(x) - h(x+u)) du + \int_B^{B+1} h(t) dt.$$

Der dritte Term hängt nicht von x ab. Der zweite Term konvergiert für $x \rightarrow \infty$ gegen 0, denn $h(x+u) - h(x) \rightarrow 0$ gleichmäßig in $u \in [0, 1]$ nach dem Satz über die gleichmäßige Konvergenz. Wir können also die Summe der beiden Terme mit $d(x)$ bezeichnen, dann ist (4.4.5) erfüllt. \square

Für regulär variierende Funktionen nimmt die Karamata-Darstellung die folgende Form an:

Aufgabe 4.4.3. Zeigen Sie: Eine Funktion $R : (A, \infty) \rightarrow (0, \infty)$ ist genau dann regulär variierend mit Index $\alpha \in \mathbb{R}$, wenn es eine Darstellung der Form

$$(4.4.6) \quad R(x) = \tilde{c}(x) \exp \left\{ \int_a^x \frac{\tilde{\varepsilon}(u)}{u} du \right\}, \quad x \geq a,$$

gibt, wobei $a > A$ und $\tilde{c}(x) > 0$, $\tilde{\varepsilon}(x)$ messbare Funktionen sind mit

$$(4.4.7) \quad \tilde{c} := \lim_{x \rightarrow \infty} \tilde{c}(x) \in (0, \infty), \quad \lim_{x \rightarrow \infty} \tilde{\varepsilon}(x) = \alpha.$$

4.5. Abschätzungen für regulär variierende Funktionen

Satz 4.5.1. Sei L langsam variierend, dann gilt für jedes $\delta > 0$

$$\lim_{x \rightarrow \infty} x^{-\delta} L(x) = 0, \quad \lim_{x \rightarrow \infty} x^{\delta} L(x) = +\infty.$$

BEWEIS. Das folgt aus dem Darstellungssatz (Übung). \square

Korollar 4.5.2. Sei R regulär variierend mit Index α , dann gilt für jedes $\delta > 0$

$$\lim_{x \rightarrow \infty} \frac{R(x)}{x^{\alpha+\delta}} = 0, \quad \lim_{x \rightarrow \infty} \frac{R(x)}{x^{\alpha-\delta}} = 0.$$

BEWEIS. Betrachte die langsam variierende Funktion $L(x) := R(x)/x^{\alpha}$ und wende Satz 4.5.1 an. \square

Aufgabe 4.5.3. Sei R regulär variierend mit Index α . Zeigen Sie, dass

$$\lim_{x \rightarrow +\infty} R(x) = \begin{cases} +\infty, & \text{falls } \alpha > 0, \\ 0, & \text{falls } \alpha < 0. \end{cases}$$

Aufgabe 4.5.4. Sei R_1 regulär variierend mit Index α_1 und R_2 regulär variierend mit Index α_2 . Zeigen Sie, dass die Funktion $R_1 + R_2$ regulär variierend mit Index $\max\{\alpha_1, \alpha_2\}$ ist.

Aufgabe 4.5.5. Sei L langsam variierend. Zeigen Sie, dass

$$\lim_{x \rightarrow +\infty} \frac{\log L(x)}{\log x} = 0.$$

Satz 4.5.6 (Potter, 1942). Sei L langsam variierend. Für alle $A > 1$, $\delta > 0$ gibt es ein $K = K(A, \delta)$ mit

$$\frac{L(y)}{L(x)} \leq A \max \left\{ \left(\frac{x}{y} \right)^{\delta}, \left(\frac{y}{x} \right)^{\delta} \right\}, \quad \text{für alle } x, y > K.$$

BEWEIS. Mit dem Darstellungssatz erhalten wir

$$\frac{L(y)}{L(x)} = \frac{c(y)}{c(x)} \exp \left\{ \int_x^y \frac{\varepsilon(u)}{u} du \right\}.$$

Da $c(x) \rightarrow c$, $c(y) \rightarrow c$, $\varepsilon(u) \rightarrow 0$ für $x, y, u \rightarrow +\infty$, können wir ein K finden, so dass $c(x) > c/\sqrt{A}$, $c(y) < c\sqrt{A}$ und $\varepsilon(u) < \delta$ für $x, y, u > K$. Somit ist

$$\frac{c(y)}{c(x)} < A, \quad \left| \int_x^y \frac{\varepsilon(u)}{u} du \right| \leq \delta \left| \log \frac{y}{x} \right|.$$

Daraus folgt die Behauptung. \square

4.6. Charakterisierung der Grenzwertfunktionen

In der Definition einer regulär variierenden Funktion f wird verlangt, dass der Grenzwert

$$g(\lambda) := \lim_{x \rightarrow \infty} \frac{f(\lambda x)}{f(x)}$$

für jedes $\lambda > 0$ existiert und eine bestimmte Form, nämlich λ^α , hat. Erstaunlicherweise reicht es viel weniger zu verlangen: Wenn der Grenzwert $g(\lambda)$ für alle λ aus einer Menge mit positivem Lebesgue-Maß existiert, dann existiert er für alle $\lambda > 0$ und hat automatisch die Form λ^α .

Satz 4.6.1. Sei $f : (A, \infty) \rightarrow (0, \infty)$ eine messbare Funktion, für die der Grenzwert

$$(4.6.1) \quad g(\lambda) := \lim_{x \rightarrow +\infty} \frac{f(\lambda x)}{f(x)} \in (0, \infty)$$

für alle λ aus einer messbaren Menge $K \subset (0, \infty)$ mit strikt positivem Lebesgue-Maß existiert. Dann existiert der Grenzwert $g(\lambda)$ sogar für alle $\lambda > 0$. Weiterhin hat die Funktion g die Form $g(\lambda) = \lambda^\alpha$ für ein $\alpha \in \mathbb{R}$. Die Funktion f ist somit regulär variierend.

BEWEIS. SCHRITT 1. Es sei Λ die Menge aller $\lambda > 0$, für die der Grenzwert (4.6.1) in $(0, \infty)$ existiert. Die Menge Λ ist eine multiplikative Untergruppe von $(0, \infty)$, nämlich

- (1) für alle $\lambda_1, \lambda_2 \in \Lambda$ ist auch $\lambda_1 \lambda_2 \in \Lambda$;
- (2) für alle $\lambda \in \Lambda$ ist auch $1/\lambda \in \Lambda$.

Wir zeigen die erste Eigenschaft: Für $\lambda_1, \lambda_2 \in \Lambda$ gilt

$$(4.6.2) \quad \frac{f(\lambda_1 \lambda_2 x)}{f(x)} = \frac{f(\lambda_1 \lambda_2 x)}{f(\lambda_2 x)} \cdot \frac{f(\lambda_2 x)}{f(x)} \rightarrow g(\lambda_1)g(\lambda_2) \quad \text{für } x \rightarrow +\infty.$$

Der Beweis der zweiten Eigenschaft ist eine Übungsaufgabe.

SCHRITT 2. Somit ist die Menge $S := \log \Lambda = \{\log \lambda : \lambda \in \Lambda\}$ eine additive Untergruppe von \mathbb{R} . Diese Untergruppe enthält die Menge $\log K$, die (genauso wie K) messbar ist und ein strikt positives Lebesgue-Maß hat. Nach Satz 2.3.7 gilt also $S = \mathbb{R}$ und somit $\Lambda = (0, \infty)$. Die Existenz des Grenzwerts $g(\lambda)$ in (4.6.1) ist somit für alle $\lambda > 0$ bewiesen.

SCHRITT 3. Aus (4.6.2) folgt, dass g eine multiplikative Funktion ist. Außerdem ist g als Grenzwert von messbaren Funktionen, siehe (4.6.1), messbar. Nach dem Satz von Ostrowski hat g die Form $g(\lambda) = \lambda^\alpha$ für ein $\alpha \in \mathbb{R}$. \square

Aufgabe 4.6.2 (Satz von Landau, 1911). Eine Funktion $L : (A, \infty) \rightarrow (0, \infty)$ sei *monoton* (steigend oder fallend) mit der Eigenschaft, dass

$$\lim_{x \rightarrow +\infty} \frac{L(2x)}{L(x)} = 1.$$

Zeigen Sie, dass L langsam variierend ist.

Aufgabe 4.6.3. Konstruieren Sie eine nicht langsam variierende Funktion $L : (0, \infty) \rightarrow (0, \infty)$ mit der Eigenschaft, dass

$$\lim_{x \rightarrow +\infty} \frac{L(2x)}{L(x)} = 1.$$

4.7. Integrale von regulär variierenden Funktionen

Definition 4.7.1. Seien $f, g : (A, \infty) \rightarrow \mathbb{R}$ zwei Funktionen mit $g(x) \neq 0$. Wir sagen, dass f und g für $x \rightarrow +\infty$ **asymptotisch äquivalent** sind, wenn

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = 1.$$

Notation: $f \sim g$ für $x \rightarrow \infty$.

Beispiel 4.7.2. Für $x \rightarrow \infty$ gilt $x \sim x + 2$.

Beispiel 4.7.3. Für jedes $\alpha > -1$ gilt

$$\int_1^x t^\alpha dt = \frac{x^{\alpha+1} - 1}{\alpha + 1} \sim \frac{x^{\alpha+1}}{\alpha + 1}, \quad x \rightarrow +\infty.$$

Wie kann man aber zum Beispiel die Asymptotik des Integrals $\int_1^x t^\alpha (\log t)^3 dt$ für $x \rightarrow +\infty$ berechnen? Man kann wie folgt argumentieren. Die Funktion $(\log t)^3$ ist langsam variierend, also unterscheidet sich ihr Wert auf dem ganzen Intervall $[x/100, x]$ nur unwesentlich von $(\log x)^3$ (Satz über die gleichmäßige Konvergenz). Auf dem Intervall $[0, x/100]$ kann der Unterschied wesentlich sein, allerdings ist dieses Intervall klein und sein Beitrag zum Integral sollte auch klein sein. Somit kann man vermuten, dass man $(\log t)^3$ durch $(\log x)^3$ ersetzen kann:

$$\int_1^x t^\alpha (\log t)^3 dt \sim (\log x)^3 \frac{x^{\alpha+1}}{\alpha + 1}, \quad x \rightarrow +\infty.$$

Hier ist eine allgemeine Aussage darüber.

Satz 4.7.4 (Karamata, 1930). Sei L eine langsam variierende Funktion und B so groß, dass L auf jedem Intervall der Form $[B, B+c]$ beschränkt ist. Dann gilt für jedes $\alpha > -1$

$$(4.7.1) \quad \int_B^x t^\alpha L(t) dt \sim L(x) \frac{x^{\alpha+1}}{\alpha + 1}, \quad x \rightarrow +\infty.$$

Beweis. Sei $\delta \in (0, \alpha + 1)$. Nach dem Satz von Potter können wir $B' \geq B$ so groß wählen, dass

$$\frac{L(y)}{L(x)} \leq 2 \max \left\{ \left(\frac{x}{y} \right)^\delta, \left(\frac{y}{x} \right)^\delta \right\}, \quad x, y > B'.$$

Wir können die untere Grenze im Integral (4.7.1) auf B' verschieben, denn dadurch ändert sich das Integral nur um eine Konstante. Sei im Folgenden $x > B'$. Betrachte

$$\frac{\int_{B'}^x t^\alpha L(t) dt}{x^{\alpha+1} L(x)/(\alpha+1)} = \frac{1}{\alpha+1} \int_0^1 \frac{L(ux)}{L(x)} u^\alpha \mathbb{1}_{[B'/x, 1]}(u) du,$$

wobei wir die Variable $u = t/x$ eingeführt haben. Nun gilt es für $u \in (0, 1)$

$$\lim_{x \rightarrow +\infty} \left(\frac{L(ux)}{L(x)} u^\alpha \mathbb{1}_{[B'/x, 1]}(u) \right) = u^\alpha, \quad 0 \leq \frac{L(ux)}{L(x)} u^\alpha \mathbb{1}_{[B'/x, 1]}(u) \leq 2u^{\alpha-\delta}.$$

Es liegt also eine dominierte Konvergenz vor, denn $\int_0^1 u^{\alpha-\delta} du < \infty$ wegen $\delta < \alpha + 1$. Mit dem Satz über die dominierte Konvergenz erhalten wir, dass

$$\lim_{x \rightarrow +\infty} \frac{\int_{B'}^x t^\alpha L(t) dt}{x^{\alpha+1} L(x)/(\alpha+1)} = \frac{1}{\alpha+1} \int_0^1 u^\alpha du = 1,$$

was die Behauptung beweist.

Beispiel 4.7.5. Für $\alpha = 0$ erhalten wir, dass

$$\int_B^x L(t) dt \sim xL(x), \quad x \rightarrow +\infty.$$

Max–Anziehungsbereiche

Es seien X_1, X_2, \dots unabhängige identisch verteilte Zufallsvariablen mit Verteilungsfunktion F . Sei

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

Wir erinnern daran, dass F im Max–Anziehungsbereich einer nichtdegenerierten Verteilungsfunktion G liegt, wenn es Folgen $a_n > 0$ und $b_n \in \mathbb{R}$ gibt, so dass

$$\frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} G.$$

Eine äquivalente Bedingung: Für alle $t \in \mathbb{R}$ gilt

$$\lim_{n \rightarrow \infty} F^n(a_n t + b_n) = G(t).$$

Als G kommen nur Extremwertverteilungen in Frage. Diese wurden im Satz von Fisher–Tippett–Gnedenko beschrieben.

In diesem Kapitel werden wir die Max–Anziehungsbereiche der Verteilungen Φ_α, Ψ_α und Λ beschreiben. Nicht alle Beweise in diesem Kapitel sind vollständig. Für mehr Einzelheiten verweisen wir auf die Bücher von S. Resnick “*Extreme Values, Regular Variation and Point Processes*”, L. de Haan, A. Ferreira “*Extreme Value Theory: An Introduction*”, N. H. Bingham, C. M. Goldie, J. L. Teugels “*Regular Variation*”.

5.1. Max–Anziehungsbereich der Fréchet–Verteilung Φ_α

Der nächste Satz beschreibt den Max–Anziehungsbereich der Fréchet–Verteilung Φ_α , $\alpha > 0$.

Satz 5.1.1 (Gnedenko, 1943). Eine Verteilungsfunktion F mit rechtem Endpunkt x^* liegt im Max–Anziehungsbereich der Fréchet–Verteilung Φ_α mit Parameter $\alpha > 0$ genau dann, wenn folgende zwei Bedingungen erfüllt sind:

- (1) $x^* = +\infty$.
- (2) Die Tailfunktion \bar{F} ist regulär variierend mit Index $-\alpha$, d.h.

$$\lim_{x \rightarrow +\infty} \frac{1 - F(\lambda x)}{1 - F(x)} = \lambda^{-\alpha} \text{ für alle } \lambda > 0.$$

Beispiel 5.1.2. Pareto–Verteilung mit der Tailfunktion $\bar{F}(x) = x^{-\alpha}$, $x > 1$, liegt im Max–Anziehungsbereich von Φ_α , denn $x^* = +\infty$ und $\bar{F} \in \text{RV}_{-\alpha}$.

Beispiel 5.1.3. Eine beliebige Verteilungsfunktion, für die $\bar{F}(x) \sim Kx^{-\alpha}$ für $x \rightarrow +\infty$ gilt (wobei $K > 0$ und $\alpha > 0$), liegt im Max–Anziehungsbereich von Φ_α .

Wir beweisen zuerst die Rückrichtung von Satz 5.1.1. Dies geschieht im folgenden Satz.

Satz 5.1.4. Es seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit Verteilungsfunktion F , für die $x^* = +\infty$ und $\bar{F} \in \text{RV}_{-\alpha}$ gilt. Weiterhin sei a_n eine beliebige Folge mit

$$(5.1.1) \quad \lim_{n \rightarrow \infty} n\bar{F}(a_n) = 1.$$

Dann gilt

$$\frac{M_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha.$$

BEWEIS. Gegeben ist, dass $x^* = +\infty$ und $\bar{F} \in \text{RV}_{-\alpha}$. Wir zeigen, dass für alle $t \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} F^n(a_n t) = \Phi_\alpha(t).$$

SCHRITT 1. Zuerst zeigen wir durch Widerspruch, dass

$$\lim_{n \rightarrow \infty} a_n = +\infty.$$

Hätte a_n eine nach oben beschränkte Teilfolge, so wäre entlang dieser Teilfolge $\bar{F}(a_n)$ wegbeschränkt von 0 (wegen $x^* = +\infty$) und wir hätten dann $\lim_{n \rightarrow \infty} n\bar{F}(a_n) = +\infty$. Widerspruch zu (5.1.1). Also gilt $\lim_{n \rightarrow \infty} a_n = +\infty$.

SCHRITT 2. Sei $t > 0$. Da \bar{F} regulär variierend mit Index $-\alpha$ ist, ergibt sich unter Berücksichtigung von (5.1.1), dass

$$n\bar{F}(a_n t) = n\bar{F}(a_n) \cdot \frac{\bar{F}(a_n t)}{\bar{F}(a_n)} \xrightarrow[n \rightarrow \infty]{} 1 \cdot t^{-\alpha} = t^{-\alpha}.$$

Dadurch folgt:

$$F^n(a_n t) = (1 - \bar{F}(a_n t))^n \xrightarrow[n \rightarrow \infty]{} e^{-t^{-\alpha}} = \Phi_\alpha(t).$$

SCHRITT 3. Sei $t \leq 0$. Es gilt für hinreichend großes n , dass $a_n > 0$ (Schritt 1) und folglich

$$F^n(a_n t) \leq F^n(0) \xrightarrow[n \rightarrow \infty]{} 0 = \Phi_\alpha(t),$$

wobei wir benutzt haben, dass $F(0) < 1$ wegen $x^* = +\infty$. □

Wir haben allerdings nicht gesagt, wie man eine Folge a_n konstruiert, die (5.1.1) erfüllt. Am einfachsten definiert man a_n als eine Lösung der Gleichung $\bar{F}(a_n) = \frac{1}{n}$. Leider kann es bei einer unstetigen Verteilungsfunktion F sein, dass \bar{F} den Wert $\frac{1}{n}$ überspringt und es somit keine Lösung gibt. Für eine Konstruktion, die immer funktioniert, benötigen wir den Begriff der Quantilfunktion.

Definition 5.1.5. Die **Quantilfunktion** (oder die **linksstetige Inverse**) einer Verteilungsfunktion F ist die Funktion

$$F^{\leftarrow}(a) := \inf\{t \in \mathbb{R} : F(t) \geq a\}, \quad a \in (0, 1).$$

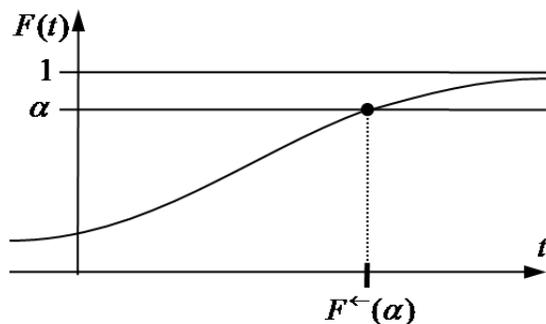


ABBILDUNG 1. Veranschaulichung von $F^{\leftarrow}(a)$.

Ist die Funktion F streng monoton steigend und stetig, so ist F^{\leftarrow} die inverse Funktion zu F . Im Allgemeinen können aber zwei Arten von Problemen auftreten:

- (1) Die Funktion F kann auf einem Intervall konstant bleiben.
- (2) Die Funktion F kann Sprünge haben.

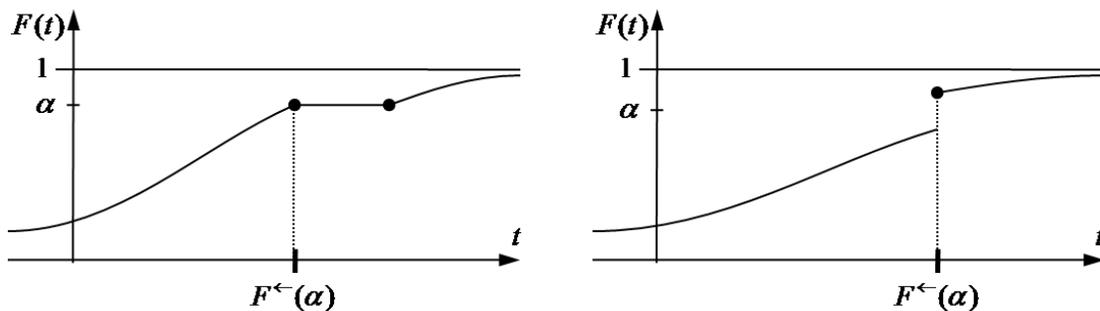


ABBILDUNG 2. Problemfälle

In beiden Fällen ist die inverse Funktion zu F nicht wohldefiniert. Die Quantilfunktion existiert aber trotzdem.

Aufgabe 5.1.6. Sei F eine Verteilungsfunktion mit rechtem Endpunkt x^* .

- (1) Zeigen Sie, dass F^{\leftarrow} linksstetig und monoton nicht-fallend ist.
- (2) Zeigen Sie, dass $\lim_{y \uparrow 1} F^{\leftarrow}(y) = x^*$.

Abbildung 2, rechts, zeigt, dass $F(F^{\leftarrow}(y))$ nicht immer gleich y sein muss. Es gilt lediglich eine einseitige Abschätzung:

Lemma 5.1.7. Es gilt $F(F^{\leftarrow}(y)) \geq y$ für alle $y \in (0, 1)$.

BEWEIS. Sei $x = F^{\leftarrow}(y)$. Für jedes $\varepsilon > 0$ ist $F(x + \varepsilon) \geq y$ nach Definition von F^{\leftarrow} . Lassen wir ε gegen Null gehen, so gilt $F(x + \varepsilon) \rightarrow F(x)$, weil F als Verteilungsfunktion rechtsstetig ist. Daraus ergibt sich, dass $F(x) \geq y$. \square

Nun können wir eine Normierungsfolge a_n angeben, die der Bedingung aus Satz 5.1.4 genügt:

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

Lemma 5.1.8. Sei $x^* = +\infty$ und $\bar{F} \in \text{RV}_{-\alpha}$. Mit der obigen Wahl von a_n gilt

$$\lim_{n \rightarrow \infty} n\bar{F}(a_n) = 1.$$

BEWEIS. Ist F streng monoton steigend und stetig, so gilt $F(a_n) = 1 - \frac{1}{n}$, denn F^{\leftarrow} ist dann die inverse Funktion von F . In diesem Fall ist die Aussage des Lemmas gültig, denn es ist sogar $n\bar{F}(a_n) = 1$. Im Fall eines beliebigen F müssen wir anders argumentieren.

SCHRITT 1. Aus Lemma 5.1.7 folgt, dass $\bar{F}(a_n) \leq \frac{1}{n}$, woraus sich direkt ergibt, dass

$$(5.1.2) \quad \limsup_{n \rightarrow \infty} n\bar{F}(a_n) \leq 1.$$

SCHRITT 2. Es bleibt also noch zu zeigen, dass

$$(5.1.3) \quad \liminf_{n \rightarrow \infty} n\bar{F}(a_n) \geq 1.$$

Sei dazu $x \in (0, 1)$. Für n groß genug gilt $xa_n > 0$, denn $a_n \rightarrow \infty$. Es gilt außerdem $F(xa_n) < 1 - \frac{1}{n}$ nach Definition von a_n . Somit gilt:

$$n\bar{F}(xa_n) = n(1 - F(xa_n)) > n \cdot \frac{1}{n} = 1$$

Damit folgt unmittelbar:

$$n\bar{F}(a_n) = n\bar{F}(xa_n) \cdot \frac{\bar{F}(a_n)}{\bar{F}(xa_n)} > \frac{\bar{F}(a_n)}{\bar{F}(xa_n)} \rightarrow x^\alpha, \quad n \rightarrow \infty,$$

da \bar{F} regulär variierend mit Index $-\alpha$ ist. Es ergibt sich also, dass $\liminf_{n \rightarrow \infty} n\bar{F}(a_n) \geq x^\alpha$ für alle $x \in (0, 1)$. Wenn man nun x gegen 1 gehen lässt, ergibt sich (5.1.3). Damit ist insgesamt $\lim_{n \rightarrow \infty} n\bar{F}(a_n) = 1$ und das Lemma ist bewiesen. \square

Hier sind einige Beispiele von Verteilungen aus dem Max-Anziehungsbereich der Fréchet-Verteilung.

Aufgabe 5.1.9. In der Versicherungsmathematik wird für die Modellierung der Schadenhöhen manchmal die sogenannte Burr-Verteilung mit Verteilungsfunktion

$$F(t) = 1 - \left(\frac{C}{C + t^\beta} \right)^\alpha, \quad t \geq 0,$$

verwendet. Dabei sind $C > 0, \alpha > 0, \beta > 0$ Parameter. Zeigen Sie, dass die Burr-Verteilung im Max-Anziehungsbereich der Fréchet-Verteilung $\Phi_{\alpha\beta}$ liegt, und geben Sie explizit eine Folge a_n an, für die M_n/a_n gegen $\Phi_{\alpha\beta}$ konvergiert.

Aufgabe 5.1.10. Sei X eine Zufallsvariable mit Dichte f und Verteilungsfunktion F . Beweisen Sie: Gilt $f(t) \sim Kt^{-\alpha}, t \rightarrow +\infty$, mit $K > 0$ und $\alpha > 1$, so folgt

$$\bar{F}(t) \sim \frac{K}{\alpha - 1} t^{1-\alpha}, \quad t \rightarrow +\infty.$$

Zeigen Sie, dass X im Max-Anziehungsbereich von $\Phi_{\alpha-1}$ liegt.

Aufgabe 5.1.11. Seien X_1, X_2, \dots unabhängig und Cauchy-verteilt mit der Dichte

$$f(t) = \frac{1}{\pi(1+t^2)}, \quad t \in \mathbb{R}.$$

Zeigen Sie, dass $\frac{\pi}{n} M_n \xrightarrow[n \rightarrow \infty]{d} \Phi_1$.

Auch diskrete Verteilungen können im Max-Anziehungsbereich der Fréchet-Verteilung liegen:

Aufgabe 5.1.12. Eine Zufallsvariable X heißt Zeta-verteilt mit Parameter $\alpha > 1$, wenn

$$\mathbb{P}[X = k] = \frac{1}{\zeta(\alpha)k^\alpha}, \quad k = 1, 2, \dots,$$

wobei $\zeta(\alpha) := \sum_{k=1}^{\infty} k^{-\alpha}$ die Zeta-Funktion ist. Zeigen Sie, dass die Zeta-Verteilung im Max-Anziehungsbereich von $\Phi_{\alpha-1}$ liegt.

Nun beweisen wir die Hinrichtung von Satz 5.1.1.

BEWEIS VON SATZ 5.1.1: “ \Rightarrow ”. Es sei F eine Verteilungsfunktion und $a_n > 0, b_n \in \mathbb{R}$ Folgen, so dass für alle $t \in \mathbb{R}$,

$$(5.1.4) \quad \lim_{n \rightarrow \infty} F^n(a_n t + b_n) = \Phi_\alpha(t).$$

Wir zeigen, dass $x^* = +\infty$ und $\bar{F} \in \text{RV}_{-\alpha}$.

SCHRITT 1. Zuerst müssen wir (5.1.4) auf nichtganzzahlige Werte von n erweitern. Für ein nicht notwendigerweise ganzzahliges $s \geq 0$ definiere $a_s = a_{\lfloor s \rfloor}$ und $b_s = b_{\lfloor s \rfloor}$. Dann gilt für alle $t \in \mathbb{R}$

$$(5.1.5) \quad \lim_{s \rightarrow \infty} F^s(a_s t + b_s) = \Phi_\alpha(t),$$

wobei s nicht ganzzahlig sein muss. In der Tat,

$$F^s(a_s t + b_s) = (F^{\lfloor s \rfloor}(a_{\lfloor s \rfloor} t + b_{\lfloor s \rfloor}))^{s/\lfloor s \rfloor} \xrightarrow[s \rightarrow \infty]{} \Phi_\alpha(t),$$

wobei wir (5.1.4) mit $n = \lfloor s \rfloor$ und die Relation $\lim_{s \rightarrow +\infty} s/\lfloor s \rfloor = 1$ benutzt haben.

SCHRITT 2. Sei $\lambda > 0$. Wegen (5.1.5) mit λs anstelle von s gilt

$$(5.1.6) \quad \lim_{s \rightarrow \infty} F^s(a_{\lambda s}t + b_{\lambda s}) = \lim_{s \rightarrow \infty} (F^{\lambda s}(a_{\lambda s}t + b_{\lambda s}))^{1/\lambda} = \Phi_\alpha^{1/\lambda}(t) = \Phi_\alpha(\lambda^{1/\alpha}t).$$

Indem wir nun (5.1.5) mit (5.1.6) vergleichen und das Lemma von Chintschin benutzen, erhalten wir, dass

$$(5.1.7) \quad \lim_{s \rightarrow \infty} \frac{a_{\lambda s}}{a_s} = \lambda^{1/\alpha}, \quad \lim_{s \rightarrow \infty} \frac{b_{\lambda s} - b_s}{a_s} = 0.$$

Somit ist die Funktion $s \mapsto a_s$ regulär variierend mit Index $1/\alpha$.

SCHRITT 3. Wir werden nun (5.1.7) benutzen, um zu zeigen, dass

$$(5.1.8) \quad \lim_{s \rightarrow \infty} \frac{b_s}{a_s} = 0$$

Sei $\varepsilon > 0$ fest. Sei A so groß, dass

$$(5.1.9) \quad \frac{a_{x/\lambda}}{a_x} < 2\lambda^{-\alpha/2} \text{ für alle } x \geq A, \quad \lambda > 1$$

(das folgt aus der Potter-Schranke für die regulär variierende Funktion $x \mapsto a_x$) und

$$(5.1.10) \quad \left| \frac{b_{2x} - b_x}{a_x} \right| < \varepsilon \text{ für alle } x \geq A$$

(das folgt aus der zweiten Relation in (5.1.7)). Für ein $s > A$ können wir ein $n = n(s) \in \mathbb{N}_0$ finden mit $s/2^{n+1} \leq A < s/2^n$. Wir erhalten die Abschätzung

$$\left| \frac{b_s}{a_s} \right| \leq \sum_{k=1}^n \left| \frac{b_{s/2^{k-1}} - b_{s/2^k}}{a_{s/2^k}} \right| \frac{a_{s/2^k}}{a_s} + \frac{b_{s/2^n}}{a_s} \leq \varepsilon \sum_{k=1}^n 2 \cdot 2^{-\alpha k/2} + \frac{C_1}{a_s},$$

wobei wir im zweiten Schritt (5.1.9) und (5.1.10) benutzt haben, sowie die Tatsache, dass $s/2^n \in [A, 2A]$ und somit $|b_{s/2^n}| < C_1$ für eine Konstante C_1 . Da $a_s \rightarrow \infty$ für $s \rightarrow \infty$ (denn a_s ist regulär variierend mit positivem Index), ergibt sich

$$\limsup_{s \rightarrow \infty} \left| \frac{b_s}{a_s} \right| \leq 2\varepsilon \sum_{k=1}^{\infty} 2^{-\alpha k/2}.$$

Da $\varepsilon > 0$ beliebig und die Summe auf der rechten Seite endlich ist, erhalten wir (5.1.8).

SCHRITT 4. Wegen (5.1.8) und des Chintschin-Lemmas können wir nun (5.1.4) wie folgt vereinfachen: Für alle $t \in \mathbb{R}$ gilt

$$(5.1.11) \quad \lim_{n \rightarrow \infty} F^n(a_n t) = \Phi_\alpha(t).$$

Wir zeigen, dass $x^* = +\infty$. Sei $t > 0$. Wäre x^* endlich, so wäre $a_n t > x^*$ für n groß genug und wir hätten $F^n(a_n t) = 1$ für n groß genug, was in einem Widerspruch zu (5.1.11) steht. Also ist $x^* = +\infty$.

SCHRITT 5. Durch Logarithmieren ergibt sich aus (5.1.11), dass für alle $t > 0$,

$$\lim_{n \rightarrow \infty} n \log F(a_n t) = -t^{-\alpha}.$$

Dies kann man auch wie folgt umschreiben: Für alle $t > 0$ gilt

$$\lim_{n \rightarrow \infty} n \log(1 - \bar{F}(a_n t)) = -t^{-\alpha}.$$

Da $a_n \rightarrow \infty$ (wegen der regulären Variation) und somit $\bar{F}(a_n t) \rightarrow 0$, können wir die Formel $\log(1 - x) = -x + o(x)$ verwenden. Es ergibt sich, dass für alle $t > 0$

$$(5.1.12) \quad \lim_{n \rightarrow \infty} n \bar{F}(a_n t) = t^{-\alpha}.$$

SCHRITT 6. Schließlich zeigen wir unter Benutzung von (5.1.12), dass \bar{F} regulär variierend mit Index $-\alpha$ ist. Sei dazu $\lambda > 0$. Für $x > 0$ definiere

$$n(x) := \inf\{m \in \mathbb{N} : a_{m+1} > x\}.$$

Wegen $a_n \rightarrow +\infty$ ist $n(x)$ wohldefiniert. Es gilt $a_{n(x)} \leq x < a_{n(x)+1}$ und $\lim_{x \rightarrow \infty} n(x) = \infty$. Da \bar{F} außerdem monoton nichtsteigend ist, folgt daraus die Abschätzung

$$\frac{\bar{F}(\lambda x)}{\bar{F}(x)} \leq \frac{\bar{F}(\lambda a_{n(x)})}{\bar{F}(a_{n(x)+1})} = \frac{\bar{F}(\lambda a_{n(x)})n(x)}{\bar{F}(a_{n(x)+1})(n(x) + 1)} \cdot \frac{n(x) + 1}{n(x)} \xrightarrow{n \rightarrow \infty} \frac{\lambda^{-\alpha}}{1} \cdot 1 = \lambda^{-\alpha},$$

wobei wir im letzten Schritt (5.1.12) zweimal benutzt haben. Daraus ergibt sich

$$\limsup_{x \rightarrow \infty} \frac{\bar{F}(\lambda x)}{\bar{F}(x)} \leq \lambda^{-\alpha}.$$

Der Beweis der unteren Abschätzung ist analog. □

5.2. Max–Anziehungsbereich der Weibull–Verteilung Ψ_α

Der Max–Anziehungsbereich der Weibull–Verteilung Ψ_α hat eine ähnliche Charakterisierung wie der Max–Anziehungsbereich der Fréchet–Verteilung. Der Unterschied ist, dass im Fall der Weibull–Verteilung der rechte Endpunkt x^* endlich sein muss. Damit eine Verteilungsfunktion F im Max–Anziehungsbereich von Ψ_α liegt, muss \bar{F} an der Stelle x^* regulär variierend sein. Wir geben nun eine präzise Definition.

