

WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER

EXTREMWERTTHEORIE UND POISSON-PROZESSE

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 und Sommersemester 2018) 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 Jürgen Kampf, 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.

7. Juli 2020

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
Einleitung	1
Kapitel 1. Extremwertverteilungen	2
1.1. Grundbegriffe	2
1.2. Extremwertverteilungen: Gumbel, Fréchet, Weibull	4
1.3. Extremwertverteilungen und deren Max-Anziehungsbereiche: Definition	9
1.4. Satz von Fisher-Tippett-Gnedenko	10
1.5. Jenkinson-von Mises-Darstellung	12
Kommentare	12
Kapitel 2. Cauchy-Funktionalgleichung	13
2.1. Additive und multiplikative Funktionen	13
2.2. Stetige Lösungen	14
2.3. Satz von Steinhaus	15
2.4. Messbare Lösungen	17
2.5. Hamel-Funktionen	18
Kommentare	19
Kapitel 3. Beweis des Satzes von Fisher-Tippett-Gnedenko	20
3.1. Eindeutigkeit der Normierungskonstanten	20
3.2. Max-stabile Verteilungen	22
3.3. Charakterisierung der max-stabilen Verteilungen	24
Kommentare	26
Kapitel 4. Regulär variierende Funktionen	27
4.1. Langsam variierende Funktionen	27
4.2. Regulär variierende Funktionen	29
4.3. Satz über die gleichmäßige Konvergenz	30
4.4. Darstellungssatz von Karamata	32
4.5. Abschätzungen für regulär variierende Funktionen	34
4.6. Charakterisierung der Grenzwertfunktionen	35
4.7. Asymptotische Äquivalenz	36
4.8. Integrale von regulär variierenden Funktionen	36
Kommentare	38
Kapitel 5. Max-Anziehungsbereiche	39

5.1. Max-Anziehungsbereich der Fréchet-Verteilung Φ_α	39
5.2. Max-Anziehungsbereich der Weibull-Verteilung Ψ_α	46
5.3. Max-Anziehungsbereich der Gumbel-Verteilung Λ	48
5.4. Beispiel: Normalverteilung	50
5.5. Von Mises-Bedingungen	54
Kommentare	59
Kapitel 6. Statistik der Extremwertverteilungen	60
6.1. Statistik der Blockmaxima: GEV-Verteilungen	60
6.2. Peaks over Threshold: Statistik der GP-Verteilungen	64
Kommentare	67
Kapitel 7. Ordnungsstatistiken	68
7.1. Allgemeine Eigenschaften der Ordnungsstatistiken	68
7.2. Extreme Ordnungsstatistiken	72
7.3. Ordnungsstatistiken der exponentialverteilten Zufallsvariablen	76
7.4. Ordnungsstatistiken der gleichverteilten Zufallsvariablen	80
7.5. Asymptotische Unabhängigkeit des Maximums und des Minimums	83
Kommentare	86
Kapitel 8. Rekorde	87
8.1. Satz von Rényi	87
8.2. Anzahl der Rekorde	93
8.3. Rekordzeiten	97
8.4. Zentrale Grenzwertsätze	102
8.5. Rekordwerte	105
8.6. Grenzwertverteilungen der Rekordwerte	111
8.7. Rekorde für diskrete Zufallsvariablen	114
Kommentare	115
Kapitel 9. Anhang: Markov-Prozesse	116
9.1. Markov-Ketten mit diskretem Zustandsraum	116
9.2. Markov-Ketten mit allgemeinem Zustandsraum	117
9.3. Markov-Prozesse mit stetiger Zeit	117
Kapitel 9. Poisson-Punktprozesse	118
9.1. Poisson-Verteilung	118
9.2. Beispiel zu Poisson-Prozessen	118
9.3. Definition von Zählmaßen und Punktprozessen	119
9.4. Definition der Poisson-Punktprozesse	122
9.5. Superpositionssatz	124
9.6. Abbildungssatz	126
9.7. Laplace-Funktionale	129
9.8. Färbungs- und Markierungssatz für Poisson-Punktprozesse	131
9.9. Simulation	135
9.10. Poisson-Linienprozess	135
Kommentare	135

Kapitel 10. Konvergenz von Punktprozessen	137
10.1. Vage Konvergenz	137
10.2. Verteilungskonvergenz von Punktprozessen	138
10.3. Bernoulli-Experimente mit kleiner Erfolgswahrscheinlichkeit	139
10.4. Konvergenz der Binomialpunktprozesse gegen die Poisson-Punktprozesse	142
10.5. Konvergenz der extremen Ordnungsstatistiken: Gumbel-Fall	143
10.6. Konvergenz der extremen Ordnungsstatistiken: Fréchet und Weibull-Fall	147
10.7. Konvergenz der Rekordzeiten gegen einen Poisson-Punktprozess	150
10.8. Konvergenz gegen den Extremwertprozess	151
10.9. Allgemeiner Extremwertprozess	153
Kommentare	154
Literaturverzeichnis	155

Einleitung

KAPITEL 1

Extremwertverteilungen

In diesem ersten Kapitel werden wir Extremwertverteilungen einführen.

1.1. Grundbegriffe

Wasserstände, Niederschlagsmengen, Log>Returns einer Aktie, Leistungen von Sportlern und andere Größen lassen sich unter Umständen durch unabhängige identisch verteilte (u.i.v.) Zufallsvariablen modellieren. Wir bezeichnen also mit X_1, X_2, \dots eine Folge von u.i.v. Zufallsvariablen.

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

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

Es sei bemerkt, dass alle Zufallsvariablen die gleiche Verteilungsfunktion haben, da sie als identisch verteilt vorausgesetzt werden. In der Extremwerttheorie ist es oft einfacher, die sogenannte Tailfunktion \bar{F} zu betrachten.

Definition 1.1.2. Die **Tailfunktion** von X_i ist definiert durch

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

Stellt man sich X_i als den Wasserstand während des i -ten Sturmes und t als die Deichhöhe vor, so ist $\bar{F}(t)$ die Wahrscheinlichkeit, dass der Deich während eines Sturmes überflutet wird.

Wir fassen die wichtigsten Eigenschaften der Verteilungsfunktion zusammen:

- F ist monoton nichtfallend, d.h. es gilt $F(s) \leq F(t)$ für alle $s \leq t$.
- $\lim_{t \rightarrow -\infty} F(t) = 0$ und $\lim_{t \rightarrow +\infty} F(t) = 1$.
- F ist rechtsstetig, d.h. es gilt $\lim_{t \downarrow t_0} F(t) = F(t_0)$ für alle $t \in \mathbb{R}$.

Für die Tailfunktion gilt dementsprechend

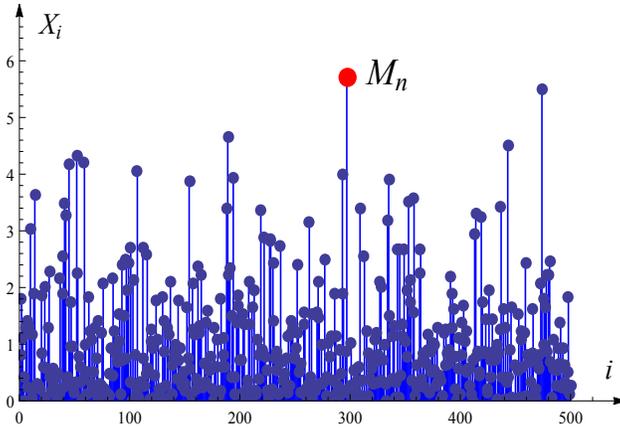


ABBILDUNG 1. Veranschaulichung von M_n .

- \bar{F} ist monoton nichtsteigend, d.h. es gilt $\bar{F}(s) \geq \bar{F}(t)$ für alle $s \leq t$.
- $\lim_{t \rightarrow -\infty} \bar{F}(t) = 1$ und $\lim_{t \rightarrow +\infty} \bar{F}(t) = 0$.
- \bar{F} ist rechtsstetig.

Bezeichnen X_1, X_2, \dots, X_n die Wasserstände während der n Stürme, die innerhalb eines Jahres beobachtet wurden, so interessieren wir uns für das Jahresmaximum. Dafür führen wir nun eine Notation ein:

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

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

Der nächste Satz, auf den wir im Folgenden sehr oft zurückgreifen werden, beschreibt die Verteilungsfunktion von M_n .

Satz 1.1.3. 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. Das Maximum M_n bleibt genau dann unterhalb von t , wenn alle Zufallsvariablen X_1, \dots, X_n unterhalb von t bleiben, d.h.

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

Indem wir zuerst die Unabhängigkeit der Zufallsvariablen X_1, \dots, X_n benutzen und dann die Verteilungsfunktion F einsetzen, erhalten wir

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

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

Aufgabe 1.1.4. Seien X_1, \dots, X_n u.i.v. Zufallsvariablen mit Verteilungsfunktion F .

- (a) 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.
- (b) Bestimmen Sie die gemeinsame Verteilungsfunktion von m_n und M_n , d.h. berechnen Sie die Wahrscheinlichkeit $\mathbb{P}[m_n \leq s, M_n \leq t]$.

Aufgabe 1.1.5. Es seien N, X_1, X_2, \dots unabhängige Zufallsvariablen. Die Zufallsvariable N sei Poisson-verteilt mit Parameter $\lambda > 0$. Die Zufallsvariablen X_1, X_2, \dots seien exponentialverteilt mit Parameter 1, d.h. $\mathbb{P}[X_i \geq t] = e^{-t}$ für $t > 0$. Bestimmen Sie die Verteilungsfunktion von

$$M = \max\{X_1, \dots, X_N\}.$$

Für $N = 0$ sei das Maximum als 0 definiert.

1.2. Extremwertverteilungen: Gumbel, Fréchet, Weibull

Wir werden uns für das Verhalten der Verteilungsfunktion von M_n für $n \rightarrow \infty$ interessieren. Als Vorbild soll uns der **zentrale Grenzwertsatz** dienen. Dieser besagt, dass für u.i.v. Zufallsvariablen X_1, X_2, \dots mit endlichem Erwartungswert μ und endlicher Varianz $\sigma^2 > 0$ die Verteilungskonvergenz

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

gilt, wobei $\mathcal{N}(0, 1)$ 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 $a_n \in \mathbb{R}$, $b_n > 0$ und eine Verteilungsfunktion G gibt, so dass für $n \rightarrow \infty$ die folgende Verteilungskonvergenz gilt:

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

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, wie wir später sehen werden, in vielen interessanten Fällen der Erwartungswert und die Varianz gar nicht existieren.

Eine Verteilung, die als Grenzwert in (1.2.1) auftreten kann, heißt **Extremwertverteilung**. Die Verteilungskonvergenz in (1.2.1) bedeutet, dass

$$(1.2.2) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n - a_n}{b_n} \leq t \right] = G(t)$$

für alle Stetigkeitspunkte t von G . Im Folgenden werden wir drei Beispiele von (1.2.2) Extremwertverteilungen (oder sogar Familien von Extremwertverteilungen) konstruieren. Später werden wir zeigen, dass es bis auf lineare Transformationen keine weiteren Extremwertverteilungen gibt.

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

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

Satz 1.2.1. Die Zufallsvariablen X_1, X_2, \dots seien unabhängig und exponentialverteilt mit Parameter 1, d.h.

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

Dann gilt

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

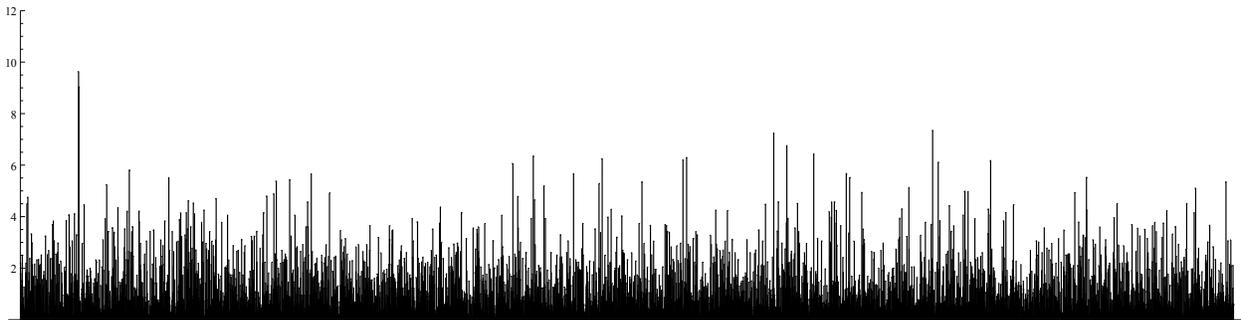


ABBILDUNG 2. Eine mit Parameter 1 exponentialverteilte Stichprobe vom Umfang $n = 5000$.

BEWEIS. Sei $t \in \mathbb{R}$ beliebig. Mit Satz 1.1.3 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 $M_n - \log n \xrightarrow[n \rightarrow \infty]{d} \Lambda$. □

Definition 1.2.2. Eine Zufallsvariable hat **Gumbel-Verteilung**, wenn ihre Verteilungsfunktion die folgende Gestalt hat:

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

Bemerkung 1.2.3. Gemäß Satz 1.2.1 liegt die Exponentialverteilung im Max-Anziehungsbereich der Gumbel-Verteilung:

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

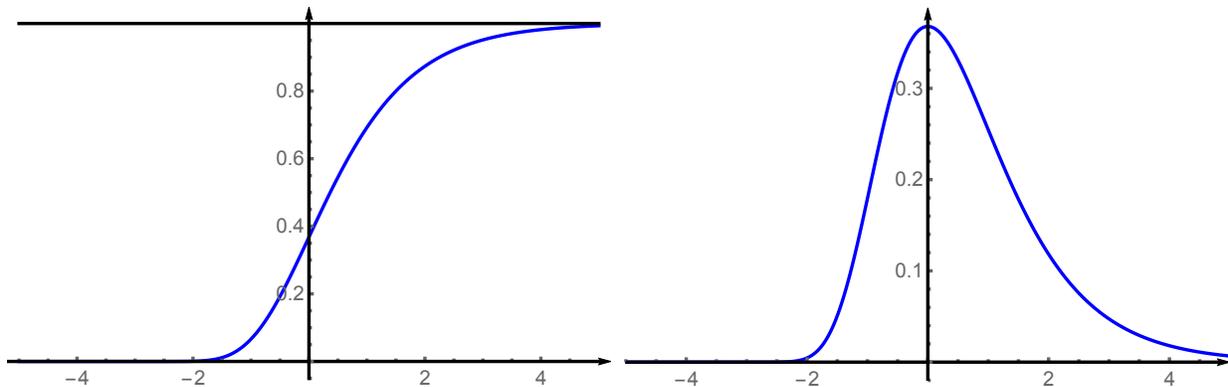


ABBILDUNG 3. Links: Verteilungsfunktion der Gumbel-Verteilung. Rechts: Dichte der Gumbel-Verteilung.

Die Gumbel-Verteilung ist somit eine Extremwertverteilung. Man kann Satz 1.2.1 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.

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

Definition 1.2.4. Eine Zufallsvariable heißt Fréchet-verteilt mit Parameter $\alpha > 0$, wenn ihre Verteilungsfunktion die folgende Gestalt hat:

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

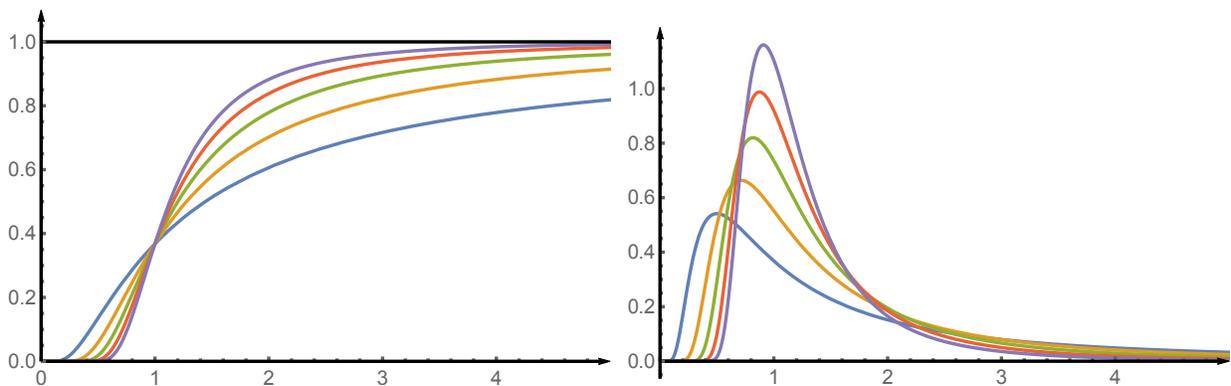


ABBILDUNG 4. Verteilungsfunktionen (links) und Dichten (rechts) der Fréchet-Verteilungen mit $\alpha = 1.0, 1.5, 2.0, 2.5, 3.0$.

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

Satz 1.2.5. 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 1.1.3 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 1.2.6. Man kann Satz 1.2.5 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 1.2.7. Eine Zufallsvariable heißt **Weibull-verteilt** mit Parameter $\alpha > 0$, wenn ihre Verteilungsfunktion die folgende Form hat:

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

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

Satz 1.2.8. Seien die Zufallsvariablen X_1, X_2, \dots unabhängig mit der Verteilungsfunktion

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

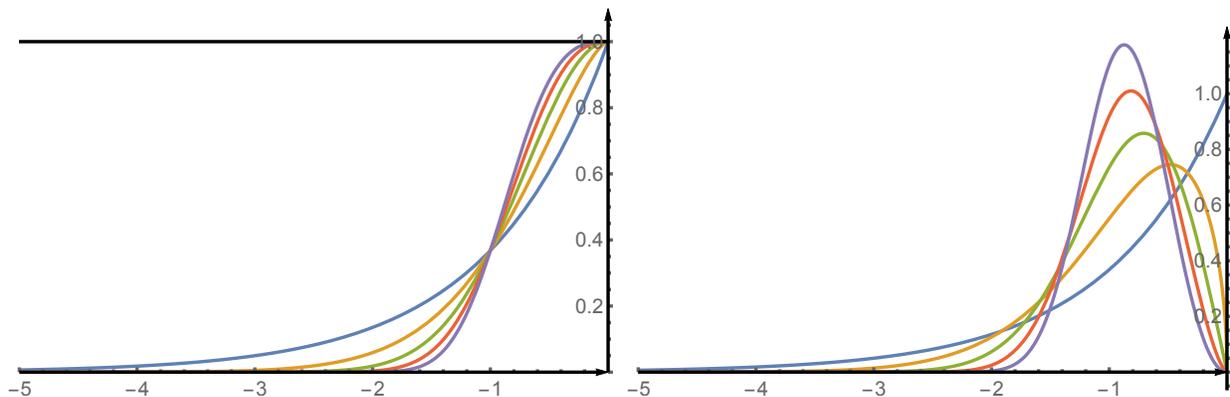


ABBILDUNG 5. Verteilungsfunktionen (links) und Dichten (rechts) der Weibull-Verteilungen mit $\alpha = 1.0, 1.5, 2.0, 2.5, 3.0$.

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

$$(1.2.3) \quad \lim_{n \rightarrow \infty} \mathbb{P}[n^{1/\alpha}M_n \leq t] = \begin{cases} e^{-(-t)^\alpha}, & t \leq 0, \\ 1, & t \geq 0. \end{cases}$$

BEWEIS. Sei $t \leq 0$ beliebig. Mit Satz 1.1.3 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 1.2.9. Man kann Satz 1.2.8 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 1.2.10. 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.

1.3. Extremwertverteilungen und deren Max-Anziehungsbereiche: Definition

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

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

Wenn (1.2.1) bzw. (1.3.1) 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.