Definition 5.2.1. Eine messbare Funktion $f : (0, A) \rightarrow (0, \infty)$ heißt **regulär variierend in 0** mit Index $\alpha \in \mathbb{R}$, falls

$$\lim_{x \downarrow 0} \frac{f(\lambda x)}{f(x)} = \lambda^\alpha \text{ für alle } \lambda > 0.$$

Bezeichnung: $f \in \text{RV}_\alpha(0)$.

Beispiel 5.2.2. Die Funktion $f(x) = x^\alpha$ ist regulär variierend in 0 mit Index α .

Aufgabe 5.2.3. Zeigen Sie: $f(x)$ ist regulär variierend mit Index α an der Stelle 0 genau dann, wenn $1/f(1/x)$ regulär variierend mit Index α (an der Stelle $+\infty$) ist.

Der nächste Satz charakterisiert den Max–Anziehungsbereich der Weibull–Verteilung.

Satz 5.2.4 (Gnedenko, 1943). Eine Verteilungsfunktion F mit rechtem Endpunkt x^* liegt im Max–Anziehungsbereich der Weibull–Verteilung Ψ_α mit Parameter $\alpha > 0$ genau dann, wenn folgende zwei Bedingungen erfüllt sind:

- (1) $x^* < \infty$.
- (2) Die Funktion $x \mapsto \bar{F}(x^* - x)$, $x > 0$, ist regulär variierend in 0 mit Index α , d.h.

$$\lim_{x \downarrow 0} \frac{1 - F(x^* - \lambda x)}{1 - F(x^* - x)} = \lambda^\alpha \text{ für alle } \lambda > 0.$$

Bemerkung 5.2.5. Sind die beiden Bedingungen von Theorem 5.2.4 erfüllt, so werden wir zeigen, dass

$$\frac{M_n - x^*}{a_n} \xrightarrow[n \rightarrow \infty]{d} \Psi_\alpha,$$

wobei a_n eine beliebige Folge mit

$$(5.2.1) \quad \lim_{n \rightarrow \infty} n\bar{F}(x^* - a_n) = 1$$

ist. Ein Beispiel einer solchen Folge a_n ist gegeben durch

$$a_n = x^* - F^{\leftarrow} \left(1 - \frac{1}{n} \right).$$

Beispiel 5.2.6. Betrachte die Verteilungsfunktion $F(x) = 1 - (x^* - x)^\alpha$ mit $x \in (x^* - 1, x^*)$, wobei $\alpha > 0$. Dann ist $1 - F(x^* - x) = x^\alpha \in \text{RV}_\alpha(0)$. Somit liegt F im $\text{MDA}(\Psi_\alpha)$.

Beispiel 5.2.7. Eine Verteilungsfunktion F mit endlichem rechten Endpunkt x^* , für die $\bar{F}(x^* - x) \sim Kx^\alpha$ für $x \downarrow 0$ gilt (wobei $K > 0$, $\alpha > 0$), liegt im Max–Anziehungsbereich von Ψ_α .

Wir beweisen nur die Rückrichtung von Satz 5.2.4. Der Beweis der Hinrichtung benutzt ähnliche Ideen wie im Fréchet–Fall.

BEWEIS VON SATZ 5.2.4: “ \Leftarrow ”. Sei $x < 0$. Es gilt

$$\mathbb{P} \left[\frac{M_n - x^*}{a_n} \leq x \right] = F^n(a_n x + x^*) = (1 - \bar{F}(a_n x + x^*))^n \xrightarrow[n \rightarrow \infty]{} e^{-(x)^\alpha} = \Psi_\alpha(x),$$

denn

$$\lim_{n \rightarrow \infty} n\bar{F}(a_n x + x^*) = \lim_{n \rightarrow \infty} n\bar{F}(x^* - a_n) \cdot \frac{\bar{F}(x^* - a_n(-x))}{\bar{F}(x^* - a_n)} = (-x)^\alpha.$$

Dabei haben wir die reguläre Variation von $\bar{F}(x^* - x)$ an der Stelle $x = 0$ und (5.2.1) benutzt.

Sei $x \geq 0$. Dann gilt

$$\mathbb{P} \left[\frac{M_n - x^*}{a_n} \leq x \right] = F^n(a_n x + x^*) = 1 \xrightarrow[n \rightarrow \infty]{} 1 = \Psi_\alpha(x),$$

denn $M_n \leq x^*$ f.s. □

Hier sind einige Beispiele von Verteilungen aus dem Max-Anziehungsbereich der Weibull-Verteilung.

Aufgabe 5.2.8. Seien X_1, X_2, \dots unabhängige Zufallsvariablen mit Beta(α, β)-Verteilung, d.h. die Dichte von X_i sei gegeben durch

$$f(t) = \begin{cases} \frac{1}{B(\alpha, \beta)} t^{\alpha-1} (1-t)^{\beta-1}, & \text{für } t \in [0, 1], \\ 0, & \text{sonst.} \end{cases}$$

Dabei ist $B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt$ die Eulersche Beta-Funktion und $\alpha > 0, \beta > 0$ sind Parameter. Geben Sie explizit eine Folge $c_n > 0$ an mit

$$c_n(M_n - 1) \xrightarrow[n \rightarrow \infty]{d} \Psi_\beta.$$

Aufgabe 5.2.9. Seien X_1, X_2, \dots unabhängig und standardnormalverteilt. Zeigen Sie, dass die Zufallsvariablen $-|X_i|$ im Max-Anziehungsbereich von Ψ_1 liegen. Geben Sie explizit eine Folge $c_n > 0$ an mit

$$c_n \min\{|X_1|, \dots, |X_n|\} \xrightarrow[n \rightarrow \infty]{d} \text{Exp}(1).$$

5.3. Max-Anziehungsbereich der Gumbel-Verteilung Λ

Eine Theorie des Max-Anziehungsbereiches der Gumbel-Verteilung wurde von L. de Haan entwickelt. In diesem Skript werden wir auf diese Theorie nicht eingehen und verweisen stattdessen auf die Bücher von L. de Haan “*On Regular Variation and its Application to the Weak Convergence of Sample Extremes*”, N. H. Bingham, C. M. Goldie, J. L. Teugels “*Regular Variation*” und S. Resnick “*Extreme Values, Regular Variation and Point Processes*”. Wir beweisen nur ein einfaches Resultat.

Satz 5.3.1 (Gnedenko, 1943). Eine Verteilungsfunktion F mit rechtem Endpunkt x^* liegt im Max-Anziehungsbereich der Gumbel-Verteilung $\Lambda(x) = e^{-e^{-x}}$ genau dann, wenn es eine positive und messbare Funktion $g(x)$ gibt mit

$$(5.3.1) \quad \lim_{x \uparrow x^*} \frac{\bar{F}(x + ug(x))}{\bar{F}(x)} = e^{-u} \text{ für alle } u \in \mathbb{R}.$$

Bemerkung 5.3.2. x^* kann im Gumbel-Fall endlich oder unendlich sein, Beispiele werden unten gegeben.

Bemerkung 5.3.3. Wir werden zeigen: Ist die Bedingung (5.3.1) erfüllt, so gilt

$$\frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}},$$

wobei a_n und b_n Folgen sind, die die folgenden Bedingungen erfüllen:

$$(5.3.2) \quad \lim_{n \rightarrow \infty} n\bar{F}(b_n) = 1, \quad a_n = g(b_n).$$

Eine mögliche Wahl von b_n ist $b_n = F^{\leftarrow} \left(1 - \frac{1}{n}\right)$ (Beweis ähnlich wie in Lemma 5.1.8).

BEWEIS VON SATZ 5.3.1. Es wird hier nur ein Beweis für die Rückrichtung gegeben. Es seien also (5.3.1) und (5.3.2) erfüllt. Es gilt $\lim_{n \rightarrow \infty} b_n = x^*$, denn hätte b_n eine von x^* wegbeschränkte Teilfolge, so würde entlang dieser Teilfolge (5.3.2) verletzt sein. Man betrachte nun

$$\lim_{n \rightarrow \infty} n\bar{F}(a_n u + b_n) = \lim_{n \rightarrow \infty} \frac{\bar{F}(b_n + g(b_n)u)}{\bar{F}(b_n)} \cdot n\bar{F}(b_n) = e^{-u},$$

wobei wir (5.3.1) und (5.3.2) benutzt haben. Es folgt

$$\mathbb{P} \left[\frac{M_n - b_n}{a_n} \leq u \right] = F^n(a_n u + b_n) = (1 - \bar{F}(a_n u + b_n))^n \xrightarrow[n \rightarrow \infty]{} e^{-e^{-u}}.$$

Und dadurch ergibt sich $\frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} \Lambda$. □

Beispiel 5.3.4. Die Exponentialverteilung $\text{Exp}(\lambda)$ mit Tailfunktion

$$\bar{F}(x) = e^{-\lambda x}, \quad x > 0,$$

liegt im Max-Anziehungsbereich von Λ . Man kann nachrechnen, dass Bedingung (5.3.1) mit $g(x) = \frac{1}{\lambda}$ erfüllt ist:

$$\frac{\bar{F}(x + ug(x))}{\bar{F}(x)} = e^{-\lambda ug(x)} = e^{-u}.$$

Aus (5.3.2) ergibt sich (als eine mögliche Wahl) $b_n = \frac{\log n}{\lambda}$ und $a_n = \frac{1}{\lambda}$, so dass

$$\lambda M_n - \log n \xrightarrow[n \rightarrow \infty]{d} \Lambda.$$

Einige weitere Beispiele von Verteilungen aus dem Max-Anziehungsbereich der Gumbel-Verteilung Λ finden sich in den nachfolgenden Aufgaben.

Aufgabe 5.3.5. Seien X_1, X_2, \dots unabhängige identisch verteilte Zufallsvariablen mit Tailfunktion

$$\bar{F}(t) = e^{-t^\alpha}, \quad t > 0,$$

wobei $\alpha > 0$. Geben Sie explizit $a_n > 0$ und $b_n \in \mathbb{R}$ mit $(M_n - b_n)/a_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ an.

Aufgabe 5.3.6. Seien X_1, X_2, \dots unabhängige identisch verteilte Zufallsvariablen mit Tailfunktion

$$\bar{F}(t) = e^{-(\log t)^\alpha}, \quad t > 1,$$

wobei $\alpha > 1$. Geben Sie explizit $a_n > 0$ und $b_n \in \mathbb{R}$ mit $(M_n - b_n)/a_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ an.

Die nächste Aufgabe zeigt, dass im Max-Anziehungsbereich der Gumbel-Verteilung der rechte Endpunkt auch endlich sein kann.

Aufgabe 5.3.7. Seien X_1, X_2, \dots unabhängige identisch verteilte Zufallsvariablen mit Tailfunktion

$$\bar{F}(t) = \begin{cases} e^{1/t}, & t < 0, \\ 0, & t \geq 0. \end{cases}$$

Geben Sie explizit $a_n > 0$ und $b_n \in \mathbb{R}$ mit $(M_n - b_n)/a_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ an.

Alle drei Max-Anziehungsbereiche sind abgeschlossen bzgl. der asymptotischen Äquivalenz:

Aufgabe 5.3.8. Es seien F und G zwei Verteilungsfunktionen mit dem gleichen rechten Endpunkt x^* und

$$\lim_{x \uparrow x^*} \frac{1 - F(x)}{1 - G(x)} = c,$$

wobei $0 < c < \infty$. Zeigen Sie: Liegt F im Max-Anziehungsbereich von Φ_α , Ψ_α oder Λ , so liegt auch G in demselben Max-Anziehungsbereich.

Es gibt Verteilungen, die in keinem Max-Anziehungsbereich liegen:

Aufgabe 5.3.9. Die Zufallsvariable X sei geometrisch verteilt mit Parameter $1/2$, d.h.

$$\mathbb{P}[X = k] = 1/2^k, \quad k = 1, 2, \dots$$

Zeigen Sie, dass X in keinem der drei Max-Anziehungsbereiche liegt.

Aufgabe 5.3.10. Die Zufallsvariable Y besitze die Tailfunktion

$$\bar{F}(t) = 1/\log t, \quad t > e.$$

Zeigen Sie, dass Y in keinem der drei Max-Anziehungsbereiche liegt.

5.4. Beispiel: Normalverteilung

Im Folgenden werden wir zeigen, dass die Normalverteilung zum Max-Anziehungsbereich der Gumbel-Verteilung gehört. Dazu benötigen wir ein Lemma, das die Asymptotik der Tailfunktion der Standardnormalverteilung beschreibt. Wir erinnern daran, dass die Notation $f(t) \sim g(t)$ für $t \rightarrow +\infty$ bedeutet, dass

$$\lim_{t \rightarrow +\infty} \frac{f(t)}{g(t)} = 1.$$

Lemma 5.4.1. Für die Tailfunktion \bar{F} und die Dichte f der Standardnormalverteilung gilt

$$\bar{F}(t) \sim \frac{f(t)}{t} = \frac{1}{\sqrt{2\pi}} \frac{1}{t} e^{-t^2/2} \text{ für } t \rightarrow +\infty.$$

BEWEIS. Es gilt

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(t)}{\frac{1}{\sqrt{2\pi}} e^{-t^2/2}} = \lim_{t \rightarrow \infty} \frac{\int_t^\infty e^{-s^2/2} ds}{\frac{1}{t} e^{-t^2/2}} = \lim_{t \rightarrow \infty} \frac{-e^{-t^2/2}}{-\frac{1}{t^2} e^{-t^2/2} - t \frac{1}{t} e^{-t^2/2}} = \lim_{t \rightarrow \infty} \frac{1}{\frac{1}{t^2} + 1} = 1,$$

wobei wir den Satz von L'Hospital für den Fall „ $\frac{0}{0}$ “ angewendet haben. □

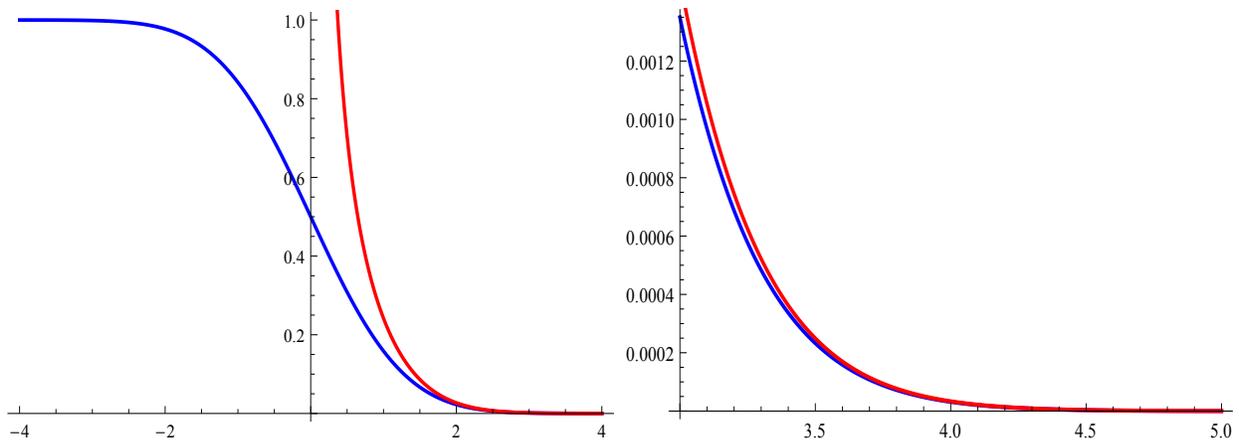


ABBILDUNG 3. Veranschaulichung von Lemma 5.4.1. Blaue Kurve: Die Tailfunktion \bar{F} der Standardnormalverteilung. Rote Kurve: Die Approximation. Das rechte Bild zeigt die beiden Kurven auf dem Intervall $[3, 5]$.

Satz 5.4.2. Die Standardnormalverteilung liegt im Max-Anziehungsbereich der Gumbel-Verteilung Λ .

BEWEIS. Wir werden zeigen, dass Bedingung (5.3.1) des Satzes 5.3.1 mit $g(t) = 1/t$ gilt. Mit Lemma 5.4.1 ergibt sich:

$$\lim_{t \rightarrow +\infty} \frac{\bar{F}(t + xg(t))}{\bar{F}(t)} = \lim_{t \rightarrow +\infty} \frac{\frac{1}{t+xg(t)} e^{-(t+xg(t))^2/2}}{\frac{1}{t} e^{-t^2/2}} = \lim_{t \rightarrow \infty} \frac{\frac{1}{t+x/t} e^{-t^2/2 - x - \frac{x^2}{2t^2}}}{\frac{1}{t} e^{-t^2/2}} = e^{-x}$$

für alle $x \in \mathbb{R}$. □

Es seien X_1, X_2, \dots unabhängige und standardnormalverteilte Zufallsvariablen und $M_n = \max\{X_1, \dots, X_n\}$. Wir werden nun die Folgen $a_n > 0$ und $b_n \in \mathbb{R}$ so bestimmen, dass

$$\frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

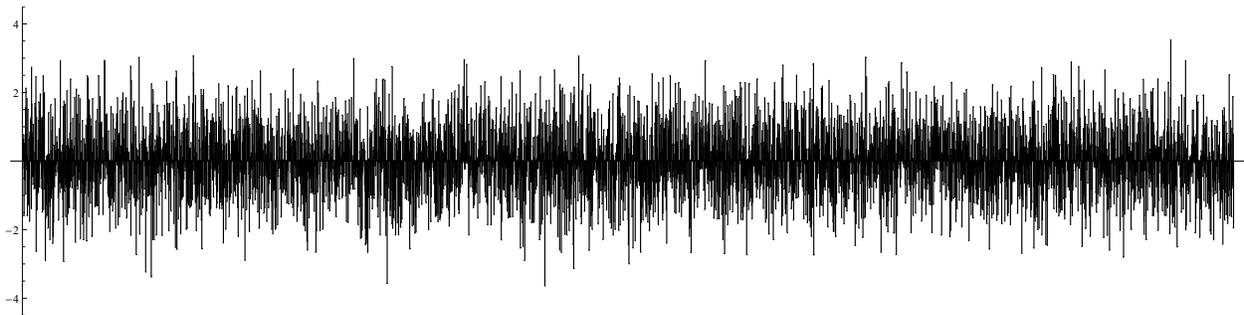


ABBILDUNG 4. Eine standardnormalverteilte Stichprobe vom Umfang $n = 5000$.

Laut Satz 5.3.1 sollten wir b_n so wählen, dass

$$(5.4.1) \quad \lim_{n \rightarrow \infty} n\bar{F}(b_n) = 1.$$

Damit das gilt, muss die Folge b_n gegen $+\infty$ divergieren. Hätte nämlich die Folge b_n eine nach oben beschränkte Teilfolge, so wäre $\bar{F}(b_n)$ entlang dieser Teilfolge von 0 wegbeschränkt sein und $n\bar{F}(b_n)$ würde entlang dieser Teilfolge gegen unendlich gehen. Widerspruch, also geht b_n gegen $+\infty$.

Wir können nun Lemma 5.4.1 benutzen und (5.4.1) in der folgenden Form schreiben:

$$(5.4.2) \quad \sqrt{2\pi}b_n e^{b_n^2/2} \sim n, \quad n \rightarrow \infty.$$

Auf der linken Seite ist $e^{b_n^2/2}$ derjenige Term, der am schnellsten gegen $+\infty$ geht. Wir können also als eine erste Annäherung zu b_n eine Folge w_n mit $e^{w_n^2/2} = n$ wählen, d.h.

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

Mit der Folge w_n sind wir aber noch nicht am Ziel, denn

$$\sqrt{2\pi}w_n e^{w_n^2/2} = \sqrt{2\pi} \sqrt{2 \log n} n \approx n.$$

Wir machen also den Ansatz $b_n = \sqrt{2 \log n} + \delta_n$, wobei δ_n noch genauer spezifiziert werden muss. Mit diesem Ansatz gilt:

$$(5.4.3) \quad \begin{aligned} \sqrt{2\pi}b_n e^{b_n^2/2} &= \sqrt{2\pi}(\sqrt{2 \log n} + \delta_n) e^{\log n + \sqrt{2 \log n} \delta_n + \frac{\delta_n^2}{2}} \\ &= n \cdot \sqrt{2\pi}(\sqrt{2 \log n} + \delta_n) e^{\sqrt{2 \log n} \delta_n} e^{\frac{\delta_n^2}{2}}. \end{aligned}$$

Wir wollen δ_n so bestimmen, dass alle Terme auf der rechten Seite außer n asymptotisch äquivalent zu 1 sind. Wähle δ_n so dass

$$\sqrt{2\pi} \sqrt{2 \log n} e^{\sqrt{2 \log n} \delta_n} = 1.$$

Dann folgt durch Umformungen, dass

$$\delta_n = -\frac{\log(4\pi \log n)}{2\sqrt{2 \log n}}.$$

Es sei bemerkt, dass ein so gewähltes δ_n gegen 0 geht. Somit gilt

$$e^{\delta_n^2} \sim 1, \quad \sqrt{2 \log n} + \delta_n \sim \sqrt{2 \log n}.$$

Also ist die rechte Seite von (5.4.3) tatsächlich asymptotisch äquivalent zu 1, wie gewünscht. Wir kommen also zum Schluss, dass b_n wie folgt gewählt werden sollte:

$$(5.4.4) \quad b_n = \sqrt{2 \log n} - \frac{\log(4\pi \log n)}{2\sqrt{2 \log n}}.$$

Als a_n wählt man schließlich $a_n = g(b_n) = \frac{1}{b_n}$. Da aber $b_n \sim \sqrt{2 \log n}$, kann man mit dem Lemma von Chintschin zeigen, dass auch die folgende einfachere Wahl von a_n reicht:

$$(5.4.5) \quad a_n = \frac{1}{\sqrt{2 \log n}}.$$

Wir haben somit den folgenden Satz bewiesen:

Satz 5.4.3. Seien X_1, X_2, \dots unabhängige und standardnormalverteilte Zufallsvariablen. Dann gilt für alle $x \in \mathbb{R}$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sqrt{2 \log n} \left\{ M_n - \left(\sqrt{2 \log n} - \frac{\log(4\pi \log n)}{2\sqrt{2 \log n}} \right) \right\} \leq x \right] = e^{-e^{-x}}.$$

Bemerkung 5.4.4. Der Satz lässt sich wie folgt interpretieren: Das Maximum M_n nimmt Werte an, die sehr nahe bei $b_n = \sqrt{2 \log n} - \frac{\log(4\pi \log n)}{2\sqrt{2 \log n}}$ sind. Die Differenz zwischen M_n und b_n ist zufällig und hat die Größenordnung $a_n = \frac{1}{\sqrt{2 \log n}}$. Multipliziert man $M_n - b_n$ mit dem Faktor $\sqrt{2 \log n}$, so erhält man eine approximativ Gumbel-verteilte Zufallsvariable.

Zum Schluss werden wir noch zeigen, dass auch die sogenannte Log-Normalverteilung im Max-Anziehungsbereich der Gumbel-Verteilung liegt.

Definition 5.4.5. Es sei X eine standardnormalverteilte Zufallsvariable. Dann heißt die Zufallsvariable $Y := e^X$ **log-normalverteilt**.

Die Begriff log-normal erklärt sich dadurch, dass die Zufallsvariable $X = \log Y$ normalverteilt ist.

Aufgabe 5.4.6. Zeigen Sie, dass die Dichte einer log-normalverteilten Zufallsvariable durch die folgende Formel gegeben ist:

$$f(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2} \log^2 y}, \quad y > 0.$$

Satz 5.4.7. Es seien Y_1, Y_2, \dots unabhängige, log-normalverteilte Zufallsvariablen. Sei

$$L_n := \max\{Y_1, \dots, Y_n\}.$$

Mit a_n und b_n wie in Formeln (5.4.4) und (5.4.5) gilt für alle $x \in \mathbb{R}$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{L_n - e^{b_n}}{a_n e^{b_n}} \leq x \right] = e^{-e^{-x}}.$$

BEWEIS. Wir können annehmen, dass $Y_1 = e^{X_1}, Y_2 = e^{X_2}, \dots$, wobei X_1, X_2, \dots unabhängig und standardnormalverteilt sind. Sei $M_n = \max\{X_1, \dots, X_n\}$, dann ist $L_n = e^{M_n}$. Wir wissen bereits, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n - b_n \leq a_n x] = e^{-e^{-x}}.$$

Durch die Anwendung der Exponentialfunktion folgt, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}[e^{-b_n} L_n \leq e^{a_n x}] = e^{-e^{-x}}.$$

Da $\lim_{n \rightarrow \infty} a_n = 0$, können wir die Entwicklung $e^{a_n x} = 1 + a_n x + o(a_n)$ benutzen und es folgt (Übungsaufgabe: Zeigen Sie, dass man den o -term weglassen kann), dass

$$\lim_{n \rightarrow \infty} \mathbb{P} [e^{-b_n} L_n \leq 1 + a_n x] = e^{-e^{-x}}.$$

Daraus ergibt sich die Behauptung. □

Die Gammaverteilung und die Log-Gammaverteilung liegen ebenfalls im Max-Anziehungsbereich der Gumbel-Verteilung.

Aufgabe 5.4.8. Seien X_1, X_2, \dots unabhängige $\text{Gamma}(\alpha, \lambda)$ -verteilte Zufallsvariablen, d.h. die Dichte von X_i sei

$$f(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \quad t > 0,$$

wobei $\alpha > 0, \lambda > 0$. Geben Sie explizit $a_n > 0$ und $b_n \in \mathbb{R}$ mit $(M_n - b_n)/a_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ an.

Aufgabe 5.4.9. Es seien X_1, X_2, \dots unabhängig und Log-Gammaverteilt, d.h. $\log X_i$ seien $\text{Gamma}(\alpha, \lambda)$ -verteilt. Geben Sie explizit $a_n > 0$ und $b_n \in \mathbb{R}$ mit $(M_n - b_n)/a_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ an.

5.5. Von Mises-Bedingungen

Wir haben notwendige und hinreichende Bedingungen dafür hergeleitet, dass eine Verteilung in einem Max-Anziehungsbereich liegt. Leider sind diese Bedingungen manchmal schwer zu überprüfen. In diesem Abschnitt leiten wir einfach zu überprüfende hinreichende Bedingungen, die allerdings nicht notwendig sind.

Zuerst müssen wir den Begriff ‘Ausfallrate’ einführen. Man betrachte ein Gerät, dessen Lebensdauer als eine Zufallsvariable $Z \geq 0$ modelliert werde. Wir nehmen an, dass Z absolut stetig mit Dichte f ist. Wir bezeichnen die Verteilungsfunktion von Z mit F . Es sei dx ein sehr kleines Zeitintervall. Wir stellen uns die folgende Frage:

Gegeben, dass das Gerät zum Zeitpunkt x noch funktioniert, wie groß ist die Wahrscheinlichkeit, dass es zwischen x und $x + dx$ ausfällt?

Diese bedingte Wahrscheinlichkeit kann wie folgt berechnet werden:

$$\mathbb{P}[Z \in (x, x + dx) | Z > x] = \frac{\mathbb{P}[Z \in (x, x + dx), Z > x]}{\mathbb{P}[Z > x]} = \frac{\mathbb{P}[Z \in (x, x + dx)]}{\mathbb{P}[Z > x]} \approx \frac{f(x)}{\bar{F}(x)} dx.$$

Die Funktion $h(x) = f(x)/\bar{F}(x)$ heißt die **Ausfallrate** des Geräts.

Im obigen Beispiel ist die Lebensdauer des Geräts nicht-negativ. Die Definition der Ausfallrate kann allerdings auch für Zufallsvariablen verwendet werden, die negative Werte annehmen dürfen. Für die folgende Definition sei Z eine Zufallsvariable mit rechtem Endpunkt x^* , deren Verteilungsfunktion F auf einem Intervall (x_0, x^*) eine stetige Ableitung $f = F'$ besitzt.

Definition 5.5.1. Die **Ausfallrate** von Z ist die Funktion

$$h(x) = \frac{f(x)}{\bar{F}(x)} = -(\log \bar{F})'(x), \quad x \in (x_0, x^*).$$

Beispiel 5.5.2. Sei Z exponentialverteilt mit Tailfunktion $\bar{F}(x) = e^{-\lambda x}$, $x > 0$, und Dichte $f(x) = \lambda e^{-\lambda x}$, $x > 0$. Die Ausfallrate ist dann gegeben durch

$$h(x) = \frac{f(x)}{\bar{F}(x)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda, \quad x > 0.$$

Die Exponentialverteilung hat also eine *konstante* Ausfallrate. Das hat mit der Vergessenseigenschaft der Exponentialverteilung in Zusammenhang gebracht werden: Ein Gerät, das zum Zeitpunkt x noch funktioniert, “erinnert” sich an sein Alter nicht und fällt mit der gleichen Rate aus, wie ein neues Gerät gleich nach Inbetriebnahme.

Im Folgenden werden wir eine Formel benutzen, die die Tailfunktion durch die Ausfallrate darstellt:

Lemma 5.5.3. Für beliebiges $a \in (x_0, x^*)$ gilt

$$\bar{F}(x) = \bar{F}(a) \exp \left\{ - \int_a^x h(u) du \right\}, \quad x \in [a, x^*).$$

BEWEIS. Mit der Newton–Leibniz–Formel ergibt sich

$$\bar{F}(x) = e^{\log \bar{F}(x)} = \exp \left\{ \int_a^x (\log \bar{F})'(u) du + \log \bar{F}(a) \right\} = \bar{F}(a) \exp \left\{ - \int_a^x h(u) du \right\},$$

wobei wir im letzten Schritt die Definition der Ausfallrate benutzt haben. □

Von Mises–Bedingung für den Fréchet–Max–Anziehungsbereich

Bevor wir die von Mises–Bedingung für den Max–Anziehungsbereich der Fréchet–Verteilung formulieren, betrachten wir ein Beispiel.

Beispiel 5.5.4. Sei X Pareto–verteilt mit Tailfunktion $\bar{F}(x) = x^{-\alpha}$, $x > 1$. Dabei sei $\alpha > 0$ ein Parameter. Die Dichte ist $f(x) = \alpha x^{-\alpha-1}$. Die Ausfallrate ist somit

$$h(x) = \frac{\alpha}{x}, \quad x > 1.$$

Bekanntlich gilt $X \in \text{MDA}(\Phi_\alpha)$.

Wir werden nun zeigen, dass alle Verteilungen mit unendlichem rechten Endpunkt und einer zu $\frac{\alpha}{x}$ asymptotisch äquivalenten Ausfallrate in $\text{MDA}(\Phi_\alpha)$ liegen.

Satz 5.5.5 (von Mises, 1936). Sei X eine Zufallsvariable mit rechtem Endpunkt $x^* = +\infty$ und Ausfallrate $h(x)$, $x > x_0$. Es gelte

$$(5.5.1) \quad h(x) \sim \frac{\alpha}{x}, \quad x \rightarrow +\infty,$$

mit einem Parameter $\alpha > 0$. Dann liegt X im Max-Anziehungsbereich der Fréchet-Verteilung Φ_α .

BEWEIS. Wähle ein $a > x_0$. Es gilt nach Lemma 5.5.3

$$\bar{F}(x) = \bar{F}(a) \exp \left\{ - \int_a^x h(u) du \right\} = \bar{F}(a) \exp \left\{ \int_a^x \frac{\varepsilon(u)}{u} du \right\},$$

wobei die Funktion $\varepsilon(u) := -h(u)u$ wegen (5.5.1) die Bedingung $\lim_{u \rightarrow \infty} \varepsilon(u) = -\alpha$ erfüllt. Es folgt aus dem Darstellungssatz von Karamata, dass \bar{F} regulär variierend mit Index $-\alpha$ ist. Somit liegt F im Max-Anziehungsbereich der Fréchet-Verteilung Φ_α . \square

Von Mises-Bedingung für den Weibull-Max-Anziehungsbereich

Der Weibull-Fall ist dem Fréchet-Fall sehr ähnlich, ein wesentlicher Unterschied ist allerdings, dass im Weibull-Fall der rechte Endpunkt endlich ist.

Beispiel 5.5.6. Sei X eine Zufallsvariable mit rechtem Endpunkt $x^* < +\infty$ und Tailfunktion $\bar{F}(x) = (x^* - x)^\alpha$, $x \in [x^* - 1, x^*]$. Dabei sei $\alpha > 0$ ein Parameter. Die Dichte ist $f(x) = \alpha(x^* - x)^{\alpha-1}$, $x \in [x^* - 1, x^*]$. Die Ausfallrate ist somit

$$h(x) = \frac{\alpha}{x^* - x}, \quad x \in (x^* - 1, x^*).$$

Bekanntlich liegt X im Max-Anziehungsbereich der Weibull-Verteilung Ψ_α .

Satz 5.5.7 (von Mises, 1936). Sei X eine absolut stetige Zufallsvariable mit endlichem rechtem Endpunkt $x^* < +\infty$ und Ausfallrate $h(x)$, $x \in (x_0, x^*)$. Es gelte

$$(5.5.2) \quad h(x) \sim \frac{\alpha}{x^* - x}, \quad x \uparrow x^*,$$

wobei $\alpha > 0$. Dann liegt X im Max-Anziehungsbereich der Weibull-Verteilung Ψ_α .

BEWEIS. Ähnlich wie im Fréchet-Fall. \square

Von Mises-Bedingung für den Gumbel-Max-Anziehungsbereich

Um die von Mises-Bedingung im Gumbel-Fall zu formulieren, erinnern wir uns zuerst daran, dass wir den Gumbel-Fall als eine Art "Grenzwert" des Fréchet- und des Weibull-Falls für $\alpha \rightarrow +\infty$ betrachten können. Der rechte Endpunkt kann im Gumbel-Fall sowohl endlich als auch unendlich sein. Betrachten wir den Fall eines endlichen rechten Endpunktes. Aus der

von Mises-Bedingung für den Fréchet-Fall $h(x) \sim \frac{\alpha}{x}$ bzw. $\frac{1}{h(x)} \sim \frac{x}{\alpha}$ für $x \rightarrow +\infty$ wird bei einem informellen Übergang zum Grenzwert $\alpha \rightarrow +\infty$ die Bedingung

$$\frac{1}{h(x)} = o(x), \quad x \rightarrow +\infty.$$

Leider braucht man im Gumbel-Fall eine etwas stärkere Bedingung, die man durch formales Ableiten der obigen Bedingung erhält:

Satz 5.5.8 (von Mises, 1936). Sei X eine Zufallsvariable mit rechtem Endpunkt x^* der endlich oder unendlich sein darf. Die Ausfallrate $h(x)$ existiere und sei differenzierbar und strikt positiv auf einem Intervall (x_0, x^*) , und es gelte

$$(5.5.3) \quad \lim_{x \uparrow x^*} \left(\frac{1}{h} \right)'(x) = 0.$$

Dann liegt X im Max-Anziehungsbereich der Gumbel-Verteilung Λ .

BEWEIS. Wir überprüfen die Bedingung von Satz 5.3.1 mit

$$g(x) := \frac{1}{h(x)}.$$

Wir zeigen, dass für jedes $u \in \mathbb{R}$,

$$(5.5.4) \quad \lim_{x \uparrow x^*} \frac{\bar{F}(x + ug(x))}{\bar{F}(x)} = e^{-u}.$$

Sei $u > 0$, denn für negatives u ist der Beweis analog. Mit Lemma 5.5.3 gilt

$$\frac{\bar{F}(x + ug(x))}{\bar{F}(x)} = \exp \left\{ - \int_x^{x+ug(x)} h(y) dy \right\} = \exp \left\{ - \int_0^u \frac{g(x)}{g(x + zg(x))} dz \right\},$$

wobei wir in der zweiten Gleichung den Ansatz $z := \frac{y-x}{g(x)} \in [0, u]$ gemacht haben. Für den Beweis von (5.5.4) reicht es zu zeigen, dass gleichmäßig in $z \in [0, u]$ gilt

$$(5.5.5) \quad \lim_{x \uparrow x^*} \frac{g(x)}{g(x + zg(x))} = 1.$$

Sei $\varepsilon > 0$. Wegen der von Mises-Bedingung $\lim_{x \uparrow x^*} g'(x) = 0$ gibt es ein $s_0 = s_0(\varepsilon) \in (x_0, x^*)$, so dass $|g'(s)| < \frac{\varepsilon}{u}$ für alle $s \in (s_0, x^*)$. Es folgt, dass für alle $z \in [0, u]$

$$|g(x + zg(x)) - g(x)| = \left| \int_x^{x+zg(x)} g'(s) ds \right| \leq \frac{\varepsilon}{u} zg(x) \leq \varepsilon g(x).$$

Somit gilt für alle $x \in (s_0, x^*)$

$$\left| \frac{g(x + zg(x)) - g(x)}{g(x)} \right| \leq \varepsilon.$$

Das beweist die Behauptung (5.5.5). □

Beispiel 5.5.9. Für die Exponentialverteilung mit Parameter λ gilt $h(x) = \lambda$ und somit

$$g(x) = \frac{1}{h(x)} = \frac{1}{\lambda}.$$

Die von-Mises Bedingung ist erfüllt, denn $(1/h)'(x) = 0$ für alle $x > 0$.

Beispiel 5.5.10. Die Dichte f der Standardnormalverteilung erfüllt die Differenzialgleichung $f'(x) = -xf(x)$. Somit ergibt sich

$$\left(\frac{1}{h}\right)'(x) = \left(\frac{\bar{F}}{f}\right)'(x) = \frac{-f^2(x) + \bar{F}(x)f'(x)}{f^2(x)} = -1 + \frac{\bar{F}(x)x}{f(x)} \rightarrow 0, \quad x \rightarrow +\infty,$$

denn wir haben in Lemma 5.4.1 gezeigt, dass $\bar{F}(x) \sim \frac{1}{x}f(x)$.

A. A. Balkema und L. de Haan (1972) haben bewiesen, dass eine Verteilungsfunktion F genau dann im Max-Anziehungsbereich der Gumbel-Verteilung liegt, wenn es eine Verteilungsfunktion G mit demselben rechten Endpunkt x^* wie F gibt, die die von Mises-Bedingung erfüllt und für die $\bar{F}(t) \sim \bar{G}(t)$ für $t \uparrow x^*$ gilt. In diesem Sinne ist die von Mises-Bedingung nicht nur hinreichend, sondern auch “bis auf asymptotische Äquivalenz” sogar notwendig.

Statistik der Extremwertverteilungen

In diesem Kapitel beschäftigen wir uns mit statistischen Anwendungen der Extremwertverteilungen. Wir werden zwei verschiedene Zugänge zur Modellierung von Extremwerten betrachten.

- Der erste Zugang basiert auf der Modellierung von *Blockmaxima* durch die bereits bekannten Extremwertverteilungen, die hier GEV-Verteilungen (Generalized Extreme-Value Distributions) genannt werden.
- Der zweite Zugang (*Peaks Over Threshold Method*) benutzt die verallgemeinerten Pareto-Verteilungen (GPD, Generalized Pareto Distributions).

Wir werden hier nur auf einige grundlegende Ideen der statistischen Modellierung von Extremwerten eingehen. Für mehr Einzelheiten verweisen wir auf die Bücher von S. Coles “*An introduction to statistical modeling of extreme values*”, E. Gumbel “*Statistics of extremes*”, J. Beirlant, Y. Goegebeur, J. Teugels, J. Segers “*Statistics of extremes*”.

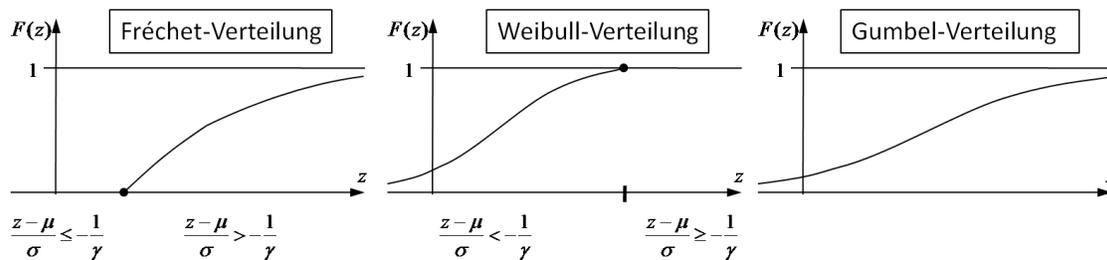
6.1. Statistik der Blockmaxima: GEV-Verteilungen

Wir haben bisher gesehen, dass Extremwertverteilungen folgende Form haben:

$$G_{\gamma,\mu,\sigma}(z) = \exp \left\{ - \left(1 + \gamma \frac{z - \mu}{\sigma} \right)^{-\frac{1}{\gamma}} \right\} \text{ für } 1 + \gamma \frac{z - \mu}{\sigma} > 0.$$

Extremwertverteilungen bilden also eine dreiparametrische Familie: $\gamma \in \mathbb{R}$ ist der formgebende Parameter, $\mu \in \mathbb{R}$ ist der Lageparameter und $\sigma > 0$ ist der Skalenparameter. Für γ gilt:

- $\gamma > 0$: G ist eine Fréchet-Verteilung (definiert für $\frac{z-\mu}{\sigma} > -\frac{1}{\gamma}$ wie oben, sonst 0).
- $\gamma < 0$: G ist eine Weibull-Verteilung (definiert für $\frac{z-\mu}{\sigma} < -\frac{1}{\gamma}$ wie oben, sonst 1).
- $\gamma = 0$: G ist eine Gumbel-Verteilung (definiert für $z \in \mathbb{R}$ wie oben).



Extremwertverteilungen, die in der obigen Form dargestellt werden, werden auch General Extreme-Value distributions (GEV-Verteilungen) genannt.

Beispiel 6.1.1 (Wasserstände an einem Deich). Am Tag $j \in \{1, \dots, 365\}$ im Jahr $i \in \{1, \dots, n\}$ wurde an einem Deich der Wasserstand x_{ij} gemessen. Wir betrachten die jährlichen Maxima (“Blockmaxima”)

$$x_i = \max_{j=1, \dots, 365} x_{ij}$$

und wollen aus diesen Daten die Deichhöhe Z_p bestimmen, bei der eine Überflutung mit einer gegebenen Wahrscheinlichkeit p in einem Jahr stattfindet. Dabei ist die Wahrscheinlichkeit p sehr klein (viel kleiner als $1/n$, zum Beispiel), so dass alle gemessenen Wasserstände sicherlich viel kleiner als die gesuchte Höhe Z_p sind.

Dazu betrachten wir folgendes Modell: x_1, \dots, x_n sind Realisierungen von X_1, \dots, X_n , die u.i.v. Zufallsvariablen mit einer GEV-Verteilung $G_{\gamma, \mu, \sigma}$ mit Parametervektor $\theta = (\gamma, \mu, \sigma) \in \mathbb{R}^2 \times \mathbb{R}_+$ sind.

Bemerkung 6.1.2. Den jährlichen Maxima eine GEV-Verteilung zu unterstellen, ist eine natürliche Wahl, da jedes X_i ein Maximum von vielen u.i.v. Zufallsvariablen ist. Wir haben in früheren Kapiteln gezeigt, dass solche Maxima unter sehr allgemeinen Bedingungen gegen Extremwertverteilungen konvergieren. Natürlich braucht man für die Konvergenz Normierungskonstanten, in unserem Fall kann man aber annehmen, dass die Normierungskonstanten bereits in den Parametern μ und σ enthalten sind.

Bemerkung 6.1.3. Da wir im obigen Modell voraussetzen, dass die X_i identisch verteilt sind, kann das Modell nur auf stationäre Daten angewendet werden, d.h. Daten, die keinen Trend aufweisen. Werden die jährlichen Maxima mit der Zeit immer größer (kleiner), muss ein anderes Modell verwendet werden, siehe unten.

Unser Problem besteht nun darin, den Parametervektor θ zu schätzen. Wir werden die *Maximum-Likelihood-Methode* (ML-Methode) benutzen. Dazu benötigt man die Dichte $f_{\gamma, \mu, \sigma}(z)$ der GEV-Verteilung. Durch Ableiten der Verteilungsfunktion $G_{\gamma, \mu, \sigma}$ erhält man, dass für $\gamma \neq 0$

$$f_{\theta}(z) = f_{\gamma, \mu, \sigma}(z) = \begin{cases} \frac{1}{\sigma} \left(1 + \gamma \frac{z - \mu}{\sigma}\right)^{-\frac{1}{\gamma} - 1} \exp\left\{-\left(1 + \gamma \frac{z - \mu}{\sigma}\right)^{-\frac{1}{\gamma}}\right\}, & 1 + \gamma \frac{z - \mu}{\sigma} > 0, \\ 0, & \text{sonst} \end{cases}$$

während für $\gamma = 0$

$$f_{\theta}(z) = f_{0, \mu, \sigma} = \frac{1}{\sigma} e^{-\frac{z - \mu}{\sigma}} \exp\left\{-e^{-\frac{z - \mu}{\sigma}}\right\}, \text{ für } z \in \mathbb{R}.$$

Mit Hilfe der Dichten kann man die Log-Likelihoodfunktion aufstellen:

$$l(\theta) := l(\theta | x_1, \dots, x_n) := \sum_{i=1}^n \log f_{\theta}(x_i).$$

Für $\gamma \neq 0$ gilt:

$$l(\theta) = -n \log \sigma - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^n \log \left(1 + \gamma \frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^n \left(1 + \gamma \frac{x_i - \mu}{\sigma}\right)^{-\frac{1}{\gamma}},$$

falls $1 + \gamma \frac{x_i - \mu}{\sigma} > 0$ für alle $i = 1, \dots, n$, und $l(\theta) = -\infty$ sonst. Für $\gamma = 0$ gilt:

$$l(\theta) = -n \log \sigma - \sum_{i=1}^n \frac{x_i - \mu}{\sigma} - \sum_{i=1}^n e^{-\frac{x_i - \mu}{\sigma}}.$$

Mit der log-Likelihoodfunktion lässt sich der Maximum-Likelihood-Schätzer

$$\hat{\theta} = (\hat{\gamma}, \hat{\mu}, \hat{\sigma}) = \operatorname{argmax} l(\gamma, \mu, \sigma)$$

herleiten. Hier kann $\hat{\theta}$ nicht analytisch bestimmt werden, sondern muss numerisch ermittelt werden.

Nachdem der Parameter θ geschätzt wurde, können wir die Deichhöhe Z_p schätzen. Wir erinnern, dass Z_p die Deichhöhe ist, bei der eine Überflutung mit Wahrscheinlichkeit p in einem Jahr stattfindet. Das Problem besteht also darin, dass $(1 - p)$ -Quantil des jährlichen Maximums zu schätzen. Wir schätzen Z_p indem wir die Gleichung

$$G_{\hat{\gamma}, \hat{\mu}, \hat{\sigma}}(\hat{Z}_p) = 1 - p$$

lösen (falls es mehrere Lösungen gibt, betrachten wir die kleinste):

$$\hat{Z}_p = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\gamma}} \{1 - (-\log(1 - p))^{-\frac{1}{\hat{\gamma}}}\}, & \hat{\gamma} \neq 0, \\ \hat{\mu} - \hat{\sigma} \log(-\log(1 - p)), & \hat{\gamma} = 0. \end{cases}$$

Für $\hat{\gamma} < 0$ (im Fall der Weibull-Verteilung) besitzt die Verteilung $G_{\hat{\gamma}, \hat{\mu}, \hat{\sigma}}$ einen endlichen rechten Endpunkt, der übrigens per Definition Z_0 ist. In diesem Fall gehen wir davon aus, dass es einen absolut höchsten Wasserstand gibt, der niemals überschritten wird. Der Schätzer für Z_0 ist dann gegeben durch:

$$\hat{Z}_0 = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\gamma}}.$$

Nachdem nun das Problem gelöst wurde, stellt sich die Frage, wie wir die Lösung verifizieren können. Wie können wir überprüfen, ob die Daten x_1, \dots, x_n durch die Verteilungsfunktion $\hat{G} = G_{\hat{\gamma}, \hat{\mu}, \hat{\sigma}}$ tatsächlich gut beschrieben werden? Zur Verifikation des Modells gibt es mehrere Methoden, die wir im Folgenden betrachten.

Ordnen wir die Stichprobe x_1, \dots, x_n monoton aufsteigend an, so erhalten wir die Ordnungsstatistiken

$$x_{(1)} \leq \dots \leq x_{(n)}.$$

Definition 6.1.4 (Probability-Plot). Der **PP-Plot** ist die Menge

$$\left\{ \left(\hat{G}(x_{(i)}), \frac{i}{n+1} \right) : i = 1, \dots, n \right\} \subset [0, 1]^2.$$

Trifft die Annahme, dass die Daten x_1, \dots, x_n gemäß \hat{G} verteilt sind zu, so sollte

$$\hat{G}(x_{(i)}) \approx \frac{i}{n+1}$$

gelten bzw. sollten die Punkte in einem Probability-Plot auf der Winkelhalbierenden liegen (etwa wie in Grafik 1).

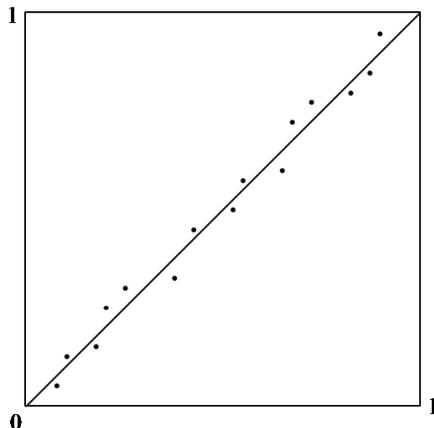


ABBILDUNG 1. PP-Plot

Der PP-Plot besitzt einen Nachteil: Für $i \approx n$ gilt $\hat{G}(x_{(i)}) \approx 1$ und $\frac{i}{n+1} \approx 1$, egal ob \hat{G} die Daten gut beschreibt oder nicht. Mit anderen Worten, auch wenn \hat{G} die Daten im Bereich der großen Werte nicht gut beschreibt, sieht man das in einem Probability-Plot möglicherweise nicht. Dabei sind gerade die großen Werte besonders interessant für uns. Wir betrachten deshalb eine andere Methode, die dieser Überlegung Rechnung trägt.

Das q -Quantil $\hat{G}^{\leftarrow}(q)$, wobei $q \in (0, 1)$, einer Verteilungsfunktion \hat{G} ist definiert als (die kleinste) Lösung z der Gleichung

$$\hat{G}(z) = q.$$

Definition 6.1.5 (Quantil-Plot). Der **QQ-Plot** ist die Menge

$$\left\{ \left(\hat{G}^{\leftarrow} \left(\frac{i}{n+1} \right), x_{(i)} \right) : i = 1, \dots, n \right\} \subset \mathbb{R}^2.$$

Beim QQ-Plot werden auf der horizontalen Achse die $\frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}$ -Quantile der Verteilung \hat{G} abgetragen und auf der vertikalen Achse die Ordnungsstatistiken $x_{(1)}, \dots, x_{(n)}$. Wenn \hat{G} die Daten gut beschreibt, sollte

$$\hat{G}^{\leftarrow} \left(\frac{i}{n+1} \right) \approx x_{(i)}$$

gelten bzw. sollten die Punkte in Abbildung 2 auf der Winkelhalbierenden liegen.

Es kann vorkommen, dass die Daten x_1, \dots, x_n einen Trend aufweisen (z.B. werden die jährlichen Maxima höher). Wir betrachten nun ein Modell, das der Nichtstationarität der Daten Rechnung trägt. Die beobachteten Blockmaxima x_1, \dots, x_n seien Realisierungen von

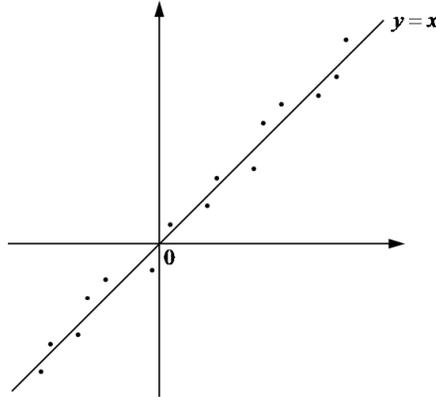


ABBILDUNG 2. QQ-Plot

Zufallsvariablen X_1, \dots, X_n , die unabhängig aber nicht identisch verteilt seien mit

$$X_i \sim G_{\gamma(i), \sigma(i), \mu(i)}, \quad i = 1, \dots, n.$$

Dabei ist der Parametervektor $(\gamma(i), \sigma(i), \mu(i))$ eine Funktion der Zeit i . Für diese Funktion kann man z.B. den folgenden Ansatz verwenden:

$$\gamma(i) = \gamma, \quad \sigma(i) = \sigma, \quad \mu(i) = \beta_0 + \beta_1 \cdot i.$$

Wir gehen also von einem konstanten Formparameter γ , einem konstanten Skalenparameter σ und einem linearen Trend, der im Lageparameter μ berücksichtigt wird, aus. Die Parameter $(\gamma, \sigma, \beta_0, \beta_1)$ lassen sich wieder mit der ML-Methode schätzen und somit lässt sich das Problem mit den bereits im Fall von stationären Daten betrachteten Methoden lösen. Möchte man das Modell verifizieren, so kann man PP- oder QQ-Plots erstellen. Davor muss man aber die Stichprobe x_1, \dots, x_n von dem Trend bereinigen:

$$x'_i := x_i - \hat{\beta}_0 - \hat{\beta}_1 i.$$

Die bereinigte Stichprobe x'_1, \dots, x'_n sollte man dann mit der Verteilungsfunktion $G_{\hat{\gamma}, 0, \hat{\sigma}}$ vergleichen.

Bemerkung 6.1.6. Der Ansatz kann verallgemeinert werden, ohne dass sich das Modell grundsätzlich ändert. So ist es zum Beispiel problemlos möglich, einen exponentiellen Trend zu modellieren:

$$\gamma(i) = \gamma, \quad \sigma(i) = \sigma, \quad \mu(i) = e^{\beta_0 + \beta_1 i}.$$

6.2. Peaks over Threshold: Statistik der GP-Verteilungen

Die oben beschriebene Methode basiert auf der Betrachtung von Blockmaxima (z.B. von jährlichen Maxima). Es gibt eine andere Methode (Peaks over Threshold), bei der man nur Beobachtungen berücksichtigt, die einen Schwellenwert überschreiten. Im Folgenden beschäftigen wir uns mit dieser Methode.

Wir fangen damit an, dass wir die verallgemeinerten Pareto-Verteilungen definieren. Es sei X eine Zufallsvariable, die man sich z.B. als eine Schadenhöhe vorstellen kann. Wir interessieren uns nur für die großen Werte von X und stellen die folgende Frage:

Wie ist der sogenannte **Exzess** $X - u$ asymptotisch verteilt, gegeben dass $X > u$? Dabei geht $u \rightarrow \infty$.

Wir betrachten drei Beispiele.

Beispiel 6.2.1. Sei X exponentialverteilt mit Parameter $\lambda > 0$, d.h. $\bar{F}(t) = e^{-\lambda t}$, $t > 0$. Dann gilt

$$\mathbb{P}[X - u > t | X > u] = \frac{\mathbb{P}[X > u + t, X > u]}{\mathbb{P}[X > u]} = \frac{\mathbb{P}[X > u + t]}{\mathbb{P}[X > u]} = \frac{e^{-\lambda(u+t)}}{e^{-\lambda u}} = e^{-\lambda t}.$$

Es gilt also: Die bedingte Verteilung von $X - u$ gegeben, dass $X > u$, ist die Exponentialverteilung mit Parameter λ . Dies ist die Gedächtnislosigkeit der Exponentialverteilung.

Beispiel 6.2.2. Sei X aus dem Max-Anziehungsbereich der Fréchet-Verteilung Φ_α , $\alpha > 0$. D.h., $\bar{F} \in RV_{-\alpha}$. Dann gilt für alle $t > 0$:

$$\mathbb{P}\left[\frac{X - u}{u} > t \mid X > u\right] = \frac{\mathbb{P}[X > u + ut]}{\mathbb{P}[X > u]} = \frac{\bar{F}(u(t+1))}{\bar{F}(u)} \rightarrow (1+t)^{-\alpha}$$

für $u \rightarrow \infty$. Es gilt also: gegeben, dass $X > u$, konvergiert die Verteilung von $(X - u)/u$ gegen die Verteilungsfunktion $1 - (1+t)^{-\alpha}$, $t > 0$.

Beispiel 6.2.3. Sei X standardnormalverteilt. Folgende Relation wurde in Lemma 5.4.1 mit der Regel von L'Hôpital bewiesen:

$$\mathbb{P}[X > u] \sim \frac{1}{\sqrt{2\pi}u} e^{-u^2/2} \text{ für } u \rightarrow \infty.$$

Unter Verwendung dieser Relation erhalten wir für jedes $t > 0$:

$$\mathbb{P}[u(X - u) > t | X > u] = \frac{\mathbb{P}[X > u + \frac{t}{u}]}{\mathbb{P}[X > u]} \sim \frac{\exp\{-\frac{u^2}{2} - t - \frac{t^2}{2u^2}\}}{\exp\{-\frac{u^2}{2}\}} \rightarrow e^{-t}$$

für $u \rightarrow \infty$. Es gilt also: gegeben, dass $X > u$, konvergiert die Verteilung von $u(X - u)$ gegen die Exponentialverteilung mit Parameter 1.

In allen drei Beispielen konnte die bedingte Verteilung von $X - u$ gegeben, dass $X > u$, durch eine Verteilung approximiert werden. Wir werden nun ein allgemeines Resultat formulieren, das die drei Beispiele als Spezialfälle beinhaltet.

Definition 6.2.4. Die **verallgemeinerte Pareto-Verteilung** (GPD, Generalized Pareto Distribution) mit Index $\gamma \in \mathbb{R}$ und Skalenparameter $\sigma > 0$ ist definiert durch die Verteilungsfunktion

$$P_{\gamma,\sigma}(t) = 1 - \left(1 + \frac{\gamma t}{\sigma}\right)^{-\frac{1}{\gamma}} \text{ mit } \begin{cases} t > 0, & \text{falls } \gamma > 0, \\ t \in [0, -\frac{\sigma}{\gamma}], & \text{falls } \gamma < 0. \end{cases}$$

Bemerkung 6.2.5. Für $\gamma = 0$ interpretieren wir die Formel als Grenzwert:

$$P_{0,\sigma}(t) = \lim_{\gamma \rightarrow 0} \left(1 - \left(1 + \frac{\gamma t}{\sigma} \right)^{-\frac{1}{\gamma}} \right) = 1 - e^{-t/\sigma}, \quad t > 0.$$

Somit stimmt $P_{0,\sigma}$ mit der Exponentialverteilung mit Parameter $1/\sigma$ überein.

Satz 6.2.6 (Pickands–Balkema–de Haan, 1974). Sei X eine Zufallsvariable mit Verteilungsfunktion F , die im rechten Endpunkt x^* stetig ist. Dann liegt F im Max–Anziehungsbereich von G_γ genau dann, wenn es eine positive messbare Funktion $\beta(u)$ gibt mit

$$\lim_{u \uparrow x^*} \sup_{t \in [0, x^* - u]} |\mathbb{P}[X - u \leq t | X > u] - P_{\gamma, \beta(u)}(t)| = 0.$$

Grob gesagt gilt die Approximation

$$\mathbb{P}[X - u \leq t | X > u] \approx P_{\gamma, \beta(u)}(t),$$

falls X im Max–Anziehungsbereich von G_γ liegt.

Nun werden wir die GP–Verteilungen in der Statistik anwenden. Es seien x_1, \dots, x_n unabhängige identisch verteilte Beobachtungen, z.B. Wasserstände an einem Deich an n Tagen. Wir interessieren uns nur für die extrem großen Beobachtungen. Das heißt, wir wählen einen Schwellenwert u und betrachten nur die Beobachtungen x_{i_1}, \dots, x_{i_k} , die u überschreiten. Wir definieren die Exzesse

$$y_1 = x_{i_1} - u, \dots, y_k = x_{i_k} - u$$

und ignorieren alle anderen Daten. Der Satz von Pickands–Balkema–de Haan macht folgendes Modell plausibel: Die Exzesse y_1, \dots, y_k sind Realisierungen von unabhängigen und identisch verteilten Zufallsvariablen Y_1, \dots, Y_k , die gemäß einer verallgemeinerten Pareto–Verteilung $P_{\gamma, \sigma}$ verteilt sind. Dabei sind $\gamma \in \mathbb{R}$ und $\sigma > 0$ unbekannte Parameter. Die Dichte von $P_{\gamma, \sigma}$ ist gegeben durch

$$f_{\gamma, \sigma}(t) = \frac{1}{\sigma} \left(1 + \frac{\gamma t}{\sigma} \right)^{-\frac{1}{\gamma} - 1} \quad \text{mit} \quad \begin{cases} t > 0, & \text{falls } \gamma > 0, \\ t \in [0, -\frac{\sigma}{\gamma}], & \text{falls } \gamma < 0, \end{cases}$$

und $f_{0, \sigma}(t) = \sigma e^{-t/\sigma}$, $t > 0$, für $\gamma = 0$. Damit ergibt sich für die Log–Likelihoodfunktion

$$l(\gamma, \sigma) := \sum_{i=1}^k \log f_{\gamma, \sigma}(y_i) = -k \log \sigma - \left(1 + \frac{1}{\gamma} \right) \sum_{i=1}^k \log \left(1 + \frac{\gamma y_i}{\sigma} \right),$$

zumindest für $\gamma \neq 0$. Der ML–Schätzer

$$(\hat{\gamma}, \hat{\sigma}) = \underset{\gamma, \sigma}{\operatorname{argmax}} l(\gamma, \sigma)$$

muss numerisch berechnet werden.

Nun werden wir für ein gegebenes kleines p die Deichhöhe Z_p schätzen, die an einem Tag mit Wahrscheinlichkeit p überflutet wird. Es sei X die Zufallsvariable, die den Wasserstand

an einem Tag beschreibt. Mit Berücksichtigung des Satzes von Pickands–Balkema–de Haan gehen wir davon aus, dass für großes u :

$$\mathbb{P}[X - u > t | X > u] \approx \left(1 + \frac{\gamma t}{\sigma}\right)^{-\frac{1}{\gamma}}.$$

Mit $t = Z_p - u$ folgt also:

$$\mathbb{P}[X > Z_p] \approx \mathbb{P}[X > u] \left(1 + \gamma \frac{Z_p - u}{\sigma}\right)^{-\frac{1}{\gamma}}.$$

Nun setzen wir die rechte Seite gleich p . Wenn die Gleichung nach Z_p umgestellt wird, erhält man schließlich

$$Z_p \approx u + \frac{\sigma}{\gamma} \left[\left(\frac{\mathbb{P}[X > u]}{p} \right)^\gamma - 1 \right].$$

Dabei haben wir Schätzer für σ und γ bereits hergeleitet. Die Wahrscheinlichkeit $\mathbb{P}[X > u]$ kann durch $\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{x_i > u}$ geschätzt werden. Es ergibt sich der Schätzer

$$\hat{Z}_p = u + \frac{\hat{\sigma}}{\hat{\gamma}} \left[\left(\frac{1}{np} \sum_{i=1}^n \mathbb{1}_{x_i > u} \right)^{\hat{\gamma}} - 1 \right].$$

Ordnungsstatistiken

Seien X_1, \dots, X_n unabhängige und identisch verteilte (u.i.v.) Zufallsvariablen mit Verteilungsfunktion F . Ordnen wir die Stichprobe X_1, \dots, X_n monoton aufsteigend an, so erhalten wir die sogenannten **Ordnungsstatistiken**

$$X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}.$$

Zum Beispiel ist $X_{1:n} = \min\{X_1, \dots, X_n\}$ und $X_{n:n} = \max\{X_1, \dots, X_n\}$.

7.1. Allgemeine Eigenschaften der Ordnungsstatistiken

Zuerst berechnen wir die Verteilungsfunktion der Ordnungsstatistik $X_{k:n}$.

Satz 7.1.1. Für alle $1 \leq k \leq n$ und $t \in \mathbb{R}$ gilt

$$\mathbb{P}[X_{k:n} \leq t] = \sum_{m=k}^n \binom{n}{m} F(t)^m \bar{F}(t)^{n-m}.$$

BEWEIS. Es sei $L := \sum_{i=1}^n \mathbb{1}_{X_i \leq t}$ die Anzahl der Elemente der Stichprobe X_1, \dots, X_n , die unterhalb von t liegen. Dann gilt $L \sim \text{Bin}(n, F(t))$ und somit

$$\mathbb{P}[X_{k:n} \leq t] = \mathbb{P}[L \geq k] = \sum_{m=k}^n \mathbb{P}[L = m] = \sum_{m=k}^n \binom{n}{m} F(t)^m \bar{F}(t)^{n-m}.$$

Im letzten Schritt haben wir die Formel für die Zähldichte einer Binomialverteilung benutzt. \square

Sind die u.i.v. Zufallsvariablen X_1, \dots, X_n absolut stetig, so können wir die Dichte der Ordnungsstatistik $X_{k:n}$ berechnen.

Satz 7.1.2. Es seien X_1, \dots, X_n u.i.v. Zufallsvariablen mit Dichte f und Verteilungsfunktion F . Dann ist für alle $1 \leq k \leq n$ die Dichte von $X_{k:n}$ gegeben durch

$$f_{X_{k:n}}(t) = n \binom{n-1}{k-1} f(t) F(t)^{k-1} \bar{F}(t)^{n-k} = k \binom{n}{k} f(t) F(t)^{k-1} \bar{F}(t)^{n-k}.$$

BEWEISIDEE. Man kann den Satz beweisen, indem man die Formel aus Satz 7.1.1 ableitet (siehe z.B. Skript "Mathematische Statistik", Satz 1.6.1). Dieser Weg führt zu komplizierten

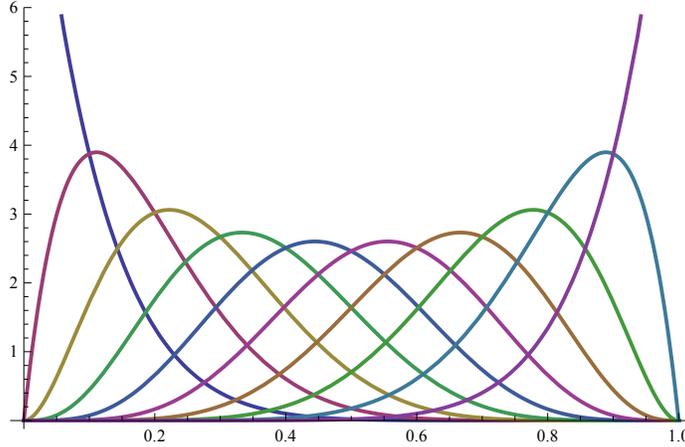


ABBILDUNG 1. Dichten der Ordnungsstatistiken $X_{1:10}, \dots, X_{10:10}$ einer unabhängigen und auf $[0, 1]$ gleichverteilten Stichprobe X_1, \dots, X_{10} .

Berechnungen. Wir geben hier einen anderen Beweis, der viel eleganter (allerdings nicht ganz streng) ist.

Damit $X_{k:n} = t$ ist, muss Folgendes passieren:

1. Eine der Zufallsvariablen, z.B. X_i , muss den Wert t annehmen. Es gibt n Möglichkeiten, das i auszuwählen. Die "Dichte" des Ereignisses $X_i = t$ ist $f(t)$.
2. Unter den restlichen $n - 1$ Zufallsvariablen müssen genau $k - 1$ Zufallsvariablen Werte kleiner als t annehmen. Wir haben $\binom{n-1}{k-1}$ Möglichkeiten, die $k - 1$ Zufallsvariablen auszuwählen. Die Wahrscheinlichkeit, dass die ausgewählten Zufallsvariablen allesamt kleiner als t sind, ist $F(t)^{k-1}$.
3. Die verbliebenen $n - k$ Zufallsvariablen müssen allesamt größer als t sein. Die Wahrscheinlichkeit davon ist $(1 - F(t))^{n-k}$.

Indem wir nun alles ausmultiplizieren, erhalten wir die Behauptung des Satzes. Den kombinatorischen Faktor kann man auch anders berechnen: Zuerst wählen wir aus n Zufallsvariablen k Zufallsvariablen, die $\leq t$ sind (dafür gibt es $\binom{n}{k}$ Möglichkeiten), und dann wählen aus diesen k Zufallsvariablen eine Zufallsvariable, die gleich t sein soll (dafür gibt es k Möglichkeiten). \square

Beispiel 7.1.3. Seien X_1, \dots, X_n unabhängig und gleichverteilt auf $[0, 1]$, d.h. die Dichte von X_i sei $f(t) = \mathbb{1}_{[0,1]}(t)$. Dann ist die Dichte von $X_{k:n}$ gegeben durch

$$f_{X_{k:n}}(t) = \begin{cases} k \binom{n}{k} t^{k-1} (1-t)^{n-k}, & t \in [0, 1], \\ 0, & \text{sonst.} \end{cases}$$

Das heißt, $X_{k:n}$ hat eine Beta-Verteilung $\text{Beta}(k, n - k + 1)$.

Aufgabe 7.1.4. Zeigen Sie, dass $\mathbb{E}X_{k:n} = \frac{k}{n+1}$ für alle $1 \leq k \leq n$.

Aufgabe 7.1.5. Seien X_1, \dots, X_n unabhängige und identisch verteilte Zufallsvariablen mit Dichte f und Verteilungsfunktion F . Bestimmen Sie für $1 \leq i < j \leq n$ die gemeinsame Dichte $f_{X_{i:n}, X_{j:n}}(t, s)$ der Ordnungsstatistiken $X_{i:n}$ und $X_{j:n}$.

Im nächsten Satz bestimmen wir die gemeinsame Dichte *aller* n Ordnungsstatistiken.