Eine äquivalente Formulierung ist diese:

$$(1.3.1) \quad \lim_{n \rightarrow \infty} F^n(a_n + b_n t) = G(t)$$

für alle Stetigkeitspunkte t von G .

Bemerkung 1.3.1. Es gibt einen Spezialfall von (1.2.1) und (1.3.1), 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.

$$(1.3.2) \quad G(t) = \begin{cases} 0, & t < c, \\ 1, & t \geq c \end{cases}$$

gibt. Für jede Verteilungsfunktion F kann man durch die “falsche” Wahl der Konstanten a_n, b_n erreichen, dass (1.2.1) bzw. (1.3.1) mit einer degenerierten Verteilungsfunktion G gilt. Man kann zum Beispiel $a_n = 0$ und b_n derart schnell steigend wählen, dass M_n/b_n gegen 0 in Verteilung konvergiert (Übungsaufgabe). Deshalb werden wir im Folgenden die degenerierten Verteilungsfunktionen G der Form (1.3.2) aus unseren Definitionen ausschließen.

Definition 1.3.2. Der **Max-Anziehungsbereich** einer nichtdegenerierten Verteilungsfunktion G besteht aus allen Verteilungsfunktionen F , für die es zwei Folgen $a_n \in \mathbb{R}$ und $b_n > 0$ gibt, so dass

$$(1.3.3) \quad \lim_{n \rightarrow \infty} F^n(a_n + b_n t) = G(t)$$

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 1.3.3. 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 $a_n \in \mathbb{R}$ und $b_n > 0$ gibt, so dass (1.3.3) gilt.

Bemerkung 1.3.4. Später werden wir sehen, dass alle Extremwertverteilungsfunktionen stetig sind. Deshalb kann man Definition 1.3.2 vereinfachen, indem man die Einschränkung auf die Stetigkeitspunkte von t weglässt und stattdessen verlangt, dass (1.3.3) für alle $t \in \mathbb{R}$ gilt.

Bemerkung 1.3.5. Wir machen in diesem Buch 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”.

1.4. 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: $Z_1 \asymp Z_2$.

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: $F_1 \asymp F_2$.

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 \asymp eine Äquivalenzrelation ist, d. h.

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

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 $a_n \in \mathbb{R}, b_n > 0$ gibt, so dass

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

Daraus folgt, dass

$$\frac{\max\{cX_1, \dots, cX_n\} - (ca_n - db_n)}{b_n} \xrightarrow[n \rightarrow \infty]{d} cZ + d.$$

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 einer der drei Familien (1.4.1), (1.4.2), (1.4.3) an.

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*).

Kommentare

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. Im Folgenden 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 Messbarkeit¹ 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 einer Menge A mit $\lambda(A)$.

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

BEWEIS. SCHRITT 1. Ohne Einschränkung der Allgemeinheit können wir annehmen, dass $\lambda(A) < \infty$. 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

¹Messbarkeit wird hier immer im Sinne von Lebesgue verstanden.

$A_i - A_i \subset A - A$. Wegen $\lambda(A_i) \leq 1$ reicht es, den Satz von Steinhaus nur für Mengen mit endlichem Lebesgue-Maß zu beweisen.

SCHRITT 2. Sei also $\lambda(A) \neq 0, \infty$. Wir zeigen, dass es ein Intervall I mit $\lambda(A \cap I) > \frac{9}{10}\lambda(I)$ gibt. Somit wird das Intervall I zu mindestens 90% von A überdeckt. Das Lebesgue-Maß von A wird wie folgt definiert:

$$\lambda(A) = \inf \sum_{k=1}^{\infty} \lambda(I_k),$$

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

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

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

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

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

SCHRITT 3. Sei $\delta = \frac{1}{2}\lambda(I)$. Wir zeigen, dass $(-\delta, \delta) \subset A - A$. Sei also x eine reelle Zahl mit $|x| < \delta$. Wir wollen zeigen, dass $x \in A - A$. Dies ist aber äquivalent zu $A \cap (A + x) \neq \emptyset$ (Übungsaufgabe!), wobei $A + x = \{a + x : a \in A\}$ die um x verschobene Menge A bezeichnet. Es reicht also zu zeigen, dass

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

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

Aufgabe 2.3.3. Sei $A \subset \mathbb{R}$ eine messbare Menge mit $\lambda(A) > 0$. 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 $\lambda(A) > 0, \lambda(B) > 0$. 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} und die Menge der algebraischen Zahlen sind additive Untergruppen von \mathbb{R} . Auch $\{0\}$ und ganz \mathbb{R} sind additive Untergruppen.

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 A mit positivem Lebesgue-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 $(-\delta, \delta) \subset A - A \subset S - S$. Also ist $(-\delta, \delta) \subset S$. Sei nun $x \in \mathbb{R}$ beliebig. Dann gilt für hinreichend großes $n \in \mathbb{N}$, dass $\frac{x}{n} \in (-\delta, \delta) \subset S$. Aus der Darstellung $x = \frac{x}{n} + \dots + \frac{x}{n}$ und der Untergruppeneigenschaft folgt, dass $x \in S$. 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 $\lambda(S) = 0$ oder $S = \mathbb{R}$.

Beweis. Ist $\lambda(S) = 0$, so gibt es nichts zu beweisen. Sei also $\lambda(S) > 0$. Aus Satz 2.3.7 mit $A = S$ (was messbar ist) folgt, dass $S = \mathbb{R}$.

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 Lebesgue-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 $\lambda(A_n) > 0$. 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. Mit der Additivität und der Dreiecksungleichung gilt

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

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

SCHRITT 3. Wir zeigen, dass f an der Stelle 0 stetig ist. Wegen der Additivität gilt $f(0) = 0$. Es reicht also zu zeigen, dass

$$(2.4.1) \quad \lim_{m \rightarrow \infty} \sup_{|t| < \delta/m} |f(t)| = 0.$$

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 (2.4.1). 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.

Aufgabe 2.4.3.

- (a) Bestimmen Sie alle messbaren Funktionen $f : (0, \infty) \rightarrow \mathbb{R}$ mit $f(xy) = f(x) + f(y)$ für alle $x, y > 0$.
- (b) Bestimmen Sie alle messbaren Funktionen $f : \mathbb{R} \rightarrow (0, \infty)$ mit $f(x+y) = f(x)f(y)$ für alle $x, y \in \mathbb{R}$.

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 weglassen? 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.

Bekanntlich ist die Dimension von \mathbb{R} als Vektorraum über \mathbb{R} gleich 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 ein beliebiger Vektorraum über einem beliebigen Körper eine Basis. Das heißt, es gibt eine Menge B (die sogenannte **Hamel-Basis**) 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. Es sei bemerkt, dass die Hamel-Basis unendlich und sogar überabzählbar ist (sonst wäre \mathbb{R} abzählbar).

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

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

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

Additive Funktionen, die nicht die Form cx haben, heißen **Hamel-Funktionen**. Sie sind 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 Lebesgue-Maß (das folgt aus dem Beweis des Satzes von Ostrowski). Man kann zeigen, dass der Graph einer Hamel-Funktion f , d.h. die Menge $\{(x, f(x)) : x \in \mathbb{R}\}$, überall dicht in der Ebene \mathbb{R}^2 ist. Hamel-Funktionen sind wahre Monster, die uns in diesem Buch allerdings nie wieder begegnen werden. Wir verweisen den interessierten Leser auf die Bücher von H. Herrlich [14] und M. Kuczma [16].

Aufgabe 2.5.2. Zeigen Sie, dass jede Hamel-Basis gleichmächtig mit \mathbb{R} ist. Beweisen Sie weiterhin, dass die Menge der Hamel-Funktionen gleichmächtig mit der Menge aller Funktionen von \mathbb{R} nach \mathbb{R} ist.

Kommentare

Dieses Kapitel folgt im Wesentlichen dem Buch von Bingham, Goldie and Teugels [7].

Beweis des Satzes von Fisher-Tippett-Gnedenko

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

Satz 3.0.1 (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 und Z Zufallsvariablen, sowie $a_n \in \mathbb{R}$ und $b_n > 0$ Folgen von Normierungskonstanten mit

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

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 $\tilde{a}_n \in \mathbb{R}$ und $\tilde{b}_n > 0$ zwei weitere Folgen mit

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

Dann gilt auch

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

BEWEIS. Es gilt

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

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 $\tilde{a}_n \in \mathbb{R}$ und $\tilde{b}_n > 0$ zwei weitere Folgen mit

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

Dann gilt auch

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

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 $a_n \in \mathbb{R}, b_n \in \mathbb{R}$ sowie $\tilde{a}_n \in \mathbb{R}, \tilde{b}_n > 0$ Normierungsfolgen mit

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

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

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

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

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

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

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

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

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 $c_n \in \mathbb{R}$ und $d_n > 0$ gibt mit

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

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 $n \in \mathbb{N}$

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

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 $c_n \in \mathbb{R}$, $d_n > 0$, so dass $G^n(c_n + d_n t) = G(t)$. Es gilt also für alle $t \in \mathbb{R}$

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

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 $a_n \in \mathbb{R}$, $b_n > 0$, so dass

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

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

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

Indem wir die k -te Wurzel ziehen, erhalten wir

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

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 $c_k \in \mathbb{R}$ und $d_k > 0$ mit

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

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 $G^n(c_n + d_n t) = G(t)$ 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 $s > 0$ gilt:

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

Dabei muss s nicht notwendigerweise ganzzahlig sein.

BEWEIS. Wir bezeichnen mit $\lfloor t \rfloor$ die Abrundungsfunktion (oder die Gaußklammer) einer reellen Zahl t :

$$(3.2.4) \quad \lfloor t \rfloor = \max\{n \in \mathbb{Z} : n \leq t\}.$$

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

$$(3.2.5) \quad G^n(c_n + d_n t) = G(t).$$

Für beliebiges $s > 0$ folgt daraus, dass

$$G^{\lfloor ns \rfloor}(c_{\lfloor ns \rfloor} + d_{\lfloor ns \rfloor} t) = G(t).$$

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

$$(3.2.6) \quad G^n(c_{\lfloor ns \rfloor} + d_{\lfloor ns \rfloor} t) = (G^{\lfloor ns \rfloor}(c_{\lfloor ns \rfloor} + d_{\lfloor ns \rfloor} t))^{\frac{n}{\lfloor ns \rfloor}} = G^{\frac{n}{\lfloor ns \rfloor}}(t) \xrightarrow{n \rightarrow \infty} G^{1/s}(t).$$

Gleichzeitig gilt aber wegen (3.2.5) auch

$$(3.2.7) \quad G^n(c_n + d_n t) \xrightarrow{n \rightarrow \infty} G(t).$$

Mit Satz 3.1.3 folgt aus (3.2.6) und (3.2.7), dass die folgenden Grenzwerte existieren:

$$(3.2.8) \quad d(s) := \lim_{n \rightarrow \infty} \frac{d_{\lfloor ns \rfloor}}{d_n} \in (0, \infty), \quad c(s) := \lim_{n \rightarrow \infty} \frac{c_{\lfloor ns \rfloor} - c_n}{d_n} \in \mathbb{R},$$

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

Außerdem folgt aus der Darstellung (3.2.8), 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 $c(s) \in \mathbb{R}$, $d(s) > 0$ mit

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

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

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

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 $s_1, s_2 > 0$ und $t \in \mathbb{R}$ gilt

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

Indem wir (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 $\varphi_s(t) = c(s) + d(s)t$ erhalten wir

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

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

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

Die erste Gleichung ist eine Cauchy-Funktionalgleichung in multiplikativer Form, siehe Korollar 2.4.2. Da d eine messbare Funktion ist, gibt es ein $\rho \in \mathbb{R}$ mit

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

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

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

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

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

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 $c(s) = a \log s$. Im Fall $\rho = 0$ haben wir also die folgende Lösung erhalten:

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

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 $d(s) = s^\rho$ und die Gleichung für c nimmt die folgende Gestalt an: Für alle $s_1, s_2 > 0$

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

Indem wir s_1 und s_2 vertauschen, erhalten wir

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

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

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

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

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

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

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

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 $s > 0$, $t \in \mathbb{R}$,

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

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

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

Es sei bemerkt, dass $\log H(1) < 0$. Es folgt, dass H und somit auch G vom gleichen Typ wie die 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. \square

Aufgabe 3.3.2 (Grenzwertverteilungen für Minima). Eine nichtdegenerierte Verteilungsfunktion H heißt **Extremwertverteilung für Minima** oder **min-stabil**, wenn es u.i.v. Zufallsvariablen X_1, X_2, \dots und Normierungskonstanten $a_n \in \mathbb{R}$, $b_n > 0$ gibt mit

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

Zeigen Sie: Verteilungsfunktion H ist min-stabil genau dann, wenn H vom gleichen Typ wie $\Lambda(-t)$, $\Phi_\alpha(-t)$ oder $\Psi_\alpha(-t)$ ist, wobei $\alpha > 0$.

Hinweis: Zeigen Sie mit Hilfe der Identität $\min\{X_1, \dots, X_n\} = -\max\{-X_1, \dots, -X_n\}$, dass eine Zufallsvariable Z genau dann min-stabil ist, wenn $-Z$ max-stabil ist.

Kommentare

[18], [22], [11]

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.

4.1. Langsam variierende Funktionen

Zuerst definieren wir eine Klasse von Funktionen, die sich “nicht sehr schnell verändern”.

Definition 4.1.1 (Karamata, 1930). Eine messbare Funktion $L : (A, \infty) \rightarrow (0, \infty)$ heißt **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 messbare 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.

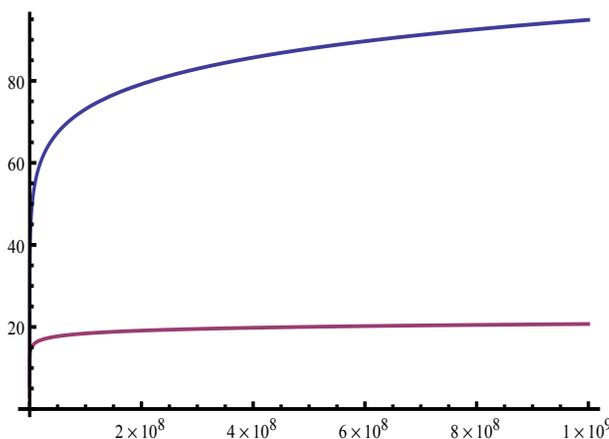


ABBILDUNG 1. Zwei langsam variierende Funktionen: $\exp\{\sqrt[3]{\log x}\}$ (blau) und $\log x$ (rot).

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 $0 \leq \beta < 1$ langsam variierend (Übungsaufgabe). Für $\beta = 1$ ergibt sich die Funktion $L(x) = e^{\log x} = x$, die nicht langsam variierend ist. Allerdings ist die Funktion

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

langsam variierend (Übungsaufgabe).

Beispiel 4.1.8. Nicht jede von 0 und $+\infty$ wegbeschränkte Funktion ist langsam variierend. Beispiel: $f(x) = 2 + \sin x$ 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. Sei L eine langsam variierende Funktion. Zeigen Sie, dass L^α für $\alpha \in \mathbb{R}$ ebenfalls langsam variierend ist.

Aufgabe 4.1.11. Seien L_1 und L_2 langsam variierende Funktionen. Zeigen Sie, dass $L_1 L_2$, L_1/L_2 und $L_1 + L_2$ ebenfalls langsam variierend sind.

4.2. Regulär variierende Funktionen

Nun definieren wir eine Klasse von Funktionen, die sich “so ähnlich wie x^α ” verhalten.

Definition 4.2.1 (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: $R \in \text{RV}_\alpha$.

Bemerkung 4.2.2. Eine Funktion ist langsam variierend genau dann, wenn sie regulär variierend mit Index $\alpha = 0$ ist.

Beispiel 4.2.3. Die Funktion $R(x) = x^\alpha$, wobei $\alpha \in \mathbb{R}$, ist regulär variierend mit Index α , denn

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

Auch die Funktion $R(x) = cx^\alpha$ mit $c > 0$ ist regulär variierend.

Beispiel 4.2.4. Die Funktion $f(x) = e^{\lfloor \log x \rfloor}$ (wobei $\lfloor \cdot \rfloor$ 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 4.2.5. 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, wie der folgende Satz zeigt.

Satz 4.2.6. 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. Für beliebiges $\lambda > 0$ gilt

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

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

Es gibt auch den Begriff der **schnell variierenden** Funktionen. Eine Funktion $f : (A, \infty) \rightarrow (0, \infty)$ heißt schnell variierend mit Index $\alpha = +\infty$, wenn für jedes $\lambda > 0$

$$\lim_{x \rightarrow +\infty} \frac{f(\lambda x)}{f(x)} = \lambda^{+\infty} := \begin{cases} +\infty, & \text{falls } \lambda > 1, \\ 1, & \text{falls } \lambda = 1, \\ 0, & \text{falls } \lambda < 1. \end{cases}$$

Analog werden schnell variierende Funktionen mit Index $\alpha = -\infty$ definiert. So ist z.B. die Funktion $f(x) = e^x$ schnell variierend mit Index $+\infty$, während die Funktion $f(x) = e^{-x}$ schnell variierend mit Index $-\infty$ ist.

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

4.3. Satz über die gleichmäßige Konvergenz

In der Definition der langsamen Variation wird verlangt, dass

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

für jedes feste $\lambda > 0$. Erstaunlicherweise gilt diese Konvergenz sogar gleichmäßig in λ , solange es von 0 und ∞ wegbeschränkt bleibt.

Satz 4.3.1 (Über die gleichmäßige Konvergenz, Karamata, 1930). Sei $L : (A, \infty) \rightarrow (0, \infty)$ 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. Zuallererst 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

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

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

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

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 $\lambda(Z_n) = \lambda(I_k) > 0$ 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:

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

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. Sei $f \in \text{RV}_\alpha$. Zeigen Sie, dass

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

Aufgabe 4.3.4. 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 dann die Komposition $f(g(x))$ regulär variierend mit Index $\alpha\beta$ ist.

Satz 4.3.5 (Über die lokale Beschränktheit). Sei $L : (A, \infty) \rightarrow (0, \infty)$ eine langsam variierende Funktion. Dann gibt es ein $B > A$, so dass L auf jedem Intervall der Form $[B, B + c]$ beschränkt ist, d.h.

$$\sup_{x \in [B, B+c]} L(x) < \infty \text{ für alle } c > 0.$$

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 4.3.6. 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

$$(4.4.1) \quad L(x) = c(x) \exp \left\{ \int_{A_1}^x \frac{\varepsilon(u)}{u} du \right\}, \quad x \geq A_1,$$

gibt, wobei $A_1 > A$ 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. Geben Sie eine solche Darstellung für die Funktionen $L_1(x) = \log x$, $L_2(x) = \log \log x$, $L_3(x) = e^{(\log x)^\beta}$, $0 \leq \beta < 1$, explizit an.