Satz 7.1.6. Seien X_1, \dots, X_n unabhängig und identisch verteilt mit Dichte f . Dann gilt für die gemeinsame Dichte von $X_{1:n}, \dots, X_{n:n}$:

$$f_{X_{1:n}, \dots, X_{n:n}}(t_1, \dots, t_n) = \begin{cases} n! \cdot f(t_1) \cdot \dots \cdot f(t_n), & t_1 < t_2 < \dots < t_n, \\ 0, & \text{sonst.} \end{cases}$$

BEWEISIDEE. Da die Ordnungsstatistiken per Definition aufsteigend sind, ist die Dichte gleich 0, wenn die Bedingung $t_1 < \dots < t_n$ nicht erfüllt ist. Sei also die Bedingung $t_1 < \dots < t_n$ erfüllt. Damit $X_{1:n} = t_1, \dots, X_{n:n} = t_n$ ist, muss eine der Zufallsvariablen (für deren Wahl es n Möglichkeiten gibt) gleich t_1 sein, eine andere (für deren Wahl es $n - 1$ Möglichkeiten gibt) gleich t_2 , usw. Wir haben also $n!$ Möglichkeiten für die Wahl der Reihenfolge der Variablen. Zum Beispiel tritt für $n = 2$ das Ereignis $\{X_{1:2} = t_1, X_{2:2} = t_2\}$ genau dann ein, wenn entweder $\{X_1 = t_1, X_2 = t_2\}$ oder $\{X_1 = t_2, X_2 = t_1\}$ eintritt, was 2 Möglichkeiten ergibt. Da alle Möglichkeiten sich nur durch Permutationen unterscheiden und somit die gleiche "Dichte" besitzen, betrachten wir nur eine Möglichkeit und multiplizieren dann das Ergebnis mit $n!$. Die einfachste Möglichkeit ist, dass $\{X_{1:n} = t_1, \dots, X_{n:n} = t_n\}$ eintritt. Diesem Ereignis entspricht die "Dichte" $f(t_1) \cdot \dots \cdot f(t_n)$, da die Zufallsvariablen X_1, \dots, X_n unabhängig sind. Multiplizieren wir nun diese Dichte mit $n!$, so erhalten wir das gewünschte Ergebnis. \square

Auch wenn die Zufallsvariablen X_1, \dots, X_n unabhängig sind, sind die Ordnungsstatistiken $X_{1:n}, \dots, X_{n:n}$ im Allgemeinen nicht unabhängig. Es gilt jedoch eine schwächere Eigenschaft, die sogenannte Markov-Eigenschaft.

Satz 7.1.7 (Bedingte Unabhängigkeit von Ordnungsstatistiken). Seien X_1, \dots, X_n unabhängige und identisch verteilte Zufallsvariablen mit Dichte f . Die bedingte Dichte $f(u|x_1, \dots, x_k)$ von $X_{k+1:n}$ gegeben, dass $X_{1:n} = x_1, \dots, X_{k:n} = x_k$, stimmt mit der bedingten Dichte $f(u|x_k)$ von $X_{k+1:n}$ gegeben, dass $X_{k:n} = x_k$, überein.

Bemerkung 7.1.8. Angenommen, die ersten k Ordnungsstatistiken der Stichprobe X_1, \dots, X_n sind bekannt: $X_{1:n} = x_1, \dots, X_{k:n} = x_k$. Wo liegt nun die nächste Ordnungsstatistik $X_{k+1:n}$? Der obige Satz behauptet, dass für die Beantwortung dieser Frage nur der Wert $X_{k:n} = x_k$ relevant ist. Die Werte der vorherigen Ordnungsstatistiken x_1, \dots, x_{k-1} tauchen in der bedingten Verteilung von $X_{k+1:n}$ nicht auf.

BEWEIS. Sei $f_{1,\dots,k:n}$ die gemeinsame Dichte von $X_{1:n}, \dots, X_{k:n}$. Dann gilt

$$(7.1.1) \quad f(u|x_1, \dots, x_k) = \frac{f_{1,\dots,k+1:n}(x_1, \dots, x_k, u)}{f_{1,\dots,k:n}(x_1, \dots, x_k)}.$$

Sei $f_{k,k+1:n}$ die gemeinsame Dichte von $(X_{k:n}, X_{k+1:n})$ und $f_{k:n}$ die Dichte von $X_{k:n}$. Dann gilt

$$(7.1.2) \quad f(u|x_k) = \frac{f_{k,k+1:n}(x_k, u)}{f_{k:n}(x_k)}.$$

Für die gemeinsame Dichte von $X_{1:n}, \dots, X_{k:n}$ erhält man

$$f_{1,\dots,k:n}(x_1, \dots, x_k) = f(x_1) \cdot \dots \cdot f(x_k) \cdot (1 - F(x_k))^{n-k} \cdot n \cdot (n-1) \cdot \dots \cdot (n-k+1),$$

falls $x_1 < \dots < x_k$ und 0 sonst.

Analog erhält man für die gemeinsame Dichte von $X_{1:n}, \dots, X_{k+1:n}$

$$f_{1,\dots,k+1:n}(x_1, \dots, x_{k+1}) = f(x_1) \cdot \dots \cdot f(x_{k+1}) \cdot (1 - F(x_{k+1}))^{n-k-1} \cdot n \cdot (n-1) \cdot \dots \cdot (n-k),$$

falls $x_1 < \dots < x_{k+1}$ und 0 sonst. Außerdem gilt:

$$f_{k,k+1:n}(x_k, u) = f(x_k) \cdot f(u) \cdot F(x_k)^{k-1} (1 - F(u))^{n-k-1} \cdot n(n-1) \binom{n-2}{k-1}$$

bzw.

$$f_{k:n}(x_k) = f(x_k) \cdot F(x_k)^{k-1} (1 - F(x_k))^{n-k} \cdot n \binom{n-1}{k-1}.$$

Einsetzen in (7.1.1) bzw. (7.1.2) ergibt

$$(7.1.1) = (7.1.2) = \frac{(n-k)(1 - F(u))^{n-k-1} f(u)}{(1 - F(x_k))^{n-k}}.$$

Somit stimmen die beiden bedingten Dichten $f(u|x_1, \dots, x_k)$ und $f(u|x_k)$ überein. □

Aufgabe 7.1.9. Seien X_1, \dots, X_n unabhängig und exponentialverteilt mit Parameter $\lambda > 0$. Welche Verteilung besitzt $X_{k+1:n} - x_k$ gegeben, dass $X_{1:n} = x_1, \dots, X_{k:n} = x_k$?

7.2. Extreme Ordnungsstatistiken

Es seien X_1, X_2, \dots unabhängige identisch verteilte Zufallsvariablen. Wir bezeichnen mit $X_{1:n} \leq \dots \leq X_{n:n}$ die Ordnungsstatistiken von X_1, \dots, X_n . Außerdem benutzen wir die Notation

$$M_n^{(k)} = X_{n-k+1:n}, \quad k = 1, \dots, n.$$

Somit ist $M_n^{(1)} = X_{n:n}$ das größte Element der Stichprobe, $M_n^{(2)} = X_{n-1:n}$ das zweitgrößte Element und so weiter.

Wir haben die möglichen Grenzwertverteilungen von $M_n^{(1)} = M_n$ für $n \rightarrow \infty$ bereits in früheren Kapiteln beschrieben (Extremwertverteilungen). In diesem Abschnitt beschreiben

wir die Grenzverteilungen der sogenannten *extremen Ordnungsstatistiken* $M_n^{(k)}$, wobei $k \in \mathbb{N}$ fest und $n \rightarrow \infty$.

Zuerst benötigen wir einen Satz, der die Grenzwertverteilung für die Anzahl der Überschreitungen eines hohen Schwellenwerts beschreibt.

Satz 7.2.1. Sei $u_n \in \mathbb{R}$ eine Folge mit $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau$, wobei $0 \leq \tau < \infty$. Dann gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sum_{i=1}^n \mathbb{1}_{X_i > u_n} = k \right] = e^{-\tau} \frac{\tau^k}{k!}, \quad k = 0, 1, \dots$$

BEWEIS. Wir bezeichnen mit $S_n := \sum_{i=1}^n \mathbb{1}_{X_i > u_n}$ die Anzahl der Beobachtungen, die oberhalb von u_n liegen. Es gilt $S_n \sim \text{Bin}(n, \bar{F}(u_n))$. Nach Voraussetzung gilt $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau$. Da dies gilt, darf der Poissongrenzwertsatz angewendet werden:

$$S_n \xrightarrow[n \rightarrow \infty]{d} \text{Poi}(\tau).$$

Mit anderen Worten, $\lim_{n \rightarrow \infty} \mathbb{P}[S_n = k] = e^{-\tau} \frac{\tau^k}{k!}$, für alle $k = 0, 1, \dots$ □

Bemerkung 7.2.2. Der Satz behält seine Gültigkeit für $\tau = +\infty$ (niedriger Schwellenwert, sehr viele Überschreitungen) im folgenden Sinne:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sum_{i=1}^n \mathbb{1}_{X_i > u_n} = k \right] = 0, \quad k = 0, 1, \dots$$

Eine $\text{Poi}(+\infty)$ -verteilte Zufallsvariable interpretiert man dabei als eine Zufallsvariable, die mit Wahrscheinlichkeit 1 den Wert $+\infty$ annimmt.

Satz 7.2.3. Es gebe Normierungsfolgen $a_n > 0$ und $b_n \in \mathbb{R}$, so dass

$$\frac{M_n^{(1)} - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} G,$$

wobei G eine Extremwertverteilung ist. Dann gilt für alle $k = 1, 2, \dots$ und $x \in \mathbb{R}$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n^{(k)} - b_n}{a_n} \leq x \right] = \begin{cases} G(x) \cdot \sum_{s=0}^{k-1} \frac{(-\log G(x))^s}{s!}, & G(x) \neq 0, \\ 0, & G(x) = 0. \end{cases}$$

BEWEIS. Setzt man $u_n := a_n x + b_n$ und $S_n = \sum_{i=1}^n \mathbb{1}_{X_i > u_n}$, so ergibt sich:

$$\mathbb{P} [M_n^{(k)} \leq a_n x + b_n] = \mathbb{P}[S_n \leq k - 1] = \sum_{s=0}^{k-1} \mathbb{P}[S_n = s].$$

Mit Satz 7.2.1 folgt:

$$(7.2.1) \quad \lim_{n \rightarrow \infty} \sum_{s=0}^{k-1} \mathbb{P}[S_n = s] = \sum_{s=0}^{k-1} e^{-\tau} \frac{\tau^s}{s!}$$

mit $\tau = \lim_{n \rightarrow \infty} n\bar{F}(u_n)$. Es bleibt also noch die Bestimmung von τ . Nach Voraussetzung gilt:

$$\lim_{n \rightarrow \infty} F^n(ax + b_n) = G(x).$$

Für $G(x) \neq 0$ ist dies äquivalent zu

$$\lim_{n \rightarrow \infty} n \log(1 - \bar{F}(u_n)) = \log G(x).$$

Unter Verwendung der Taylorentwicklung des Logarithmus $\log(1 - x) = -x + o(x)$, $x \rightarrow 0$, folgt

$$\tau = \lim_{n \rightarrow \infty} nF(u_n) = -\log G(x)$$

Einsetzen von τ in (7.2.1) liefert die Behauptung. Für $G(x) = 0$ ist der Beweis analog, liefert allerdings $\tau = +\infty$. \square

Im Folgenden beschreiben wir die Grenzwertverteilung des Vektors $(M_n^{(1)}, \dots, M_n^{(r)})$, mit r fest und $n \rightarrow \infty$. Dazu benötigen wir eine Verallgemeinerung von Satz 7.2.1, die r verschiedene Folgen von Schwellenwerten zulässt. Betrachte die Schwellenwerte $u_n^{(1)} \geq u_n^{(2)} \geq \dots \geq u_n^{(r)}$ mit

$$\lim_{n \rightarrow \infty} n\bar{F}(u_n^{(1)}) = \tau_1, \quad \lim_{n \rightarrow \infty} n\bar{F}(u_n^{(2)}) = \tau_1 + \tau_2, \quad \dots, \quad \lim_{n \rightarrow \infty} n\bar{F}(u_n^{(r)}) = \tau_1 + \dots + \tau_r,$$

wobei $\tau_1, \dots, \tau_r \in [0, \infty)$. Sei

$$S_n^{(l)} := \sum_{i=1}^n \mathbb{1}_{X_i > u_n^{(l)}}$$

die Anzahl der Überschreitungen des Schwellenwerts $u_n^{(l)}$, $1 \leq l \leq r$.

Satz 7.2.4. Für alle $k_1, k_2, \dots, k_r \in \{0, 1, 2, \dots\}$ gilt

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \dots, S_n^{(r)} = k_1 + \dots + k_r] = e^{-(\tau_1 + \dots + \tau_r)} \frac{\tau_1^{k_1} \tau_2^{k_2} \dots \tau_r^{k_r}}{k_1! k_2! \dots k_r!}.$$

Bemerkung 7.2.5. Mit anderen Worten, es besteht die folgende Verteilungskonvergenz

$$(S_n^{(1)}, S_n^{(2)} - S_n^{(1)}, \dots, S_n^{(r)} - S_n^{(r-1)}) \xrightarrow[n \rightarrow \infty]{d} (\text{Poi}(\tau_1), \dots, \text{Poi}(\tau_r)),$$

wobei die Komponenten der Grenzwertverteilung unabhängig sind.

BEWEIS. Definiere $p_{n,l} = \bar{F}(u_n^{(l)}) - \bar{F}(u_n^{(l-1)})$ für $1 \leq l \leq r$, wobei $\bar{F}(u_n^{(0)})$ als 0 interpretiert wird. Setzt man

$$P_n(k_1, \dots, k_r) = \mathbb{P}[S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \dots, S_n^{(r)} = k_1 + \dots + k_r],$$

dann gilt

$$(7.2.2) \quad P_n(k_1, \dots, k_n) = \binom{n}{k_1} p_{n,1}^{k_1} \cdot \binom{n-k_1}{k_2} p_{n,2}^{k_2} \cdot \dots \cdot \binom{n-\sum_{i=1}^{r-1} k_i}{k_r} p_{n,r}^{k_r} \cdot (1-p_{n,1}-\dots-p_{n,r})^{n-\sum_{i=1}^r k_i}.$$

Damit kann der Beweis faktorweise durchgeführt werden. Für den ersten Faktor gilt

$$\binom{n}{k_1} p_{n,1}^{k_1} = \frac{n(n-1)\dots(n-k_1+1)}{k_1!} \cdot \bar{F}(u_n^{(1)})^{k_1} \sim \frac{n^{k_1} \bar{F}(u_n^{(1)})^{k_1}}{k_1!} \xrightarrow{n \rightarrow \infty} \frac{\tau_1^{k_1}}{k_1!}.$$

Ähnlich gilt für den zweiten Faktor

$$\binom{n-k_1}{k_2} p_{n,2}^{k_2} \sim \frac{n^{k_2} (\bar{F}(u_n^{(2)}) - \bar{F}(u_n^{(1)}))^{k_2}}{k_2!} \xrightarrow{n \rightarrow \infty} \frac{\tau_2^{k_2}}{k_2!},$$

was sich nun analog für sämtliche Faktoren, außer dem letzten, fortführen lässt. Für den letzten Faktor gilt schließlich:

$$(1-p_{n,1}-\dots-p_{n,r})^{n-k_1-\dots-k_r} \xrightarrow{n \rightarrow \infty} e^{-(\tau_1+\dots+\tau_r)},$$

denn es gilt:

$$(n-k_1-\dots-k_r)(p_{n,1}+\dots+p_{n,r}) = (n-k_1-\dots-k_r) \bar{F}(u_n^{(r)}) \xrightarrow{n \rightarrow \infty} \tau_1+\dots+\tau_r.$$

Setzt man alles zusammen, so erhält man die Behauptung. \square

Nun können wir die gemeinsame Verteilung der oberen r extremen Ordnungsstatistiken $(M_n^{(1)}, \dots, M_n^{(r)})$ beschreiben. Es gilt nämlich

$$\mathbb{P}[M_n^{(1)} \leq u_n^{(1)}, \dots, M_n^{(r)} \leq u_n^{(r)}] = \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} \leq 1, S_n^{(3)} \leq 2, \dots, S_n^{(r)} \leq r-1].$$

Der Grenzwert der rechten Seite kann mit dem obigen Satz berechnet werden, was zu einer langen Formel führt. Im nächsten Satz betrachten wir den Spezialfall $r = 2$, d.h. wir beschreiben die gemeinsame Grenzwertverteilung der zwei größten Werte der Stichprobe.

Satz 7.2.6. Es gebe Normierungsfolgen $a_n > 0$ und $b_n \in \mathbb{R}$, so dass

$$\frac{M_n^{(1)} - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} G,$$

wobei G eine Extremwertverteilung ist. Dann gilt für alle $x_1 > x_2$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n^{(1)} - b_n}{a_n} \leq x_1, \frac{M_n^{(2)} - b_n}{a_n} \leq x_2 \right] = G(x_2)(\log G(x_1) - \log G(x_2) + 1),$$

falls $G(x_2) \neq 0$ und 0 sonst.

BEWEIS. Setze $u_n^{(1)} := a_n x_1 + b_n$ bzw. $u_n^{(2)} := a_n x_2 + b_n$. Es ergibt sich aus

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x) \Rightarrow \lim_{n \rightarrow \infty} n \bar{F}(a_n x + b_n) = -\log G(x),$$

dass

$$\begin{aligned}\tau_1 &:= \lim_{n \rightarrow \infty} n\bar{F}(u_n^{(1)}) = -\log G(x_1), \\ \tau_2 &:= \lim_{n \rightarrow \infty} n\bar{F}(u_n^{(2)}) = -\log G(x_2).\end{aligned}$$

Es gilt:

$$\begin{aligned}\mathbb{P}[M_n^{(1)} \leq u_n^{(1)}, M_n^{(2)} \leq u_n^{(2)}] &= \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} \leq 1] \\ &= \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 0] + \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 1].\end{aligned}$$

Mit Satz 7.2.4 folgt dann

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n^{(1)} \leq u_n^{(1)}, M_n^{(2)} \leq u_n^{(2)}] = e^{-\tau_2} + (\tau_2 - \tau_1)e^{-\tau_2}.$$

Setzt man die Werte von τ_1 und τ_2 ein, so erhält man die Behauptung. \square

Bemerkung 7.2.7. Für ein $r > 2$ ist die Formel für die Verteilungsfunktion der Grenzwertverteilung von $(M_n^{(1)}, \dots, M_n^{(r)})$ sehr kompliziert. Allerdings ist die Formel für die Dichte dieser Verteilung sehr einfach; wir werden diese später unter Verwendung von Poisson-Prozessen herleiten.

7.3. Darstellungen der Ordnungsstatistiken

Eine Zufallsvariable Z heißt exponentialverteilt mit Parameter $\lambda > 0$, falls

$$\mathbb{P}[Z > t] = e^{-\lambda t} \text{ für } t \geq 0.$$

Die Ordnungsstatistiken der exponentialverteilten Zufallsvariablen besitzen eine besonders schöne Darstellung.

Satz 7.3.1. Seien Z_1, \dots, Z_n unabhängige und mit Parameter $\lambda = 1$ exponentialverteilte Zufallsvariablen. Dann gilt

$$(Z_{1:n}, Z_{2:n}, \dots, Z_{n:n}) \stackrel{d}{=} \left(\frac{\nu_1}{n}, \frac{\nu_1}{n} + \frac{\nu_2}{n-1}, \dots, \frac{\nu_1}{n} + \frac{\nu_2}{n-1} + \dots + \frac{\nu_n}{1} \right),$$

wobei ν_1, \dots, ν_n unabhängig und exponentialverteilt mit Parameter $\lambda = 1$ sind.

Bemerkung 7.3.2. Die Abstände $Z_{1:n}, Z_{2:n} - Z_{1:n}, \dots, Z_{n:n} - Z_{n-1:n}$ sind somit unabhängig und (nicht identisch) exponentialverteilt:

$$Z_{k:n} - Z_{k-1:n} \stackrel{d}{=} \frac{\nu_k}{n-k+1} \sim \text{Exp}(n-k+1).$$

BEWEIS VON SATZ 7.3.1. Die Dichte von $(Z_{1:n}, \dots, Z_{n:n})$ ist gegeben durch

$$f(x_1, \dots, x_n) = \begin{cases} n! \cdot e^{-x_1} \cdot \dots \cdot e^{-x_n}, & 0 < x_1 < \dots < x_n, \\ 0, & \text{sonst.} \end{cases}$$

Wegen der Unabhängigkeit ist die Dichte von (ν_1, \dots, ν_n) gegeben durch

$$g(y_1, \dots, y_n) = \begin{cases} e^{-y_1} \cdot \dots \cdot e^{-y_n}, & y_i > 0 \text{ für alle } i \in \{1, \dots, n\}, \\ 0, & \text{sonst.} \end{cases}$$

Wir betrachten die lineare Transformation

$$T : (y_1, \dots, y_n) \mapsto \left(\frac{y_1}{n}, \frac{y_1}{n} + \frac{y_2}{n-1}, \dots, \frac{y_1}{n} + \frac{y_2}{n-1} + \dots + \frac{y_n}{1} \right) =: (z_1, z_2, \dots, z_n).$$

Die Jacobi–Determinante dieser Transformation ist

$$J := \begin{vmatrix} \frac{1}{n} & 0 & 0 & \dots & 0 \\ \frac{1}{n} & \frac{1}{n-1} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n-1} & \dots & \frac{1}{2} & 0 \\ \frac{1}{n} & \frac{1}{n-1} & \dots & \frac{1}{2} & 1 \end{vmatrix} = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \dots \cdot \frac{1}{2} \cdot 1 = \frac{1}{n!}.$$

Mit der Transformationsformel können wir die Dichte von $T(\nu_1, \dots, \nu_n)$ berechnen:

$$h(z_1, \dots, z_n) = n!g(y_1, \dots, y_n) = \begin{cases} n!e^{-(y_1+\dots+y_n)} = n!e^{-(z_1+\dots+z_n)}, & 0 < z_1 < \dots < z_n, \\ 0, & \text{sonst.} \end{cases}$$

Bei obiger Transformation wurde die Dichte von (ν_1, \dots, ν_n) durch die Jacobi–Determinante geteilt, um zur Dichte von $T(\nu_1, \dots, \nu_n)$ zu gelangen. Man sieht, dass die Dichte f von $(Z_{1:n}, \dots, Z_{n:n})$ und die Dichte h von $T(\nu_1, \dots, \nu_n)$ übereinstimmen, was die Behauptung impliziert. \square

Bemerkung 7.3.3. Es seien $Z_1, Z_2, \dots \sim \text{Exp}(1)$ unabhängig. Für das Maximum $M_n = \max\{Z_1, \dots, Z_n\}$ gilt mit Satz 7.3.1

$$\mathbb{E}M_n = \mathbb{E}Z_{n:n} = \sum_{k=1}^n \frac{1}{k} = \log n + \gamma + o(1), \quad n \rightarrow \infty,$$

wobei $\gamma \approx 0.5772$ die Euler–Konstante ist. Auf der anderen Seite gilt aber auch

$$M_n - \log n \xrightarrow[n \rightarrow \infty]{d} G,$$

wobei G eine Gumbel–verteilte Zufallsvariable ist. Wendet man nun auf beide Seiten den Erwartungswert an (die Begründung lassen wir weg), so erhält man eine Formel für der Erwartungswert der Gumbel–Verteilung:

$$\mathbb{E}G = \gamma.$$

Analog lässt sich zeigen, dass $\text{Var } G = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Aufgabe 7.3.4. Berechnen Sie mit Satz 7.3.1 $\mathbb{E}Z_{k:n}$, $\text{Var } Z_{k:n}$ und $\text{Cov}(Z_{k:n}, Z_{l:n})$.

Die Ordnungsstatistiken der Zufallsvariablen, die gemäß einer Gleichverteilung verteilt sind, besitzen ebenfalls eine schöne Darstellung.

Satz 7.3.5. Die Zufallsvariablen U_1, \dots, U_n seien unabhängig und gleichverteilt auf dem Intervall $[0, 1]$. Dann gilt:

$$(U_{1:n}, U_{2:n}, \dots, U_{n:n}) \stackrel{d}{=} \left(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}} \right),$$

wobei $S_k = \nu_1 + \dots + \nu_k$ und ν_1, \dots, ν_{n+1} unabhängige, mit Parameter 1 exponentialverteilte Zufallsvariablen sind.

Bemerkung 7.3.6. Die Folge $S_1 < S_2 < \dots$ heißt der **Poisson-Prozess** auf $(0, \infty)$ mit Intensität 1. Intuitiv kann man die Aussage des Satzes so verstehen, dass die ersten n Punkte S_1, \dots, S_n des Poisson-Prozesses auf dem Intervall $[0, S_{n+1}]$ gleichverteilt sind.

BEWEIS. Die Dichte von $(U_{1:n}, \dots, U_{n:n})$ ist gegeben durch:

$$f(x_1, \dots, x_n) = \begin{cases} n!, & 0 < x_1 < x_2 < \dots < x_n < 1, \\ 0, & \text{sonst.} \end{cases}$$

Die Dichte von $(\nu_1, \dots, \nu_{n+1})$ ist wegen der Unabhängigkeit gegeben durch:

$$g(y_1, \dots, y_{n+1}) = \begin{cases} e^{-y_1} \cdot \dots \cdot e^{-y_{n+1}}, & y_i > 0 \text{ für alle } 1 \leq i \leq n+1, \\ 0, & \text{sonst.} \end{cases}$$

Wir betrachten die lineare Transformation

$$T : (y_1, \dots, y_{n+1}) \mapsto (y_1, y_1 + y_2, \dots, y_1 + y_2 + \dots + y_{n+1}) =: (z_1, z_2, \dots, z_{n+1}).$$

Die Jacobi-Determinante von T ist

$$J_T := \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{vmatrix} = 1.$$

Mit der Transformationsformel errechnen wir die Dichte von $(S_1, S_2, \dots, S_{n+1})$:

$$h(z_1, \dots, z_{n+1}) = \begin{cases} e^{-(y_1 + \dots + y_{n+1})} = e^{-z_{n+1}}, & 0 < z_1 < z_2 < \dots < z_{n+1}, \\ 0, & \text{sonst.} \end{cases}$$

Nun betrachten wir eine weitere (diesmal nichtlineare) Transformation

$$S : (z_1, \dots, z_{n+1}) \mapsto \left(\frac{z_1}{z_{n+1}}, \frac{z_2}{z_{n+1}}, \dots, \frac{z_n}{z_{n+1}}, z_{n+1} \right) =: (w_1, w_2, \dots, w_n, w_{n+1}).$$

Die Jacobi-Determinante von S ist

$$J_S(z_1, \dots, z_{n+1}) = \begin{vmatrix} \frac{1}{z_{n+1}} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{z_{n+1}} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{z_{n+1}} & 0 \\ \frac{-z_1}{z_{n+1}^2} & \frac{-z_2}{z_{n+1}^2} & \dots & \frac{-z_n}{z_{n+1}^2} & 1 \end{vmatrix} = \frac{1}{z_{n+1}^n} = \frac{1}{w_{n+1}^n}.$$

Die Dichte von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}}, S_{n+1})$ kann nun mit der Transformationsformel wie folgt berechnet werden:

$$p(w_1, \dots, w_{n+1}) = \begin{cases} e^{-z_{n+1}} z_{n+1}^n = e^{-w_{n+1}} w_{n+1}^n, & 0 < w_1 < w_2 < \dots < w_n < 1, w_{n+1} > 0, \\ 0, & \text{sonst.} \end{cases}$$

Die Dichte von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$ ist eine Marginaldichte von p :

$$r(w_1, \dots, w_n) = \begin{cases} \int_{\mathbb{R}} p(w_1, \dots, w_{n+1}) dw_{n+1} = n!, & 0 < w_1 < w_2 < \dots < w_n < 1, \\ 0, & \text{sonst.} \end{cases}$$

Man sieht, dass die Dichte f von $(U_{1:n}, \dots, U_{n:n})$ und die Dichte r von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$ übereinstimmen, woraus die Behauptung folgt. \square

Aufgabe 7.3.7. Es seien U_1, \dots, U_n unabhängige, auf dem Intervall $[0, 1]$ gleichverteilte Zufallsvariablen. Zeigen Sie, dass für jedes $k \in \mathbb{N}$ die folgende Verteilungskonvergenz gilt:

$$(nU_{1:n}, \dots, nU_{k:n}) \xrightarrow[n \rightarrow \infty]{d} (\nu_1, \nu_1 + \nu_2, \dots, \nu_1 + \dots + \nu_k),$$

wobei ν_1, ν_2, \dots unabhängige Zufallsvariablen mit $\mathbb{P}[\nu_i > t] = e^{-t}$, $t > 0$, sind (Standard-exponentialverteilung). Zeigen Sie auch, dass

$$(n(1 - U_{n:n}), \dots, n(1 - U_{n-k+1:n})) \xrightarrow[n \rightarrow \infty]{d} (\nu_1, \nu_1 + \nu_2, \dots, \nu_1 + \dots + \nu_k).$$

7.4. Wiederkehrperiode und Gumbel-Überschreitungsmethode

Seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen und sei $u \in \mathbb{R}$ ein Schwellenwert. Es sei $p := \mathbb{P}[X_i > u] \neq 0$. Wir definieren

$$T_1 := \min\{i \in \mathbb{N} : X_i > u\} \text{ und } T_{n+1} := \min\{i > T_n : X_i > u\}.$$

Man kann sich T_1, T_2, \dots als Zeitpunkte denken, zu denen ein bestimmtes Ereignis (eine Überflutung) eintritt. Wir definieren außerdem $W_n = T_n - T_{n-1}$, wobei $T_0 = 0$ gesetzt wird.

Satz 7.4.1. Für alle $n \in \mathbb{N}$ gilt $W_n \sim \text{Geo}(p)$, d.h. $\mathbb{P}[W_n = k] = p(1-p)^{k-1}$ für alle $k = 1, 2, \dots$

Bemerkung 7.4.2. Es folgt $\mathbb{E}W_n = \frac{1}{p}$. Die Zahl $\frac{1}{p}$ heißt die **Wiederkehrperiode**. Es ist der erwartete Abstand zwischen zwei aufeinanderfolgenden Überflutungen.

BEWEIS VON SATZ 7.4.1. Wir beweisen den Satz nur für $n = 1$:

$$\mathbb{P}[W_1 = k] = \mathbb{P}[X_1 \leq u, \dots, X_{k-1} \leq u, X_k > u] = (1 - p)^{k-1}p.$$

□

Beispiel 7.4.3. Per Definition ist die Wahrscheinlichkeit, dass ein 100-Jahresereignis in einem Jahr eintritt, gleich $p = 1/100$. Die Wiederkehrperiode ist somit 100 Jahre. Wie groß ist die Wahrscheinlichkeit, dass ein 100-Jahresereignis in den nächsten 100 Jahren nicht eintritt? Diese Wahrscheinlichkeit ist

$$\left(1 - \frac{1}{100}\right)^{100} \approx \frac{1}{e} \approx 0,3679.$$

Außerdem können wir direkt angeben, wieviele 100-Jahresereignisse etwa in 100 Jahren stattfinden. Diese Zahl ist nämlich $\text{Bin}(100, \frac{1}{100})$, also approximativ $\text{Poi}(1)$ -verteilt.

Wir werden nun die **Gumbel–Überschreitungsmethode** vorstellen, mit der man die Deichhöhe berechnen kann, für die die Überflutungswahrscheinlichkeit einen vorgegebenen Wert annimmt. Seien X_1, X_2, \dots u.i.v Zufallsvariablen mit Verteilungsfunktion F . Die absteigend angeordneten Ordnungsstatistiken von X_1, \dots, X_n seien mit $M_n^{(1)} \geq M_n^{(2)} \geq \dots \geq M_n^{(n)}$ bezeichnet. Wir definieren

$$S_{k,r,n} = S = \sum_{i=1}^r \mathbb{1}_{X_{n+i} > M_n^{(k)}}.$$

Somit ist $S_{k,r,n}$ die Anzahl der Überschreitungen des Schwellenwerts $M_n^{(k)}$ im Zeitintervall $n + 1, \dots, n + r$.

Satz 7.4.4. Die Verteilung von $S_{k,r,n}$ ist gegeben durch

$$\mathbb{P}[S = j] = \frac{\binom{r+n-k-j}{n-k} \binom{j+k-1}{k-1}}{\binom{r+n}{n}}, \quad j = 0, \dots, r.$$

BEWEIS. Unter Anwendung der totalen Wahrscheinlichkeit gilt:

$$\mathbb{P}[S = j] = \int_{\mathbb{R}} \mathbb{P}[S = j | M_n^{(k)} = u] dF_{n,k}(u).$$

Dabei ist $F_{n,k}$ die Verteilungsfunktion von $M_n^{(k)}$. Es ist zu beachten, dass

$$S | M_n^{(k)} = u \sim \text{Bin}(r, \bar{F}(u))$$

beziehungsweise

$$dF_{n,k}(u) = \frac{n!}{(k-1)!(n-k)!} F^{n-k}(u) \bar{F}^{k-1}(u) dF(u).$$

Es folgt:

$$\mathbb{P}[S = j] = \int_{\mathbb{R}} \binom{r}{j} \bar{F}^j(u) F^{r-j}(u) \frac{n!}{(k-1)!(n-k)!} F^{n-k}(u) \bar{F}^{k-1}(u) dF(u),$$

was sich unter der Substitution $F(u) = t$ wie folgt schreiben lässt:

$$\mathbb{P}[S = j] = \int_0^1 \binom{r}{j} \frac{n!}{(k-1)!(n-k)!} t^{r-j+n-k} (1-t)^{j+k-1} dt,$$

was sich schließlich unter Beachtung der bekannten Formel für die Beta-Funktion

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt = \frac{(\alpha-1)!(\beta-1)!}{(\alpha+\beta-1)!}$$

zu Folgendem umschreiben lässt:

$$\mathbb{P}[S = j] = \binom{r}{j} \frac{n!}{(k-1)!(n-k)!} \frac{(k+j-1)!(r-j+n-k)!}{(r+n)!}.$$

□

Beispiel 7.4.5. Die maximalen jährlichen Wasserstände an einem Fluss seien seit $n = 100$ Jahren bekannt und mit X_1, \dots, X_n bezeichnet. Wie hoch muss ein Deich sein, bei dem die Wahrscheinlichkeit einer Überflutung in den nächsten $r = 12$ Jahren $P = 0,3$ sein soll? Wir setzen die gesuchte Deichhöhe als $M_n^{(k)}$ mit einem unbekanntem k an. Im obigen Satz wurde gezeigt, dass

$$\mathbb{P}[S_{k,r,n} = 0] = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{(r+n) \cdot (r+n-1) \cdot \dots \cdot (r+n-k+1)} \stackrel{!}{=} 0,7.$$

Als Lösung der obigen Gleichung mit $n = 100$, $r = 12$ ergibt sich $k = 3$. Die Deichhöhe kann als die dritte Ordnungsstatistik $M_n^{(3)}$ gewählt werden. Der Vorteil dieser Methode besteht darin, dass die (unbekannte) Verteilungsfunktion von X_i keine Rolle spielt (und nicht geschätzt werden muss). Der Nachteil ist, dass die Methode bei einer kleineren Überflutungswahrscheinlichkeit P nicht funktioniert.

KAPITEL 8

Rekorde

Seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit stetiger Verteilungsfunktion F . Wir setzen $M_n = \max\{X_1, \dots, X_n\}$.

Definition 8.0.6. Wir sagen, dass zum Zeitpunkt $n \in \mathbb{N}$ ein **Rekord** aufgestellt wird, wenn der Wert X_n größer als alle vorherigen Werte X_1, \dots, X_{n-1} ist.

Der zum Zeitpunkt $n = 1$ beobachtete Wert gilt per Definition immer als ein Rekord. Wir definieren deshalb die **Rekord-Indikatoren** ξ_1, ξ_2, \dots durch $\xi_1 = 1$ und

$$\xi_n = \mathbb{1}_{X_n > M_{n-1}}, \quad n = 2, 3, \dots$$

Somit ist ξ_n die Indikatorvariable des Ereignisses, dass zum Zeitpunkt n ein neuer Rekord aufgestellt wird.

8.1. Satz von Rényi

Der folgende Satz beschreibt die gemeinsame Verteilung der Zufallsvariablen ξ_1, ξ_2, \dots

Satz 8.1.1 (Rényi, 1962, und Dwass, 1960). Es gilt $\mathbb{P}[\xi_n = 1] = \frac{1}{n}$ für alle $n \in \mathbb{N}$. Außerdem sind die Zufallsvariablen ξ_1, \dots, ξ_n, M_n unabhängig.

BEWEIS. SCHRITT 1. Wir zeigen, dass $\mathbb{P}[\xi_n = 1] = \frac{1}{n}$. Aus der Stetigkeit der Verteilungsfunktion F folgt, dass $\mathbb{P}[X_i = X_j] = 0$ für $i \neq j$. Somit sind alle Werte X_1, X_2, \dots unterschiedlich mit Wahrscheinlichkeit 1. Es gilt

$$1 = \mathbb{P}[M_n = X_1] + \mathbb{P}[M_n = X_2] + \dots + \mathbb{P}[M_n = X_n] = n\mathbb{P}[M_n = X_n] = n\mathbb{P}[\xi_n = 1],$$

wobei die erste Gleichheit wegen der Disjunktheit der Ereignisse gilt und die zweite Gleichheit aus Symmetriegründen Bestand hat. (Jede der Beobachtungen X_1, \dots, X_n ist mit gleicher Wahrscheinlichkeit $\frac{1}{n}$ das Maximum). Es folgt, dass $\mathbb{P}[\xi_n = 1] = \frac{1}{n}$.

SCHRITT 2. Wir zeigen die Unabhängigkeit von ξ_1, \dots, ξ_n, M_n . Seien dazu $1 \leq \alpha(1) < \alpha(2) < \dots < \alpha(s) \leq n$ und $x \in \mathbb{R}$ beliebig. Es reicht zu zeigen, dass

$$\mathbb{P}[\xi_{\alpha(1)} = 1, \dots, \xi_{\alpha(s)} = 1, M_n < x] = \mathbb{P}[\xi_{\alpha(1)} = 1] \cdot \dots \cdot \mathbb{P}[\xi_{\alpha(s)} = 1] \cdot \mathbb{P}[M_n < x].$$

Es sei zuerst $s = 1$. Schreibe $M_{k,l} = \max\{X_k, \dots, X_l\}$ mit $k \leq l$. Um die Notation zu vereinfachen, schreiben wir α anstelle von $\alpha(1)$. Es gilt

$$\mathbb{P}[\xi_\alpha = 1, M_n < x] = \mathbb{P}[M_{\alpha-1} < X_\alpha < x, M_{\alpha+1,n} < x] = \mathbb{P}[M_{\alpha-1} < X_\alpha < x](F(x))^{n-\alpha},$$

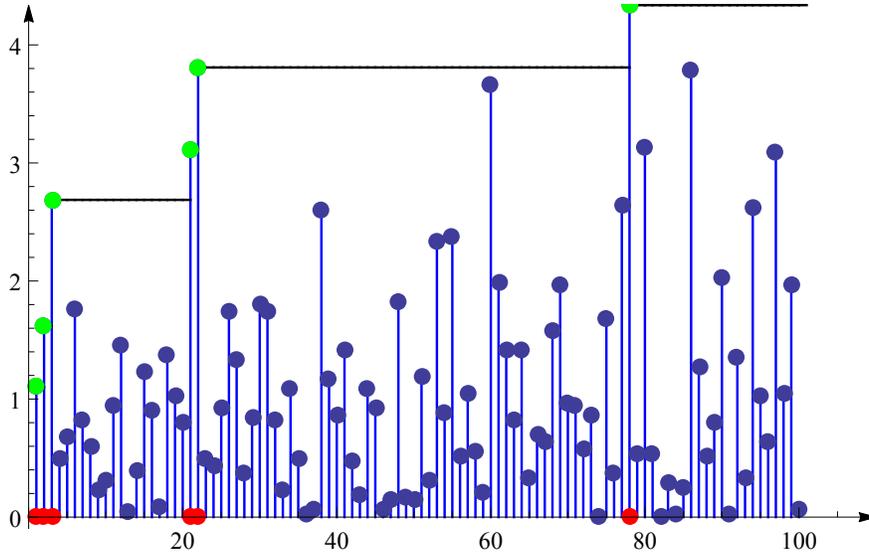


ABBILDUNG 1. Rekorde. Rote Punkte: Rekordzeiten $L(n)$. Grüne Punkte: Rekordwerte $X(n)$.

wobei in der letzten Gleichheit die Unabhängigkeit benutzt wurde. Indem man nun auf $X_\alpha = u \in (-\infty, x)$ bedingt und die Formel der totalen Wahrscheinlichkeit unter Berücksichtigung von $\mathbb{P}[M_{\alpha-1} < u] = (F(u))^{\alpha-1}$ verwendet, kann man das wie folgt umschreiben:

$$\mathbb{P}[\xi_\alpha = 1, M_n < x] = (F(x))^{n-\alpha} \cdot \int_{-\infty}^x (F(u))^{\alpha-1} dF(u).$$

Setzt man nun $w = F(u)$, so erhält man, dass

$$\mathbb{P}[\xi_\alpha = 1, M_n < x] = (F(x))^{n-\alpha} \cdot \int_0^{F(x)} w^{\alpha-1} dw = \frac{(F(x))^n}{\alpha} = \mathbb{P}[M_n < x] \cdot \mathbb{P}[\xi_\alpha = 1],$$

was die Aussage im Fall $s = 1$ beweist.

Sei nun $s \in \mathbb{N}$ beliebig. Es gilt

$$\mathbb{P}[\xi_{\alpha(1)} = 1, \dots, \xi_{\alpha(s)} = 1, M_n < x] = \mathbb{P}[M_{\alpha(1)-1} < X_{\alpha(1)} < x, M_{\alpha(1), \dots, \alpha(2)-1} < X_{\alpha(2)} < x, \dots, M_{\alpha(s-1), \dots, \alpha(s)-1} < X_{\alpha(s)} < x, M_{\alpha(s)+1, n} < x].$$

Indem man nun auf $X_{\alpha(1)} = u_1, \dots, X_{\alpha(s)} = u_s$ bedingt, kann man den obigen Ausdruck ähnlich wie im Fall $s = 1$ schreiben als

$$(F(x))^{n-\alpha(s)} \int_{u_1 < u_2 < \dots < u_s < x} (F(u_1))^{\alpha(1)-1} \dots (F(u_s))^{\alpha(s)-\alpha(s-1)-1} dF(u_1) \dots dF(u_s).$$

Setzt man $F(u_1) = w_1, \dots, F(u_s) = w_s$, so erhält man

$$(F(x))^{n-\alpha(s)} \int_{0 < w_1 < w_2 < \dots < w_s < F(x)} w_1^{\alpha(1)-1} \dots w_s^{\alpha(s)-\alpha(s-1)-1} dw_1 \dots dw_s.$$

Als Übungsaufgabe bleibt zu zeigen, dass sich obiges Integral zu Folgendem errechnen lässt:

$$\frac{F^n(x)}{\alpha(1) \cdot \dots \cdot \alpha(s)} = \mathbb{P}[M_n < x] \cdot \mathbb{P}[\xi_{\alpha(1)=1}] \cdot \dots \cdot \mathbb{P}[\xi_{\alpha(s)=1}].$$

(Man kann z.B. Induktion nach s verwenden). □

Bemerkung 8.1.2. Nach dem Satz von Rényi hängt die gemeinsame Verteilung von ξ_1, ξ_2, \dots nicht von der Verteilungsfunktion F ab.

8.2. Anzahl der Rekorde

Es sei $N(n)$ die **Anzahl der Rekorde** im Intervall $1, \dots, n$, d.h.

$$N(n) = \sum_{k=1}^n \xi_k.$$

Nach dem Satz von Rényi gilt

$$\begin{aligned} \mathbb{E}N(n) &= \sum_{k=1}^n \mathbb{E}\xi_k = \sum_{k=1}^n \frac{1}{k} = \log n + \gamma + o(1), \\ \text{Var } N(n) &= \sum_{k=1}^n \text{Var } \xi_k = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k^2} \right) = \log n + \gamma - \frac{\pi^2}{6} + o(1), \end{aligned}$$

für $n \rightarrow \infty$. Dabei ist γ die **Euler–Mascheroni–Konstante**:

$$\gamma = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \log n \right) = 0,57721\dots$$

Es werden also unter den ersten n Beobachtungen lediglich ungefähr $\log n$ Rekorde (was sehr wenig ist!) erwartet.

Im nächsten Satz werden wir die komplette Verteilung von $N(n)$ mit Hilfe von Stirling–Zahlen erster Art beschreiben.

Definition 8.2.1. Die **Stirling–Zahlen erster Art** sind definiert als Koeffizienten in der Formel

$$x(x+1)\dots(x+n-1) = \sum_{k=1}^n \left[\begin{matrix} n \\ k \end{matrix} \right] x^k.$$

Der nächste Satz beschreibt die Verteilung der Anzahl der Rekorde $N(n)$.

Satz 8.2.2. Für die Verteilung der Anzahl der Rekorde gilt

$$\mathbb{P}[N(n) = k] = \frac{1}{n!} \left[\begin{matrix} n \\ k \end{matrix} \right], \quad k = 1, \dots, n.$$

Bemerkung 8.2.3. Setzt man $x = 1$ in die Definition der Stirling–Zahlen ein, so erhält man

$$\sum_{k=1}^n \left[\begin{matrix} n \\ k \end{matrix} \right] = n!.$$

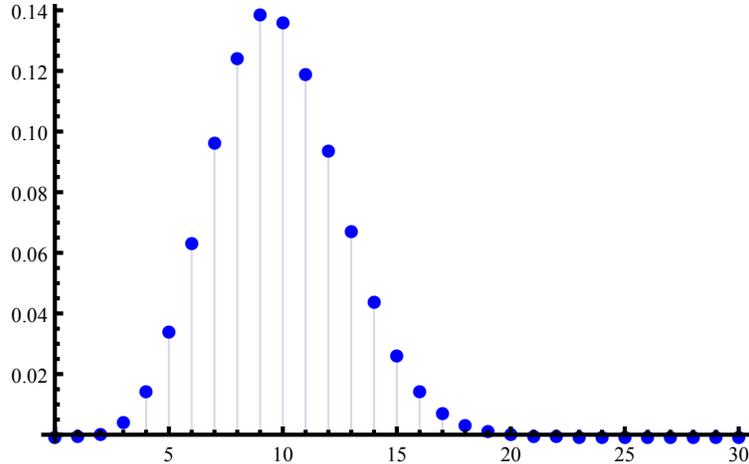


ABBILDUNG 2. Zähldichte der Zufallsvariable $N(n)$ für $n = 10000$.

Also summieren sich die Wahrscheinlichkeiten zu 1.

BEWEIS. Die Zufallsvariablen ξ_1, \dots, ξ_n sind nach dem Satz von Rényi unabhängig und Bernoulli-verteilt mit

$$\mathbb{P}[\xi_k = 1] = \frac{1}{k}, \quad \mathbb{P}[\xi_k = 0] = 1 - \frac{1}{k}.$$

Die erzeugende Funktion einer Zufallsvariable Z mit Werten in $\{0, 1, \dots\}$ ist definiert durch

$$g_Z(t) = \mathbb{E}[t^Z] = \sum_{k=0}^{\infty} \mathbb{P}[Z = k]t^k, \quad |t| < 1.$$

Im Folgenden werden wir die erzeugende Funktion von $N(n)$ angeben. Die erzeugende Funktion von ξ_k ist

$$g_{\xi_k}(t) = \left(1 - \frac{1}{k}\right) + \frac{t}{k}.$$

Es gilt $N(n) = \xi_1 + \dots + \xi_n$ (mit unabhängigen Summanden) und deshalb ist

$$g_{N(n)}(t) = g_{\xi_1}(t) \dots g_{\xi_n}(t) = \prod_{k=1}^n \frac{k-1+t}{k} = \frac{t(t+1) \dots (t+n-1)}{n!} = \frac{1}{n!} \sum_{k=1}^n \begin{bmatrix} n \\ k \end{bmatrix} t^k,$$

wobei im letzten Schritt die Definition der Stirling-Zahlen verwendet wurde. Auf der anderen Seite gilt definitionsgemäß

$$g_{N(n)}(t) = \sum_{k=1}^n t^k \mathbb{P}[N(n) = k].$$

Durch Vergleich der Koeffizienten erhalten wir die gewünschte Formel. □

Aufgabe 8.2.4. Zeigen Sie, dass die Stirling-Zahlen erster Art die folgende Rekursionsformel erfüllen:

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = n \begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix}.$$

Zum Vergleich: Für Binomialkoeffizienten gilt eine ähnliche Formel ohne den Faktor n :

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

Aufgabe 8.2.5. Beweisen Sie für die Stirling-Zahl erster Art die Formel

$$\frac{1}{n!} \left[\begin{matrix} n+1 \\ k+1 \end{matrix} \right] = \sum \frac{1}{i_1 \dots i_k},$$

wobei über alle ganzzahligen $1 \leq i_1 < i_2 < \dots < i_k \leq n$ summiert wird.

8.3. Rekordzeiten

Wir definieren nun die **Rekordzeiten** $L(1) < L(2) < \dots$ durch: $L(1) = 1$, $L(2) = \min\{j > 1 : \xi_j = 1\}$ und allgemein

$$L(n+1) = \min\{j > L(n) : \xi_j = 1\}, \quad n = 2, 3, \dots$$

Somit ist $L(n)$ der Zeitpunkt, zu dem der n -te Rekord aufgestellt wird. Im nächsten Satz beschreiben wir die gemeinsame Verteilung des Vektors $(L(1), \dots, L(n))$.

Satz 8.3.1. Für beliebige natürliche Zahlen $1 = j(1) < j(2) < \dots < j(n)$ gilt

$$\mathbb{P}[L(1) = j(1), L(2) = j(2), \dots, L(n) = j(n)] = \frac{1}{j(n)(j(2) - 1) \dots (j(n) - 1)}.$$

BEWEIS. Es gilt:

$$\begin{aligned} \mathbb{P}[L(1) = j(1), \dots, L(n) = j(n)] &= \mathbb{P}[\xi_2 = \xi_3 = \dots = \xi_{j(2)-1} = 0, \\ &\quad \xi_{j(2)} = 1, \xi_{j(2)+1} = \dots = \xi_{j(3)-1} = 0, \xi_{j(3)} = 1, \dots, \xi_{j(n)} = 1]. \end{aligned}$$

Wegen der sich aus Satz 8.1.1 ergebenden Unabhängigkeit kann man dies in folgenden Ausdruck umschreiben:

$$\mathbb{P}[\xi_2 = 0] \cdot \dots \cdot \mathbb{P}[\xi_{j(n)} = 0] \cdot \frac{\mathbb{P}[\xi_{j(2)} = 1]}{\mathbb{P}[\xi_{j(2)} = 0]} \cdot \dots \cdot \frac{\mathbb{P}[\xi_{j(n)} = 1]}{\mathbb{P}[\xi_{j(n)} = 0]},$$

was sich ebenfalls, wegen des Satzes von Rényi, wie folgt darstellen lässt:

$$\left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \dots \left(1 - \frac{1}{j(n)}\right) \frac{1/j(2)}{1 - 1/j(2)} \dots \frac{1/j(n)}{1 - 1/j(n)}.$$

Durch geschicktes Umformen lässt sich das wie folgt darstellen:

$$\frac{2-1}{2} \frac{3-1}{3} \dots \frac{j(n)-1}{j(n)} \frac{1}{j(2)-1} \dots \frac{1}{j(n)-1} = \frac{1}{j(n)} \frac{1}{j(2)-1} \dots \frac{1}{j(n)-1},$$

wobei sich die Gleichheit ergibt, da die ersten $j(n) - 1$ Faktoren ein Teleskopprodukt bilden. \square

Bemerkung 8.3.2. Die gemeinsame Verteilung der Rekordzeiten $L(1), L(2), \dots$ ist (abgesehen von der Stetigkeitsannahme an die Verteilungsfunktion F) unabhängig von der Verteilung der Zufallsvariablen X_1, X_2, \dots

Bemerkung 8.3.3. Die Verteilung von $L(2)$ sieht somit folgendermaßen aus:

$$\mathbb{P}[L(2) = j] = \frac{1}{j(j-1)}, \quad j = 2, 3, \dots$$

Insbesondere gilt $\mathbb{E}L(2) = \sum_{j=1}^{\infty} \frac{1}{j-1} = +\infty$. Die *mittlere* Wartezeit auf den zweiten Rekord ist somit unendlich. (Was erstaunlich ist!) Auf der anderen Seite ist die Wartezeit auf den zweiten Rekord *fast sicher* endlich.

Satz 8.3.4. Für die Verteilung der Rekordzeit $L(n)$ gilt

$$P[L(n) = k] = \frac{1}{k!} \binom{n-1}{k-1}, \quad k = n, n+1, \dots$$

BEWEIS. Übung: Benutzen Sie Satz 8.2.2. □

Im nächsten Satz werden wir zeigen, dass die Rekordzeiten $L(1), L(2), \dots$ eine **Markov-Kette** bilden.

Satz 8.3.5. Die Folge $L(1), L(2), \dots$ ist eine Markov-Kette mit Anfangszustand $L(1) = 1$ und Übergangswahrscheinlichkeiten

$$p_{ij} = \frac{i}{j(j-1)}$$

für $i = 1, 2, \dots$ und $j = i+1, i+2, \dots$

BEWEIS. Zu zeigen ist, dass für alle $1 = j(1) < j(2) < \dots < j(n)$ gilt

$$\mathbb{P}[L(1) = j(1), L(2) = j(2), \dots, L(n) = j(n)] = p_{j(1)j(2)} p_{j(2)j(3)} \cdots p_{j(n-1)j(n)}.$$

Mit der Formel aus Satz 8.3.1 gilt

$$\begin{aligned} & \mathbb{P}[L(1) = j(1), L(2) = j(2), \dots, L(n) = j(n)] \\ &= \frac{1}{j(n)(j(2)-1) \cdots (j(n)-1)} \\ &= \frac{j(1)}{j(2)(j(2)-1)} \cdot \frac{j(2)}{j(3)(j(3)-1)} \cdots \frac{j(n-1)}{j(n)(j(n)-1)}, \end{aligned}$$

was die gewünschte Formel ergibt. □

Angenommen, die ersten n Rekordzeiten sind bekannt: $L(1) = 1, L(2) = i(2), \dots, L(n) = i(n)$. Wo liegt nun die nächste Rekordzeit $L(n+1)$? Wegen der Markov-Eigenschaft der Folge $L(1), L(2), \dots$ stellt es sich heraus, dass man für die Bestimmung von $L(n+1)$ lediglich den Wert $L(n) = i(n)$ benötigt, die Werte von $L(1), \dots, L(n-1)$ sind hingegen irrelevant.

Satz 8.3.6. Für alle $1 = i(1) < i(2) < \dots < i(n) = i < j$ gilt die Markov-Eigenschaft der Rekordzeiten:

$$\mathbb{P}[L(n+1) = j | L(n) = i] = \mathbb{P}[L(n+1) = j | L(n) = i, L(n-1) = i(n-1), \dots, L(2) = i(2)].$$

Außerdem gilt:

$$\mathbb{P}[L(n+1) = j | L(n) = i] = \frac{i}{j(j-1)}.$$

BEWEIS. Folgt aus Satz 8.3.5. □

Der obige Satz zeigt, wie man die Folge der Rekordzeiten am Rechner simulieren kann, ohne dafür die Variablen X_1, X_2, \dots erzeugen zu müssen. Man startet mit $L(1) = 1$ und geht induktiv vor. Sind die Werte $L(1), \dots, L(n)$ mit $L(n) = i$ bekannt, so erzeugt man eine Zufallsvariable auf $\{i+1, i+2, \dots\}$, indem man den Wert j mit Wahrscheinlichkeit $\frac{j}{i(i-1)}$ auswählt. Dieser Wert ist dann der Wert von $L(n+1)$. Danach wiederholt man das Ganze.

Der nächste Satz gibt einen viel einfacheren Algorithmus zur Simulation der Rekordzeiten. Definiere die Gauß-Klammer durch

$$\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}, \quad x \in \mathbb{R}.$$

Satz 8.3.7 (Williams, 1973). Seien U_1, U_2, \dots unabhängig und gleichverteilt auf dem Intervall $[0, 1]$. Definiere $R(1) = 1$ und $R(n+1) = \lfloor \frac{R(n)}{U_n} \rfloor + 1$ für $n \in \mathbb{N}$. Dann gilt die Gleichheit der Verteilungen:

$$(L(1), L(2), \dots, L(n)) \stackrel{d}{=} (R(1), R(2), \dots, R(n)).$$

BEWEIS. Wegen der Markov-Eigenschaft reicht es zu zeigen, dass für alle $i \in \mathbb{N}$ und $j \in \{i+1, i+2, \dots\}$ gilt

$$\mathbb{P}[L(n+1) = j | L(n) = i] = \mathbb{P}[R(n+1) = j | R(n) = i].$$

Die Wahrscheinlichkeit auf der linken Seite ist gleich $\frac{i}{j(j-1)}$ nach Satz 8.3.6. Wir berechnen die Wahrscheinlichkeit auf der rechten Seite. Es ist

$$\mathbb{P}[R(n+1) = j | R(n) = i] = \mathbb{P}\left[\left\lfloor \frac{R(n)}{U_n} \right\rfloor + 1 = j | R(n) = i\right] = \mathbb{P}\left[\left\lfloor \frac{i}{U_n} \right\rfloor + 1 = j | R(n) = i\right].$$

Die Zufallsvariable $R(n)$ hängt nur von U_1, \dots, U_{n-1} ab. Die Ereignisse $\{\lfloor \frac{i}{U_n} \rfloor + 1 = j\}$ und $\{R(n) = i\}$ sind also unabhängig und somit vereinfacht sich das Ganze zu folgendem Ausdruck:

$$\mathbb{P}[R(n+1) = j | R(n) = i] = \mathbb{P}\left[\left\lfloor \frac{i}{U_n} \right\rfloor = j-1\right] = \mathbb{P}\left[\frac{i}{U_n} \in [j-1, j)\right] = \frac{i}{j(j-1)},$$

da U_n gleichverteilt auf $[0, 1]$ ist. □

Satz 8.3.8 (Tata, 1969). Sei $x > 1$ eine natürliche Zahl, dann gilt:

$$\mathbb{P} \left[\frac{L(n+1)}{L(n)} > x \right] = \frac{1}{x}.$$

Sei $x > 1$ eine beliebige reelle Zahl, dann gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{L(n+1)}{L(n)} > x \right] = \frac{1}{x}.$$

Bemerkung 8.3.9. Man sieht hier noch einmal, dass Rekorde mit $n \rightarrow \infty$ immer seltener auftreten. Zum Beispiel ist mit Wahrscheinlichkeit $1/2$ der Abstand zwischen der $(n+1)$ -ten und der n -ten Rekordzeit größer als die n -te Rekordzeit selbst. Mit Wahrscheinlichkeit $1/3$ ist der Abstand zwischen der $(n+1)$ -ten und der n -ten Rekordzeit mindestens doppelt so groß wie die n -te Rekordzeit selbst, usw. Die Tatsache, dass Rekorde immer seltener auftreten ist ziemlich natürlich: die Rekordwerte steigen nämlich mit der Zeit und es wird immer schwieriger neue Rekorde aufzustellen.

BEWEIS. Sei $x > 1$. Dann gilt wegen des Gesetzes der totalen Wahrscheinlichkeit:

$$\begin{aligned} \mathbb{P}[L(n+1) > xL(n)] &= \sum_{i=n}^{\infty} \mathbb{P}[L(n+1) > xi | L(n) = i] \cdot \mathbb{P}[L(n) = i] \\ &= \sum_{i=n}^{\infty} \mathbb{P}[L(n+1) > [xi] | L(n) = i] \cdot \mathbb{P}[L(n) = i], \end{aligned}$$

denn $L(n+1)$ ist ganzzahlig und somit ist $L(n+1) > xi$ zu $L(n+1) > [xi]$ äquivalent. Im Beweis von Satz 8.3.6 wurde die Wahrscheinlichkeit $\mathbb{P}[L(n+1) > xi | L(n) = i]$ bereits berechnet. Wir wollen das Ergebnis hier verwenden. So lässt sich obiger Ausdruck zu folgendem vereinfachen:

$$(8.3.1) \quad \mathbb{P}[L(n+1) > xL(n)] = \sum_{i=n}^{\infty} \frac{i}{[xi]} \mathbb{P}[L(n) = i].$$

Sei zuerst $x \in \mathbb{N}$. Man sieht:

$$\sum_{i=n}^{\infty} \frac{i}{[xi]} \mathbb{P}[L(n) = i] = \frac{1}{x} \sum_{i=n}^{\infty} \mathbb{P}[L(n) = i] = \frac{1}{x}.$$

Sei nun $x > 1$ beliebig reell. Aus der Definition der Gauß-Klammer folgt, dass $[xi] \leq xi < [xi] + 1$. Durch leichte Umformungen folgt:

$$\frac{1}{x} \leq \frac{i}{[xi]} < \frac{1}{x} + \frac{1}{[xi]x}.$$

Deshalb kann man (8.3.1) wie folgt nach unten abschätzen:

$$\sum_{i=n}^{\infty} \frac{i}{[xi]} \mathbb{P}[L(n) = i] \geq \frac{1}{x} \sum_{i=n}^{\infty} \mathbb{P}[L(n) = i] = \frac{1}{x}.$$

Außerdem kann man daher (8.3.1) wie folgt nach oben abschätzen:

$$\sum_{i=n}^{\infty} \frac{i}{[xi]} \mathbb{P}[L(n) = i] < \sum_{i=n}^{\infty} \left(\frac{1}{x} + \frac{1}{[xi]x} \right) \mathbb{P}[L(n) = i] \leq \frac{1}{x} + \frac{1}{x} \frac{1}{[xn]},$$

da $[xi] \geq [xn]$ für $i \geq n$. Insgesamt folgt mit dem ‘Sandwich-Prinzip’ die Behauptung. \square

Korollar 8.3.10. Sei U gleichverteilt auf $[0, 1]$, dann gilt

$$\frac{L(n+1)}{L(n)} \xrightarrow[n \rightarrow \infty]{d} \frac{1}{U}.$$

BEWEIS. Die Zufallsvariable $\frac{1}{U}$ ist Pareto-verteilt mit Tailfunktion $\frac{1}{x}$, $x > 1$. Die Behauptung folgt nun aus Satz 8.3.8. \square

Bemerkung 8.3.11. Shorrock, 1972, hat gezeigt, dass für beliebiges $k \in \mathbb{N}$ und für unabhängige, auf $[0, 1]$ gleichverteilte Zufallsvariablen U_1, \dots, U_k , gilt

$$\left(\frac{L(n+1)}{L(n)}, \frac{L(n+2)}{L(n+1)}, \dots, \frac{L(n+k)}{L(n+k-1)} \right) \xrightarrow[n \rightarrow \infty]{d} \left(\frac{1}{U_1}, \frac{1}{U_2}, \dots, \frac{1}{U_k} \right).$$

8.4. Zentrale Grenzwertsätze

Nun zeigen wir, dass für großes n die Anzahl der Rekorde $N(n)$ approximativ normalverteilt ist, siehe Abbildung 2. Mit $\mathcal{N}(0, 1)$ bezeichnen wir die Standardnormalverteilung mit Verteilungsfunktion

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}z^2} dz.$$

Satz 8.4.1 (Zentraler Grenzwertsatz für die Anzahl der Rekorde). Es gilt

$$\frac{N(n) - \log n}{\sqrt{\log n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1) \text{ für } n \rightarrow \infty.$$

Der Beweis basiert auf einer Verallgemeinerung des zentralen Grenzwertsatzes, die wir ohne Beweis angeben.

Satz 8.4.2 (Zentraler Grenzwertsatz von Ljapunow). Für jedes $n \in \mathbb{N}$ seien Z_{n1}, \dots, Z_{nn} unabhängige Zufallsvariablen mit $\mathbb{E}Z_{nk} = 0$, $\sigma_{nk}^2 := \text{Var } Z_{nk} \in (0, \infty)$ für $k = 1, \dots, n$

und $\sum_{k=1}^n \sigma_{nk}^2 = 1$. Außerdem gelte die folgende *Ljapunow-Bedingung*: Für ein $\delta > 0$

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E}|Z_{nk}|^{2+\delta} \rightarrow 0.$$

Dann gilt:

$$Z_{n1} + \dots + Z_{nn} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Bemerkung 8.4.3. Die identische Verteiltheit der Zufallsvariablen wird im zentralen Grenzwertsatz von Ljapunow nicht vorausgesetzt.

BEWEIS VON SATZ 8.4.1. Es ist $N(n) = \xi_1 + \dots + \xi_n$, wobei die ξ_1, \dots, ξ_n nach dem Satz von Rényi unabhängig aber nicht identisch verteilt sind. Setze

$$Z_{nk} = \frac{\xi_k - \frac{1}{k}}{\sigma_n}, \quad k = 1, \dots, n,$$

mit

$$\sigma_n^2 = \text{Var } N(n) = \sum_{k=1}^n \text{Var } \xi_k = \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{k^2} \sim \log n, \quad n \rightarrow \infty,$$

denn $\sum_{k=1}^n \frac{1}{k} \sim \log n$ und $\sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$. Definitionsgemäß gilt

$$\mathbb{E}Z_{nk} = 0, \quad \sum_{k=1}^n \mathbb{E}Z_{nk}^2 = 1.$$

Wir zeigen, dass die Ljapunow-Bedingung mit $\delta = 1$ gilt. Die Zufallsvariable $\frac{\xi_k - 1/k}{\sigma_n}$ nimmt nur zwei Werte an, und zwar $\frac{1/k}{\sigma_n}$ mit Wahrscheinlichkeit $1 - \frac{1}{k}$ und $\frac{1-1/k}{\sigma_n}$ mit Wahrscheinlichkeit $\frac{1}{k}$. Es folgt

$$\sum_{k=1}^n \mathbb{E} \left| \frac{\xi_k - \frac{1}{k}}{\sigma_n} \right|^3 \leq \sum_{k=1}^n \left(\frac{1}{k\sigma_n} \right)^3 \cdot 1 + \left(\frac{1}{\sigma_n} \right)^3 \cdot \frac{1}{k} = \frac{1}{\sigma_n^3} \sum_{k=1}^n \left(\frac{1}{k^3} + \frac{1}{k} \right) \rightarrow 0 \text{ für } n \rightarrow \infty,$$

da $\sigma_n^3 \sim (\log n)^{3/2}$ und $\sum_{k=1}^n \left(\frac{1}{k^3} + \frac{1}{k} \right) \sim \log n$ für $n \rightarrow \infty$. Es sind also alle Voraussetzungen für den zentralen Grenzwertsatz von Ljapunow gegeben und damit folgt:

$$\frac{N(n) - \sum_{k=1}^n \frac{1}{k}}{\sigma_n} = \sum_{k=1}^n Z_{nk} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1) \text{ für } n \rightarrow \infty.$$

Unter Beachtung des bereits gezeigten Zusammenhangs $\sigma_n^2 \sim \log n$ führt das zur behaupteten Grenzaussage:

$$\frac{N(n) - \log n}{\sqrt{\log n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1) \text{ für } n \rightarrow \infty.$$

(Man kann z.B. das Lemma von Chintschin verwenden). □

Aufgabe 8.4.4. Zeigen Sie, dass für die Anzahl der Rekorde auch das Gesetz der großen Zahlen gilt:

$$\lim_{n \rightarrow \infty} \frac{N(n)}{\log n} = 1 \quad \text{f.s.}$$

Nun benutzen wir den obigen Satz um auch einen zentralen Grenzwertsatz für die Rekordzeiten herzuleiten.

Satz 8.4.5 (Zentraler Grenzwertsatz für Rekordzeiten). Es gilt

$$\frac{\log L(n) - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

BEWEIS. Sei $x \in \mathbb{R}$ fest. Mit $n(x) = e^{n+x\sqrt{n}}$ gilt

$$\mathbb{P} \left[\frac{\log L(n) - n}{\sqrt{n}} \leq x \right] = \mathbb{P}[L(n) \leq n(x)] = \mathbb{P}[N(n(x)) \geq n].$$

Zur Vereinfachung sei $n(x) \in \mathbb{Z}$, dabei verliert der Beweis nicht an Allgemeingültigkeit, da für $n(x) \notin \mathbb{Z}$ einfach $[n(x)]$ betrachtet werden kann. Nach Satz 8.4.1 gilt

$$\frac{N(n(x)) - \log n(x)}{\sqrt{\log n(x)}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Es folgt, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}[N(n(x)) \geq n] = \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{N(n(x)) - \log n(x)}{\sqrt{\log n(x)}} \geq \frac{n - (n + x\sqrt{n})}{\sqrt{n + x\sqrt{n}}} \right] = 1 - \Phi(-x) = \Phi(x),$$

da $\lim_{n \rightarrow \infty} \frac{n - (n + x\sqrt{n})}{\sqrt{n + x\sqrt{n}}} = -x$. Dabei bezeichnet $\Phi(x)$ die Verteilungsfunktion von $\mathcal{N}(0, 1)$ und die Behauptung folgt. \square

Aufgabe 8.4.6. Zeigen Sie, dass für die Rekordzeiten auch das Gesetz der großen Zahlen gilt:

$$\lim_{n \rightarrow \infty} \frac{\log L(n)}{n} = 1 \quad \text{f.s.}$$

Aufgabe 8.4.7. Ist es richtig, dass

$$\lim_{n \rightarrow \infty} \frac{L(n)}{e^n} = 1 \quad \text{f.s.}?$$

8.5. Rekordwerte

Definition 8.5.1. Die **Rekordwerte** sind definiert als

$$X(n) = M_{L(n)} = X_{L(n)}.$$

Anders als bei Rekordindikatoren oder Rekordzeiten, hängt die Verteilung der Rekordwerte von der Verteilungsfunktion der Zufallsvariablen X_i ab. Im Spezialfall der exponentialverteilten Zufallsvariablen besitzen die Rekordwerte eine schöne Darstellung.