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. Zuerst 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 h die folgende Darstellung zulässt:

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

wobei $B > \log A$ und $d(x), f(x)$ messbare Funktionen sind mit

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

Daraus ergibt sich für L die Darstellung

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

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

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

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 $f(t) = h'(t)$ definieren, das geht allerdings nicht, denn h muss nicht differenzierbar sein. Wir setzen deshalb $f(t) = h(t+1) - h(t)$. Es sei bemerkt, dass $\lim_{x \rightarrow +\infty} f(x) = 0$ nach (4.4.3). Für $x > B$ gilt

$$(4.4.6) \quad \int_B^x f(t) dt = \int_B^x (h(t+1) - h(t)) dt = \int_x^{x+1} h(t) dt - \int_B^{B+1} h(t) dt,$$

wobei B so groß sei, dass alle beteiligten Funktionen auf $[B, x+1]$ beschränkt (und somit integrierbar) seien. Könnten wir in der obigen Gleichung das Integral $\int_x^{x+1} h(t) dt$ durch den Wert $h(x)$ ersetzen (wobei der dadurch entstehende Fehler nach (4.4.3) "klein" sein sollte), so würde sofort die behauptete Darstellung für h folgen. Um diese Überlegung streng zu begründen, schreiben wir

$$\int_x^{x+1} h(t) dt = h(x) + \int_x^{x+1} (h(t) - h(x)) dt = h(x) + \int_0^1 (h(x+u) - h(x)) du,$$

wobei wir die Substitution $t = x+u$ mit $u \in [0, 1]$ gemacht haben. Indem wir das in (4.4.6) einsetzen, erhalten wir nach einer Umstellung

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

Der zweite Term auf der rechten Seite hängt nicht von x ab. Der dritte 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 und dessen Beweis. Wir können also die Differenz 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

$$(4.4.7) \quad R(x) = c(x) \exp \left\{ \int_{A_1}^x \frac{\varepsilon(u)}{u} du \right\}, \quad x \geq A_1,$$

gibt, wobei $A_1 > A$ und $c(x) > 0$, $\varepsilon(x)$ messbare Funktionen sind mit

$$(4.4.8) \quad c := \lim_{x \rightarrow +\infty} c(x) \in (0, \infty), \quad \lim_{x \rightarrow +\infty} \varepsilon(x) = \alpha.$$

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). □

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

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

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

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 regulär variierend mit Index α . Zeigen Sie, dass

$$\lim_{x \rightarrow +\infty} \frac{\log R(x)}{\log x} = \alpha.$$

Aufgabe 4.5.5. 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.

Satz 4.5.6 (Potter-Schranke, 1942). Sei $L : (A, \infty) \rightarrow (0, \infty)$ langsam variierend. Für vorgegebene Konstanten $B > 1$ und $\delta > 0$ gibt es ein $K = K(B, \delta) > A$ mit

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

BEWEIS. Mit dem Darstellungssatz von Karamata 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 hinreichend großes $K > A$ finden, so dass $c(x) > c/\sqrt{B}$, $c(y) < c\sqrt{B}$ und $|\varepsilon(u)| < \delta$ für $x, y, u > K$. Somit ist

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

Daraus folgt die Behauptung. □

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

$$(4.6.2) \quad \frac{f(\lambda_1 \lambda_2 x)}{f(x)} = \frac{f(\lambda_1 \lambda_2 x)}{f(\lambda_2 x)} \cdot \frac{f(\lambda_2 x)}{f(x)} \rightarrow g(\lambda_1) g(\lambda_2) \in (0, \infty) \quad \text{für } x \rightarrow +\infty.$$

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 (Übung). 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. Asymptotische Äquivalenz

Definition 4.7.1. Wir sagen, dass zwei Funktionen $f, g : (A, \infty) \rightarrow (0, \infty)$ für $x \rightarrow +\infty$ **asymptotisch äquivalent** sind, wenn

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

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

Beispiel 4.7.2. Es gilt $x \sim x + 3$ und $x^2 + x \sim x^2 + 2x$ für $x \rightarrow +\infty$.

Aufgabe 4.7.3. 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 4.7.4. Es gelte $f \sim g$ und $f \in \text{RV}_\alpha$. Zeigen Sie, dass dann auch $g \in \text{RV}_\alpha$ ist.

4.8. Integrale von regulär variierenden Funktionen

In diesem Abschnitt werden wir uns mit dem asymptotischen Verhalten der Integrale der Form $\int_B^x t^\alpha L(t) dt$ für $x \rightarrow +\infty$ beschäftigen, wobei L langsam variierend sei. Im einfachsten

Fall wenn $L = 1$, lässt sich dieses Integral explizit berechnen. Für jedes $\alpha > -1$ gilt nämlich

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

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

$$(4.8.1) \quad \int_B^x t^\alpha L(t) dt \sim L(x) \frac{x^{\alpha+1}}{\alpha + 1}, \quad \text{für } x \rightarrow +\infty.$$

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 (4.8.1) auf B' verschieben, denn dadurch ändert sich das Integral nur um eine Konstante. Es gilt nämlich für $x > B'$, dass

$$\frac{\int_B^x t^\alpha L(t) dt}{x^{\alpha+1} L(x) / (\alpha + 1)} = \frac{\int_{B'}^x t^\alpha L(t) dt}{x^{\alpha+1} L(x) / (\alpha + 1)} + \frac{\int_B^{B'} t^\alpha L(t) dt}{x^{\alpha+1} L(x) / (\alpha + 1)},$$

und da der zweite Summand wegen $\alpha > -1$ gegen 0 geht (siehe Aufgabe 4.5.3), werden wir uns im Folgenden nur mit dem ersten Summanden beschäftigen. Betrachte also

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

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

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

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

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

was die Behauptung beweist.

Beispiel 4.8.2. Für $\alpha = 0$ erhalten wir, dass

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

Kommentare

Regulär variierende Funktionen wurden von Jovan Karamata um 1930 eingeführt und fanden seitdem zahlreiche Anwendungen in der Wahrscheinlichkeitstheorie. In diesem Kapitel konnten wir nur auf einige Aspekte der regulären Variation eingehen. Eine umfassende Darstellung des Gebiets findet sich im exzellenten Buch von Bingham, Goldie und Teugels [7]. Weitere Referenz??? [?].

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 $a_n \in \mathbb{R}$ und $b_n > 0$ gibt, so dass

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

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

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

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 b_n eine beliebige Folge mit

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

Dann gilt

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

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

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

SCHRITT 1. Zuerst zeigen wir durch Widerspruch, dass

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

Wäre nämlich $b_n < C$ für unendlich viele n , so hätten wir $n\bar{F}(b_n) \geq n\bar{F}(C)$ für unendlich viele n . Es gilt aber $\bar{F}(C) > 0$ (wegen $x^* = +\infty$), was zu einem Widerspruch zu (5.1.1) führt. Also gilt $\lim_{n \rightarrow \infty} b_n = +\infty$.

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

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

Dadurch folgt:

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

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

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

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

Wir haben allerdings nicht gesagt, wie man eine Folge b_n konstruiert, die (5.1.1) erfüllt. Am einfachsten definiert man b_n als eine Lösung der Gleichung $\bar{F}(b_n) = \frac{1}{n}$. 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

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

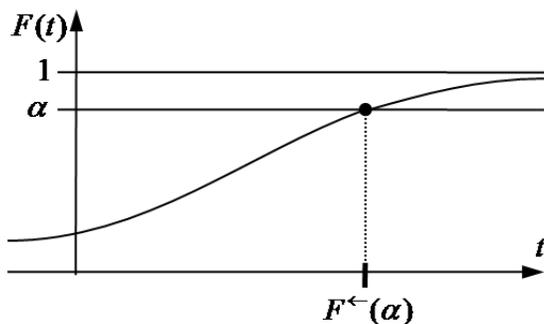


ABBILDUNG 1. Veranschaulichung von $F^{\leftarrow}(a)$.

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.

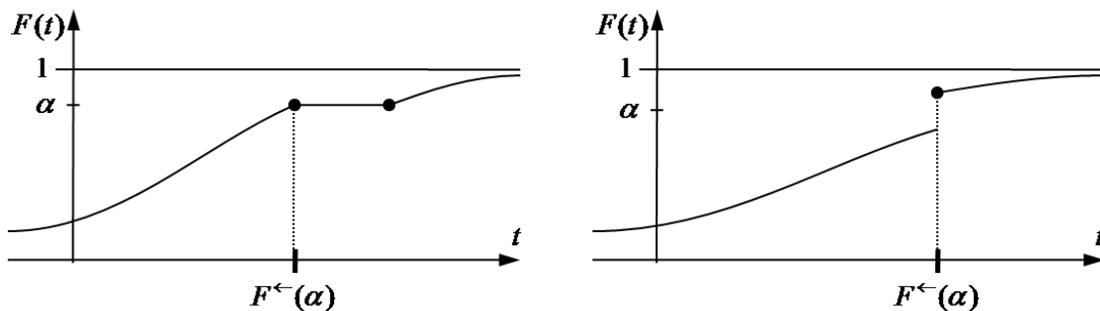


ABBILDUNG 2. Problemfälle

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 b_n angeben, die der Bedingung aus Satz 5.1.4 genügt:

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

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

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

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

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

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

SCHRITT 2. Es bleibt also noch zu zeigen, dass

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

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

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

Damit folgt unmittelbar:

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

da \bar{F} regulär variierend mit Index $-\alpha$ ist. Es ergibt sich also, dass $\liminf_{n \rightarrow \infty} n\bar{F}(b_n) \geq x^\alpha$ für alle $x \in (0, 1)$. Wenn man nun x gegen 1 gehen lässt, ergibt sich (5.1.3). Damit ist insgesamt $\lim_{n \rightarrow \infty} n\bar{F}(b_n) = 1$ 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 b_n an, für die M_n/b_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 $a_n \in \mathbb{R}$, $b_n > 0$ Folgen, so dass für alle $t \in \mathbb{R}$,

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

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 eine nicht notwendigerweise ganze Zahl $s \geq 0$ definiere $a_s = a_{[s]}$ und $b_s = b_{[s]}$, wobei $[s]$ die Abrundungsfunktion bezeichnet, siehe (3.2.4). Wir behaupten, dass

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

für alle $t \in \mathbb{R}$, wobei s nicht ganzzahlig sein muss. In der Tat,

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

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

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

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

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

Somit ist die Funktion $s \mapsto b_s$ regulär variierend mit Index $1/\alpha$. Außerdem suggeriert die zweite Relation, dass die Funktion a_s "langsamer" als b_s steigen sollte.

SCHRITT 3. Ausgehend von (5.1.7) soll gezeigt werden, dass

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

Zu diesem Zweck benutzen wir die folgende teleskopische Darstellung:

$$\frac{a_s}{b_s} = \sum_{k=1}^n \frac{a_{s/2^{k-1}} - a_{s/2^k}}{b_s} + \frac{a_{s/2^n}}{b_s} = \sum_{k=1}^n \frac{a_{s/2^{k-1}} - a_{s/2^k}}{b_{s/2^k}} \cdot \frac{b_{s/2^k}}{b_s} + \frac{a_{s/2^n}}{b_s},$$

wobei n noch zu wählen ist. Nun sollen einzelne Summanden abgeschätzt werden. Sei $\varepsilon > 0$ fest. Sei $A > 0$ so groß gewählt, dass

$$(5.1.9) \quad \left| \frac{a_{2x} - a_x}{b_x} \right| < \varepsilon \text{ für alle } x \geq A$$

(was aus der zweiten Relation in (5.1.7) folgt) und

$$(5.1.10) \quad \frac{b_{x/\lambda}}{b_x} < 2\lambda^{-1/(2\alpha)} \text{ für alle } x \geq A, \quad \lambda > 1$$

(was aus der Potter-Schranke für die regulär variierende Funktion $x \mapsto b_x$ folgt). Für ein $s > A$ können wir ein $n = n(s) \in \mathbb{N}_0$ mit $s/2^{n+1} \leq A < s/2^n$ finden. Wir erhalten die Abschätzung

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

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

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

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

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

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

SCHRITT 5. In diesem Schritt soll (5.1.11) vereinfacht werden. Durch Logarithmieren ergibt sich aus (5.1.11), dass

$$\lim_{n \rightarrow \infty} n \log F(b_n t) = -t^{-\alpha} \quad \text{für alle } t > 0.$$

Dies kann man auch wie folgt umschreiben:

$$\lim_{n \rightarrow \infty} n \log(1 - \bar{F}(b_n t)) = -t^{-\alpha} \quad \text{für alle } t > 0.$$

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

$$(5.1.12) \quad \lim_{n \rightarrow \infty} n \bar{F}(b_n t) = t^{-\alpha} \quad \text{für alle } t > 0.$$

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

$$\lim_{x \rightarrow +\infty} \frac{\bar{F}(\lambda x)}{\bar{F}(x)} = \lambda^{-\alpha}.$$

Zuerst eine Vorüberlegung. Hätte x die Form $x = b_n$, so könnten wir schreiben

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

wobei (5.1.12) zweimal benutzt wurde. Im allgemeinen kann aber x auch Werte annehmen, die nicht die Form b_n haben. Wir werden deshalb x zwischen zwei aufeinanderfolgenden Termen der Folge b_n einquetschen. Definiere dazu

$$n(x) := \inf\{m \in \mathbb{N} : b_{m+1} > x\}, \quad x > 0.$$

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

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

wobei bei der Grenzwertbildung die Relation (5.1.12) zweimal zum Einsatz kam. 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 Satz 5.2.4 erfüllt, so werden wir zeigen, dass

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

wobei b_n eine beliebige Folge mit

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

ist. Ein Beispiel einer solchen Folge b_n ist gegeben durch

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

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

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

denn

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

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

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

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 $\text{Beta}(\alpha, \beta)$ -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 Buch werden wir auf diese Theorie nicht eingehen und verweisen stattdessen auf die Bücher von de Haan [10], Bingham, Goldie, Teugels [7] und Resnick [22]. 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

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

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

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

Eine mögliche Wahl von a_n ist $a_n = F^{\leftarrow}\left(1 - \frac{1}{n}\right)$ (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 $\lim_{n \rightarrow \infty} a_n = x^*$, denn hätte a_n eine von x^* wegbeschränkte Teilfolge, so würde entlang dieser Teilfolge (5.3.2) verletzt sein. Man betrachte nun

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

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

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

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

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) $a_n = \frac{\log n}{\lambda}$ und $b_n = \frac{1}{\lambda}$, 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 $a_n \in \mathbb{R}$ und $b_n > 0$ mit $(M_n - a_n)/b_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ 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 $a_n \in \mathbb{R}$ und $b_n > 0$ mit $(M_n - a_n)/b_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ 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 $a_n \in \mathbb{R}$ und $b_n > 0$ mit $(M_n - a_n)/b_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ an.

Alle drei Max-Anziehungsbereiche sind abgeschlossenes 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

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

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

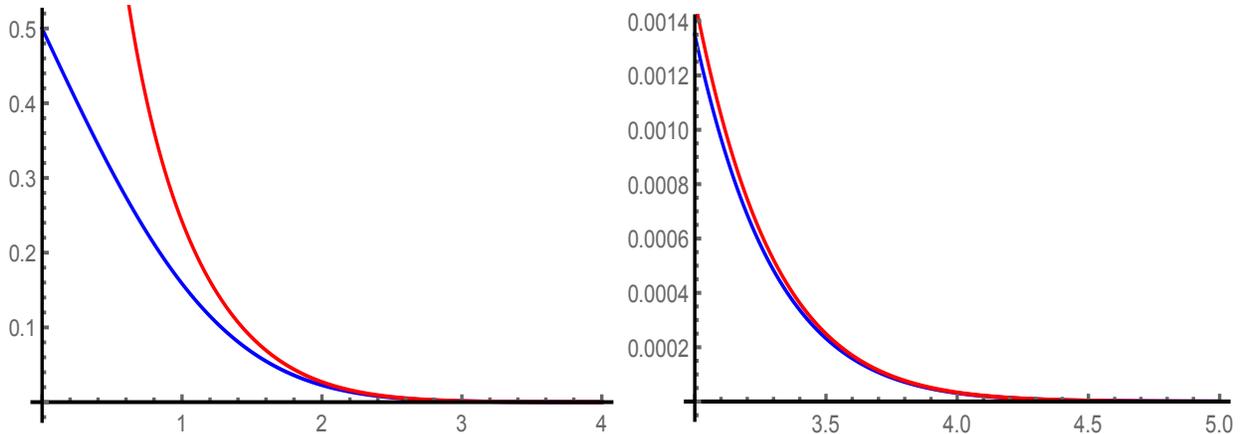


ABBILDUNG 3. Veranschaulichung von Lemma 5.4.1. Blaue Kurve: Die Tailfunktion \bar{F} der Standardnormalverteilung. Rote Kurve: Die Approximation. Das rechte Bild zeigt die beiden Kurven auf dem Intervall $[3, 5]$.

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 $a_n \in \mathbb{R}$ und $b_n > 0$ derart bestimmen, dass

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

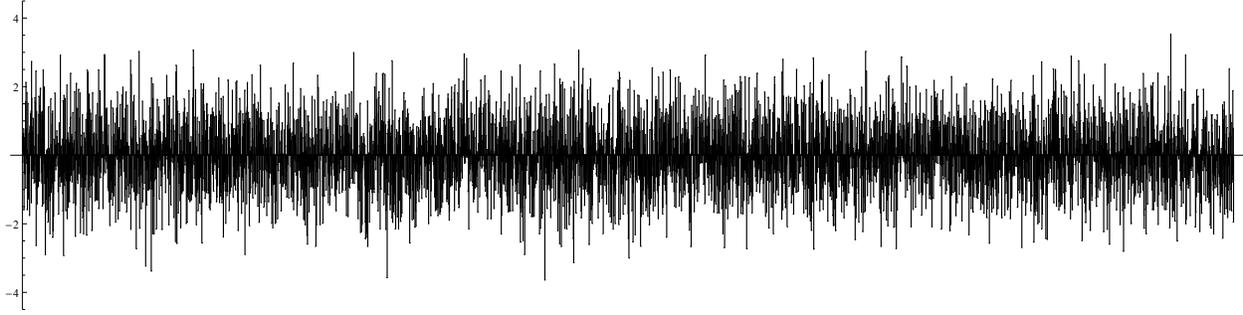


ABBILDUNG 4. Eine standardnormalverteilte Stichprobe vom Umfang $n = 5000$.

Laut Satz 5.3.1 sollten wir a_n so wählen, dass

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

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

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

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

Auf der linken Seite ist $e^{a_n^2/2}$ derjenige Term, der am schnellsten gegen $+\infty$ geht. Wir können also als eine erste Annäherung zu a_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 $a_n = \sqrt{2 \log n} + \delta_n$, wobei δ_n noch genauer spezifiziert werden muss. Mit diesem Ansatz gilt:

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

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 a_n wie folgt gewählt werden sollte:

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

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

$$(5.4.5) \qquad \qquad \qquad b_n = \frac{1}{\sqrt{2 \log n}}.$$

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 $a_n = \sqrt{2 \log n} - \frac{\log(4\pi \log n)}{2\sqrt{2 \log n}}$ sind. Die Differenz zwischen M_n und a_n ist zufällig und hat die Größenordnung $b_n = \frac{1}{\sqrt{2 \log n}}$. Multipliziert man $M_n - a_n$ 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 $x \in \mathbb{R}$:

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

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

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

Durch die Anwendung der Exponentialfunktion folgt, dass

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

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

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

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 $a_n \in \mathbb{R}$ und $b_n > 0$ mit $(M_n - a_n)/b_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ 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 $a_n \in \mathbb{R}$ und $b_n > 0$ mit $(M_n - a_n)/b_n \xrightarrow[n \rightarrow \infty]{d} \Lambda$ 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 Gerat, dessen Lebensdauer als eine Zufallsvariable $Z \geq 0$ modelliert werde. Wir betrachten ein sehr kleines Zeitintervall $(x, x + dx)$ und stellen uns die folgende Frage:

Gegeben, dass das Gerat zum Zeitpunkt x noch funktioniert, wie gro ist die Wahrscheinlichkeit, dass es zwischen x und $x + dx$ ausfallt?

Dabei nehmen wir an, dass Z absolut stetig mit Dichte f ist. Die Verteilungsfunktion von Z sei mit F bezeichnet. 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 $h(x) = f(x)/\bar{F}(x)$ heit die **Ausfallrate**¹ des Gerats.