Satz 8.5.2 (Tata, 1969). Seien X_1, X_2, \dots unabhängig und exponentialverteilt mit Parameter 1. Dann gilt:

$$(X(1), X(2), \dots, X(n)) \stackrel{d}{=} (\nu_1, \nu_1 + \nu_2, \dots, \nu_1 + \dots + \nu_n),$$

wobei ν_1, ν_2, \dots unabhängig und exponentialverteilt mit Parameter 1 sind.

Bemerkung 8.5.3. Mit anderen Worten, die Folge $X(1), X(2), \dots$ ist ein Poisson-Prozess mit Intensität 1.

BEWEISIDEE. Der Beweis basiert auf der Gedächtnislosigkeit der Exponentialverteilung. Ist nämlich $X \sim \text{Exp}(1)$, so ist für jedes $t > 0$ die bedingte Verteilung von $X - t$ gegeben, dass $X > t$, ebenfalls eine Exponentialverteilung mit Parameter 1.

Nach Voraussetzung ist $X(1) = X_1 \sim \text{Exp}(1)$. Wir halten nun $X_1 = a_1$ fest und warten auf den zweiten Rekordwert $X(2)$. Der Exzess $X(2) - a_1$ hat die gleiche Verteilung wie $X - a_1$ gegeben, dass $X > a_1$, also die Exponentialverteilung mit Parameter 1. Nun halten wir $X(2) = a_2$ fest und warten auf den dritten Rekordwert $X(3)$. Der Exzess $X(3) - a_2$ hat die gleiche Verteilung, wie $X - a_2$ gegeben, dass $X > a_2$, also wieder die Exponentialverteilung mit Parameter 1, usw.

Somit sind die Zuwächse $X(1), X(2) - X(1), X(3) - X(2), \dots$ standard exponentialverteilt und unabhängig. \square

Korollar 8.5.4. Seien X_1, X_2, \dots unabhängig und exponentialverteilt mit Parameter 1. Dann gilt

$$\frac{X(n) - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1) \text{ für } n \rightarrow \infty.$$

BEWEIS. Aus Satz 8.5.2 wissen wir, dass $X(n) \stackrel{d}{=} \nu_1 + \dots + \nu_n$. Dabei ist $\mathbb{E}\nu_i = \text{Var } \nu_i = 1$. Der zentrale Grenzwertsatz ergibt, dass

$$\frac{\nu_1 + \dots + \nu_n - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1) \text{ für } n \rightarrow \infty.$$

Daraus folgt die Behauptung. \square

Bemerkung 8.5.5. In Satz 1.3.2 haben wir nachgewiesen, dass im Fall der standard exponentialverteilten Zufallsvariablen, die Zufallsvariable $M_n - \log n$ gegen die Gumbel-Verteilung Λ konvergiert. Es gilt also

$$M_n - \log n \xrightarrow[n \rightarrow \infty]{d} \Lambda \quad \text{und} \quad \frac{M_{L(n)} - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Die Grenzverteilung wird also offenbar durch den Umstand, dass $L(n)$ zufällig ist, völlig verändert.

Aufgabe 8.5.6. Seien X_1, X_2, \dots unabhängige identisch verteilte Zufallsvariablen mit stetiger Verteilungsfunktion F . Es sei $X(n)$ der n -te Rekordwert.

(1) Zeigen Sie, dass

$$\mathbb{P}[X(n) < x] = Q_n(F(x)), \quad x \in \mathbb{R},$$

wobei $Q_n(s) = \mathbb{E}s^{L(n)}$ die erzeugende Funktion von $L(n)$ sei.

(2) Zeigen Sie, dass

$$Q_n(s) = \frac{1}{(n-1)!} \int_0^{-\log(1-s)} t^{n-1} e^{-t} dt.$$

Hinweis: Zu Teil (2): Teil (1) gilt auch für exponentialverteilte Zufallsvariablen.

Poisson–Punktprozesse

Poisson–Punktprozesse sind natürliche Modelle für zufällige Konfigurationen von Punkten im Raum. Als Beispiele von zufälligen Konfigurationen, die mit Poisson–Prozessen modelliert werden können, kann man sich die Positionen der Sterne im Himmel, die Regentropfen auf dem Boden, die Meteoriteinschläge auf dem Mond, oder die Zeitpunkte, zu denen bei einer Telefonzentrale die Anrufe ankommen, vorstellen. Wie schon der Name sagt, spielt die Poisson–Verteilung eine entscheidende Rolle. Wir werden also mit der Definition der Poisson–Verteilung beginnen.

9.1. Poisson–Verteilung

Man betrachte n Bernoulli–Experimente mit Erfolgswahrscheinlichkeit p . Bezeichnet man mit $S_{n,p}$ die Anzahl der Erfolge in diesen Experimenten, so ist $S_{n,p}$ eine binomialverteilte Zufallsvariable, d.h.

$$\mathbb{P}[S_{n,p} = k] = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \dots, n.$$

Der Poisson–Grenzwertsatz behandelt die Situation, in der die Anzahl der Experimente sehr groß, die Erfolgswahrscheinlichkeit jedoch sehr gering ist.

Satz 9.1.1 (Poisson–Grenzwertsatz). Es sei $p_1, p_2, \dots \in [0, 1]$ eine Folge mit $\lim_{n \rightarrow \infty} np_n = \lambda$, wobei $\lambda \geq 0$. Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_{n,p_n} = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots$$

Definition 9.1.2. Eine Zufallsvariable S heißt **Poisson–verteilt** mit Parameter $\lambda > 0$, falls

$$\mathbb{P}[S = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots$$

Wir benutzen die Schreibweise $S \sim \text{Poi}(\lambda)$. Die obige Definition kann man etwas erweitern. Wir sagen, dass $S \sim \text{Poi}(0)$, falls $S = 0$ fast sicher. Außerdem sagen wir, dass $S \sim \text{Poi}(+\infty)$, falls $S = +\infty$ fast sicher.

9.2. Beispiel zu Poisson-Prozessen

Wir betrachten nun wieder eine sehr große Zahl von unabhängigen Experimenten mit sehr kleinen Erfolgswahrscheinlichkeiten. Diesmal stellen wir uns aber vor, dass jedes Experiment eine Position im Raum besitzt. Die Positionen der Experimente, die mit einem Erfolg ausgehen, bilden eine zufällige Konfiguration von Punkten im Raum. Diese Konfiguration beschreibt man mit einem Poisson-Punktprozess.

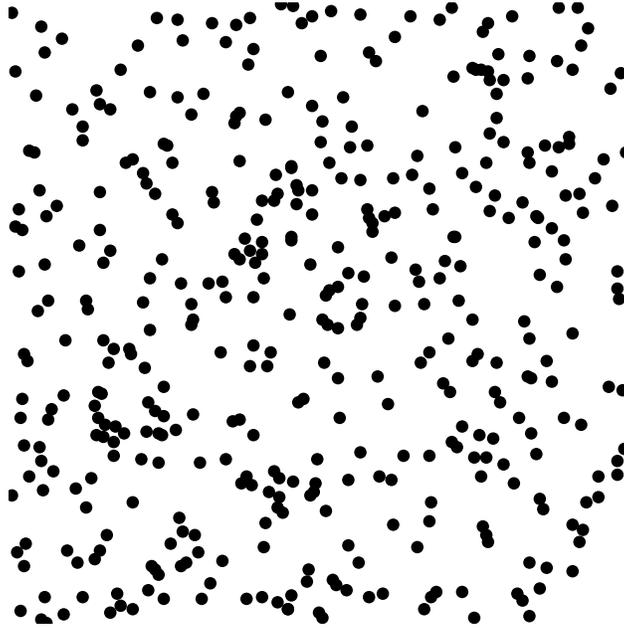


ABBILDUNG 1. Homogener Poisson-Punktprozess.

Zum Beispiel können wir versuchen, ein stochastisches Modell für die Verteilung der Sterne im Himmel zu finden. Ist A ein Gebiet (ein Teil des Himmels), so bezeichnen wir mit $\pi(A)$ die Anzahl der Sterne in A . Diese Anzahl fassen wir als eine Zufallsvariable mit Werten in \mathbb{N}_0 auf. Wir gehen davon aus, dass die folgenden natürlichen Eigenschaften gelten:

- (1) Für disjunkte Gebiete A_1, \dots, A_n sind die Zufallsvariablen $\pi(A_1), \dots, \pi(A_n)$ unabhängig.
- (2) Die Wahrscheinlichkeit, dass es in einem kleinen Gebiet Q mit Fläche $\epsilon \approx 0$ einen Stern gibt, ist $\approx \lambda\epsilon$.

Der Koeffizient $\lambda > 0$ beschreibt dabei die Intensität der Sterne im Himmel.

Wie ist nun die Anzahl der Sterne $\pi(A)$ in einem beliebigen Gebiet A verteilt? Um dies herauszufinden, unterteilen wir A in kleine Gebiete mit Fläche ϵ . Die Anzahl dieser Gebiete ist $\approx \text{Fläche}(A)/\epsilon$. Aus unseren Voraussetzungen und dem Poisson-Grenzwertsatz folgt, dass

$$\pi(A) \approx \text{Bin} \left(\frac{\text{Fläche}(A)}{\epsilon}, \lambda\epsilon \right) \approx \text{Poi}(\lambda \cdot \text{Fläche}(A)) \text{ für } \epsilon \downarrow 0.$$

Es gilt also:

Für jedes Gebiet A ist $\pi(A)$ Poisson-verteilt mit Parameter $\lambda \cdot \text{Fläche}(A)$.

Eine zufällige Konfiguration von Punkten im Raum, die die beiden oben genannten Eigenschaften hat, bezeichnen wir als einen **homogenen Poisson–Punktprozess** mit Intensität λ .

9.3. Definition der Zählmaße und Punktprozesse

Wir werden nun eine mathematische Definition der Poisson–Punktprozesse geben. Die erste Frage ist, wie man eine “Punktekongfiguration” definiert. Zuerst müssen wir einige Forderungen an den Raum formulieren, wo die Punkte leben.

Lokal kompakte separable metrische Räume. Sei (E, ρ) ein metrischer Raum mit Metrik ρ . Als Beispiel kann man sich $E = \mathbb{R}^d$ mit der üblichen Euklid’schen Metrik vorstellen.

Definition 9.3.1. Eine Teilmenge $A \subset E$ heißt **kompakt**, wenn jede Folge $x_1, x_2, \dots \in A$ eine Teilfolge besitzt, die gegen ein Element aus A konvergiert.

Zum Beispiel ist jede abgeschlossene und beschränkte Teilmenge von \mathbb{R}^d (versehen mit der Euklid’schen metrik) kompakt. Es ist bekannt, dass kompakte Mengen immer abgeschlossen sind.

Definition 9.3.2. Der Raum (E, ρ) heißt **lokal kompakt**, wenn man für jeden Punkt $x \in E$ ein $r > 0$ finden kann, so dass der abgeschlossene Ball

$$\bar{B}_r(x) = \{y \in E : \rho(x, y) \leq r\}$$

kompakt ist.

Zum Beispiel ist jede abgeschlossene und jede offene Teilmenge von \mathbb{R}^d lokal kompakt. Nicht lokal kompakt ist z.B. der Raum $C[0, 1]$ der stetigen Funktionen auf $[0, 1]$ versehen mit der Supremumsmetrik.

Definition 9.3.3. Der Raum (E, ρ) heißt **separabel**, wenn es eine abzählbare Teilmenge $\{x_1, x_2, \dots\} \subset E$ gibt, die in E überall dicht liegt. Das heißt, für alle $x \in E$ und $r > 0$ enthält der Ball $\bar{B}_r(x)$ mindestens einen Punkt der Form x_i .

Beispiel 9.3.4. Nicht jeder lokal kompakter Raum ist separabel: Betrachte eine beliebige überabzählbare Menge E (z.B. $E = \mathbb{R}$) versehen mit der diskreten Metrik

$$\rho(x, y) = \begin{cases} 1, & \text{falls } x \neq y, \\ 0, & \text{falls } x = y. \end{cases}$$

Dieser Raum ist lokal kompakt aber nicht separabel.

Zählmaße. Sei im Folgenden (E, ρ) ein metrischer Raum. Als Beispiel kann man sich immer eine abgeschlossene oder eine offene Menge von \mathbb{R}^d mit der Euklid'schen Metrik vorstellen.

Definition 9.3.5. Die **Borel**- σ -Algebra $\mathcal{B} = \mathcal{B}(E)$ ist die durch die Familie aller offenen Teilmengen von E erzeugte σ -Algebra.

Man kann \mathcal{B} auch als die durch die Familie aller abgeschlossenen Teilmengen erzeugte σ -Algebra definieren, denn eine Menge A ist offen genau dann, wenn A^c abgeschlossen ist.

Definition 9.3.6. Ein Maß μ auf (E, \mathcal{B}) heißt **lokal endlich** (oder **Radon-Maß**), wenn für jede kompakte Menge $B \subset E$ deren Maß $\mu(B)$ endlich ist.

Zum Beispiel ist das Lebesgue-Maß auf \mathbb{R}^d ein Radon-Maß.

Beispiel 9.3.7. Sei $x \in E$. Das **Dirac-Maß** δ_x ist definiert durch

$$\delta_x(A) = \begin{cases} 1, & x \in A, \\ 0, & \text{sonst.} \end{cases}$$

Offensichtlich ist δ_x ein Radon-Maß. Man kann sich δ_x als eine Punktconfiguration vorstellen, die aus einem Punkt x besteht.

Beispiel 9.3.8. Das Maß $\mu := \sum_{x \in \mathbb{Q}} \delta_x$ auf \mathbb{R} (wobei \mathbb{Q} die Menge der rationalen Zahlen ist) ist kei Radon-Maß, denn $\mu([0, 1]) = +\infty$.

Definition 9.3.9. Ein Radon-Maß μ auf (E, \mathcal{B}) heißt **Zählmaß**, falls für jede kompakte Menge $A \subset E$, $\mu(A)$ eine nicht-negative ganze Zahl ist.

Beispiel 9.3.10. Jedes Dirac-Maß δ_x ist ein Zählmaß. Jede endliche Summe von Dirac-Maßen der Form $\mu = \sum_{i=1}^n \delta_{x_i}$ mit $x_1, \dots, x_n \in E$ ist ein Zählmaß. Auch eine abzählbar unendliche Summe $\mu = \sum_{i=1}^{\infty} \delta_{x_i}$ ist ein Zählmaß, vorausgesetzt, dass die Folge $x_1, x_2, \dots \in E$ keine Häufungspunkte besitzt. Der folgende Satz besagt, dass auf einem lokal kompakten separablen Raum jedes Zählmaß diese Form besitzt.

Satz 9.3.11. Sei E ein lokal kompakter separabler metrischer Raum. Dann lässt sich jedes Zählmaß μ auf E als

$$\mu = \sum_{i=1}^n \delta_{x_i}$$

darstellen, wobei $n \in \mathbb{N} \cup \{\infty\}$ und x_1, x_2, \dots eine endliche oder abzählbar unendliche Folge von Punkten in E ist, die keine Häufungspunkte besitzt.

BEWEIS. Weggelassen. □

Auf lokal kompakten separablen Räumen können wir also Zählmaße mit höchstens abzählbaren Konfigurationen von Punkten identifizieren, die keine Häufungspunkte besitzen.

Beispiel 9.3.12. Das Maß $\mu = \sum_{i=1}^{\infty} \delta_{1/i}$ ist ein Zählmaß auf $E = (0, \infty)$, denn die Folge $1, \frac{1}{2}, \frac{1}{3}, \dots$ besitzt keine Häufungspunkte in $(0, \infty)$. Auf dem Raum $E = [0, \infty)$ ist jedoch das gleiche Maß μ kein Zählmaß, denn 0 ist ein Häufungspunkt.

Punktprozesse. Wir wollen nun definieren, was eine *zufällige* Punktconfiguration ohne Häufungspunkte (ein zufälliges Zählmaß) ist. Solche zufälligen Punktconfigurationen heißen Punktprozesse. Es sei (E, ρ) ein lokal kompakter separabler metrischer Raum. Wir bezeichnen mit $\mathbb{M} = \mathbb{M}(E)$ die Menge aller Zählmaße auf E .

Definition 9.3.13. Sei $\mathcal{M} \subset 2^{\mathbb{M}}$ die von allen Mengen der Form

$$U_A^k := \{\mu \in \mathbb{M} : \mu(A) = k\} \subset \mathbb{M},$$

erzeugte σ -Algebra, wobei $A \subset E$ eine Borel-Menge und $k \in \mathbb{N}_0 \cup \{+\infty\}$ ist.

Die σ -Algebra \mathcal{M} enthält auch Mengen der Form

$$U_{A_1, \dots, A_n}^{k_1, \dots, k_n} := \{\mu \in \mathbb{M} : \mu(A_1) = k_1, \dots, \mu(A_n) = k_n\} = U_{A_1}^{k_1} \cap \dots \cap U_{A_n}^{k_n},$$

wobei $n \in \mathbb{N}$, $k_1, \dots, k_n \in \mathbb{N}_0 \cup \{+\infty\}$ und $A_1, \dots, A_n \subset E$ Borel-Mengen.

Definition 9.3.14. Ein **Punktprozess** ist eine messbare Abbildung π von einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ nach $(\mathbb{M}, \mathcal{M})$.

Für jedes $\omega \in \Omega$ ist somit $\pi(\omega)$ ein Zählmaß (=eine Punktconfiguration) auf E .

Notation 9.3.15. Für eine Borel-Menge $A \subset E$ bezeichnen wir mit $\pi(A; \omega)$ oder auch abgekürzt mit $\pi(A)$ die Anzahl der Punkte dieser Punktconfiguration in der Menge A .

Dann ist $\pi(A) : \Omega \rightarrow \mathbb{N}_0 \cup \{+\infty\}$, $\omega \mapsto \pi(A; \omega)$, eine Zufallsvariable, denn für jedes $k \in \mathbb{N}_0 \cup \{+\infty\}$ ist das Urbild

$$\{\omega \in \Omega : \pi(A; \omega) = k\} = \pi^{-1}(U_A^k)$$

\mathcal{A} -messbar.

Beispiel 9.3.16. Sei $N \in \mathbb{N}$ fest und es seien X_1, \dots, X_N u.i.v. Zufallsvektoren mit Werten in \mathbb{R}^d . Der Punktprozess

$$\pi = \sum_{i=1}^N \delta_{X_i}$$

wird der **Binomialpunktprozess** genannt, denn die Anzahl der Punkte von π in einer Borel-Menge $A \subset \mathbb{R}^d$ ist binomialverteilt:

$$\pi(A) \sim \text{Bin}(N, \mathbb{P}[X_1 \in A]).$$

9.4. Definition der Poisson-Punktprozesse

Sei μ ein Radon-Maß auf einem lokal kompakten separablen metrischen Raum (E, ρ) . Als Beispiel von E kann man sich immer eine abgeschlossene oder eine offene Teilmenge von \mathbb{R}^d vorstellen.

Definition 9.4.1. Ein Punktprozess π auf E heißt **Poisson-Punktprozess** mit **Intensitätsmaß** μ , falls folgende zwei Bedingungen gelten:

- (1) Für alle Borel-Mengen $A \subset E$ ist $\pi(A)$ Poisson-verteilt mit Parameter $\mu(A)$.
- (2) Für alle disjunkten Borel-Mengen $A_1, \dots, A_n \subset E$ sind die Zufallsvariablen $\pi(A_1), \dots, \pi(A_n)$ unabhängig.

Wir benutzen die Schreibweise $\pi \sim \text{PPP}(\mu)$.

Bemerkung 9.4.2. Aus der ersten Eigenschaft folgt, dass $\mathbb{E}\pi(A) = \mu(A)$. Das Intensitätsmaß beschreibt also die erwartete Anzahl der Punkte in einem Poisson-Punktprozess.

In den Fällen, die für uns von Interesse sind, hat das Maß μ eine Dichte bezüglich des Lebesgue-Maßes.

Definition 9.4.3. Eine messbare Funktion $f : E \rightarrow \mathbb{R}$ heißt **lokal integrierbar**, falls $\int_B |f(t)| dt < \infty$ für jede kompakte Menge $B \subset E$.

Ist nun f eine nicht-negative lokal integrierbare Funktion auf \mathbb{R}^d (oder allgemeiner auf einer offenen oder abgeschlossenen Teilmenge von \mathbb{R}^d), so kann man ein Radon-Maß μ mit

$$\mu(B) = \int_B f(t) dt$$

für alle Borel-Mengen $B \subset E$ definieren. Die Funktion f heißt die **Dichte** von μ und wir schreiben dann $\mu(dt) = f(t)dt$. Einen Poisson-Punktprozess $\pi \sim \text{PPP}(\mu)$ werden wir dann auch mit $\text{PPP}(f(t)dt)$ bezeichnen. Die Funktion f nennen wir dann die **Intensität** von π .

Beispiel 9.4.4. Im Beispiel mit dem Sternenhimmel haben wir einen Poisson-Punktprozess mit einer konstanten Intensität $f(t) = \lambda > 0$ betrachtet. Ein solcher Poisson-Punktprozess heißt **homogen**; siehe Abbildung 1.

Beispiel 9.4.5. Einen Poisson–Punktprozess auf $(0, \infty)$ mit konstanter Intensität $\lambda > 0$ kann man folgendermaßen konstruieren. Seien E_1, E_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit $E_i \sim \text{Exp}(\lambda)$, d.h.

$$\mathbb{P}[E_i > t] = e^{-\lambda t}, \quad t > 0.$$

Dann bilden die Punkte $E_1, E_1 + E_2, E_1 + E_2 + E_3, \dots$ einen Poisson–Punktprozess mit Intensität λ (ohne Beweis).

Durch eine zweiseitige Version des obigen Verfahrens kann man auch einen Poisson–Punktprozess auf ganz \mathbb{R} mit Intensität λ konstruieren. Dazu seien $E_1, E'_1, E_2, E'_2, \dots$ unabhängige und mit Parameter λ exponentialverteilte Zufallsvariablen. Dann ist

$$\sum_{n=1}^{\infty} \delta_{E_1+\dots+E_n} + \sum_{n=1}^{\infty} \delta_{-(E'_1+\dots+E'_n)}$$

ein Poisson–Punktprozess auf \mathbb{R} mit konstanter Intensität λ .

9.5. Superpositionssatz

Eine wichtige Eigenschaft der Poisson–Verteilung ist ihre Faltungstabilität. Sind nämlich $X_1 \sim \text{Poi}(\lambda_1), \dots, X_n \sim \text{Poi}(\lambda_n)$ unabhängige Zufallsvariablen, so gilt

$$X_1 + \dots + X_n \sim \text{Poi}(\lambda_1 + \dots + \lambda_n).$$

Wir beweisen, dass diese Eigenschaft auf unendliche Summen erweitert werden kann.

Satz 9.5.1. Seien $X_i \sim \text{Poi}(\lambda_i)$, $i \in \mathbb{N}$, unabhängige Zufallsvariablen mit $\lambda_i \in [0, \infty]$. Dann gilt:

$$S := \sum_{i=1}^{\infty} X_i \sim \text{Poi} \left(\sum_{i=1}^{\infty} \lambda_i \right).$$

Bemerkung 9.5.2. Wenn $\sum_{i=1}^{\infty} \lambda_i = \infty$, dann gilt $S = \infty$ fast sicher.

BEWEIS. Wenn mindestens ein λ_i unendlich ist, dann ist $X_i = S = +\infty$ fast sicher und die Aussage stimmt. Seien also alle λ_i endlich. Sei $S_n = X_1 + \dots + X_n$ und $\sigma_n = \lambda_1 + \dots + \lambda_n$. Bekannt ist bereits, dass $S_n \sim \text{Poi}(\sigma_n)$. Sei $r \in \mathbb{N}_0$, dann gilt

$$\{S_1 \leq r\} \supseteq \{S_2 \leq r\} \supseteq \dots \quad \text{und} \quad \bigcap_{i=1}^{\infty} \{S_i \leq r\} = \{S \leq r\}$$

Deshalb gilt wegen der Stetigkeit der Wahrscheinlichkeit:

$$\mathbb{P}[S \leq r] = \lim_{n \rightarrow \infty} \mathbb{P}[S_n \leq r] = \lim_{n \rightarrow \infty} \sum_{k=0}^r e^{-\sigma_n} \frac{\sigma_n^k}{k!},$$

wobei im letzten Schritt benutzt wurde, dass $S_n \sim \text{Poi}(\sigma_n)$. Man kann hier zwei Fälle unterscheiden:

Fall 1: $\sigma_n \rightarrow \sigma$ mit $\sigma < \infty$. Dann gilt: $\mathbb{P}[S \leq r] = \sum_{k=0}^r e^{-\sigma} \frac{\sigma^k}{k!}$ und daher ist $S \sim \text{Poi}(\sigma)$.

Fall 2: $\sigma_n \rightarrow \infty$. Dann gilt: $\mathbb{P}[S \leq r] = 0$ für alle $r \in \mathbb{N}$ und somit ist $S = \infty$ fast sicher.

□

Es seien π_1, π_2, \dots Punktprozesse, die auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ definiert sind. Die **Superposition** $\pi = \sum_{i=1}^{\infty} \pi_i$ der Punktprozesse π_1, π_2, \dots ist die Vereinigung aller Punkte dieser Punktprozesse:

$$\pi(A) = \sum_{i=1}^{\infty} \pi_i(A).$$

Hierbei ist π nicht immer ein Punktprozess, denn die unendliche Summe $\pi(A)$ kann auch für eine kompakte Menge A unendlich sein. Im nächsten Satz beweisen wir, dass die Superposition von unabhängigen Poisson-Punktprozessen wieder ein Poisson-Punktprozess ist, wenn die Summe der Intensitätsmaße wieder ein Radon-Maß ist.

Satz 9.5.3. Seien $\pi_i \sim \text{PPP}(\mu_i)$, $i \in \mathbb{N}$, unabhängige Poisson-Punktprozesse, die auf einem gemeinsamen Wahrscheinlichkeitsraum definiert sind. Ist $\mu = \sum_{i=1}^{\infty} \mu_i$ ein Radon-Maß, so gilt:

$$\pi := \sum_{i=1}^{\infty} \pi_i \sim \text{PPP}(\mu).$$

BEWEIS. Sei $A \subset E$ eine Borel-Menge. Es gilt $\pi_i(A) \sim \text{Poi}(\mu_i(A))$, weil $\pi_i \sim \text{PPP}(\mu_i)$. Mit Satz 9.5.1 folgt

$$\pi(A) = \sum_{i=1}^{\infty} \pi_i(A) \sim \text{Poi}\left(\sum_{i=1}^{\infty} \mu_i(A)\right) = \text{Poi}(\mu(A)).$$

Damit ist die Eigenschaft 1 aus der Definition der Poisson-Punktprozesse gezeigt. Dabei sei bemerkt, dass für eine kompakte Menge A die Zahl $\mu(A)$ endlich ist (denn μ ist ein Radon-Maß), somit ist π ein Zählmaß.

Seien nun $A_1, \dots, A_n \subset E$ disjunkte Borel-Mengen. Es gilt für jedes $i \in \mathbb{N}$:

$$\pi_i(A_1), \dots, \pi_i(A_n) \text{ sind unabhängig, da } \pi_i \sim \text{PPP}(\mu_i).$$

Außerdem sind Punktprozesse π_1, π_2, \dots unabhängig. Es folgt, dass die Zufallsvariablen

$$\pi(A_1) = \sum_{i=1}^{\infty} \pi_i(A_1), \quad \dots, \quad \pi(A_n) = \sum_{i=1}^{\infty} \pi_i(A_n) \text{ unabhängig sind.}$$

Daher ist auch die zweite Eigenschaft aus der Definition nachgewiesen, woraus die Behauptung folgt. □

9.6. Abbildungssatz

Der Abbildungssatz behauptet grob gesagt, dass das Bild eines Poisson-Punktprozesses unter einer Abbildung wieder ein Poisson-Punktprozess ist. Diese Behauptung gilt allerdings nicht für jede Abbildung, wie das nächste Beispiel zeigt.

Beispiel 9.6.1. Betrachte einen homogenen Poisson–Punktprozess π auf \mathbb{R}^2 mit Intensität 1. Projiziere nun alle Punkte auf die x -Achse. Die Projektionen bilden allerdings gar keinen Punktprozess, denn sie liegen überall dicht in \mathbb{R} . In der Tat, jeder Streifen der Form $[a, b] \times \mathbb{R}$ enthält unendlich viele Punkte von π , denn das Lebesgue–Maß dieses Streifens ist unendlich. Somit werden unendlich viele Punkte in das Intervall $[a, b]$ projiziert.

Wir werden deshalb die Klasse der Abbildungen einschränken müssen.

Definition 9.6.2. Eine stetige Abbildung $T : E_1 \rightarrow E_2$ zwischen zwei lokal kompakten metrischen Räumen E_1 und E_2 heißt **eigentlich**, wenn für jede kompakte Menge $K \subset E_2$ das Urbild $T^{-1}(K)$ ebenfalls kompakt ist.

Beispiel 9.6.3. Die Projektionsabbildung $T : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $T(x, y) = x$ ist nicht eigentlich.

Definition 9.6.4. Sei $T : E_1 \rightarrow E_2$ eine eigentliche Abbildung zwischen zwei lokal kompakten metrischen Räumen. Sei μ ein Radon–Maß auf E_1 . Definiere das sogenannte **Bildmaß** $T\mu$ von μ unter T als ein Maß auf E_2 mit

$$(T\mu)(A) = \mu(T^{-1}(A)), \quad A \in \mathcal{B}(E_2).$$

Proposition 9.6.5. Das Bildmaß $T\mu$ ist ein Radon–Maß auf E_2 . Ist μ sogar ein Zählmaß, dann ist $T\mu$ ebenfalls ein Zählmaß.

BEWEIS. Für jede kompakte Menge $K \subset E_2$ gilt $(T\mu)(K) = \mu(T^{-1}(K)) < \infty$, denn $T^{-1}(K)$ ist kompakt und μ ist ein Radon–Maß. Ist μ sogar ein Zählmaß, dann ist $(T\mu)(K) = \mu(T^{-1}(K)) \in \mathbb{N}_0$ und somit $T\mu$ ebenfalls ein Zählmaß. \square

Beispiel 9.6.6. Sei μ ein Zählmaß auf E_1 mit $\mu = \sum_i \delta_{x_i}$. Dann ist $T\mu = \sum_i \delta_{Tx_i}$.

Wir können nun den Abbildungssatz formulieren.

Satz 9.6.7 (Abbildungssatz). Sei $T : E_1 \rightarrow E_2$ eine eigentliche Abbildung zwischen zwei lokal kompakten metrischen Räumen E_1 und E_2 . Sei $\pi \sim \text{PPP}(\mu)$ ein Poisson–Punktprozess auf E_1 mit Intensitätsmaß μ . Dann gilt $T\pi \sim \text{PPP}(T\mu)$.

BEWEIS. Wir überprüfen, ob die Bedingungen aus Definition 9.4.1 erfüllt sind. Sei $A \subset E_2$ eine Borel–Menge. Es gilt:

$$(T\pi)(A) = \pi(T^{-1}(A)) \sim \text{Poi}(\mu(T^{-1}(A))) = \text{Poi}((T\mu)(A)).$$

Wenn $A \subset E_2$ sogar kompakt ist, dann folgt außerdem, dass $(T\pi)(A) \sim \text{Poi}((T\mu)(A)) < \infty$ fast sicher, weshalb $T\pi$ ein Punktprozess ist.

Seien nun $A_1, \dots, A_n \subset E_2$ disjunkte Borel-Mengen. Dann sind auch die Urbilder dieser Mengen $T^{-1}(A_1), \dots, T^{-1}(A_n)$ disjunkt. Es folgt, dass die Zufallsvariablen

$$(T\pi)(A_1) = \pi(T^{-1}(A_1)), \dots, (T\pi)(A_n) = \pi(T^{-1}(A_n))$$

unabhängig sind, da π ein Poisson-Punktprozess ist. □

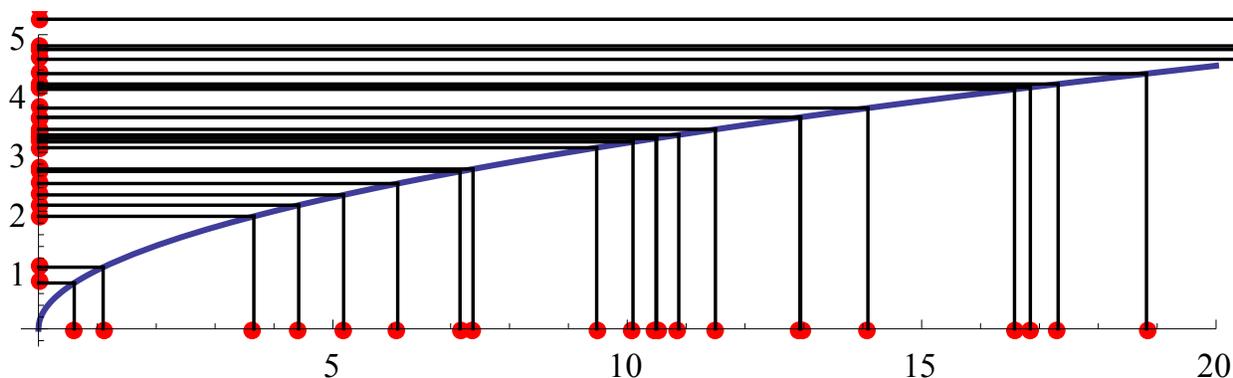


ABBILDUNG 2. Abbildungssatz.

Beispiel 9.6.8 (Siehe Abbildung 2). Seien $P_1 < P_2 < \dots$ die Punkte eines Poisson-Punktprozesses mit Intensität 1 auf $(0, \infty)$. Betrachte die Punkte $\sqrt{P_1}, \sqrt{P_2}, \dots$. Diese bilden ebenfalls einen Poisson-Punktprozess auf $(0, \infty)$, denn die Abbildung $T : (0, \infty) \rightarrow (0, \infty)$ mit $T(x) = \sqrt{x}$ ist eigentlich. Sei μ das Lebesgue-Maß auf $(0, \infty)$, dann gilt

$$(T\mu)((0, x)) = \mu(T^{-1}((0, x))) = \mu((0, x^2)) = x^2.$$

Somit hat $T\mu$ die Dichte $f(x) = 2x$, $x > 0$, und die Intensität des Poisson-Punktprozesses $\sqrt{P_1}, \sqrt{P_2}, \dots$ ist $2x$.

9.7. Laplace-Funktionale

Sei E ein lokal kompakter separabler metrischer Raum.

Definition 9.7.1. Es sei $B(E)$ die Menge aller Borel-Funktionen $f : E \rightarrow \mathbb{R}$. Es sei $B_+(E)$ die Menge aller nicht-negativen Borel-Funktionen $f : E \rightarrow \mathbb{R}$.