Im obigen Beispiel ist die Lebensdauer des Gerats nicht-negativ. Die Definition der Ausfallrate kann allerdings auch fur Zufallsvariablen verwendet werden, die negative Werte annehmen durfen. Fur 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 kann mit der Vergessenseigenschaft der Exponentialverteilung in Zusammenhang gebracht werden: Ein Gerat, dass

¹In der Literatur kommen auch Begriffe Sterblichkeitsrate, Hazardrate, usw. vor

zum Zeitpunkt x noch funktioniert, “erinnert” sich an sein Alter nicht und fällt mit der gleichen Rate aus, wie ein neues Gerät unmittelbar nach Inbetriebnahme.

Beispiel 5.5.3 (Additionsregel für Ausfallraten). Wir betrachten ein System, das aus n stochastisch unabhängigen Komponenten besteht, die mit Raten $h_1(x), \dots, h_n(x)$ ausfallen. Das ganze System falle genau dann aus, wenn mindestens eine der Komponenten ausfällt. Wir behaupten, dass die Ausfallrate des Systems durch $h_1(x) + \dots + h_n(x)$ gegeben ist. Es seien Z_1, \dots, Z_n Zufallsvariablen, die die Lebensdauer der einzelnen Komponenten angeben, und F_1, \dots, F_n die entsprechenden Verteilungsfunktionen. Die Lebensdauer des Systems ist dann die Zufallsvariable

$$Z = \min\{Z_1, \dots, Z_n\}.$$

Für die Tailfunktion von Z gilt

$$\bar{F}(x) = \mathbb{P}[Z > x] = \bar{F}_1(x) \cdot \dots \cdot \bar{F}_n(x).$$

Somit gilt für die Ausfallrate des Systems

$$\begin{aligned} h(x) &= -(\log \bar{F})'(x) = -(\log \bar{F}_1 + \dots + \log \bar{F}_n)'(x) \\ &= -(\log \bar{F}_1)'(x) - \dots - (\log \bar{F}_n)'(x) = h_1(x) + \dots + h_n(x), \end{aligned}$$

wobei wir die Formel aus Definition 5.5.1 mehrmals benutzt haben.

Wir betrachten nun den Spezialfall, in dem die Lebensdauer der i -ten Komponente exponentialverteilt mit Parameter $\lambda_i > 0$ ist. Die Ausfallrate der i -ten Komponente ist somit konstant und gleich λ_i , woraus folgt, dass die Ausfallrate des Gesamtsystems ebenfalls konstant und gleich $\lambda_1 + \dots + \lambda_n$ ist. Also ist die Lebensdauer des Systems exponentialverteilt mit Parameter $\lambda_1 + \dots + \lambda_n$. Symbolisch können wir das Gezeigte wie folgt darstellen:

$$\min\{\text{Exp}(\lambda_1), \dots, \text{Exp}(\lambda_n)\} \stackrel{d}{=} \text{Exp}(\lambda_1 + \dots + \lambda_n).$$

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

Lemma 5.5.4. 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. \square

Für eine ausführliche Behandlung der mathematischen Zuverlässigkeitstheorie verweisen wir auf das Buch [?].

In den obigen Überlegungen haben wir absolut stetige Verteilungen betrachtet. In der nachfolgenden Aufgabe betrachten wir den Fall der diskreten Zufallsvariablen.

Aufgabe 5.5.5 (Ausfallrate für diskrete Verteilungen). Für eine diskrete Zufallsvariable Z mit Werten in $\mathbb{N}_0 = \{0, 1, \dots\}$ kann man die Ausfallrate wie folgt definieren:

$$H_Z(k) = \mathbb{P}[Z = k | Z \geq k] = \frac{\mathbb{P}[Z = k]}{\mathbb{P}[Z \geq k]}, \quad k \in \mathbb{N}_0.$$

- (a) Zeigen Sie: Sind Z_1, \dots, Z_n unabhängige Zufallsvariablen mit Werten in \mathbb{N}_0 , so gilt für die Ausfallrate von $\min\{Z_1, \dots, Z_n\}$:

$$1 - H_{\min\{Z_1, \dots, Z_n\}} = \prod_{i=1}^n (1 - H_{Z_i}(k)), \quad k \in \mathbb{N}_0.$$

- (b) Sei Z geometrisch verteilt mit Parameter $p \in (0, 1)$, d.h. $\mathbb{P}[Z = k] = p(1-p)^k$ für $k = 0, 1, \dots$. Zeigen Sie: Die Ausfallrate von Z ist konstant und gleich p .
- (c) Seien Z_1, \dots, Z_n unabhängig und geometrisch verteilt mit Parametern p_1, \dots, p_n . Zeigen Sie, dass $\min\{Z_1, \dots, Z_n\}$ geometrisch verteilt mit Parameter $1 - \prod_{i=1}^n (1 - p_i)$ ist.

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 5.5.6. 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 5.5.7 (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 5.5.4

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

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 (siehe Aufgabe 4.4.3), 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 5.5.8. 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 5.5.9 (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. □

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 unendlichen 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 5.5.10 (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 5.5.4 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 $s \in (s_0, x^*)$. Für alle $x \in (s_0, x^*)$ und $z \in [0, u]$ ergibt sich die Abschätzung

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

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 5.5.11. 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 5.5.12. 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 [5] 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.

Kommentare

Zuverlässigkeitstheorie [?].

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*”.

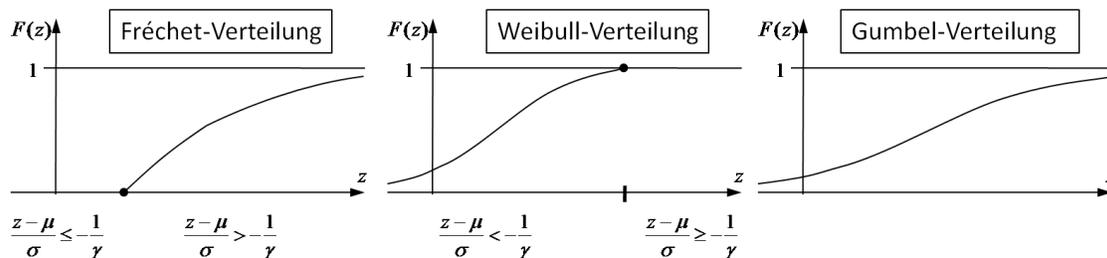
6.1. Statistik der Blockmaxima: GEV-Verteilungen

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.¹

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):

$$\hat{Z}_p = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\gamma}} \{1 - (-\log(1 - p))^{-\hat{\gamma}}\}, & \hat{\gamma} \neq 0, \\ \hat{\mu} - \hat{\sigma} \log(-\log(1 - p)), & \hat{\gamma} = 0. \end{cases}$$

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

$$\left\{ \left(\hat{G}(x_{(i)}), \frac{i}{n+1} \right) : i = 1, \dots, n \right\} \subset [0, 1]^2.$$

¹Leider treten bei der Anwendung der Maximum-Likelihood-Methode auf die GEV-Verteilungen einige Schwierigkeiten auf. Z.B. kann die Likelihood-Funktion unbeschränkt sein. Außerdem sind die üblichen Regularitätsbedingungen, die eine Konvergenzrate von $n^{-1/2}$ sicherstellen, nicht erfüllt, s. Richard L. Smith, "Maximum likelihood estimation in a class of nonregular cases", *Biometrika* (1985) 72 (1): 67-90.

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).

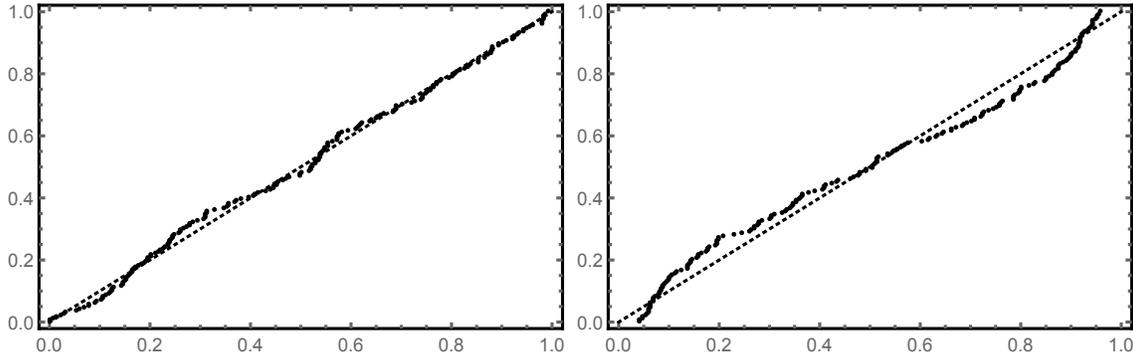


ABBILDUNG 1. PP-Plot

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

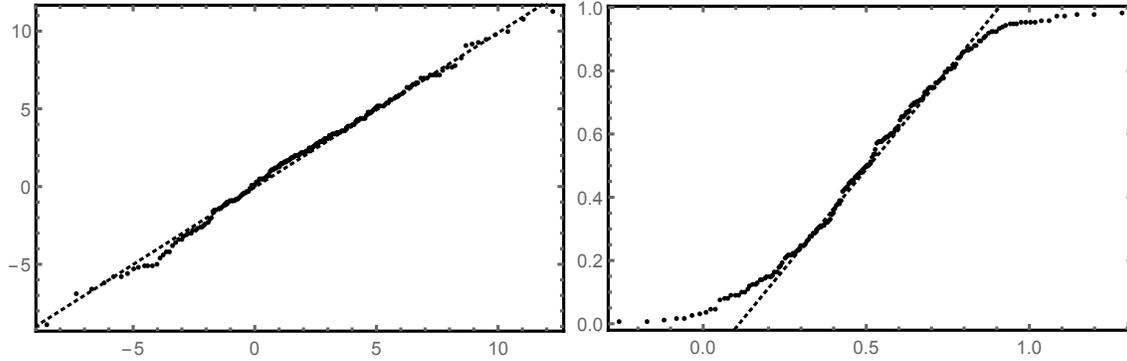


ABBILDUNG 2. Zwei QQ-Plots. Links: Ein “guter” Quantilplot (eine normalverteilte Stichprobe wurde gegen Quantile der Normalverteilung geplottet). Rechts: ein schlechter Quantilplot (eine gleichverteilte Stichprobe wurde gegen Quantile der Normalverteilung geplottet)

Daten Rechnung trägt. Die beobachteten Blockmaxima x_1, \dots, x_n seien Realisierungen von 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

$$f_{\gamma,\sigma}(t) = \frac{1}{\sigma} \left(1 + \frac{\gamma t}{\sigma}\right)^{-\frac{1}{\gamma}-1} \text{ mit } \begin{cases} t > 0, & \text{falls } \gamma > 0, \\ t \in [0, -\frac{\sigma}{\gamma}], & \text{falls } \gamma < 0, \end{cases}$$

und $f_{0,\sigma}(t) = (1/\sigma)e^{-t/\sigma}$, $t > 0$, 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

$$\hat{Z}_p = u + \frac{\hat{\sigma}}{\hat{\gamma}} \left[\left(\frac{1}{np} \sum_{i=1}^n \mathbb{1}_{\{x_i > u\}} \right)^{\hat{\gamma}} - 1 \right].$$

Kommentare

[8], [13], [6].

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 $X_{n:n} = \max\{X_1, \dots, X_n\}$. In diesem Kapitel werden wir Formeln für die Verteilung der Ordnungsstatistiken herleiten.

7.1. Allgemeine Eigenschaften der Ordnungsstatistiken

Zuerst berechnen wir die Verteilungsfunktion einer einzelnen Ordnungsstatistik $X_{k:n}$.

Satz 7.1.1. Für alle $k \in \{1, \dots, n\}$ 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 der 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 $k \in \{1, \dots, n\}$ die Dichte von $X_{k:n}$ gegeben durch

$$f_{X_{k:n}}(t) = n \binom{n-1}{k-1} f(t) F(t)^{k-1} \bar{F}(t)^{n-k} = k \binom{n}{k} f(t) F(t)^{k-1} \bar{F}(t)^{n-k}.$$

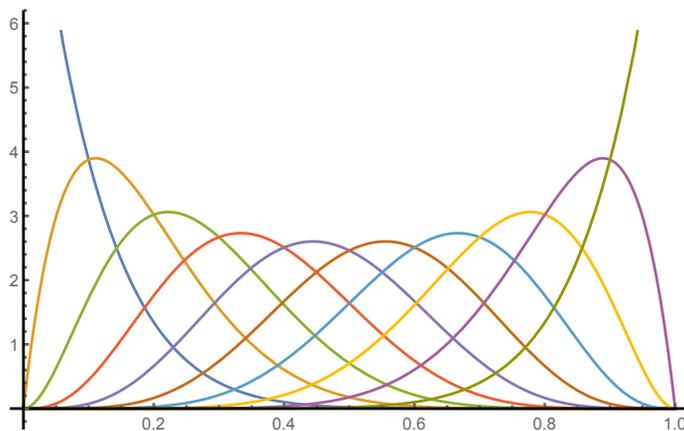


ABBILDUNG 1. Dichten der Ordnungsstatistiken $X_{1:10}, \dots, X_{10:10}$ einer unabhängigen und auf $[0, 1]$ gleichverteilten Stichprobe X_1, \dots, X_{10} .

BEWEISIDEE. Man kann den Satz beweisen, indem man die Formel aus Satz 7.1.1 ableitet. Dieser Weg führt zu komplizierten Berechnungen. Wir geben hier einen anderen Beweis, der viel eleganter (allerdings nicht ganz rigoros) ist.

Wir bestimmen die Wahrscheinlichkeit, dass $X_{k:n}$ in einem unendlich kleinen Intervall $(t, t + dt)$ liegt. Damit das eintritt, muss Folgendes passieren:

1. Eine der Zufallsvariablen, z.B. X_i , muss einen Wert im Intervall $(t, t + dt)$ annehmen. Für jedes i hat das Wahrscheinlichkeit $f(t)dt$. Außerdem gibt es n Möglichkeiten, das i auszuwählen.
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 $t + dt$ sein. Die Wahrscheinlichkeit davon ist $(1 - F(t + dt))^{n-k}$, was für $dt \rightarrow 0$ gegen $(1 - F(t))^{n-k}$ konvergiert.

Indem wir nun alles ausmultiplizieren, erhalten wir

$$f_{X_{k:n}}(t) = \lim_{dt \rightarrow 0} \frac{\mathbb{P}[X_{k:n} \in (t, t + dt)]}{dt} = n \binom{n-1}{k-1} f(t) F(t)^{k-1} \bar{F}(t)^{n-k}.$$

Den kombinatorischen Faktor kann man auch anders berechnen: Zuerst wählen wir aus n Zufallsvariablen k Zufallsvariablen, die $< t + dt$ sind (dafür gibt es $\binom{n}{k}$ Möglichkeiten), und dann wählen aus diesen k Zufallsvariablen eine Zufallsvariable, die im Intervall $(t, t + dt)$ liegen 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 mit Parametern $(k, n - k + 1)$; siehe Abbildung 1.

Aufgabe 7.1.4. Zeigen Sie, dass $\mathbb{E}X_{k:n} = \frac{k}{n+1}$ für alle $k \in \{1, \dots, n\}$.

Interpretation: Die n Ordnungsstatistiken teilen $[0, 1]$ in $n + 1$ kleine Intervalle. Diese Intervalle haben die gleiche erwartete Länge, nämlich $\frac{1}{n+1}$.

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 $X_{j:n}$.

Aufgabe 7.1.6. Seien X_1, \dots, X_n unabhängige Zufallsvariablen mit der Cauchy-Dichte $f(x) = 1/(\pi(1 + x^2))$, wobei $n \geq 3$. Zeigen Sie, dass $\mathbb{E}|X_{1:n}| = \mathbb{E}|X_{n:n}| = \infty$ und (erstaunlicherweise!) $\mathbb{E}|X_{k:n}| < \infty$ für alle $k = 2, \dots, n - 1$.

Im nächsten Satz bestimmen wir die gemeinsame Dichte *aller* n Ordnungsstatistiken.

Satz 7.1.7. Seien X_1, \dots, X_n unabhängig und identisch verteilt mit Dichte f . Dann gilt für die gemeinsame Dichte von $X_{1:n}, \dots, X_{n:n}$:

$$f_{X_{1:n}, \dots, X_{n:n}}(t_1, \dots, t_n) = \begin{cases} n! \cdot f(t_1) \cdot \dots \cdot f(t_n), & t_1 < t_2 < \dots < t_n, \\ 0, & \text{sonst.} \end{cases}$$

INTERPRETATION. 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 $\{X_1 = t_1, \dots, X_n = t_n\}$ 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.

BEWEIS. □

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 andere interessante Eigenschaft, die sogenannte Markov-Eigenschaft.

Satz 7.1.8 (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.

INTERPRETATION. Angenommen, die ersten k Ordnungsstatistiken der Stichprobe X_1, \dots, X_n sind uns 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 und können ignoriert werden. Gegeben die “Gegenwart” (also $X_{k:n}$) ist die “Zukunft” (also $X_{k+1:n}$ und auch alle größeren Ordnungsstatistiken) unabhängig von der “Vergangenheit” (also $X_{1:n}, \dots, X_{k-1:n}$). Die Ordnungsstatistiken bilden somit einen Markov-Prozess.

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. \square

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 $X_{1:n} = x_1, \dots, X_{k:n} = x_k$?

Interpretation: Stellen Sie sich n Geräte vor, die $\text{Exp}(\lambda)$ -verteilte Lebenszeiten haben und unabhängig voneinander ausfallen. Wenn die ersten k Ausfälle zu den Zeitpunkten x_1, \dots, x_k stattgefunden haben, wieviele Geräte funktionieren zum Zeitpunkt x_k und wie

lange dauert es bis zum nächsten Ausfall? Siehe Abschnitt 7.3 für mehr über Ordnungsstatistiken der Exponentialverteilung.

Aufgabe 7.1.10. Seien X_1, \dots, X_n unabhängig und gleichverteilt auf $[0, 1]$. Welche Verteilung besitzt $(X_{k+1:n} - x_k)/(1 - x_k)$ gegeben, dass $X_{1:n} = x_1, \dots, X_{k:n} = x_k$?

Interpretation: Wenn $X_{1:n} = x_1, \dots, X_{k:n} = x_k$ gilt, wie sind dann die restlichen Punkte auf dem Intervall $[x_k, 1]$ verteilt? Siehe Abschnitt 7.4 für mehr über Ordnungsstatistiken der Gleichverteilung.

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 vorangegangenen Kapiteln beschrieben (Extremwertverteilungen). In diesem Abschnitt beschreiben wir die Grenzwertverteilungen der sogenannten *extremen Ordnungsstatistiken* $M_n^{(k)}$, wobei $k \in \mathbb{N}$ fest ist 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 Poisson-Grenzwertsatz 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 (Gumbel, 1935, und Smirnov, 1952). Es gebe Normierungsfolgen $a_n \in \mathbb{R}$ und $b_n > 0$, so dass

$$\frac{M_n^{(1)} - a_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} G,$$

wobei G eine Extremwertverteilung ist. Dann gilt für alle $k = 1, 2, \dots$ und $x \in \mathbb{R}$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n^{(k)} - a_n}{b_n} \leq x \right] = \begin{cases} G(x) \cdot \sum_{s=0}^{k-1} \frac{(-\log G(x))^s}{s!}, & G(x) \neq 0, \\ 0, & G(x) = 0. \end{cases}$$

BEWEIS. Setzt man $u_n := a_n + b_n x$ und $S_n = \sum_{i=1}^n \mathbb{1}_{\{X_i > u_n\}}$, so ergibt sich:

$$\mathbb{P} [M_n^{(k)} \leq a_n + b_n x] = \mathbb{P}[S_n \leq k - 1] = \sum_{s=0}^{k-1} \mathbb{P}[S_n = s].$$

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:

$$\lim_{n \rightarrow \infty} F^n(u_n) = G(x).$$

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

$$\tau = \lim_{n \rightarrow \infty} n F(u_n) = -\log G(x)$$

Einsetzen von τ in (7.2.1) liefert die Behauptung. Für $G(x) = 0$ ist der Beweis analog, liefert allerdings $\tau = +\infty$. \square

Beispiel 7.2.4. Speziell erhalten wir für die drei Typen von Extremwertverteilungen, dass

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n^{(k)} - a_n}{b_n} \leq x \right] = \begin{cases} e^{-e^{-x}} \cdot \sum_{s=0}^{k-1} e^{-sx} / s!, & x \in \mathbb{R}, \text{ falls } G = \Lambda \text{ (Gumbel-Fall)}, \\ e^{-x^{-\alpha}} \cdot \sum_{s=0}^{k-1} x^{-s\alpha} / s!, & x > 0, \text{ falls } G = \Phi_\alpha \text{ (Fréchet-Fall)}, \\ e^{-(-x)^\alpha} \cdot \sum_{s=0}^{k-1} (-x)^{s\alpha} / s!, & x < 0, \text{ falls } G = \Psi_\alpha \text{ (Weibull-Fall)}. \end{cases}$$

Durch Ableiten und Vereinfachen einer Teleskopsumme erhält man für die Dichte der Grenzverteilung die Formel

$$f_k(x) = \begin{cases} G'(x) \frac{(-\log G(x))^{k-1}}{(k-1)!}, & G(x) \neq 0, \\ 0, & G(x) = 0. \end{cases}$$

Speziell im Fall der Gumbel-Verteilung ist die Dichte gegeben durch

$$f_k(x) = e^{-e^{-x}} \frac{e^{-kx}}{(k-1)!}, \quad x \in \mathbb{R},$$

siehe Abbildung 2. In Kapitel 10 werden wir die gleichen Ergebnisse mit Hilfe von Poisson-Punktprozessen noch einmal herleiten.

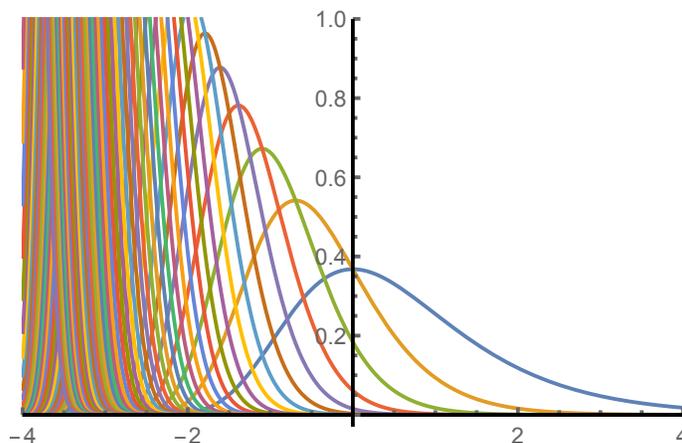


ABBILDUNG 2. Dichten f_1, f_2, \dots aus Beispiel 7.2.4.

Nachdem wir für jedes einzelne $k \in \mathbb{N}$ die Grenzwertverteilung von $M_n^{(k)}$ bestimmt haben, stellt sich die Frage nach der *gemeinsamen* Grenzwertverteilung des Zufallsvektors $(M_n^{(1)}, \dots, M_n^{(r)})$. Dabei ist die Anzahl der Ordnungsstatistiken $r \in \mathbb{N}$ fest und $n \rightarrow \infty$.

Satz 7.2.5. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit Verteilungsfunktion F . Es gebe Normierungsfolgen $a_n \in \mathbb{R}$ und $b_n > 0$, so dass

$$\frac{M_n^{(1)} - a_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} G,$$

wobei G eine Extremwertverteilung ist. Es sei $r \in \mathbb{N}$ fest. Dann konvergiert der Zufallsvektor

$$\left(\frac{M_n^{(1)} - a_n}{b_n}, \dots, \frac{M_n^{(r)} - a_n}{b_n} \right)$$

für $n \rightarrow \infty$ in Verteilung.

Für den Beweis 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 $u_n^{(l)}$, $1 \leq l \leq r$. Aus der Definition folgt unmittelbar, dass $S_n^{(1)} \leq \dots \leq S_n^{(r)}$. Der folgende Satz beschreibt die Grenzwertverteilung des Zufallsvektors $(S_n^{(1)}, \dots, S_n^{(r)})$ für $n \rightarrow \infty$.

Satz 7.2.6. Für alle $k_1, k_2, \dots, k_r \in \{0, 1, 2, \dots\}$ gilt

$$\lim_{n \rightarrow \infty} \mathbb{P} [S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \dots, S_n^{(r)} = k_1 + \dots + k_r] = e^{-(\tau_1 + \dots + \tau_r)} \frac{\tau_1^{k_1} \tau_2^{k_2} \dots \tau_r^{k_r}}{k_1! k_2! \dots k_r!}.$$

Mit anderen Worten, es besteht die 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. Stellen wir uns die Zufallsvariablen X_1, \dots, X_n als Kugeln und die Intervalle

$$(-\infty, u_n^{(r)}], (u_n^{(r-1)}, u_n^{(r)}], \dots, (u_n^{(2)}, u_n^{(1)}], (u_n^{(1)}, \infty)$$

als Fächer vor, auf die diese Kugeln zufällig verteilt werden. Die Wahrscheinlichkeit, dass X_i im Intervall $(u_n^{(l-1)}, u_n^{(l)})$ landet, ist $p_{n,l} := \bar{F}(u_n^{(l)}) - \bar{F}(u_n^{(l-1)})$ für $l \in \{1, \dots, r\}$, wobei $\bar{F}(u_n^{(0)})$ als 0 interpretiert wird. Somit ist die Wahrscheinlichkeit, dass X_i auf das erste Intervall $(-\infty, u_n^{(r)})$ verteilt wird, gleich $1 - p_{n,1} - \dots - p_{n,r}$. Wir interessieren uns für

$$P_n(k_1, \dots, k_r) = \mathbb{P} [S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \dots, S_n^{(r)} = k_1 + \dots + k_r],$$

also für die Wahrscheinlichkeit, dass im Intervall $(u_n^{(l-1)}, u_n^{(l)})$ genau k_l Kugeln landen, für alle $l \in \{1, \dots, r\}$. Mit der Formel für die Multinomialverteilung erhalten wir

$$(7.2.2) \quad P_n(k_1, \dots, k_r) = \binom{n}{k_1} p_{n,1}^{k_1} \cdot \binom{n-k_1}{k_2} p_{n,2}^{k_2} \cdot \dots \cdot \binom{n-\sum_{i=1}^{r-1} k_i}{k_r} p_{n,r}^{k_r} \cdot (1 - p_{n,1} - \dots - p_{n,r})^{n - \sum_{i=1}^r k_i}.$$

Wir schauen uns die einzelnen Faktoren an. 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

$$(n - k_1 - \dots - k_r)(p_{n,1} + \dots + p_{n,r}) = (n - k_1 - \dots - k_r)\bar{F}(u_n^{(r)}) \xrightarrow{n \rightarrow \infty} \tau_1 + \dots + \tau_r.$$

Setzt man alles zusammen, so erhält man die Behauptung. \square

BEWEIS VON SATZ 7.2.5. Wir zeigen, dass für alle $x_1 > \dots > x_r$ der Grenzwert

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n^{(1)} \leq a_n + b_n x_1, \dots, M_n^{(r)} \leq a_n + b_n x_r]$$

existiert. Definiere die Schwellenwerte $u_n^{(l)} := a_n + b_n x_l$ für $l = 1, \dots, r$. Es ergibt sich aus

$$\lim_{n \rightarrow \infty} F^n(a_n + b_n x) = G(x) \Rightarrow \lim_{n \rightarrow \infty} n\bar{F}(a_n + b_n x) = -\log G(x),$$

dass

$$\tau_l := \lim_{n \rightarrow \infty} n\bar{F}(u_n^{(l)}) = -\log G(x_l), \quad l = 1, \dots, r.$$

Außerdem gilt mit der Notation von Satz 7.2.6:

$$\mathbb{P}[M_n^{(1)} \leq u_n^{(1)}, \dots, M_n^{(r)} \leq u_n^{(r)}] = \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} \leq 1, S_n^{(3)} \leq 2, \dots, S_n^{(r)} \leq r - 1].$$

Nun stellt man die Wahrscheinlichkeit auf der rechten Seite als eine Summe der Wahrscheinlichkeiten der Form

$$\mathbb{P}[S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \dots, S_n^{(r)} = k_1 + \dots + k_r]$$

über alle $(k_1, \dots, k_r) \in \mathbb{N}_0^r$ mit $k_1 = 0, k_1 + k_2 \leq 1, \dots, k_1 + \dots + k_r \leq r - 1$ dar. Diese Summe ist endlich. Für jeden einzelnen Summanden garantiert Satz 7.2.6 die Existenz des Grenzwertes für $n \rightarrow \infty$, was die Behauptung beweist. \square

Rein theoretisch erlaubt der obige Beweis, eine Formel für die Verteilungsfunktion des Grenzwertes von $(M_n^{(1)}, \dots, M_n^{(r)})$ aufzustellen. Diese Formel ist allerdings zu kompliziert und wird hier nicht angegeben. In Kapitel 10 werden wir einen viel eleganteren Zugang zu den extremen Ordnungsstatistiken kennenlernen und eine einfache Formel für die gemeinsame Dichte der Grenzwertverteilung herleiten.

Aufgabe 7.2.7. Es seien die Bedingungen von Satz 7.2.5 erfüllt. Zeigen Sie, dass für alle $x_1 > x_2$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n^{(1)} - a_n}{b_n} \leq x_1, \frac{M_n^{(2)} - a_n}{b_n} \leq x_2 \right] = \begin{cases} G(x_2) \left(\log \frac{G(x_1)}{G(x_2)} + 1 \right), & \text{falls } G(x_2) \neq 0, \\ 0, & \text{sonst.} \end{cases}$$

7.3. Ordnungsstatistiken der exponentialverteilten Zufallsvariablen

Wir erinnern, dass eine Zufallsvariable Z exponentialverteilt mit Parameter $\lambda > 0$ ist, wenn

$$\mathbb{P}[Z > t] = e^{-\lambda t} \text{ für } t \geq 0.$$

Wir schreiben dann $Z \sim \text{Exp}(\lambda)$. Die Ordnungsstatistiken von unabhängigen exponentialverteilten Zufallsvariablen besitzen eine besonders schöne Darstellung.



Satz 7.3.1 (Sukhatme, 1937). Seien Z_1, \dots, Z_n unabhängige und mit Parameter $\lambda = 1$ exponentialverteilte Zufallsvariablen. Dann gilt

$$(7.3.1) \quad (Z_{1:n}, Z_{2:n}, \dots, Z_{n:n}) \stackrel{d}{=} \left(\frac{E_1}{n}, \frac{E_1}{n} + \frac{E_2}{n-1}, \dots, \frac{E_1}{n} + \frac{E_2}{n-1} + \dots + \frac{E_n}{1} \right),$$

wobei E_1, \dots, E_n ebenfalls unabhängig und exponentialverteilt mit Parameter $\lambda = 1$ sind.

INTERPRETATION. Den Satz kann man auch wie folgt formulieren (siehe Abbildung ??): Die Abstände zwischen den aufeinanderfolgenden Ordnungsstatistiken sind unabhängig und exponentialverteilt mit

$$Z_{1:n} \sim \text{Exp}(n), \quad Z_{2:n} - Z_{1:n} \sim \text{Exp}(n-1), \quad \dots, \quad Z_{n:n} - Z_{n-1:n} \sim \text{Exp}(1).$$

Diese Behauptung lässt eine anschauliche Interpretation zu. Wir stellen uns ein System

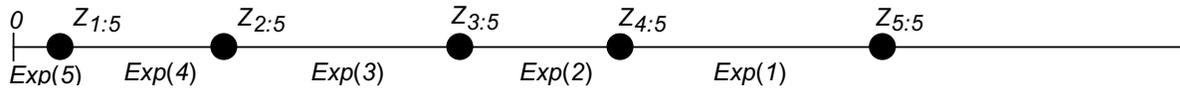


ABBILDUNG 3. Ordnungsstatistiken von $n = 5$ exponentialverteilten Zufallsvariablen.

aus n Geräten vor, die zum Zeitpunkt 0 eingeschaltet werden und unabhängig voneinander ausfallen können. Die Lebensdauer des i -ten Geräts sei $Z_i \sim \text{Exp}(1)$. Der Zeitpunkt, zu dem zum ersten Mal eines der Geräte ausfällt, ist $Z_{1:n} = \min\{Z_1, \dots, Z_n\}$. Man rechnet leicht nach, dass $Z_{1:n} \sim \text{Exp}(n)$, siehe Beispiel 5.5.3. Was passiert danach? Zum Zeitpunkt $Z_{1:n}$ sind $n - 1$ funktionierende Geräte im System geblieben, die wegen der Vergessenseigenschaft der Exponentialverteilung (siehe Abschnitt 5.5) genauso wie neu sind und nichts vom Ausfall des restlichen Geräts mitbekommen haben. Dieses System funktioniert also genauso wie das ursprüngliche System mit dem einzigen Unterschied, dass es nun am Anfang lediglich $n - 1$ Geräte gibt. Bis zum Ausfall des nächsten Geräts vergeht also die Zeit $Z_{2:n} - Z_{1:n} \sim \text{Exp}(n - 1)$. Danach startet das System mit $n - 2$ funktionierenden Geräten, usw. Zum Zeitpunkt $Z_{n-1:n}$ verbleibt ein einziges funktionierendes Gerät, das wegen der Vergessenseigenschaft genauso wie neu ist. Bis zum Ausfall dieses letzten Gerätes vergeht also die Zeit $Z_{n:n} - Z_{n-1:n} \sim \text{Exp}(1)$.

BEWEIS VON SATZ 7.3.1. Laut Satz 7.1.7 ist die Dichte von $(Z_{1:n}, \dots, Z_{n:n})$ gegeben durch

$$f(x_1, \dots, x_n) = \begin{cases} n! \cdot e^{-x_1} \cdot \dots \cdot e^{-x_n}, & 0 < x_1 < \dots < x_n, \\ 0, & \text{sonst.} \end{cases}$$

Im Rest des Beweises bestimmen wir die Dichte des Zufallsvektors auf der rechten Seite von (7.3.1). Wegen der Unabhängigkeit ist die Dichte von (E_1, \dots, E_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

$$J = \begin{vmatrix} \frac{1}{n} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{n} & \frac{1}{n-1} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{n} & \frac{1}{n-1} & \frac{1}{n-2} & \cdots & \frac{1}{2} & 0 \\ \frac{1}{n} & \frac{1}{n-1} & \frac{1}{n-2} & \cdots & \frac{1}{2} & 1 \end{vmatrix} = \frac{1}{n} \cdot \frac{1}{n-1} \cdots \frac{1}{2} \cdot 1 = \frac{1}{n!}.$$

Mit der Transformationsformel können wir die Dichte von $T(E_1, \dots, E_n)$ berechnen:

$$\begin{aligned} h(z_1, \dots, z_n) &= g(y_1, \dots, y_n) / J \\ &= n! e^{-(y_1 + \dots + y_n)} \mathbb{1}_{\{y_1 > 0, \dots, y_n > 0\}} \\ &= n! e^{-(z_1 + \dots + z_n)} \mathbb{1}_{\{0 < z_1 < \dots < z_n\}}. \end{aligned}$$

Die Dichte f von $(Z_{1:n}, \dots, Z_{n:n})$ stimmt also mit der Dichte h von $T(E_1, \dots, E_n)$ überein, was die Behauptung des Satzes beweist. \square

Aufgabe 7.3.2. Seien $Z_1, \dots, Z_n \sim \text{Exp}(1)$ unabhängig. Zeigen Sie mit Satz 7.3.1, dass

- (a) $\mathbb{E}[Z_{k:n}] = \frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{n-k+1}$.
- (b) $\text{Var}[Z_{k:n}] = \frac{1}{n^2} + \frac{1}{(n-1)^2} + \dots + \frac{1}{(n-k+1)^2}$.
- (c) $\text{Cov}(Z_{k:n}, Z_{l:n}) = \frac{1}{n^2} + \frac{1}{(n-1)^2} + \dots + \frac{1}{(n-k+1)^2}$ falls $1 \leq k \leq l \leq n$.

Aufgabe 7.3.3. Seien $Z_1, \dots, Z_n \sim \text{Exp}(1)$ unabhängig. Sei außerdem $k \in \{1, \dots, n-1\}$. Zeigen Sie:

$$(Z_{n-k+1:n} - Z_{n-k:n}, Z_{n-k+2:n} - Z_{n-k:n}, \dots, Z_{n:n} - Z_{n-k:n}) \stackrel{d}{=} (Z_{1:k}, Z_{2:k}, \dots, Z_{k:k}).$$

Ein Exkurs über die Gamma-Funktion. In diesem Abschnitt werden wir Satz 7.3.1 benutzen, um die folgende **Weierstraß-Formel** für die Gamma-Funktion $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, $x > 0$, auf probabilistische Art zu beweisen:

$$\frac{1}{\Gamma(x)} = x e^{\gamma x} \prod_{k=1}^{\infty} \left(1 + \frac{x}{k} \right) e^{-x/k}, \quad x > 0.$$

Dabei bezeichnet γ die **Euler-Mascheroni-Konstante**:

$$\gamma := \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k} - \log n \approx 0.5772.$$

Unser Ausgangspunkt ist Satz 1.2.1, der behauptet, dass für unabhängige, mit Parameter 1 exponentialverteilte Zufallsvariablen E_1, E_2, \dots

$$\max\{E_1, \dots, E_n\} - \log n \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}$$

gilt. Auf der anderen Seite folgt aus Satz 7.3.1 die Darstellung

$$E_{n:n} = \max\{E_1, \dots, E_n\} \stackrel{d}{=} \sum_{k=1}^n \frac{E_k}{k}.$$

Daraus ergibt sich

$$\sum_{k=1}^n \frac{E_k}{k} - \log n \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

Mit der Definition der Euler-Mascheroni-Konstante γ erhalten wir

$$(7.3.2) \quad \gamma + \sum_{k=1}^n \frac{E_k - 1}{k} \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

Nun ist es bekannt, dass eine Verteilungskonvergenz $X_n \rightarrow X$ im Allgemeinen *nicht* impliziert, dass $\mathbb{E}X_n \rightarrow \mathbb{E}X$ oder $\text{Var } X_n \rightarrow \text{Var } X$. Stellen wir uns aber für einen Augenblick vor, wir dürften in (7.3.2) zu den Erwartungswerten und Varianzen übergehen. Dann würden sich die folgenden schönen Formeln für die Gumbel-verteilte Zufallsvariable G ergeben:

$$(7.3.3) \quad \mathbb{E}G = \lim_{n \rightarrow \infty} \mathbb{E} \left[\gamma + \sum_{k=1}^n \frac{E_k - 1}{k} \right] = \gamma,$$

$$(7.3.4) \quad \text{Var } G = \lim_{n \rightarrow \infty} \text{Var} \left[\gamma + \sum_{k=1}^n \frac{E_k - 1}{k} \right] = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Dabei haben wir benutzt, dass $\mathbb{E}E_k = \text{Var } E_k = 1$. Weiterhin erscheint es plausibel, dass auch die Konvergenz der Laplace-Transformierten gelten sollte:

$$(7.3.5) \quad \lim_{n \rightarrow \infty} \mathbb{E} \exp \left\{ -z \left(\gamma + \sum_{k=1}^n \frac{E_k - 1}{k} \right) \right\} = \mathbb{E} \exp\{-zG\}, \quad z > 0.$$

Der Leser möge nachrechnen, dass

$$\mathbb{E} \exp\{-zG\} = \Gamma(z+1) = z\Gamma(z), \quad \mathbb{E} \exp \left\{ -z \sum_{k=1}^n \frac{E_k - 1}{k} \right\} = \prod_{k=1}^n \left(1 + \frac{z}{k} \right)^{-1} e^{-z/k}, \quad z > 0.$$

Durch Vergleich ergibt sich die gewünschte Weierstraß-Formel:

$$z\Gamma(z) = e^{-\gamma z} \prod_{k=1}^{\infty} \left(1 + \frac{z}{k} \right)^{-1} e^{z/k}, \quad z > 0.$$

Nun werden wir die obigen Überlegungen am Beispiel von (7.3.3) rechtfertigen. Es ist bekannt, dass die Konvergenz der Erwartungswerte $\mathbb{E}X_n \rightarrow \mathbb{E}X$ aus der Verteilungskonvergenz $X_n \rightarrow X$ zusammen mit der gleichgradigen Integrierbarkeit der Familie $(X_n)_{n \in \mathbb{N}}$ folgt. Die gleichgradige Integrierbarkeit liegt z.B. dann vor, wenn die Folge $\mathbb{E}X_n^2 = \text{Var } X_n + (\mathbb{E}X_n)^2$ beschränkt bleibt. In unserem Fall ist $\mathbb{E}X_n = \gamma$ konstant während die Beschränktheit von $\text{Var } X_n$ in (7.3.4) nachgerechnet wurde. Die Begründung von (7.3.4) und (7.3.5) überlassen wir dem Leser als Übung.