Definition 9.7.2. Sei π ein Punktprozess auf E . Für $f \in B_+(E)$ definiere die Zufallsvariable $S_f = \sum_{x \in \pi} f(x)$ mit Werten in $[0, \infty]$. Dann heißt die Abbildung

$$\psi_\pi : B(E) \rightarrow [0, \infty) \text{ mit } \psi_\pi(f) = \mathbb{E}[e^{-S_f}]$$

das **Laplace-Funktional** von π .

Es sei bemerkt, dass für $f \in B(E)$ (ohne die Annahme $f \geq 0$) die unendliche Summe S_f nicht immer wohldefiniert wäre. Im nächsten Satz berechnen wir das Laplace-Funktional eines Poisson-Punktprozesses.

Satz 9.7.3. Sei π ein Poisson-Punktprozess auf E mit Intensitätsmaß μ , dann ist für alle $f \in B_+(E)$

$$(9.7.1) \quad \mathbb{E}[e^{-S_f}] = \exp \left\{ - \int_E (1 - e^{-f(x)}) \mu(dx) \right\}.$$

BEWEIS. *Schritt 1.* (Indikatorfunktionen) Sei zuerst $f(x) = c \mathbb{1}_A(x)$, mit $c \in \mathbb{R}$, $A \subset E$ Borel. Dann ist $S_f = c\pi(A)$. Es gilt $\pi(A) \sim \text{Poi}(\mu(A))$. Es folgt

$$\mathbb{E}[e^{-S_f}] = \mathbb{E}[e^{-c\pi(A)}] = \sum_{k=0}^{\infty} e^{-\mu(A)} \frac{\mu(A)^k}{k!} e^{-ck} = e^{-\mu(A)} \sum_{k=0}^{\infty} \frac{(\mu(A)e^{-c})^k}{k!} = e^{-\mu(A)(1-e^{-c})}.$$

Somit gilt die Behauptung für $f(x) = c \mathbb{1}_A(x)$.

Schritt 2. (Einfache Funktionen) Sei nun $f(x) = c_1 \cdot \mathbb{1}_{A_1}(x) + \dots + c_n \cdot \mathbb{1}_{A_n}(x)$ mit disjunkten Borel-Mengen $A_1, \dots, A_n \subset E$ und $c_1, \dots, c_n \in \mathbb{R}$. Dann gilt:

$$\mathbb{E}[e^{-S_f}] = \mathbb{E}[e^{-c_1\pi(A_1)} \dots e^{-c_n\pi(A_n)}] = \prod_{i=1}^n \mathbb{E}[e^{-c_i\pi(A_i)}],$$

da π ein Poisson-Punktprozess ist und daher die Zufallsvariablen $\pi(A_1), \dots, \pi(A_n)$ unabhängig sind. Da wir die Gültigkeit der Behauptung für Indikatorfunktionen bereits im ersten Schritt gezeigt haben, folgt:

$$\mathbb{E}[e^{-S_f}] = \prod_{i=1}^n \mathbb{E}[e^{-c_i\pi(A_i)}] = \exp \left\{ - \sum_{i=1}^n \mu(A_i)(1 - e^{-c_i}) \right\} = \exp \left\{ - \int_E (1 - e^{-f(x)}) \mu(dx) \right\}.$$

Schritt 3. Sei nun $f \geq 0$ eine beliebige Borel-Funktion. Dann gibt es einfache Funktionen f_1, f_2, \dots , die punktweise von unten gegen f konvergieren. Da wir die Richtigkeit der Behauptung für einfache Funktionen bereits bewiesen haben, folgt mit dem Satz von der monotonen Konvergenz, dass sie auch für f gilt. \square

Bemerkung 9.7.4. Die Formel (9.7.1) gilt auch in folgender Form, die etwas allgemeiner ist:

$$\mathbb{E}[e^{-\theta S_f}] = \exp \left\{ - \int_{\mathbb{R}^d} (1 - e^{-\theta f(x)}) \mu(dx) \right\}, \quad \theta \geq 0.$$

Die Richtigkeit dieser Behauptung lässt sich nachweisen, indem man $\theta f(x)$ anstelle von $f(x)$ in (9.7.1) einsetzt.

Korollar 9.7.5 (Campbell). Sei $\pi \sim \text{PPP}(\mu)$ auf E und $f \in B_+(E)$, dann gilt:

$$\mathbb{E}[S_f] = \int_E f(x)\mu(dx).$$

BEWEISIDEE. Wegen Satz 9.7.3 gilt:

$$\mathbb{E}[S_f] = -\frac{d}{d\theta} \log \mathbb{E}[e^{\theta S_f}] \Big|_{\theta=0} = \frac{d}{d\theta} \int_E (1 - e^{-\theta f(x)})\mu(dx) \Big|_{\theta=0}.$$

Durch Vertauschung von Integral und Ableitung (was wir hier nicht begründen werden) lässt sich der Ausdruck wie folgt schreiben und vereinfachen:

$$\mathbb{E}[S_f] = \int_E \frac{d}{d\theta} (1 - e^{-\theta f(x)}) \Big|_{\theta=0} \mu(dx) = \int_E f(x)\mu(dx),$$

was die Behauptung beweist. □

Aufgabe 9.7.6. Zeigen Sie, dass für alle $f \in B_+(E)$

$$\mathbb{E}[S_f^2] = \int_E f^2(x)\mu(dx) + \left(\int_E f(x)\mu(dx) \right)^2.$$

Im nächsten Satz zeigen wir, dass man einen Poisson–Punktprozess an seinem Laplace–Funktional erkennen kann.

Satz 9.7.7. Sei π ein Punktprozess auf E mit $\mathbb{E}[e^{-S_f}] = e^{-\int_E (1 - e^{-f(x)})\mu(dx)}$ für alle Funktionen $f \in B_+(E)$, wobei μ ein Radon–Maß ist. Dann ist $\pi \sim \text{PPP}(\mu)$.

BEWEIS. Wir zeigen, dass $\pi(A) \sim \text{Poi}(\mu(A))$ für alle Borel–Mengen $A \subset E$. Sei $f(x) = \theta \mathbb{1}_A(x)$ mit $\theta \geq 0$. Einsetzen liefert:

$$\mathbb{E}[e^{-\theta \pi(A)}] = \exp\{-\mu(A)(1 - e^{-\theta})\}, \text{ für alle } \theta \geq 0.$$

Da es sich bei $\exp\{-\mu(A)(1 - e^{-\theta})\}$ um die Laplace–Transformierte einer $\text{Poi}(\mu(A))$ –verteilten Zufallsvariable handelt, folgt mit der Eindeutigkeit der Laplace–Transformierten, dass $\pi(A)$ Poisson–verteilt mit Parameter $\mu(A)$ ist.

Seien nun $A_1, \dots, A_n \subset E$ disjunkte Borel–Mengen. Wir zeigen, dass dann die Zufallsvariablen $\pi(A_1), \dots, \pi(A_n)$ unabhängig sind. Sei

$$f(x) = \theta_1 \mathbb{1}_{A_1}(x) + \dots + \theta_n \mathbb{1}_{A_n}(x), \text{ mit } \theta_1, \dots, \theta_n \geq 0.$$

Einsetzen liefert:

$$\mathbb{E}[e^{-\theta_1 \pi(A_1)} \cdot \dots \cdot e^{-\theta_n \pi(A_n)}] = \exp\left\{-\int_E (1 - e^{-f(x)})\mu(dx)\right\} = \exp\left\{-\sum_{i=1}^n \mu(A_i)(1 - e^{-\theta_i})\right\}.$$

Somit gilt:

$$\mathbb{E}[e^{-\theta_1\pi(A_1)} \cdot \dots \cdot e^{-\theta_n\pi(A_n)}] = \prod_{i=1}^n e^{-\mu(A_i)(1-e^{-\theta_i})} = \prod_{i=1}^n \mathbb{E}[e^{-\theta_i\pi(A_i)}],$$

woraus folgt, dass $\pi(A_1), \dots, \pi(A_n)$ unabhängig sind. \square

9.8. Färbungs- und Markierungssatz für Poisson-Punktprozesse

Man stelle sich vor, dass jeder Punkt eines Poisson-Punktprozesses auf einem Raum E_1 eine zufällige Farbe (etwa aus der Menge rot, grün, blau) bekommt; siehe Abbildung 3. Wir nehmen an, dass alle Punkte unabhängig voneinander gefärbt werden. Außerdem nehmen wir an, dass die Färbung der Punkte unabhängig von der Erzeugung der Positionen der Punkte geschieht. Bei der zweiten Annahme muss man vorsichtig sein, denn um die Punkte zu färben, muss man wissen, welche Punkte man färben soll, was wie eine Abhängigkeit aussieht. Wir werden deshalb wie folgt vorgehen: Zuerst färben wir *alle* Punkte von E_1 unabhängig voneinander, und erst dann erzeugen wir (unabhängig von der Färbungsprozedur) einen Poisson-Punktprozess. Die Farben der Punkte, die nicht zum Poisson-Punktprozess gehören, werden dann einfach ignoriert. Der Färbungssatz behauptet nun, dass alle Punkte des Poisson-Punktprozesses, die eine gegebene Farbe (etwa rot) haben, ebenfalls einen Poisson-Punktprozess bilden. Außerdem behauptet der Färbungssatz, dass die drei Poisson-Punktprozesse der roten, blauen und grünen Punkte unabhängig voneinander sind.

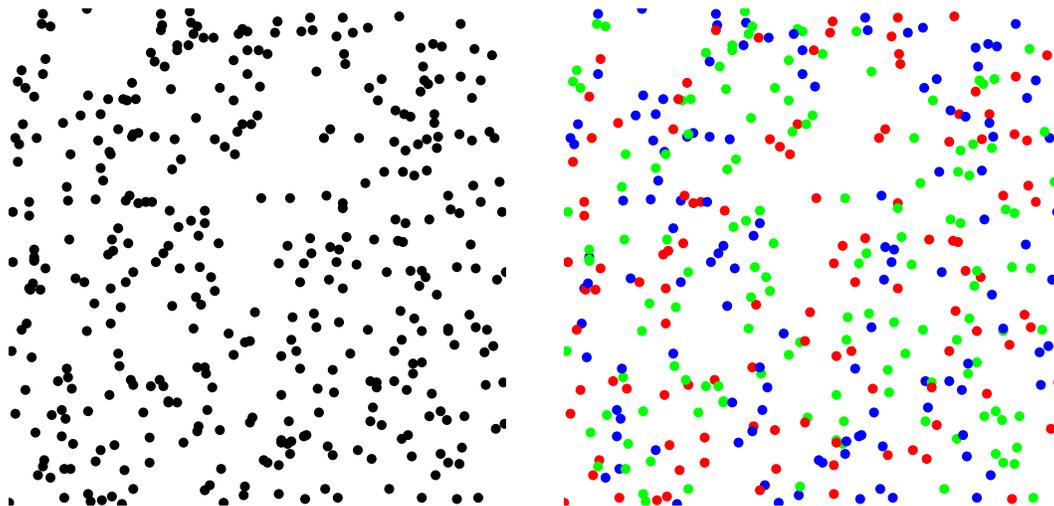


ABBILDUNG 3. Färbungssatz.

Wir werden nun die obigen Überlegungen etwas verallgemeinern. Wir werden eine beliebige Menge E_2 an Farben (z.B. eine stetige Farbpalette) zulassen. Es seien also E_1 und E_2 zwei lokal kompakte separable metrische Räume. Die Konstruktion eines markierten Punktprozesses besteht aus zwei Schritten, die unabhängig voneinander ausgeführt werden sollen.

Schritt 1: Marken. Erzeuge für jeden Punkt $x \in E_1$ eine Zufallsvariable m_x (die Marke oder die Farbe von x) mit Werten in E_2 . Wir bezeichnen die Verteilung von m_x mit λ_x . Somit ist λ_x ein Wahrscheinlichkeitsmaß auf E_2 für jedes $x \in E_1$. Das Wahrscheinlichkeitsmaß λ_x darf

im Allgemeinen eine Funktion von x sein, so dass verschiedene Punkte x nach verschiedenen Regeln gefärbt werden dürfen. Wir brauchen folgende zwei Annahmen:

- (1) $m_x, x \in E_1$, sind unabhängige Zufallsvariablen.
- (2) $x \mapsto \lambda_x(B)$ ist eine Borel-Funktion für alle Borel-Mengen $B \subset E_2$.

Beispiel 9.8.1. Sei die Menge E_2 aller möglichen Farben endlich. OEdA kann sie dann mit $\{1, \dots, n\}$ identifiziert werden. Wenn die Wahrscheinlichkeit, dass ein Punkt $x \in E_1$ mit Farbe $i \in \{1, \dots, n\}$ gefärbt wird, mit $p_i(x)$ bezeichnet wird, dann sehen die Verteilungen λ_x folgendermaßen aus:

$$\lambda_x(\{i\}) = p_i(x), \quad i = 1, \dots, n.$$

Die obigen Annahmen besagen dann, dass die Punkte unabhängig gefärbt werden und dass $p_1, \dots, p_n : E_1 \rightarrow [0, 1]$ messbare Funktionen sind.

Schritt 2: Positionen der Punkte. Sei außerdem π ein Poisson-Punktprozess auf E_1 mit Intensitätsmaß μ , das unabhängig von den Marken $\{m_x : x \in E_1\}$ erzeugt werden soll. Der Einfachheit halber nehmen wir an, dass μ keine Atome hat (d.h. $\mu(\{x\}) = 0$ für alle $x \in E_1$), so dass im Punktprozess π keine multiplen Punkte auftreten.

Wir definieren den **markierten Poisson-Punktprozess** π^* auf dem kartesischen Produkt $E_1 \times E_2$ wie folgt:

$$\pi^* = \sum_{x \in \pi} \delta_{(x, m_x)}.$$

Die Schreibweise $x \in \pi$ bedeutet, dass die Summe über alle Punkte im Punktprozess π gebildet wird. Den Markierungssatz können wir nun folgendermaßen formulieren.

Satz 9.8.2 (Markierungssatz). π^* ist Poisson-Punktprozess auf $E_1 \times E_2$. Für dessen Intensitätsmaß μ^* gilt

$$\mu^*(C) = \iint_{(x, m) \in C} \mu(dx) \lambda_x(dm), \quad \text{für } C \subset E_1 \times E_2 \text{ Borel.}$$

Bemerkung 9.8.3. Die Formel für μ^* muss man folgendermaßen verstehen: Für eine Menge C der Form $A_1 \times A_2$ mit $A_1 \subset E_1$ und $A_2 \subset E_2$ Borel, gilt

$$\mu^*(A_1 \times A_2) = \int_{A_1} \lambda_x(A_2) \mu(dx).$$

BEWEIS. Sei $f : E_1 \times E_2 \rightarrow \mathbb{R}$ mit $f \geq 0$ eine Borel-Funktion. Wir schreiben $S^* = \sum_{(x, m) \in \pi^*} f(x, m)$. Wir wollen zuerst den folgenden bedingten Erwartungswert berechnen:

$$\mathbb{E}[e^{-S^*} | \pi] = \prod_{x \in \pi} \mathbb{E}[e^{-f(x, m_x)} | \pi] = \prod_{x \in \pi} \int_{E_2} e^{-f(x, m)} \lambda_x(dm).$$

Mit Hilfe einer einfachen Transformation lässt sich der obige Ausdruck wie folgt darstellen:

$$\mathbb{E}[e^{-S^*} | \pi] = \prod_{x \in \pi} \int_{E_2} e^{-f(x, m)} \lambda_x(dm) = \exp \left\{ - \sum_{x \in \pi} - \log \int_{E_2} e^{-f(x, m)} \lambda_x(dm) \right\}.$$

Mit der Formel der totalen Erwartung folgt

$$\mathbb{E}[e^{-S^*}] = \mathbb{E}_\pi \left[\exp \left\{ - \sum_{x \in \pi} f^*(x) \right\} \right]$$

mit

$$f^*(x) = - \log \int_{E_2} e^{-f(x,m)} \lambda_x(dm).$$

Mit Hilfe von Satz 9.7.3 lässt sich das wie folgt umformen:

$$\mathbb{E}[e^{-S^*}] = \exp \left\{ - \int_{E_1} (1 - e^{-f^*(x)}) \mu(dx) \right\}.$$

Einsetzen von f^* liefert nun folgende Formel:

$$\exp \left\{ - \int_{E_1} (1 - e^{-f^*(x)}) \mu(dx) \right\} = \exp \left\{ - \int_{E_1} \int_{E_2} (1 - e^{-f(x,m)}) \mu(dx) \lambda(dm) \right\}.$$

Mit $\mu^*(dx, dm) := \mu(dx) \lambda(dm)$ und mit Satz 9.7.7 folgt schließlich, dass π^* ein Poisson-Punktprozess mit Intensitätsmaß μ^* ist. \square

Beispiel 9.8.4. Sei $A \subset E_2$ eine Menge von Farben. Betrachte nun alle Punkte von π , die eine Farbe aus E_2 bekommen. Wir zeigen, dass diese Punkte einen Poisson-Punktprozess bilden. Nach dem Markierungssatz ist $\sum_{x \in \pi} \delta_{(x, m_x)}$ ein Poisson-Punktprozess auf $E_1 \times E_2$. Da die Einschränkung dieses Poisson-Punktprozesses auf die Teilmenge $E_1 \times A$ ebenfalls ein Poisson-Punktprozess ist, folgt, dass $\sum_{x \in \pi: m_x \in A} \delta_{(x, m_x)}$ ein Poisson-Punktprozess auf $E_1 \times A$ ist. Aus dem Abbildungssatz für die Projektion $(x, m) \mapsto x$ folgt, dass $\sum_{x \in \pi: m_x \in A} \delta_x$ ebenfalls ein Poisson-Punktprozess ist.

KAPITEL 10

Konvergenz von Punktprozessen

In diesem Kapitel werden wir Verteilungskonvergenz der Punktprozesse einführen und einige Beispiele betrachten, in denen Poisson-Prozesse als Grenzwerte von Punktprozessen der “seltenen Ereignisse” entstehen. Unsere Darstellung ist sehr unvollständig, für mehr Einzelheiten verweisen wir auf das Buch von S. Resnick “*Extreme values, regular variation and point processes*”.

10.1. Vage Konvergenz

Im Folgenden sei E ein lokal kompakter separabler metrischer Raum. Zuerst werden wir einen Konvergenzbegriff für Radon-Maße auf E (und somit auch für Zählmaße auf E) einführen.

Definition 10.1.1. Eine Menge $B \subset E$ heißt **relativ kompakt**, wenn ihr Abschluss \bar{B} kompakt ist.

Beispiel 10.1.2. Eine Menge $B \subset \mathbb{R}^d$ ist relativ kompakt genau dann, wenn sie beschränkt ist.

Definition 10.1.3. Eine Folge von Radon-Maßen μ_1, μ_2, \dots auf E konvergiert **vage** gegen ein Radon-Maß μ , wenn

$$\lim_{n \rightarrow \infty} \mu_n(B) = \mu(B)$$

für alle relativ kompakten Borel-Mengen $B \subset E$ mit $\mu(\partial B) = 0$. Bezeichnung: $\mu_n \xrightarrow[n \rightarrow \infty]{v} \mu$.

Warum in obiger Definition $\mu(\partial B) = 0$ gelten muss, soll folgendes Beispiel veranschaulichen.

Beispiel 10.1.4. Betrachte folgende Radon-Maße auf \mathbb{R} : $\mu_n = \delta_{1/n}$ und $\mu = \delta_0$. Für das offene Intervall $B = (0, 2)$ gilt $\mu_n(B) = 1$ für alle $n \in \mathbb{N}$, aber $\mu(B) = 0$. Hätten wir in der Definition der vagen Konvergenz die Forderung $\mu(\partial B) = 0$ weggelassen, so würde μ_n nicht gegen μ konvergieren. Das wäre ein sehr unnatürlicher Konvergenzbegriff.

Hier sind einige einfache Beispiele der vagen Konvergenz.

Beispiel 10.1.5. Sei μ_n das Maß mit der Dichte $\mathbb{1}_{[-n, n]}$ auf \mathbb{R} . Dann konvergiert μ_n vage gegen das Lebesgue-Maß auf \mathbb{R} .

Beispiel 10.1.6. Sei $\mu_n = \delta_n$. Dann konvergiert μ_n vage gegen das Null-Maß.

Wir werden nun eine äquivalente Definition der vagen Konvergenz formulieren. Zuerst müssen wir stetige Funktionen mit kompaktem Träger definieren, die wir als “Testfunktionen” benutzen werden.

Definition 10.1.7. Eine Funktion $f : E \rightarrow \mathbb{R}$ hat **kompakten Träger**, falls es eine kompakte Menge $K \subset E$ gibt mit $f(t) = 0$ für alle $t \in E \setminus K$.

Definition 10.1.8. Die Menge aller stetigen Funktionen auf E mit kompaktem Träger sei mit $C_c(E)$ bezeichnet. Es sei $C_c^+(E)$ die Menge aller $f \in C_c(E)$ mit $f \geq 0$.

Nun formulieren wir eine äquivalente Definition der vagen Konvergenz.

Satz 10.1.9. Seien μ_1, μ_2, \dots und μ Radon-Maße auf E . Dann sind die folgenden Aussagen äquivalent:

- (1) $\mu_n \xrightarrow[n \rightarrow \infty]{v} \mu$.
- (2) Für alle $f \in C_c^+(E)$ gilt $\lim_{n \rightarrow \infty} \int_E f(x) \mu_n(dx) = \int_E f(x) \mu(dx)$.

BEWEIS. Weggelassen. □

Aufgabe 10.1.10. Zeigen Sie, dass man im obigen Satz $C_c^+(E)$ durch $C_c(E)$ ersetzen kann.

Beispiel 10.1.11. Sei $\mu_n = \frac{1}{n} \sum_{i \in \mathbb{Z}} \delta_{i/n}$. Dann konvergiert μ_n vage gegen das Lebesgue-Maß auf \mathbb{R} . In der Tat, für jedes $f \in C_c(\mathbb{R})$ gilt

$$\int_{\mathbb{R}} f(t) \mu_n(dt) = \frac{1}{n} \sum_{i \in \mathbb{Z}} f\left(\frac{i}{n}\right) \xrightarrow[n \rightarrow \infty]{} \int_{\mathbb{R}} f(t) dt,$$

da Riemann-Summen gegen das Riemann-Integral konvergieren.

10.2. Verteilungskonvergenz von Punktprozessen

Sei π ein Punktprozess (= ein zufälliges Zählmaß) auf E und seien $B_1, \dots, B_k \subset E$ kompakte Mengen. Dann ist $(\pi(B_1), \dots, \pi(B_k))$ ein k -dimensionaler Zufallsvektor mit Werten in \mathbb{N}_0^k . Vektoren dieser Art heißen auch **endlich-dimensionale Verteilungen** von π .

Wir definieren nun die Verteilungskonvergenz von Punktprozessen.

Definition 10.2.1. Seien π, π_1, π_2, \dots Punktprozesse auf E . Wir sagen, dass π_n gegen π in Verteilung konvergiert, falls für alle relativ kompakten Borel-Mengen B_1, \dots, B_k mit $\pi(\partial B_i) = 0$ f.s. gilt, dass

$$(\pi_n(B_1), \dots, \pi_n(B_k)) \xrightarrow[n \rightarrow \infty]{d} (\pi(B_1), \dots, \pi(B_k)).$$

Mit anderen Worten gilt für alle $m_1, \dots, m_k \in \mathbb{N}_0$:

$$\lim_{n \rightarrow \infty} \mathbb{P}[\pi_n(B_1) = m_1, \dots, \pi_n(B_k) = m_k] = \mathbb{P}[\pi(B_1) = m_1, \dots, \pi(B_k) = m_k].$$

Bezeichnung: $\pi_n \xrightarrow[n \rightarrow \infty]{d} \pi$.

Die Bedingung in der obigen Definition ist nicht leicht zu überprüfen. Es ist viel angenehmer, die folgende Charakterisierung der Verteilungskonvergenz von Punktprozessen zu benutzen.

Satz 10.2.2. Seien π_1, π_2, \dots und π Punktprozesse auf E , dann sind die folgenden Aussagen äquivalent:

(1) $\pi_n \xrightarrow[n \rightarrow \infty]{d} \pi$.

(2) Für alle $f \in C_c^+(E)$ gilt $\sum_{x \in \pi_n} f(x) \xrightarrow[n \rightarrow \infty]{d} \sum_{x \in \pi} f(x)$.

(3) Für alle $f \in C_c^+(E)$ gilt $\lim_{n \rightarrow \infty} \pi_{\pi_n}(f) = \psi_\pi(f)$, d.h.

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} \right] = \mathbb{E} \left[\exp \left\{ - \sum_{x \in \pi} f(x) \right\} \right].$$

BEWEIS. Weggelassen. □

Bemerkung 10.2.3. Um die Verteilungskonvergenz der Punktprozesse zu zeigen, reicht es also die Konvergenz der Laplace-Funktionale für jede Testfunktion $f \in C_c^+(E)$ nachzuweisen.

Im Folgenden werden wir eine Reihe von Beispielen der Verteilungskonvergenz von Punktprozessen betrachten.

10.3. Bernoulli-Experimente mit kleiner Erfolgswahrscheinlichkeit

Wir betrachten eine Serie aus unabhängigen Bernoulli-Experimenten mit einer sehr kleinen Erfolgswahrscheinlichkeit p . Die Erfolge in einer solchen Serie sind dementsprechend sehr selten, denn die mittlere Wartezeit auf den ersten Erfolg ist $1/p$. Wir interessieren uns für den Punktprozess der Zeitpunkte T_1, T_2, \dots , zu denen man Erfolge beobachtet. Der nächste Satz behauptet, dass der Punktprozess, der aus den Punkten pT_1, pT_2, \dots besteht, für kleines p durch einen Poisson-Punktprozess mit Intensität 1 approximiert werden kann.

Satz 10.3.1. Für jedes $n \in \mathbb{N}$ sei eine Folge ε_{ni} , $i \in \mathbb{Z}$, von unabhängigen und identisch verteilten Zufallsvariablen mit

$$\mathbb{P}[\varepsilon_{ni} = 1] = p_n, \quad \mathbb{P}[\varepsilon_{ni} = 0] = 1 - p_n$$

gegeben, wobei $\lim_{n \rightarrow \infty} np_n = \lambda \in [0, \infty)$. Dann gilt die folgende Verteilungskonvergenz von Punktprozessen auf \mathbb{R} :

$$\pi_n := \sum_{i \in \mathbb{Z}} \varepsilon_{ni} \delta_{\frac{i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\lambda dt).$$

BEWEIS. Wir werden Satz 10.2.2 benutzen. Für $f \in C_c^+(\mathbb{R})$ definieren wir die Laplace-Funktionale

$$\psi_n(f) = \mathbb{E} \left[\exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} \right], \quad \psi(f) = \mathbb{E} \left[\exp \left\{ - \sum_{x \in \pi} f(x) \right\} \right],$$

wobei π ein homogener Poisson-Punktprozess auf \mathbb{R} mit konstanter Intensität λ sei. Es reicht zu zeigen, dass $\lim_{n \rightarrow \infty} \psi_n(f) = \psi(f)$. Wegen der Unabhängigkeit der Familie ε_{ni} , $i \in \mathbb{Z}$, gilt:

$$\psi_n(f) = \mathbb{E} \left[\exp \left\{ - \sum_{i \in \mathbb{Z}} \varepsilon_{ni} f \left(\frac{i}{n} \right) \right\} \right] = \mathbb{E} \prod_{i \in \mathbb{Z}} e^{-\varepsilon_{ni} f(\frac{i}{n})} = \prod_{i \in \mathbb{Z}} \mathbb{E} e^{-\varepsilon_{ni} f(\frac{i}{n})} = \prod_{i \in \mathbb{Z}} (1 + p_n (e^{-f(\frac{i}{n})} - 1)).$$

Mit der Approximation $\log(1+x) = x + o(x)$ (für $x \rightarrow 0$) ergibt sich

$$\log \psi_n(f) = \sum_{i \in \mathbb{Z}} \log(1 + p_n (e^{-f(\frac{i}{n})} - 1)) = \sum_{i \in \mathbb{Z}} p_n (e^{-f(\frac{i}{n})} - 1) + R_n,$$

wobei R_n ein Restterm ist, für den wir später zeigen werden, dass $\lim_{n \rightarrow \infty} R_n = 0$. Es folgt, dass

$$\lim_{n \rightarrow \infty} \log \psi_n(f) = \lim_{n \rightarrow \infty} (np_n) \cdot \frac{1}{n} \sum_{i \in \mathbb{Z}} (e^{-f(\frac{i}{n})} - 1) = \lambda \int_{\mathbb{R}} (e^{-f(t)} - 1) dt,$$

wobei letzteres einfach die Definition des Riemann-Integrals ist. Zusammenfassend gilt also

$$\lim_{n \rightarrow \infty} \psi_n(f) = \exp \left\{ -\lambda \int_{\mathbb{R}} (1 - e^{-f(t)}) dt \right\},$$

was dem Laplace-Funktional eines Poisson-Punktprozesses mit Intensität λ entspricht. Mit Satz 10.2.2 folgt, dass

$$\pi_n := \sum_{i \in \mathbb{Z}} \varepsilon_{ni} \delta_{\frac{i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\lambda dt).$$

Der Restterm R_n kann wie folgt abgeschätzt werden. Es gilt

$$R_n = \sum_{i \in \mathbb{Z}} \left(\log(1 + p_n (e^{-f(\frac{i}{n})} - 1)) - p_n (e^{-f(\frac{i}{n})} - 1) \right) \leq \frac{1}{2} \sum_{i \in \mathbb{Z}} p_n^2 (e^{-f(\frac{i}{n})} - 1)^2,$$

wobei wir die Ungleichung $|\log(1+x) - x| \leq x^2/2$ für $x > 0$ benutzt haben. Die Funktion f hat einen kompakten Träger, also verschwindet sie außerhalb eines Intervalls $[-A, A]$. Es

folgt, dass höchstens $2An$ Werte von $f(i/n)$ (und somit höchstens $2An$ Summanden in der obigen Summe) ungleich 0 sind. Somit ergibt sich

$$|R_n| \leq 2AnMp_n^2 = o(1),$$

wobei M das Supremum von $(e^{-f} - 1)^2$ ist und wir benutzt haben, dass $\lim_{n \rightarrow \infty} np_n^2 = 0$. \square

Als Anwendung des obigen Satzes werden wir nun zeigen, dass die Zeitpunkte, zu denen eine Folge von u.i.v. Zufallsvariablen einen sehr hohen Schwellenwert überschreitet, nach einer Normierung gegen einen Poisson–Punktprozess konvergieren.

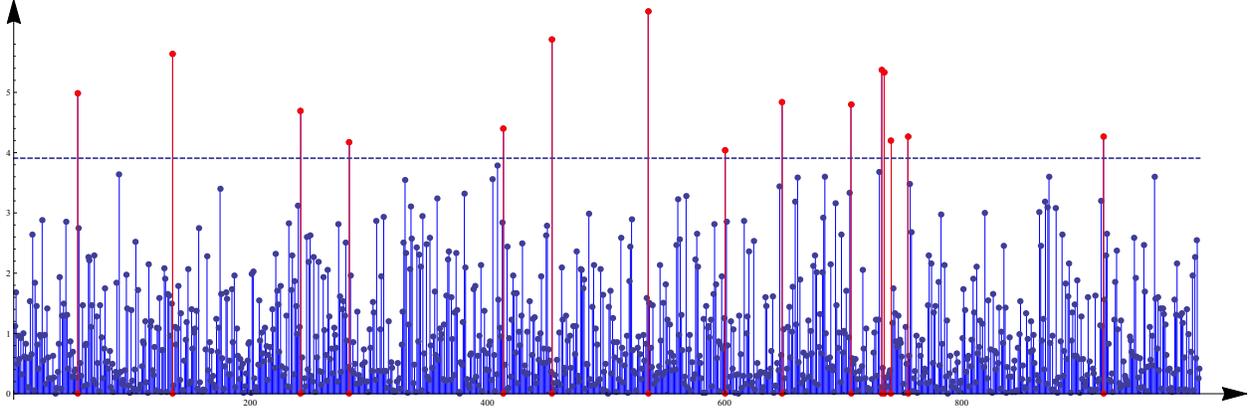


ABBILDUNG 1. Überschreitungen eines hohen Schwellenwerts in einer Folge von u.i.v. Zufallsvariablen

Satz 10.3.2. Seien X_i , $i \in \mathbb{Z}$, u.i.v. Zufallsvariablen mit Verteilungsfunktion F . Sei u_n eine Folge mit $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \lambda \in [0, \infty)$. Dann gilt die folgende Verteilungskonvergenz von Punktprozessen auf \mathbb{R} :

$$\sum_{i \in \mathbb{Z}: X_i > u_n} \delta_{\frac{i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\lambda dt).$$

BEWEIS. Betrachte die Zufallsvariablen $\varepsilon_{ni} := \mathbb{1}_{X_i > u_n}$, $i \in \mathbb{Z}$. Diese nehmen Werte 0 und 1 an, und es sei $p_n := \mathbb{P}[\varepsilon_{ni} = 1]$. Dann ist

$$\lim_{n \rightarrow \infty} np_n = \lim_{n \rightarrow \infty} n\mathbb{P}[\varepsilon_{ni} = 1] = \lim_{n \rightarrow \infty} n\bar{F}(u_n) = \lambda.$$

Damit sind die Bedingungen von Satz 10.3.1 erfüllt und die Behauptung folgt. \square

10.4. Konvergenz der Binomialpunktprozesse gegen die Poisson–Punktprozesse

In diesem Abschnitt beweisen wir einen allgemeinen Satz über die Konvergenz einer Folge von Binomialpunktprozessen gegen einen Poisson–Punktprozess.

Satz 10.4.1. Für alle $n \in \mathbb{N}$ seien $Y_{n1}, \dots, Y_{nn} : \Omega \rightarrow E$ u.i.v. Zufallselemente mit Werten in E und Verteilung μ_n . D.h. μ_n sei ein Wahrscheinlichkeitsmaß auf E mit $\mu_n(B) = \mathbb{P}[Y_{n1} \in B]$ für alle Borel-Mengen $B \subset E$. Es gelte außerdem für ein Radon-Maß μ auf E , dass

$$n\mu_n \xrightarrow[n \rightarrow \infty]{v} \mu.$$

Dann gilt die folgende Verteilungskonvergenz von Punktprozessen auf E :

$$\pi_n := \sum_{i=1}^n \delta_{Y_{ni}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\mu).$$

Beispiel 10.4.2. Seien Y_{n1}, \dots, Y_{nn} gleichverteilt auf dem Quadrat $[0, \sqrt{n}]^2$. Dann ist μ_n ein Wahrscheinlichkeitsmaß mit der Dichte $\frac{1}{n} \mathbb{1}_{[0, \sqrt{n}]^2}$ und man kann leicht zeigen, dass $n\mu_n$ vague gegen das Lebesgue-Maß auf der Viertelebene $[0, \infty)^2$ konvergiert. Es folgt, dass $\sum_{i=1}^n \delta_{Y_{ni}}$ gegen einen Poisson-Punktprozess mit Intensität 1 auf der Viertelebene konvergiert.