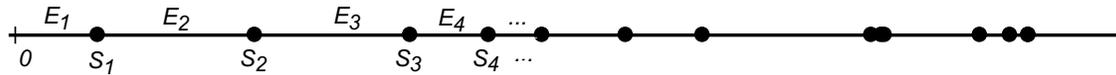
Aufgabe 7.3.4. Zeigen Sie unter Verwendung von (7.3.3) und (7.3.4), dass

$$\int_0^\infty (\log x)e^{-x} dx = \int_0^1 \log(-\log x) dx = -\gamma, \quad \int_0^\infty (\log x)^2 e^{-x} dx = \gamma^2 + \frac{\pi^2}{6}.$$

7.4. Ordnungsstatistiken der gleichverteilten Zufallsvariablen

Die Ordnungsstatistiken der Zufallsvariablen, die gemäß einer Gleichverteilung verteilt sind, besitzen ebenfalls eine schöne Darstellung.

Definition 7.4.1. Es seien E_1, E_2, \dots unabhängige, mit Parameter $\lambda > 0$ exponentialverteilte Zufallsvariablen und definiere die Partialsummen $S_k = E_1 + \dots + E_k$, $k \in \mathbb{N}$. Die Folge $S_1 < S_2 < \dots$ bildet einen **Poisson-Prozess** auf $(0, \infty)$ mit Intensität λ .



Grob gesagt behauptet der nachfolgende Satz, dass die ersten n Punkte S_1, \dots, S_n des Poisson-Prozesses auf dem Intervall $[0, S_{n+1}]$ gleichverteilt sind.

Satz 7.4.2. 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_1, S_2, \dots wie in der obigen Definition sind.

BEWEIS. SCHRITT 0. Laut Satz 7.1.7 ist die gemeinsame Dichte von $(U_{1:n}, \dots, U_{n:n})$ gegeben durch

$$f(x_1, \dots, x_n) = \begin{cases} n!, & 0 < x_1 < x_2 < \dots < x_n < 1, \\ 0, & \text{sonst.} \end{cases}$$

Im Folgenden bestimmen wir die Dichte von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$. Dieser Ausdruck ändert sich nicht, wenn man S_i überall durch λS_i ersetzt, was $\text{Exp}(1)$ -verteilt ist. Also werden wir im Folgenden den Parameter der Exponentialverteilung auf $\lambda = 1$ setzen.

SCHRITT 1. Die Dichte von (E_1, \dots, E_{n+1}) ist wegen der Unabhängigkeit gegeben durch:

$$g(y_1, \dots, y_{n+1}) = e^{-y_1} \mathbb{1}_{\{y_1 > 0\}} \cdot \dots \cdot e^{-y_{n+1}} \mathbb{1}_{\{y_{n+1} > 0\}}.$$

SCHRITT 2. Nun bestimmen wir die Dichte von $(S_1, S_2, \dots, S_{n+1})$. Dazu betrachten wir die lineare Transformation $T: \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$ mit

$$T: (y_1, \dots, y_{n+1}) \mapsto (y_1, y_1 + y_2, \dots, y_1 + y_2 + \dots + y_{n+1}) =: (z_1, z_2, \dots, z_{n+1}).$$

Die Jacobi-Determinante von T ist

$$J_T = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{vmatrix} = 1.$$

Mit der Transformationsformel errechnen wir die Dichte von $(S_1, S_2, \dots, S_{n+1})$:

$$\begin{aligned} h(z_1, \dots, z_{n+1}) &= g(y_1, \dots, y_{n+1})/J_T \\ &= e^{-(y_1 + \dots + y_{n+1})} \mathbb{1}_{\{y_1 > 0, \dots, y_{n+1} > 0\}} \\ &= e^{-z_{n+1}} \mathbb{1}_{\{0 < z_1 < z_2 < \dots < z_{n+1}\}}. \end{aligned}$$

SCHRITT 3. Nun bestimmen wir die Dichte von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}}, S_{n+1})$. Wir stellen diesen Zufallsvektor als $Q(S_1, S_2, \dots, S_{n+1})$ dar, wobei $Q : (0, \infty)^{n+1} \rightarrow (0, \infty)^{n+1}$ die nichtlineare Transformation

$$Q : (z_1, \dots, z_{n+1}) \mapsto \left(\frac{z_1}{z_{n+1}}, \frac{z_2}{z_{n+1}}, \dots, \frac{z_n}{z_{n+1}}, z_{n+1} \right) =: (w_1, w_2, \dots, w_n, w_{n+1})$$

sei. Dabei haben wir die letzte Koordinate S_{n+1} in den Zufallsvektor aufgenommen um die Transformation Q invertierbar zu machen. Die Jacobi-Determinante von Q ist

$$J_Q(z_1, \dots, z_{n+1}) = \begin{vmatrix} \frac{1}{z_{n+1}} & 0 & 0 & \cdots & 0 & -\frac{z_1}{z_{n+1}^2} \\ 0 & \frac{1}{z_{n+1}} & 0 & \cdots & 0 & -\frac{z_2}{z_{n+1}^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{z_{n+1}} & -\frac{z_n}{z_{n+1}^2} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{vmatrix} = \frac{1}{z_{n+1}^n} = \frac{1}{w_{n+1}^n}.$$

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:

$$\begin{aligned} p(w_1, \dots, w_{n+1}) &= h(z_1, \dots, z_{n+1})/J_Q(z_1, \dots, z_{n+1}) \\ &= e^{-z_{n+1}} z_{n+1}^n \mathbb{1}_{\{z_{n+1} > 0\}} \\ &= e^{-w_{n+1}} w_{n+1}^n \mathbb{1}_{\{0 < w_1 < w_2 < \dots < w_n < 1\}} \mathbb{1}_{\{w_{n+1} > 0\}}. \end{aligned}$$

SCHRITT 4. Nun müssen wir die letzte Koordinate S_{n+1} rausintegrieren. Die Dichte von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$ ist nämlich eine Marginaldichte von p :

$$r(w_1, \dots, w_n) = \int_{\mathbb{R}} p(w_1, \dots, w_{n+1}) dw_{n+1} = n! \mathbb{1}_{\{0 < w_1 < w_2 < \dots < w_n < 1\}}.$$

Man sieht, dass die Dichte f von $(U_{1:n}, \dots, U_{n:n})$ und die Dichte r von $(\frac{S_1}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$ übereinstimmen, woraus die Behauptung des Satzes folgt. \square

Aufgabe 7.4.3. Beweisen Sie die folgenden Ergänzungen zu Satz 7.4.2:

(a) S_{n+1} ist unabhängig vom Zufallsvektor $(\frac{S_1}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$.

Hinweis: Schritt 3 im Beweis von Satz 7.4.2.

(b) Die bedingte Verteilung des Zufallsvektors $(S_1/s, \dots, S_n/s)$ gegeben, dass $S_{n+1} = s > 0$ stimmt mit der Verteilung der Ordnungsstatistiken $(U_{1:n}, \dots, U_{n:n})$ überein.

Bemerkung: In der mathematischen Statistik beweist man (a) als einen Spezialfall des Satzes von Basu. Man betrachtet ein statistisches Modell, in dem E_1, \dots, E_{n+1} unabhängig und $\text{Exp}(\lambda)$ -verteilt mit einem unbekanntem Parameter $\lambda > 0$ sind. Man bemerkt, dass die Verteilung der Statistik $(\frac{S_1}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$ unabhängig von λ ist, und dass S_{n+1} vollständige und suffiziente Statistik ist. Daraus folgt die Unabhängigkeit.

Aufgabe 7.4.4 (Lücken zwischen Ordnungsstatistiken). Seien U_1, \dots, U_n unabhängig und gleichverteilt auf $[0, 1]$. Betrachte die maximale und die minimale Lücke zwischen den Ordnungsstatistiken:

$$\delta_n := \min_{i=0, \dots, n} (U_{i+1:n} - U_{i:n}), \quad \Delta_n := \max_{i=0, \dots, n} (U_{i+1:n} - U_{i:n}),$$

wobei $U_{0:n} := 0$ und $U_{n+1:n} := 1$ definiert wird. Zeigen Sie, dass

$$n^2 \delta_n \xrightarrow[n \rightarrow \infty]{d} \text{Exp}(1), \quad n \Delta_n - \log n \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

Hinweis: Benutzen Sie Satz 7.4.2 und das Lemma von Slutsky. Für S_{n+1} gilt das Gesetz der großen Zahlen und der zentrale Grenzwertsatz.

Aufgabe 7.4.5 (Empirische Verteilung der Lücken, Weiss, 1955). Seien U_1, \dots, U_n unabhängig und gleichverteilt auf $[0, 1]$. Zeigen Sie, dass für alle $x > 0$

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\{U_{k:n} - U_{k-1:n} > x/n\}} \xrightarrow[n \rightarrow \infty]{P} e^{-x}.$$

Aufgabe 7.4.6. Seien U_1, \dots, U_n unabhängig und gleichverteilt auf $[0, 1]$. Zeigen Sie, dass für alle $1 \leq k < m \leq n$, die Differenz $U_{m:n} - U_{k:n}$ eine Beta-Verteilung mit Parametern $(m - k, n - m + k + 1)$ besitzt.

Hinweis: Zeigen Sie, dass $U_{m:n} - U_{k:n} \stackrel{d}{=} U_{m-k:n}$ und benutzen Sie Beispiel 7.1.3.

Aufgabe 7.4.7. Seien U_1, \dots, U_n unabhängig und gleichverteilt auf $[0, 1]$. Sei außerdem $k \in \{1, \dots, n - 1\}$. Zeigen Sie:

$$\left(\frac{U_{1:n}}{U_{k+1:n}}, \frac{U_{2:n}}{U_{k+1:n}}, \dots, \frac{U_{k:n}}{U_{k+1:n}} \right) \stackrel{d}{=} (U_{1:k}, U_{2:k}, \dots, U_{k:k}).$$

Hinweis: Satz 7.4.2.

Mit der Darstellung aus Satz 7.4.2 lassen sich die Grenzwertverteilungen für die Ordnungsstatistiken $U_{k:n}$ für $n \rightarrow \infty$ herleiten. Wenn man dabei zulässt, dass $k = k_n$ eine Funktion von n sein darf, ergeben sich mehrere asymptotische Regime, die in den nachfolgenden Aufgaben aufgelistet werden.

Aufgabe 7.4.8 (Extreme Ordnungsstatistiken einer gleichverteilten Stichprobe). 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:

$$(nU_{1:n}, \dots, nU_{k:n}) \xrightarrow[n \rightarrow \infty]{d} (E_1, E_1 + E_2, \dots, E_1 + \dots + E_k),$$

wobei E_1, E_2, \dots unabhängige Zufallsvariablen mit $\mathbb{P}[E_i > t] = e^{-t}$, $t > 0$, sind (Standardexponentialverteilung). Zeigen Sie auch, dass

$$(n(1 - U_{n:n}), \dots, n(1 - U_{n-k+1:n})) \xrightarrow[n \rightarrow \infty]{d} (E_1, E_1 + E_2, \dots, E_1 + \dots + E_k).$$

Aufgabe 7.4.9 (Zentrale Ordnungsstatistiken einer gleichverteilten Stichprobe). Es seien U_1, \dots, U_n unabhängige, auf dem Intervall $[0, 1]$ gleichverteilte Zufallsvariablen. Sei $k_n \in \{1, \dots, n\}$ eine Folge mit $\lim_{n \rightarrow \infty} k_n/n = \alpha$, wobei $0 < \alpha < 1$. (Beispiel: $k_n = \lfloor \alpha n \rfloor$). Zeigen Sie, dass

$$\sqrt{n} \left(U_{k_n:n} - \frac{k_n}{n} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \alpha(1 - \alpha)),$$

wobei $\mathcal{N}(0, \sigma^2)$ eine Normalverteilung mit Erwartungswert 0 und Varianz σ^2 bezeichnet.

Aufgabe 7.4.10 (Intermediäre Ordnungsstatistiken einer gleichverteilten Stichprobe). Es seien U_1, \dots, U_n unabhängige, auf dem Intervall $[0, 1]$ gleichverteilte Zufallsvariablen. Sei $k_n \in \{1, \dots, n\}$ eine Folge mit $\lim_{n \rightarrow \infty} k_n/n = 0$. Zeigen Sie, dass

$$\frac{n}{\sqrt{k_n}} \left(U_{k_n:n} - \frac{k_n}{n} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1), \quad \frac{n}{\sqrt{k_n}} \left(U_{n-k_n:n} - \frac{n-k_n}{n} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Eine weitere Darstellung der Ordnungsstatistiken der gleichverteilten Zufallsvariablen wird in der folgenden Übungsaufgabe formuliert.

Aufgabe 7.4.11 (Lehmann, 1947, und Malmquist, 1953). Es seien U_1, \dots, U_n sowie T_1, \dots, T_n unabhängig und gleichverteilt auf $[0, 1]$.

(a) Zeigen Sie, dass

$$(U_{n:n}, U_{n-1:n}, \dots, U_{1:n}) \stackrel{d}{=} (T_n^{1/n}, T_n^{1/n} T_{n-1}^{1/(n-1)}, \dots, T_n^{1/n} T_{n-1}^{1/(n-1)} \dots T_1).$$

Hinweis: Die Zufallsvariablen $-\log U_i$ sind $\text{Exp}(1)$ -verteilt, was eine Verbindung zu den exponentialverteilten Ordnungsstatistiken herstellt. Die Funktion $u \mapsto -\log u$ ist monoton fallend.

(b) Zeigen Sie, dass die Zufallsvariablen $U_{2:n}/U_{1:n}, U_{3:n}/U_{2:n}, \dots, U_{n:n}/U_{n-1:n}$ unabhängig sind.

Aufgabe 7.4.12. Seien $S_1 < S_2 < \dots$ die Punkte eines homogenen Poisson-Prozesses wie in Definition 7.4.1. Zeigen Sie, dass die Zufallsvariablen $S_2/S_1, S_3/S_2, \dots$ unabhängig sind.

7.5. Asymptotische Unabhängigkeit des Maximums und des Minimums

Es seien X_1, X_2, \dots u.i.v. Zufallsvariablen. Die möglichen Grenzwertverteilungen des passend normierte Maximum $M_n = \max\{X_1, \dots, X_n\}$ wurden im Satz von Fisher-Tippett-Gnedenko beschrieben. Analog kann man auch die Grenzwertverteilungen für das Minimum $m_n = \min\{X_1, \dots, X_n\}$ beschreiben, siehe Aufgabe 3.3.2. Was sind aber die *gemeinsamen* Grenzwertverteilungen für das Paar (M_n, m_n) ? Der nächste Satz beantwortet diese Frage.

Satz 7.5.1 (Smirnov, 1949, und Homma, 1951). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit Verteilungsfunktion F , so dass

$$\frac{M_n - a_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} G \quad \text{und} \quad \frac{m_n - c_n}{d_n} \xrightarrow[n \rightarrow \infty]{d} H$$

für nichtdegenerierte Verteilungsfunktionen G und H . Dann gilt für alle $x, y \in \mathbb{R}$, dass

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n - a_n}{b_n} \leq x, \frac{m_n - c_n}{d_n} \leq y \right] = G(x)H(y).$$

INTERPRETATION. Das normierte Minimum und das normierte Maximum sind asymptotisch unabhängig für $n \rightarrow \infty$. Es sei bemerkt, dass M_n und m_n für jedes endliche n abhängig sind (es sei denn X_i ist degeneriert), was man z.B. an der Relation $m_n \leq M_n$ sieht. Satz 7.5.1 behauptet, dass diese Abhängigkeit für $n \rightarrow \infty$ "verschwindet".

BEWEIS. Es gilt

$$\begin{aligned} \mathbb{P} \left[\frac{M_n - a_n}{b_n} \leq x, \frac{m_n - c_n}{d_n} > y \right] &= \mathbb{P} [c_n + d_n y < X_i \leq a_n + b_n x \text{ für alle } i = 1, \dots, n] \\ &= (1 - \bar{F}(a_n + b_n x) - F(c_n + d_n y))^n \\ &= \left(1 + \frac{-n\bar{F}(a_n + b_n x) - nF(c_n + d_n y)}{n} \right)^n. \end{aligned}$$

Unsere Voraussetzungen sind

$$\lim_{n \rightarrow \infty} F^n(a_n + b_n x) = G(x) \quad \text{und} \quad \lim_{n \rightarrow \infty} \bar{F}^n(c_n + d_n y) = \bar{H}(y).$$

Daraus folgt mit der gleichen Begründung wie im Beweis von Satz 7.2.3, dass

$$\lim_{n \rightarrow \infty} n\bar{F}(a_n + b_n x) = -\log G(x) \quad \text{und} \quad \lim_{n \rightarrow \infty} nF(c_n + d_n y) = -\log \bar{H}(y).$$

Somit erhalten wir

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n - a_n}{b_n} \leq x, \frac{m_n - c_n}{d_n} > y \right] = e^{\log G(x) + \log \bar{H}(y)} = G(x)(1 - H(y)).$$

Schließlich ergibt sich

$$\begin{aligned} \mathbb{P} \left[\frac{M_n - a_n}{b_n} \leq x, \frac{m_n - c_n}{d_n} \leq y \right] &= \mathbb{P} \left[\frac{M_n - a_n}{b_n} \leq x \right] - \mathbb{P} \left[\frac{M_n - a_n}{b_n} \leq x, \frac{m_n - c_n}{d_n} > y \right] \\ &\xrightarrow[n \rightarrow \infty]{} G(x) - G(x)(1 - H(y)) = G(x)H(y), \end{aligned}$$

was die Behauptung beweist. □

Im nachfolgenden Satz werden wir uns die Grenzwertverteilungen für die *Spannweite* $R_n := M_n - m_n$ der Stichprobe X_1, \dots, X_n anschauen. Unser Ziel ist es, Normierungsfolgen $p_n \in \mathbb{R}$ und $q_n > 0$ derart zu finden, dass $(R_n - p_n)/q_n$ gegen eine nichtdegenerierte Verteilung konvergiert.

Satz 7.5.2. Es seien die Bedingungen von Satz 7.5.1 erfüllt. Außerdem sei (M, m) ein Zufallsvektor mit $\mathbb{P}[M \leq x, m \leq y] = G(x)H(y)$.

- (a) Im Fall, wenn $\lim_{n \rightarrow \infty} d_n/b_n = 0$, gilt $\frac{R_n - (a_n - c_n)}{b_n} \xrightarrow[n \rightarrow \infty]{d} M$.
- (b) Im Fall, wenn $\lim_{n \rightarrow \infty} d_n/b_n = +\infty$, gilt $\frac{R_n - (a_n - c_n)}{d_n} \xrightarrow[n \rightarrow \infty]{d} m$.
- (c) Im Fall, wenn $K := \lim_{n \rightarrow \infty} d_n/b_n \in (0, \infty)$, gilt $\frac{R_n - (a_n - c_n)}{d_n} \xrightarrow[n \rightarrow \infty]{d} M - Km$.

INTERPRETATION. Wir können die Verteilung von (M_n, m_n) durch die von $(a_n + b_n M, c_n + d_n m)$ approximieren, wobei Satz 7.5.1 garantiert, dass die beiden Komponenten von (M, m) unabhängig sind. Die Verteilung der Spannweite $M_n - m_n$ wird demnach wie folgt approximiert:

$$M_n - m_n \stackrel{d}{\approx} a_n - c_n + b_n M - d_n m.$$

Es bietet sich deshalb an, $M_n - m_n$ mit der Folge $p_n := a_n - c_n$ zu zentrieren. Die Wahl der Normierungsfolge q_n und die Grenzwertverteilung hängen allerdings davon ab, welche der beiden Folgen b_n oder d_n den asymptotischen Wettkampf gewinnt. So können wir im Fall (a) den Term $d_n m$ vernachlässigen, während im Fall (c) die Terme $b_n M$ und $d_n m$ die gleiche ‘‘Größenordnung’’ haben und beide berücksichtigt werden müssen.

BEWEIS VON SATZ 7.5.2. Im Fall (a) schreiben wir

$$\frac{(M_n - m_n) - (a_n - c_n)}{b_n} = \frac{M_n - a_n}{b_n} - \frac{m_n - c_n}{d_n} \cdot \frac{d_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} M - m \cdot 0 = M,$$

wobei wir die Relation $\lim_{n \rightarrow \infty} d_n/b_n = 0$ und das Lemma von Slutsky benutzt haben. Fälle (b) und (c) seien dem Leser als Übung überlassen. \square

Beispiel 7.5.3. Seien X_1, X_2, \dots unabhängig und exponentialverteilt mit Parameter 1. Dann gilt

$$M_n - \log n \xrightarrow[n \rightarrow \infty]{d} M, \quad nm_n \xrightarrow[n \rightarrow \infty]{d} m,$$

wobei M Gumbel-verteilt und $m \sim \text{Exp}(1)$ ist. Insbesondere folgt, dass m_n in Verteilung gegen 0 konvergiert. Für die Spannweite $R_n = M_n - m_n$ ergibt sich somit

$$R_n - \log n = (M_n - \log n) - m_n \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

Beispiel 7.5.4. Seien X_1, X_2, \dots unabhängig und gleichverteilt auf $[0, 1]$. Dann gilt

$$n(1 - M_n) \xrightarrow[n \rightarrow \infty]{d} E_1, \quad nm_n \xrightarrow[n \rightarrow \infty]{d} E_2,$$

wobei $E_1 \sim \text{Exp}(1)$ und $E_2 \sim \text{Exp}(1)$. Mit Satz 7.5.1 ergibt sich die gemeinsame Konvergenz

$$(n(1 - M_n), nm_n) \xrightarrow[n \rightarrow \infty]{d} (E_1, E_2),$$

wobei E_1 und E_2 unabhängig sind. Für die Spannweite $R_n = M_n - m_n$ folgt

$$n(1 - R_n) = n(1 - M_n) + nm_n \xrightarrow[n \rightarrow \infty]{d} E_1 + E_2,$$

wobei wir den Satz von der stetigen Abbildung benutzt haben.

Aufgabe 7.5.5 (Gumbel, 1947, und Cox, 1948). Seien X_1, X_2, \dots unabhängig und standardnormalverteilt. Zeigen Sie, dass für die Spannweite R_n und die sogenannte Stichprobenmitte $(M_n + m_n)/2$ gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sqrt{2 \log n} \left(R_n - \left(2\sqrt{2 \log n} - \frac{\log(4\pi \log n)}{\sqrt{2 \log n}} \right) \right) \leq x \right] = \int_{-\infty}^{+\infty} e^{-e^{y-x} - e^{-y}} dy,$$

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sqrt{8 \log n} \frac{M_n + m_n}{2} \leq x \right] = \frac{1}{1 + e^{-x}}.$$

Kommentare

Ordnungsstatistiken und ihre statistischen Anwendungen werden in vielen Büchern behandelt, s. z.B. [9, 3, 2]. Abschnitt 7.2 folgt im Wesentlichen der Darstellung Leadbetter, Lindgren und Rootzén [18, § 2.1-2.3] sowie der Originalarbeit von Weissman [?]. Wir haben uns nur Ordnungsstatistiken $M_n^{(k)}$ mit einem festen k angeschaut. Eine Übersicht über die Grenzwertverteilungen von $M_n^{(k_n)}$, wobei k_n von n abhängen darf, findet sich in [18, § 2.5-2.7] und in [21]. Abschnitte 7.3 und 7.4 basieren größtenteils auf den Büchern von Shorack und Wellner [24] (Chapter 8) und Nevzorov [20]. Das Buch von Schorack und Wellner [24] ist wohl die beste und umfangreichste Referenz zu empirischen Prozessen. Asymptotische Verteilung der Spannweite (Abschnitt 7.5) wird in [12, § 2.9] behandelt.

KAPITEL 8

Rekorde

Seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit stetiger Verteilungsfunktion F . Wir definieren $M_n = \max\{X_1, \dots, X_n\}$.

Definition 8.0.1. 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.

Insbesondere hängt die gemeinsame Verteilung von ξ_1, ξ_2, \dots nicht von der Verteilungsfunktion F ab (solange diese als stetig angenommen wird). Wir werden zwei Beweise des Satzes von Rényi präsentieren: einen direkten und einen viel eleganteren, kombinatorischen Beweis.

DIREKTER 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

$$\mathbb{P}[\xi_{\alpha(1)} = 1, \dots, \xi_{\alpha(s)} = 1, M_n < x] = \mathbb{P}[\xi_{\alpha(1)} = 1] \cdot \dots \cdot \mathbb{P}[\xi_{\alpha(s)} = 1] \cdot \mathbb{P}[M_n < x].$$

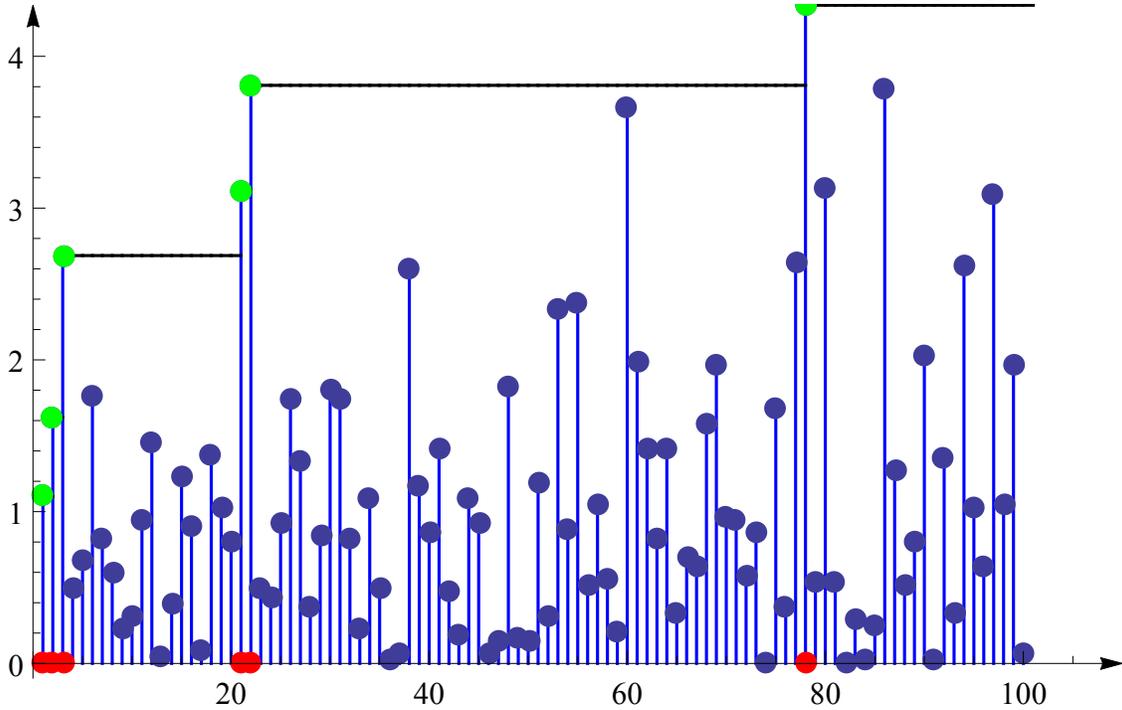


ABBILDUNG 1. Rekorde. Rote Punkte: Rekordzeiten $L(n)$. Grüne Punkte: Rekordwerte $X(n)$.

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

$$\mathbb{P}[\xi_\alpha = 1, M_n < x] = \mathbb{P}[M_{\alpha-1} < X_\alpha < x, M_{\alpha+1,n} < x] = \mathbb{P}[M_{\alpha-1} < X_\alpha < x](F(x))^{n-\alpha},$$

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

$$\mathbb{P}[\xi_\alpha = 1, M_n < x] = (F(x))^{n-\alpha} \cdot \int_0^{F(x)} w^{\alpha-1} dw = \frac{(F(x))^\alpha}{\alpha} = \mathbb{P}[M_n < x] \cdot \mathbb{P}[\xi_\alpha = 1],$$

was die Aussage im Fall $s = 1$ beweist.

Sei nun $s \in \mathbb{N}$ beliebig. Es gilt

$$\begin{aligned} \mathbb{P}[\xi_{\alpha(1)} = 1, \dots, \xi_{\alpha(s)} = 1, M_n < x] &= \mathbb{P}[M_{\alpha(1)-1} < X_{\alpha(1)} < x, M_{\alpha(1), \dots, \alpha(2)-1} < X_{\alpha(2)} < x, \\ &\quad \dots, M_{\alpha(s-1), \dots, \alpha(s)-1} < X_{\alpha(s)} < x, M_{\alpha(s)+1, n} < x]. \end{aligned}$$

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). □

Aufgabe 8.1.2 (F^α -Schema). Es sei F eine stetige Verteilungsfunktion und seien X_1, X_2, \dots unabhängige Zufallsvariablen mit $\mathbb{P}[X_k \leq t] = F^{\alpha(k)}(t)$, wobei $\alpha(1), \alpha(2), \dots$ eine Folge positiver Zahlen sei. Zeigen Sie, dass die Rekordindikatoren ξ_1, ξ_2, \dots unabhängig sind und dass

$$\mathbb{P}[\xi_n = 1] = \frac{\alpha(n)}{\alpha(1) + \dots + \alpha(n)}.$$

Zeigen Sie, dass ξ_1, \dots, ξ_n und M_n unabhängig sind. ??? Notation $\alpha(k)$ wurde im obigen Beweis benutzt.

Aufgabe 8.1.3 (Shorrock). Zeigen Sie, dass für u.i.v. Zufallsvariablen X_1, X_2, \dots mit beliebiger (möglicherweise nichtstetiger) Verteilungsfunktion die Gesamtzahl der Rekorde $\sum_{n=1}^{\infty} \xi_n$ genau dann endlich ist, wenn $\mathbb{P}[X_1 = x^*] = 0$.

KOMBINATORISCHER BEWEIS. Wir werden sehen, dass die u.i.v. Annahme im Satz von Rényi durch eine wesentlich schwächere Annahme, die Austauschbarkeit, ersetzt werden kann.

Definition 8.1.4. Ein Zufallsvektor (X_1, \dots, X_n) heißt **austauschbar**, wenn seine Verteilung invariant unter beliebigen Permutationen der Koordinaten bleibt, d.h. wenn

$$(X_1, \dots, X_n) \stackrel{d}{=} (X_{\sigma(1)}, \dots, X_{\sigma(n)})$$

für alle Permutationen $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$.

So wird zum Beispiel für $n = 3$ gefordert, dass alle 6 Vektoren

$$(X_1, X_2, X_3), (X_1, X_3, X_2), (X_2, X_1, X_3), (X_2, X_3, X_1), (X_3, X_1, X_2), (X_3, X_2, X_1)$$

die gleiche Verteilung besitzen.

Beispiel 8.1.5. Ein Vektor mit u.i.v. Komponenten ist austauschbar. Ein anderes Beispiel eines austauschbaren Vektors bietet eine zufällige Permutation der vorgegebenen deterministischen Zahlen a_1, \dots, a_n (wobei jede Permutation die gleiche Wahrscheinlichkeit $1/n!$ besitze). Noch ein Beispiel: Betrachte m Kugeln, die rein zufällig auf n Fächer verteilt werden, wobei alle Fächer gleichwahrscheinlich seien. Sei X_i die Anzahl der Kugeln im Fach i .

Dann ist der Vektor (X_1, \dots, X_n) austauschbar, und zwar sowohl im Modell mit Mehrfachbesetzungen (Multinomialverteilung) als auch im Modell ohne Mehrfachbesetzungen (zufällige m -elementige Teilmenge von $\{1, \dots, n\}$).

Definition 8.1.6. Der **Rang** von X_k unter den Beobachtungen X_1, \dots, X_k ist

$$R_k := \sum_{i=1}^k \mathbb{1}_{\{X_i \geq X_k\}} \in \{1, \dots, k\}.$$

Wir werden uns nur für den Fall interessieren, wenn in der Stichprobe keine **Bindungen** bestehen, d.h. wenn X_1, \dots, X_n mit Wahrscheinlichkeit 1 paarweise verschieden sind. In einem austauschbaren Zufallsvektor reicht es zu fordern, dass $\mathbb{P}[X_1 = X_2] = 0$, denn dann gilt auch $\mathbb{P}[X_i = X_j] = 0$ für alle $i \neq j$.

In einem Zufallsvektor ohne Bindungen gilt offenbar

$$\begin{aligned} R_k = 1 &\iff X_k \text{ ist das größte Element von } X_1, \dots, X_k, \\ R_k = 2 &\iff X_k \text{ ist das zweitgrößte Element von } X_1, \dots, X_k, \\ &\dots, \\ R_k = k &\iff X_k \text{ ist das kleinste Element von } X_1, \dots, X_k. \end{aligned}$$

Nun können wir die gemeinsame Verteilung von (R_1, \dots, R_n) bestimmen.

Satz 8.1.7. Sei (X_1, \dots, X_n) ein austauschbarer Zufallsvektor ohne Bindungen. Dann gilt

(a) Für jedes $k \in \{1, \dots, n\}$ ist R_k gleichverteilt auf $\{1, \dots, k\}$, d.h.

$$\mathbb{P}[R_k = 1] = \mathbb{P}[R_k = 2] = \dots = \mathbb{P}[R_k = k] = \frac{1}{k}.$$

(b) Die Zufallsvariablen R_1, R_2, \dots, R_n sind unabhängig.

BEWEIS. Es gibt $n!$ verschiedene mögliche Anordnungen von X_1, \dots, X_n . Für $n = 3$ gibt es z.B. die 6 Anordnungen

$$\begin{aligned} &\{X_1 < X_2 < X_3\}, \{X_1 < X_3 < X_2\}, \{X_2 < X_1 < X_3\}, \\ &\{X_2 < X_3 < X_1\}, \{X_3 < X_1 < X_2\}, \{X_3 < X_1 < X_2\}. \end{aligned}$$

Die Anordnungen sind wegen Austauschbarkeit gleichwahrscheinlich. Außerdem wird wegen Fehlen der Bindungen eine der Anordnungen mit Wahrscheinlichkeit 1 beobachtet. Somit ist die Wahrscheinlichkeit jeder Anordnung $1/n!$.

Auf der anderen Seite gibt es $n!$ verschiedene mögliche Werte des Vektors (R_1, \dots, R_n) , denn $R_1 = 1, R_2 \in \{1, 2\}, \dots, R_n \in \{1, \dots, n\}$. Jede Anordnung der X_i 's entspricht einem Wert

von (R_1, \dots, R_n) auf eine bijektive Art und Weise. Zum Beispiel gilt

$$\begin{aligned} X_1 < X_2 < X_3 &\iff R_1 = 1, R_2 = 1, R_3 = 1, \\ X_1 < X_3 < X_2 &\iff R_1 = 1, R_2 = 1, R_3 = 2, \\ X_2 < X_1 < X_3 &\iff R_1 = 1, R_2 = 2, R_3 = 1, \end{aligned}$$

usw. In der Tat, es ist klar, dass die Anordnung die Ränge eindeutig bestimmt. Nun möchten wir den Leser davon überzeugen, dass auch umgekehrt die Ränge die Anordnung eindeutig bestimmen. Man kann z.B. induktiv argumentieren. Für $k = 1$ gibt es nur eine Anordnung und die Behauptung gilt trivialerweise. Stellen wir uns vor, dass uns die Ränge (R_1, \dots, R_k) bekanntgegeben wurden und wir dadurch eindeutig die Anordnung von X_1, \dots, X_k bestimmt haben, z.B. für $k = 4$

$$X_3 < X_2 < X_1 < X_4.$$

Nun wird uns R_{k+1} , also der Rang von X_{k+1} unter (X_1, \dots, X_{k+1}) mitgeteilt, z.B. $R_5 = 3$. Dadurch können wir die Position eindeutig bestimmen, an der X_5 in die obige Reihe eingefügt werden muss. In unserem Beispiel muss X_5 das drittgrößte Element sein, also erhalten wir die Anordnung

$$X_3 < X_2 < X_5 < X_1 < X_4,$$

womit wir den Induktionsschritt begründet haben.

Es besteht also eine Bijektion zwischen den Werten von (R_1, \dots, R_n) und Anordnungen von X_1, \dots, X_n . Da alle Anordnungen gleichwahrscheinlich sind, müssen auch die Werte von (R_1, \dots, R_n) gleichwahrscheinlich sein. Somit gilt

$$\mathbb{P}[R_1 = r_1, R_2 = r_2, \dots, R_n = r_n] = \frac{1}{n!},$$

für alle $r_1 = 1, r_2 \in \{1, 2\}, \dots, r_n \in \{1, \dots, n\}$. Damit sind beide Behauptungen des Satzes bewiesen. \square

Nun sind wir imstande, den Satz von Rényi auf austauschbare Stichproben zu verallgemeinern.

Satz 8.1.8 (Rényi, 1962, und Dwass, 1960). Es sei (X_1, \dots, X_n) ein austauschbarer Zufallsvektor ohne Bindungen. Definiere ξ_k als die Indikatorvariable des Rekordereignisses $X_k > \max\{X_1, \dots, X_{k-1}\}$. Dann gilt $\mathbb{P}[\xi_k = 1] = \frac{1}{k}$ für alle $k \in \{1, \dots, n\}$ und die Zufallsvariablen ξ_1, \dots, ξ_n sind unabhängig.