BEWEIS VON SATZ 10.4.1. Sei $f \in C_c^+(E)$, dann gilt

$$\psi_n(f) := \mathbb{E} \exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} = \mathbb{E} \exp \left\{ - \sum_{i=1}^n f(Y_{ni}) \right\} = \mathbb{E} \prod_{i=1}^n e^{-f(Y_{ni})}.$$

Wegen der Unabhängigkeit der Zufallsvariablen Y_{n1}, \dots, Y_{nn} kann man diesen Ausdruck wie folgt schreiben:

$$\psi_n(f) = \prod_{i=1}^n \mathbb{E} e^{-f(Y_{ni})} = \prod_{i=1}^n \int_E e^{-f(t)} \mu_n(dt) = \prod_{i=1}^n \left(1 - \frac{\int_E (1 - e^{-f(t)}) n\mu_n(dt)}{n} \right).$$

Wegen $n\mu_n(t) \xrightarrow[n \rightarrow \infty]{v} \mu(t)$ für $n \rightarrow \infty$ folgt schließlich:

$$\psi_n(f) = \left(1 - \frac{\int_E (1 - e^{-f(t)}) n\mu_n(dt)}{n} \right)^n \xrightarrow[n \rightarrow \infty]{} \exp \left\{ - \int_E (1 - e^{-f(t)}) \mu(dt) \right\},$$

was das Laplace-Funktional eines Poisson-Punktprozesses mit Intensität μ ist. Mit Satz 10.2.2 folgt, dass π_n in Verteilung gegen $\text{PPP}(\mu)$ konvergiert. \square

10.5. Konvergenz der extremen Ordnungsstatistiken: Gumbel-Fall

In diesem Abschnitt werden wir zeigen, dass die oberen extremen Ordnungsstatistiken einer u.i.v. Stichprobe aus dem Gumbel-Max-Anziehungsbereich durch einen Poisson-Punktprozess approximiert werden können. Die beiden anderen Max-Anziehungsbereiche werden im nächsten Abschnitt behandelt.

Satz 10.5.1. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen aus dem Max-Anziehungsbereich der Gumbel-Verteilung $\Lambda(x) = e^{-e^{-x}}$, d.h. es gebe Folgen $a_n > 0$ und $b_n \in \mathbb{R}$ mit

$$(10.5.1) \quad \frac{\max\{X_1, \dots, X_n\} - a_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} \Lambda.$$

Dann gilt die folgende Verteilungskonvergenz der Punktprozesse auf \mathbb{R} :

$$\pi_n := \sum_{i=1}^n \delta_{\frac{X_i - a_n}{b_n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(e^{-t} dt).$$

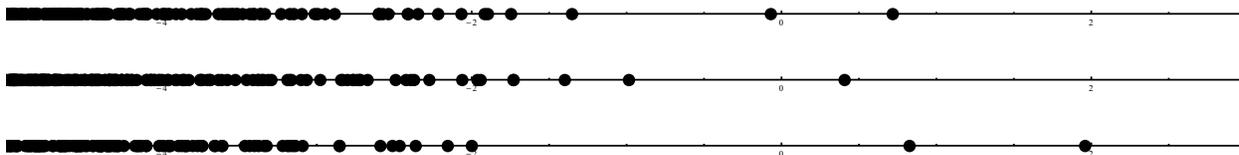


ABBILDUNG 2. Drei unabhängige Realisierungen des Poisson-Punktprozesses mit Intensität e^{-t} auf \mathbb{R} . Das Bild zeigt das Intervall $[-5, 3]$.

Bemerkung 10.5.2. Sei π der Poisson-Punktprozess mit Intensität e^{-t} auf \mathbb{R} . Für jedes $x \in \mathbb{R}$ ist die Anzahl der Punkte von π im Intervall (x, ∞) fast sicher endlich, denn es gilt

$$\pi((x, \infty)) \sim \text{Poi}\left(\int_x^\infty e^{-t} dt\right) = \text{Poi}(e^{-x}).$$

Insbesondere ist die erwartete Anzahl der Punkte auf der positiven Halbachse gleich 1. Auf der anderen Seite, ist die Anzahl der Punkte im Intervall $(-\infty, x)$ fast sicher unendlich, denn

$$\pi((-\infty, x)) \sim \text{Poi}\left(\int_{-\infty}^x e^{-t} dt\right) = \text{Poi}(\infty).$$

Wir können also die Punkte von π absteigend anordnen: $Y_1 > Y_2 > \dots$ und es gilt $\lim_{n \rightarrow \infty} Y_n = -\infty$. Wie ist nun Y_1 verteilt? Der obige Satz lässt vermuten, dass Y_1 eine Gumbel-Verteilung haben sollte. Das ist in der Tat richtig, denn

$$\mathbb{P}[Y_1 \leq x] = \mathbb{P}[\pi((x, \infty)) = 0] = e^{-e^{-x}},$$

da $\pi((x, \infty)) \sim \text{Poi}(e^{-x})$.

Aufgabe 10.5.3. Bestimmen Sie die Verteilung von Y_k , $k \in \mathbb{N}$.

Aufgabe 10.5.4. Bestimmen Sie die gemeinsame Dichte des Vektors (Y_1, \dots, Y_k) .

BEWEISIDEE VON SATZ 10.5.1. Wir wissen, dass die Konvergenz in (10.5.1) genau dann gilt, wenn

$$\lim_{n \rightarrow \infty} n\mathbb{P}[X_1 > a_n + tb_n] = e^{-t}$$

für alle $t \in \mathbb{R}$. Wir definieren $Y_{ni} := \frac{X_i - a_n}{b_n}$ für $i = 1, \dots, n$ und $\mu_n(B) := \mathbb{P}[Y_{n1} \in B]$ für $B \in \mathcal{B}(\mathbb{R})$. Dann gilt für Mengen B der Form $[t, \infty)$:

$$n\mu_n([t, \infty)) = n\mathbb{P}[Y_{n1} \geq t] = n\mathbb{P}\left[\frac{X_1 - a_n}{b_n} \geq t\right] \xrightarrow[n \rightarrow \infty]{} e^{-t}.$$

Betrachten wir als nächstes Mengen B der Form $[t_1, t_2)$:

$$n\mu_n([t_1, t_2)) = n \cdot \mu_n([t_1, \infty)) - n \cdot \mu_n([t_2, \infty)) \xrightarrow[n \rightarrow \infty]{} e^{-t_1} - e^{-t_2} = \int_{t_1}^{t_2} e^{-t} dt = \mu([t_1, t_2)).$$

Dieser Beweislogik folgend lässt sich die Konvergenz auch für beliebige relativ kompakte Borel-Mengen $B \subset \mathbb{R}$ zeigen:

$$\lim_{n \rightarrow \infty} n\mu_n(B) = \mu(B),$$

worauf wir hier verzichten wollen. Mit Satz 10.4.1 folgt die Behauptung. \square

Beispiel 10.5.5. Seien X_1, X_2, \dots unabhängige und identisch exponentialverteilte Zufallsvariablen mit Parameter 1, d.h. $\mathbb{P}[X_i > t] = e^{-t}$ für $t > 0$. Dann gilt nach Satz 1.3.2

$$\max\{X_1, \dots, X_n\} - \log n \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

Mit Satz 10.5.1 folgt nun

$$\sum_{i=1}^n \delta_{X_i - \log n} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(e^{-t} dt).$$

Beispiel 10.5.6. Seien X_1, X_2, \dots unabhängige und identisch standardnormalverteilte Zufallsvariablen. Wir haben bereits gezeigt, dass

$$\sqrt{2 \log n} (\max\{X_1, \dots, X_n\} - a_n) \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}$$

mit $a_n = \sqrt{2 \log n} - \frac{\frac{1}{2} \log \log n + \log 2\sqrt{\pi}}{\sqrt{2 \log n}}$. Es folgt mit Satz 10.5.1, dass

$$\sum_{i=1}^n \delta_{\sqrt{2 \log n} (X_i - a_n)} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(e^{-t} dt).$$

Im nächsten Satz geben wir eine explizite Konstruktion des Poisson-Punktprozesses $\pi \sim \text{PPP}(e^{-t} dt)$ als eine Transformation des homogenen Poisson-Punktprozesses $\sum_{n=1}^{\infty} \delta_{P_n}$ mit Intensität 1 an.

Proposition 10.5.7. Es sei $\sum_{n=1}^{\infty} \delta_{P_n}$ ein homogener Poisson-Punktprozess mit Intensität 1 auf $(0, \infty)$. Dann gilt

$$\sum_{n=1}^{\infty} \delta_{-\log P_n} \sim \text{PPP}(e^{-t} dt).$$

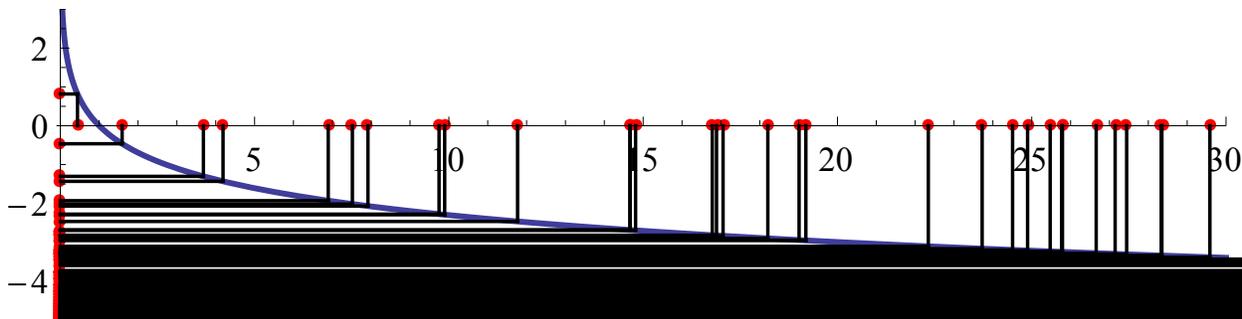


ABBILDUNG 3. Poisson–Punktprozess mit Intensität 1 (horizontale Achse) wird mit der Transformation $T(x) = -\log x$ auf den Poisson–Punktprozess mit Intensität e^{-t} (vertikale Achse) transformiert.

BEWEIS. Wir benutzen den Transformationssatz für Poisson–Punktprozesse für die Abbildung $T : (0, \infty) \rightarrow \mathbb{R}$, $T(x) = -\log x$. Sei ν das Lebesgue–Maß auf $(0, \infty)$. Mit dem Transformationssatz für Poisson–Punktprozesse folgt:

$$\sum_{n=1}^{\infty} \delta_{T(P_n)} \sim \text{PPP}(T\nu).$$

Wir müssen noch nachweisen, dass die Dichte des Bildmaßes $T\nu$ gleich e^{-t} ist. Es gilt

$$(T\nu)((a, b)) = \nu(T^{-1}((a, b))) = \nu((e^{-b}, e^{-a})) = e^{-a} - e^{-b} = \int_a^b e^{-t} dt.$$

Deshalb folgt die Behauptung. □

Aus der Konvergenz der Punktprozesse in Satz 10.5.1 folgt die Konvergenz der extremen Ordnungsstatistiken. Wir werden hier keinen exakten Beweis geben. Die Idee besteht darin, dass man die extremen Ordnungsstatistiken als ein stetiges Funktional des Punktprozesses ansehen kann. Nach dem Satz über die stetige Abbildung folgt aus der Verteilungskonvergenz der Punktprozesse die Verteilungskonvergenz der extremen Ordnungsstatistiken. Wir erinnern an die Notation $M_n^{(k)} = X_{n-k+1:n}$, $k = 1, \dots, n$.

Korollar 10.5.8. Unter Voraussetzungen von Satz 10.5.1 gilt für alle $k \in \mathbb{N}$

$$\left(\frac{M_n^{(1)} - a_n}{b_n}, \dots, \frac{M_n^{(k)} - a_n}{b_n} \right) \xrightarrow[n \rightarrow \infty]{d} (Y_1, \dots, Y_k),$$

wobei $Y_1 > Y_2 > \dots$ die Punkte des Poisson–Punktprozesses π mit Intensität e^{-t} auf \mathbb{R} seien.

10.6. Konvergenz der extremen Ordnungsstatistiken: Fréchet und Weibull–Fall

In Satz 10.5.1 haben wir Zufallsvariablen aus dem Max–Anziehungsbereich der Gumbel–Verteilung betrachtet. Nun formulieren wir einen ähnlichen Satz für den Max–Anziehungsbereich der Fréchet–Verteilung.

Satz 10.6.1. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen aus dem Max–Anziehungsbereich der Fréchet–Verteilung Φ_α , $\alpha > 0$, d.h. es gebe eine Folge $b_n > 0$ mit

$$(10.6.1) \quad \frac{\max\{X_1, \dots, X_n\}}{b_n} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha.$$

Dann gilt die folgende Konvergenz der Punktprozesse auf $(0, \infty)$:

$$\sum_{i=1}^n \delta_{\frac{X_i}{b_n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{\alpha}{t^{\alpha+1}} dt \right).$$

Dabei bedeutet die Notation \sum' , dass alle Punkte X_i mit $X_i \leq 0$ aus der Summe per Konvention ausgeschlossen werden.

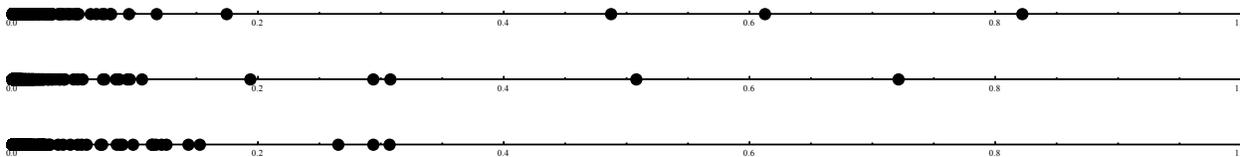


ABBILDUNG 4. Drei unabhängige Realisierungen des Poisson–Punktprozesses mit Intensität $t^{-2}dt$ auf $(0, \infty)$. Das Bild zeigt das Intervall $[0, 1]$.

Bemerkung 10.6.2. Sei π der Poisson–Punktprozess mit Intensität $\alpha t^{-(\alpha+1)}$ auf $(0, \infty)$. Für jedes $x > 0$ ist die Anzahl der Punkte von π im Intervall (x, ∞) fast sicher endlich, denn es gilt

$$\pi((x, \infty)) \sim \text{Poi} \left(\int_x^\infty \alpha t^{-(\alpha+1)} dt \right) = \text{Poi}(x^{-\alpha}).$$

Auf der anderen Seite ist die Anzahl der Punkte im Intervall $(0, x)$ fast sicher unendlich, denn

$$\pi((0, x)) \sim \text{Poi} \left(\int_0^x \alpha t^{-(\alpha+1)} dt \right) = \text{Poi}(\infty).$$

Aufgabe 10.6.3. Es sei $Y_1 > Y_2 > \dots > 0$ die absteigende Anordnung der Punkte von π . Zeigen Sie, dass Y_1 Fréchet–verteilt mit Parameter α ist. Bestimmen Sie die Verteilung von Y_k , $k \in \mathbb{N}$, und die gemeinsame Verteilung von (Y_1, \dots, Y_k) .

BEWEISIDEE VON SATZ 10.6.1. Bedingung (10.6.1) gilt genau dann, wenn

$$\lim_{n \rightarrow \infty} n \mathbb{P}[X_1 \geq b_n t] = \frac{1}{t^\alpha}$$

für alle $t > 0$. Wir definieren die Zufallsvariablen $Y_{ni} = \frac{X_i}{b_n} \mathbb{1}_{X_i > 0}$. Es sei μ_n die Verteilung von Y_{ni} . Wir betrachten zuerst Mengen der Form $[t, \infty)$:

$$n\mu_n([t, \infty)) = n\mathbb{P}[Y_{ni} \geq t] = n\mathbb{P}[X_1 \geq tb_n] \xrightarrow{n \rightarrow \infty} \frac{1}{t^\alpha}.$$

Betrachten wir nun Mengen der Form $[t_1, t_2)$ mit $0 < t_1 < t_2$:

$$n\mu_n([t_1, t_2)) = n\mu_n([t_1, \infty)) - n\mu_n([t_2, \infty)) \xrightarrow{n \rightarrow \infty} \frac{1}{t_1^\alpha} - \frac{1}{t_2^\alpha} = \int_{t_1}^{t_2} \frac{\alpha}{t^{\alpha+1}} dt.$$

Dieser Beweislogik folgend lässt sich die Konvergenz auch für beliebige relativ kompakte Borel-Mengen $B \subset (0, \infty)$ zeigen, worauf wir hier verzichten wollen. Es folgt:

$$\lim_{n \rightarrow \infty} n\mu_n(B) = \int_B \frac{\alpha}{t^{\alpha+1}} dt.$$

Damit ist Satz 10.4.1 anwendbar und die Behauptung ist bewiesen. \square

Beispiel 10.6.4. Seien X_1, X_2, \dots u.i.v. Pareto-verteilt mit Parameter $\alpha > 0$, d.h. $\mathbb{P}[X_i > t] = \frac{1}{t^\alpha}$ für alle $t > 1$. Wir haben in Satz 1.3.5 gezeigt, dass Pareto-verteilte Zufallsvariablen im Max-Anziehungsbereich der Fréchet-Verteilung liegen, nämlich

$$\frac{\max\{X_1, \dots, X_n\}}{n^{1/\alpha}} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha.$$

Mit Satz 10.6.1 gilt dann auf $(0, \infty)$:

$$\sum_{i=1}^n \delta_{\frac{X_i}{n^{1/\alpha}}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{\alpha}{t^{\alpha+1}} dt \right).$$

Beispiel 10.6.5. Es seien X_1, X_2, \dots Cauchy-verteilt mit Dichte $\frac{1}{\pi(1+t^2)}$, $t \in \mathbb{R}$. Wir wissen, dass

$$\frac{\max\{X_1, \dots, X_n\}}{n} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha.$$

Somit folgt aus Satz 10.6.1, dass auf $(0, \infty)$

$$\sum_{\substack{i=1, \dots, n \\ X_i > 0}} \delta_{\frac{X_i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{\alpha}{t^{\alpha+1}} dt \right).$$

Aus Symmetriegründen folgt aber auch, dass auf $(-\infty, 0)$

$$\sum_{\substack{i=1, \dots, n \\ X_i < 0}} \delta_{\frac{X_i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{\alpha}{(-t)^{\alpha+1}} dt \right).$$

Man kann sogar beide Fälle vereinigen und zeigen, dass auf $E = \mathbb{R} \setminus \{0\}$ gilt

$$\sum_{i=1}^n \delta_{\frac{X_i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{\alpha}{|t|^{\alpha+1}} dt \right).$$

Es sei bemerkt, dass wir den Punkt 0 ausschließen müssen, denn die Punkte des Poisson-Punktprozesses auf der rechten Seite häufen sich an der Stelle 0.

Zum Schluss betrachten wir noch den Max–Anziehungsbereich der Weibull–Verteilung Ψ_α .

Satz 10.6.6. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen aus dem Max–Anziehungsbereich der Weibull–Verteilung Ψ_α , $\alpha > 0$, d.h. der rechte Endpunkt x^* sei endlich und es gebe eine Folge $b_n > 0$ mit

$$\frac{\max\{X_1, \dots, X_n\} - x^*}{b_n} \xrightarrow[n \rightarrow \infty]{d} \Psi_\alpha.$$

Dann gilt die folgende Konvergenz von Punktprozessen auf $(-\infty, 0)$:

$$\sum_{i=1}^n \delta_{\frac{X_i - x^*}{b_n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\alpha(-t)^{\alpha-1} dt).$$

BEWEIS. Weggelassen. □

Beispiel 10.6.7. Seien X_1, X_2, \dots unabhängig und identisch gleichverteilt auf $[0, 1]$. Es gilt:

$$n(\max\{X_1, \dots, X_n\} - 1) \xrightarrow[n \rightarrow \infty]{d} e^x = \Psi_1(x), \quad x < 0.$$

Dann folgt aus Satz 10.6.6, dass

$$\sum_{i=1}^n \delta_{n(X_i - 1)} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(dt) \text{ auf } (-\infty, 0).$$

Aus Symmetriegründen gilt auch

$$\sum_{i=1}^n \delta_{nX_i} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(dt) \text{ auf } (0, \infty).$$

10.7. Konvergenz der Rekordzeiten gegen einen Poisson–Punktprozess

Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion F . Wir benutzen die Notation $M_n = \max\{X_1, \dots, X_n\}$ für das Maximum der ersten n Zufallsvariablen bzw. $\xi_j = \mathbb{1}_{X_j > M_{j-1}}$, $j = 1, 2, \dots$ für die Indikatorvariable des Ereignisses, dass zum Zeitpunkt j ein neuer Rekord aufgestellt wird. Die ebenfalls bereits behandelten Rekordzeiten $L(1), L(2), \dots$ werden weiterhin wie folgt definiert: $L(1) = 1$ und

$$L(n+1) = \min\{j > L(n) : \xi_j = 1\}, \quad n = 1, 2, \dots$$

Der nächste Satz behauptet, dass die Folge der Rekordzeiten wie ein Poisson–Punktprozess aussieht, wenn man sie aus einer sehr großen Entfernung betrachtet.

Satz 10.7.1. Es gilt

$$\pi_n := \sum_{i=1}^{\infty} \delta_{\frac{L(i)}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}\left(\frac{dt}{t}\right) \text{ auf } (0, \infty).$$

BEWEIS. Wir zeigen die Konvergenz der Laplace-Funktionale. Sei $f \in C_c^+(0, \infty)$. Es gilt

$$\psi_n(f) := \mathbb{E} \exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} = \mathbb{E} \exp \left\{ - \sum_{i=1}^{\infty} f \left(\frac{L(i)}{n} \right) \right\} = \mathbb{E} \exp \left\{ - \sum_{j=1}^{\infty} \xi_j f \left(\frac{j}{n} \right) \right\},$$

weil ξ_j immer dann 0 ist, wenn kein Rekord an Stelle j vorliegt. Man kann obiges auch wie folgt schreiben und wegen der Unabhängigkeit der ξ_j nach Satz von Rényi umformen:

$$\psi_n(f) = \mathbb{E} \prod_{j=1}^{\infty} \exp \left\{ - \xi_j f \left(\frac{j}{n} \right) \right\} = \prod_{j=1}^{\infty} \mathbb{E} \exp \left\{ - \xi_j f \left(\frac{j}{n} \right) \right\} = \prod_{j=1}^{\infty} \left(1 - \frac{1}{j} + \frac{1}{j} e^{-f(\frac{j}{n})} \right),$$

wobei sich die letzte Gleichheit ergibt, da ebenfalls mit dem Satz von Rényi $\mathbb{P}[\xi_j = 1] = 1 - \mathbb{P}[\xi_j = 0] = \frac{1}{j}$ gilt. Es folgt:

$$\log \psi_n(f) = \sum_{j=1}^{\infty} \log \left(1 + \frac{1}{j} \left(e^{-f(\frac{j}{n})} - 1 \right) \right) = \sum_{j=1}^{\infty} \frac{1}{j} \left(e^{-f(\frac{j}{n})} - 1 \right) + R_n,$$

wobei hier R_n ein Restterm ist, der, wie wir später zeigen werden, gegen 0 geht, und wir die Taylorentwicklung $\log(1+x) = x + o(x)$ im Auge behalten. Definiert man

$$g \left(\frac{j}{n} \right) = \frac{1}{j/n} \left(e^{-f(\frac{j}{n})} - 1 \right)$$

und lässt n gegen unendlich gehen, so folgt:

$$\log \psi_n(f) = \frac{1}{n} \sum_{i=1}^{\infty} \frac{n}{j} \left(e^{-f(\frac{j}{n})} - 1 \right) + R_n \xrightarrow{n \rightarrow \infty} \int_0^{\infty} g(t) dt = \int_0^{\infty} (e^{-f(t)} - 1) \frac{dt}{t},$$

wobei wir hier die Konvergenz einer Riemann-Summe gegen das Riemann-Integral benutzt haben. Schließlich ist zu beobachten, dass

$$\lim_{n \rightarrow \infty} \psi_n(t) = \exp \left\{ - \int_0^{\infty} (e^{-f(t)} - 1) \frac{dt}{t} \right\},$$

was die Laplace-Transformierte eines Poisson-Punktprozesses auf $(0, \infty)$ mit Intensität $\frac{1}{t}$ ist. Zu zeigen ist nur noch, dass $\lim_{n \rightarrow \infty} R_n = 0$. Mit der Ungleichung $|\log(1+x) - x| \leq \frac{1}{2}x^2$ ergibt sich, dass

$$|R_n| \leq \frac{1}{2} \sum_{j=1}^{\infty} \frac{1}{j^2} \left(e^{-f(\frac{j}{n})} - 1 \right)^2 = \frac{1}{2n} \frac{1}{n} \sum_{j=1}^{\infty} g^2 \left(\frac{j}{n} \right) \sim \frac{1}{2n} \int_0^{\infty} g^2(t) dt \xrightarrow{n \rightarrow \infty} 0,$$

was die Behauptung beweist. □

10.8. Konvergenz gegen den Extremwertprozess

In den vorherigen Abschnitten haben wir Sätze über die Konvergenz der oberen Ordnungsstatistiken gegen Poisson-Punktprozesse formuliert. Man kann diese Sätze erweitern indem man die Positionen der Beobachtungen berücksichtigt, wo extrem große Werte auftreten.

Satz 10.8.1. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen aus dem Max-Anziehungsbereich der Gumbel-Verteilung $\Lambda(x) = e^{-e^{-x}}$, d.h. es gebe Folgen $a_n > 0$ und $b_n \in \mathbb{R}$ mit

$$(10.8.1) \quad \frac{\max\{X_1, \dots, X_n\} - a_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} \Lambda.$$

Dann gilt die folgende Verteilungskonvergenz der Punktprozesse auf $[0, 1] \times \mathbb{R}$:

$$\pi_n := \sum_{i=1}^n \delta_{\left(\frac{i}{n}, \frac{X_i - a_n}{b_n}\right)} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(ds \times e^{-t} dt).$$

BEWEIS. Weggelassen. □

Der Punktprozess auf der rechten Seite kann wie folgt konstruiert werden. Seien $Y_1 > Y_2 > \dots$ die Punkte des Poisson-Punktprozesses mit Intensität e^{-t} auf \mathbb{R} . Unabhängig davon seien U_1, U_2, \dots u.i.v. Zufallsvariablen, die gleichverteilt auf dem Intervall $[0, 1]$ sind. Dann gilt

$$\sum_{i=1}^{\infty} \delta_{(U_i, Y_i)} \sim \text{PPP}(ds \times e^{-t} dt).$$

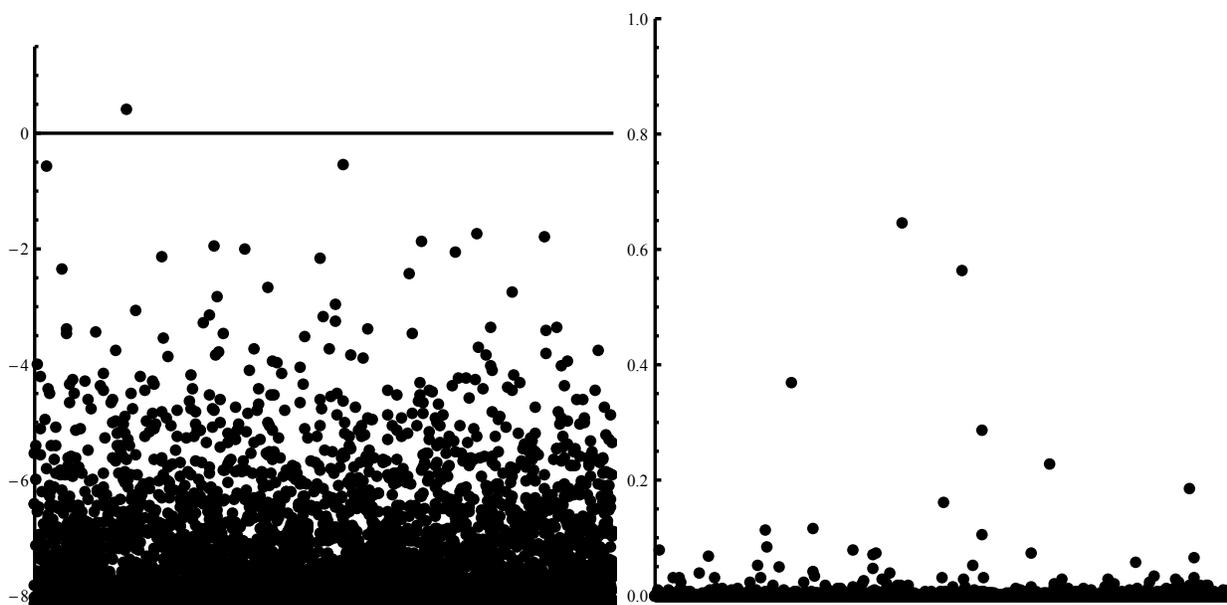


ABBILDUNG 5. Poisson-Punktprozesse aus Sätzen 10.8.1 und 10.8.4.

Der obige Satz macht den folgenden “funktionalen Grenzwertsatz” für den Maximumsprozess plausibel. Zuerst benötigen wir eine Definition.

Definition 10.8.2. Der stochastische Prozess $\{Z_t: t \in [0, 1]\}$ mit

$$Z_t = \max_{i \in \mathbb{N}: U_i \leq t} Y_i$$

heißt der **Gumbel–Extremwertprozess**.

Satz 10.8.3. Unter Voraussetzungen von Satz 10.8.1 konvergiert der Prozess

$$\{(M_{[nt]} - b_n)/a_n: t \in [0, 1]\}$$

gegen den Gumbel–Extremwertprozess im Sinne der endlich–dimensionalen Verteilungen.

D.h. für alle $0 < t_1 < \dots < t_k \leq 1$ gilt

$$\left(\frac{M_{[nt_1]} - b_n}{a_n}, \dots, \frac{M_{[nt_k]} - b_n}{a_n} \right) \xrightarrow[n \rightarrow \infty]{d} (Z_{t_1}, \dots, Z_{t_k}).$$

BEWEIS. Weggelassen. □

Ähnliche Sätze kann man auch für die beiden anderen Max–Anziehungsbereiche formulieren. Wir betrachten hier nur den Fréchet–Fall.

Satz 10.8.4. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen aus dem Max–Anziehungsbereich der Fréchet–Verteilung Φ_α , $\alpha > 0$, d.h. es gebe eine Folge $b_n > 0$ mit

$$(10.8.2) \quad \frac{\max\{X_1, \dots, X_n\}}{b_n} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha.$$

Dann gilt die folgende Konvergenz der Punktprozesse auf $[0, 1] \times (0, \infty)$:

$$\sum_{i=1}^n \delta'_{\left(\frac{i}{n}, \frac{X_i}{b_n}\right)} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(ds \times \frac{\alpha}{t^{\alpha+1}} dt \right).$$

Dabei bedeutet die Notation \sum' , dass alle Punkte X_i mit $X_i \leq 0$ aus der Summe per Konvention ausgeschlossen werden.

BEWEIS. Weggelassen. □

10.9. Allgemeiner Extremwertprozess

Seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit Verteilungsfunktion F . Sei $M_n = \max\{X_1, \dots, X_n\}$. Es gilt $M_1 \leq M_2 \leq M_3 \leq \dots$. Im nächsten Satz beschreiben wir die endlich dimensionalen Verteilungen des Prozesses M_1, M_2, \dots

Satz 10.9.1. Seien $t_1 < t_2 < \dots < t_k$, mit $t_i \in \mathbb{N}_0$ und $x_1, \dots, x_k \in \mathbb{R}$. Die Verteilung von $(M_{t_1}, \dots, M_{t_n})$ ist gegeben durch:

$$\begin{aligned} & \mathbb{P}[M_{t_1} \leq x_1, \dots, M_{t_k} \leq x_k] \\ &= F^{t_k - t_{k-1}}(x_k) F^{t_{k-1} - t_{k-2}}(\min\{x_{k-1}, x_k\}) \cdot \dots \cdot F^{t_1}(\min\{x_1, \dots, x_k\}). \end{aligned}$$

Bemerkung 10.9.2. Es gilt die Markov-Eigenschaft:

$$\mathbb{P}[M_{n+1} \leq u | M_1 = m_1, \dots, M_n = m_n] = \mathbb{P}[M_{n+1} \leq u | M_n = m_n].$$

Der nächste Satz zeigt, dass man den Maximumprozess $(M_t)_{t \in \mathbb{N}}$ in einen Extremwertprozess $(Z_t)_{t \geq 0}$ in stetiger Zeit einbetten kann. Es sei x_* bzw. x^* der linke bzw. der rechte Endpunkt der Verteilungsfunktion F .

Satz 10.9.3. Sei $\pi = \sum_{k=1}^{\infty} \delta_{(X_k, Y_k)} \sim \text{PPP}(\mu)$ auf $[0, \infty) \times (x_*, x^*]$ mit

$$\mu((a, b] \times (c, d]) = (b - a) \cdot (\log F(d) - \log F(c)).$$

Sei $Z_t = \sup\{Y_k : X_k \leq t\}$ mit $t \geq 0$. Dann gilt

$$(M_t)_{t \in \mathbb{N}} \stackrel{d}{=} (Z_t)_{t \in \mathbb{N}}.$$

BEWEIS. Seien $t_1 < t_2 < \dots < t_k$ mit $t_i \in \mathbb{N}$, dann gilt:

$$\mathbb{P}[Z_{t_1} \leq u_1, \dots, Z_{t_k} \leq u_k] = \mathbb{P}[\pi(A_j) = 0 \text{ für alle } j = 1, \dots, k] = \mathbb{P}[\pi(\cup_{j=1}^k A_j) = 0],$$

wobei hier

$$A_j = [t_{j-1}, t_j] \times (\min\{u_k, u_{k-1}, \dots, u_j\}, x^*].$$

Obiges lässt sich, da π ein Poisson-Punktprozess ist, zu Folgendem vereinfachen:

$$\mathbb{P}[Z_{t_1} \leq u_1, \dots, Z_{t_k} \leq u_k] = e^{-\mu(\cup_{j=1}^k A_j)} = \prod_{j=1}^k e^{-\mu(A_j)} = \prod_{j=1}^k e^{(t_j - t_{j-1}) \log F(\min\{u_k, u_{k-1}, \dots, u_j\})}.$$

Somit gilt

$$\mathbb{P}[Z_{t_1} \leq u_1, \dots, Z_{t_k} \leq u_k] = \mathbb{P}[M_{t_1} \leq u_1, \dots, M_{t_k} \leq u_k].$$

□

Beispiel 10.9.4. Für $F(x) = e^{-e^{-x}}$ stimmt der Prozess Z_t mit dem oben eingeführten Gumbel-Extremwertprozess überein.