BEWEIS. Es gilt $\xi_k = 1$ genau dann, wenn X_k das maximale Element unter X_1, \dots, X_k ist. Somit sind die Ereignisse $\{\xi_k = 1\}$ und $\{R_k = 1\}$ gleich. Die Wahrscheinlichkeit von $\{R_k = 1\}$ beträgt $1/k$ und diese Ereignisse sind unabhängig gemäß Satz 8.1.7. \square

Aufgabe 8.1.9. Seien X_1, \dots, X_n austauschbare Zufallsvariablen ohne Bindungen. Zeigen Sie, dass der Vektor der Rekordindikatoren (ξ_1, \dots, ξ_n) unabhängig vom Vektor der Ordnungsstatistiken $(X_{1:n}, \dots, X_{n:n})$ ist. Zum Beispiel ist (ξ_1, \dots, ξ_n) von M_n , aber auch vom Mittelwert $(X_1 + \dots + X_n)/n$ unabhängig.

Aufgabe 8.1.10 (Zweiseitige Rekorde). Seien X_1, \dots, X_n austauschbare Zufallsvariablen ohne Bindungen. Definiere $\eta_k = 1$, wenn $X_k = \max\{X_1, \dots, X_k\}$ oder $X_k = \min\{X_1, \dots, X_k\}$. Zeigen Sie, dass η_2, \dots, η_n unabhängige Zufallsvariablen mit $\mathbb{P}[\eta_k = 1] = 2/k$ sind.

Aufgabe 8.1.11. Es sei (X_1, \dots, X_{n+m}) ein austauschbarer Zufallsvektor ohne Bindungen, wobei $n, m \in \mathbb{N}$. Bestimmen Sie

$$\mathbb{P}[\max\{X_1, \dots, X_n\} \geq \max\{X_{n+1}, \dots, X_{n+m}\}].$$

Hinweis: Es sei I die Zufallsvariable mit $X_I = \max\{X_1, \dots, X_{n+m}\}$. Wie ist I verteilt?

Aufgabe 8.1.12. Es sei (X_1, \dots, X_{n+1}) ein austauschbarer Zufallsvektor ohne Bindungen. Zeigen Sie: Für jedes der Intervalle $(-\infty, X_{1:n}), (X_{1:n}, X_{2:n}), \dots, (X_{n:n}, +\infty)$ ist die Wahrscheinlichkeit, dass X_{n+1} in diesem Intervall liegt, gleich $1/(n+1)$.

Aufgabe 8.1.13 (k -te Rekorde, Dziubdziela und Kopocinsky (1976)). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion. Für ein festes $k \in \mathbb{N}$ definiere die Indikatoren der k -ten Rekorde durch

$$\xi_n(k) = \mathbb{1}_{\{X_n > M_{n-1}^{(k)}\}}, \quad n \in \{k, k+1, \dots\}.$$

Für $k = 1$ sind es die gewöhnlichen Rekordindikatoren: $\xi_n(1) = \xi_n$. Zeigen Sie: $\xi_k(k), \xi_{k+1}(k), \dots$ sind unabhängig und $\mathbb{P}[\xi_n(k) = 1] = k/n$.

Ein Exkurs über die Gumbel-Überschreitungsmethode. 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 8.1.14. 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. ???

□

Beispiel 8.1.15. 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.

Aufgabe 8.1.16 (Wilks, 1959). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion. Für ein $n \in \mathbb{N}$ sei $K = \min\{k \in \mathbb{N} : X_{n+k} > M_n\}$ die minimale Anzahl zusätzlicher Beobachtungen nach Zeitpunkt n , die nötig sind, um das Maximum M_n zu übertreffen. Zeigen Sie:

$$\mathbb{P}[K = k] = ???.$$

8.2. Anzahl der Rekorde

Sei (X_1, \dots, X_n) ein austauschbarer Zufallsvektor ohne Bindungen (z.B. können X_1, \dots, X_n u.i.v. mit stetiger Verteilungsfunktion sein). Wir bezeichnen mit $N(n)$ die **Anzahl der Rekorde** in diesem Vektor, d.h.

$$N(n) = \sum_{k=1}^n \xi_k.$$

Nach dem Satz von Rényi gilt

$$\mathbb{E}N(n) = \sum_{k=1}^n \mathbb{E}\xi_k = \sum_{k=1}^n \frac{1}{k} = \log n + \gamma + o(1),$$

$$\text{Var } N(n) = \sum_{k=1}^n \text{Var } \xi_k = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k^2} \right) = \log n + \gamma - \frac{\pi^2}{6} + o(1),$$

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 n Beobachtungen lediglich approximativ $\log n$ Rekorde erwartet (was sehr wenig ist!)

Um die komplette Verteilung von $N(n)$ zu beschreiben, müssen wir Stirling-Zahlen erster Art einführen.

Definition 8.2.1. Die **Stirling-Zahlen erster Art** $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ sind definiert als Koeffizienten in der Formel

$$x(x+1)\dots(x+n-1) = \sum_{k=1}^n \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] x^k.$$

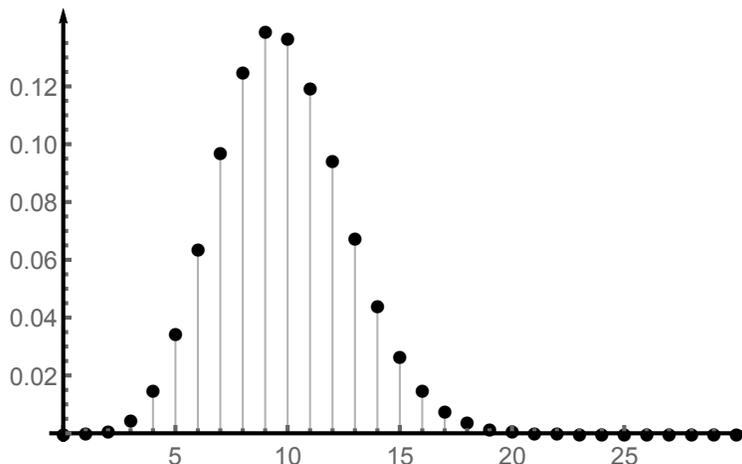


ABBILDUNG 2. Zähldichte der Zufallsvariable $N(n)$ für $n = 10000$.

Satz 8.2.2. Für die Verteilung der Anzahl der Rekorde gilt

$$\mathbb{P}[N(n) = k] = \frac{1}{n!} \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right], \quad k = 1, \dots, n.$$

Setzt man $x = 1$ in die Definition der Stirling-Zahlen ein, so erhält man

$$\sum_{k=1}^n \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] = n!.$$

Also summieren sich die Wahrscheinlichkeiten in Satz 8.2.2 tatsächlich zu 1.

BEWEIS VON SATZ 8.2.2. Definitionsgemäß gilt $N(n) = \xi_1 + \dots + \xi_n$. Dabei sind die Rekordindikatoren ξ_1, \dots, ξ_n nach dem Satz von Rényi unabhängig und Bernoulli-verteilt mit

$$\mathbb{P}[\xi_k = 1] = \frac{1}{k}, \quad \mathbb{P}[\xi_k = 0] = 1 - \frac{1}{k}.$$

Im Folgenden bestimmen wir die erzeugende Funktion von $N(n)$. Wir erinnern, dass die erzeugende Funktion einer Zufallsvariable Z mit Werten in $\{0, 1, \dots\}$ durch

$$g_Z(t) = \mathbb{E}[t^Z] = \sum_{k=0}^{\infty} \mathbb{P}[Z = k] t^k, \quad |t| < 1,$$

definiert ist. Die erzeugende Funktion von ξ_k ist

$$g_{\xi_k}(t) = \left(1 - \frac{1}{k}\right) + \frac{t}{k}.$$

Wegen der Unabhängigkeit der Summanden in der Darstellung $N(n) = \xi_1 + \dots + \xi_n$ gilt

$$g_{N(n)}(t) = g_{\xi_1}(t) \cdot \dots \cdot g_{\xi_n}(t) = \prod_{k=1}^n \frac{k-1+t}{k} = \frac{t(t+1)\dots(t+n-1)}{n!} = \frac{1}{n!} \sum_{k=1}^n \begin{bmatrix} n \\ k \end{bmatrix} t^k,$$

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 8.2.3. 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 gilt für Binomialkoeffizienten eine ähnliche Formel:

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

Aufgabe 8.2.4. Beweisen Sie für die Stirling-Zahl erster Art die Formel

$$\frac{1}{n!} \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} = \sum_{i_1 \dots i_k} \frac{1}{i_1 \dots i_k},$$

wobei über alle ganzzahligen $1 \leq i_1 < i_2 < \dots < i_k \leq n$ summiert wird.

Aufgabe 8.2.5. Zeigen Sie, dass für alle $k \in \mathbb{N}$,

$$\sum_{n=k}^{\infty} \frac{z^n}{n!} \begin{bmatrix} n \\ k \end{bmatrix} = \frac{1}{k!} \left(\log \frac{1}{1-z} \right)^k, \quad |z| < 1.$$

Ein Exkurs über die Foata-Korrespondenz. Wir möchten nun kurz auf einen interessanten Zusammenhang zwischen der Anzahl der Rekorde und der Anzahl der Zyklen in einer rein zufälligen Permutation eingehen. Wir werden nämlich zeigen, dass beide Zufallsgrößen dieselbe Verteilung besitzen.

Eine Permutation ist eine bijektive Abbildung $\varphi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$. Insgesamt gibt es $n!$ Permutationen. Wir schreiben eine Permutation in der Form

$$\varphi = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \varphi(1) & \varphi(2) & \varphi(3) & \dots & \varphi(n) \end{pmatrix}.$$

Wir sagen, dass Elemente $i_1, \dots, i_k \in \{1, \dots, n\}$ einen **Zyklus** bilden, falls

$$\varphi(i_1) = i_2, \varphi(i_2) = i_3, \dots, \varphi(i_k) = i_1.$$

Dieser Zyklus wird dann mit (i_1, \dots, i_k) bezeichnet. Jede Permutation kann als ein Produkt von disjunkten Zyklen dargestellt werden, z.B.

$$\varphi := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 9 & 5 & 8 & 2 & 7 & 6 & 10 & 1 & 4 \end{pmatrix} = (1, 3, 5, 2, 9)(4, 8, 10)(6, 7).$$

Eine **rein zufällige Permutation** ist eine Zufallsvariable Φ mit Werten in der Menge aller Permutationen und mit der Eigenschaft, dass $\mathbb{P}[\Phi = \varphi] = 1/n!$ für jede Permutation φ .

Satz 8.2.6. Es sei $Z(n)$ die Anzahl der Zyklen in einer rein zufälligen Permutation Φ von $\{1, \dots, n\}$. Dann gilt

$$\mathbb{P}[Z(n) = k] = \frac{1}{n!} \begin{bmatrix} n \\ k \end{bmatrix}, \quad k = 1, \dots, n.$$

BEWEIS. Sei Φ eine rein zufällige Permutation von $\{1, \dots, n\}$. Dann ist der Zufallsvektor $(\Phi(1), \dots, \Phi(n))$ austauschbar und für die Anzahl $N(n)$ der Rekorde in diesem Zufallsvektor gilt gemäß Satz 8.2.2

$$\mathbb{P}[N(n) = k] = \frac{1}{n!} \begin{bmatrix} n \\ k \end{bmatrix}, \quad k = 1, \dots, n.$$

Wir müssen also zeigen, dass $Z(n)$ die gleiche Verteilung wie $N(n)$ hat. Dazu werden wir uns der sogenannten **Foata-Korrespondenz** bedienen. Diese Korrespondenz ist eine bijektive Abbildung F der Menge aller Permutationen auf sich selbst mit der Eigenschaft, dass die Anzahl der Rekorde einer beliebigen Permutation φ mit der Anzahl der Zyklen der Permutation $F(\varphi)$ übereinstimmt. Wegen der bijektiven Eigenschaft von F wäre dann die Anzahl der Permutationen mit k Rekorde die gleiche wie die Anzahl der Permutationen mit k Zyklen, woraus sich $\mathbb{P}[Z(n) = k] = \mathbb{P}[N(n) = k]$ ergeben würde.

Die Foata-Korrespondenz wird wie folgt konstruiert. Betrachte die Zyklenzerlegung einer Permutation φ . Diese ist zunächst nicht eindeutig: man kann die Elemente innerhalb jedes Zyklus zyklisch permutieren. Um Eindeutigkeit zu erreichen, möchten wir jeden Zyklus mit seinem größten Element beginnen lassen, so dass die Zerlegung im obigen Beispiel wie folgt aussieht:

$$(9, 1, 3, 5, 2)(10, 4, 8)(7, 6).$$

Aber auch jetzt ist die Darstellung nicht eindeutig, denn man kann die Zyklen beliebig vertauschen. Deshalb ordnen wir die Zyklen aufsteigend nach ihrem größten Element an:

$$(7, 6)(9, 1, 3, 5, 2)(10, 4, 8).$$

Mit diesen Konventionen ist die Darstellung eindeutig geworden. Nun entfernen wir die Klammern und fassen den so konstruierten Tupel als eine neue Permutation auf, die mit $F(\varphi)$ bezeichnet wird:

$$F : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 9 & 5 & 8 & 2 & 7 & 6 & 10 & 1 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 6 & 9 & 1 & 3 & 5 & 2 & 10 & 4 & 8 \end{pmatrix}$$

Wir sehen nun, dass jeder Zyklus der Permutation φ einem Rekord in der Permutation $F(\varphi)$ entspricht. Diese Beobachtung zeigt außerdem, dass F eine Bijektion ist, denn wir können die

inverse Abbildung F^{-1} wie folgt konstruieren. Sei eine Permutation ψ gegeben. Die Rekorde im Vektor $(\psi(1), \dots, \psi(n))$ zerlegen diesen in Blöcke. Nun konstruieren wir eine Permutation φ , in der diese Blöcke zu Zyklen erklärt werden. Dann gilt $F(\varphi) = \psi$ und die Anzahl der Zyklen von φ entspricht der Anzahl der Rekorde von ψ . \square

Bemerkung 8.2.7. Wir haben gezeigt: Die Anzahl der Permutationen von $\{1, \dots, n\}$ mit k Zyklen (oder auch mit k Rekorde) ist gegeben durch die Stirling-Zahl $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$.

8.3. Rekordzeiten

Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion. 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; siehe Abbildung 1. 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:

$$\begin{aligned} \mathbb{P}[L(1) = j(1), \dots, L(n) = j(n)] &= \mathbb{P}[\xi_1 = 1, \xi_2 = \xi_3 = \dots = \xi_{j(2)-1} = 0, \\ &\quad \xi_{j(2)} = 1, \xi_{j(2)+1} = \dots = \xi_{j(3)-1} = 0, \\ &\quad \dots, \\ &\quad \xi_{j(n)} = 1]. \end{aligned}$$

Wegen der in Satz 8.1.1 bewiesenen Unabhängigkeit von ξ_1, ξ_2, \dots 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

Die gemeinsame Verteilung der Rekordzeiten $L(1), L(2), \dots$ ist somit unabhängig von der Verteilung der Zufallsvariablen X_1, X_2, \dots solange die Verteilungsfunktion F als stetig angenommen wird.

Bemerkung 8.3.2. Die Verteilung von $L(2)$ sieht somit folgendermaßen aus:

$$\mathbb{P}[L(2) = j] = \frac{1}{j(j-1)}, \quad j \in \{2, 3, \dots\}.$$

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! Dabei ist $L(2)$ eine *fast sicher* endliche Zufallsvariable.

Aufgabe 8.3.3 (Verteilung von $L(n)$). Zeigen Sie, dass für jedes $n \in \{2, 3, \dots\}$

$$\mathbb{P}[L(n) = k] = \frac{1}{k!} \binom{n-1}{k-1}, \quad k \in \{n, n+1, \dots\}.$$

Hinweis: Satz 8.2.2.

Im nächsten Satz zeigen wir, dass die Rekordzeiten $L(1), L(2), \dots$ eine Markov-Kette bilden. Für eine kurze Einführung in Markov-Ketten verweisen wir den Leser auf ????. An dieser Stelle erinnern wir, dass die endlich-dimensionalen Verteilungen einer Markov-Kette $\eta(1), \eta(2), \dots$ mit Zustandsraum E , Anfangszustand $e \in E$ und Übergangswahrscheinlichkeiten $(p_{ij})_{i,j \in E}$ durch

$$\mathbb{P}[\eta(1) = i(1), \eta(2) = i(2), \dots, \eta(n) = i(n)] = \mathbb{1}_{\{i(1)=e\}} p_{i(1)i(2)} \cdots p_{i(n-1)i(n)}$$

für alle Zustände $i(1), \dots, i(n) \in E$ und alle $n \in \mathbb{N}$ gegeben sind.

Satz 8.3.4. Die Folge $L(1), L(2), \dots$ ist eine Markov-Kette mit Zustandsraum \mathbb{N} , Anfangszustand $L(1) = 1$ und Übergangswahrscheinlichkeiten

$$p_{ij} := \frac{i}{j(j-1)}$$

für alle $i = 1, 2, \dots$ und $j = i+1, i+2, \dots$. Für $i \geq j$ ist $p_{ij} = 0$.

BEWEIS. Zu zeigen ist, dass für alle $1 = j(1), j(2), \dots, j(n) \in \mathbb{N}$ gilt

$$(8.3.1) \quad \mathbb{P}[L(1) = j(1), L(2) = j(2), \dots, L(n) = j(n)] = p_{j(1)j(2)} p_{j(2)j(3)} \cdots p_{j(n-1)j(n)}.$$

Dabei können wir $1 = j(1) < j(2) < \dots < j(n)$ annehmen, denn sonst sind beide Terme gleich 0 wegen $L(1) < L(2) < \dots < L(n)$ und $p_{ij} = 0$ für $i \geq j$. Mit der Formel aus Satz 8.3.1 gilt

$$\mathbb{P}[L(1) = j(1), L(2) = j(2), \dots, L(n) = j(n)] = \frac{1}{j(n)(j(2)-1) \cdots (j(n)-1)}.$$

Auf der anderen Seite rechnet man mit der Definition von p_{ij} nach, dass

$$\begin{aligned} p_{j(1)j(2)} p_{j(2)j(3)} \cdots p_{j(n-1)j(n)} &= \frac{j(1)}{j(2)(j(2)-1)} \cdot \frac{j(2)}{j(3)(j(3)-1)} \cdots \frac{j(n-1)}{j(n)(j(n)-1)} \\ &= \frac{1}{j(n)(j(2)-1) \cdots (j(n)-1)}. \end{aligned}$$

Die gewünschte Formel (8.3.1) ergibt sich aus dem Vergleich der beiden Ergebnisse. \square

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.

Korollar 8.3.5. Für alle $1 = i(1) < i(2) < \dots < i(n) = i < j$ gilt die Markov-Eigenschaft der Rekordzeiten:

$$\mathbb{P}[L(n+1) = j \mid L(n) = i] = \mathbb{P}[L(n+1) = j \mid L(n) = i, L(n-1) = i(n-1), \dots, L(2) = i(2)].$$

Außerdem gilt:

$$\mathbb{P}[L(n+1) = j \mid L(n) = i] = \frac{i}{j(j-1)}.$$

Aufgabe 8.3.6. Zeigen Sie: $\mathbb{P}[L(n+1) > j \mid L(n) = i] = i/j$ für alle $1 \leq i < j$.

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.

Aufgabe 8.3.7 (Rekordzeiten im F^α -Schema). Es sei F eine stetige Verteilungsfunktion und seien X_1, X_2, \dots unabhängige Zufallsvariablen mit $\mathbb{P}[X_k \leq t] = F^{\alpha(k)}(t)$, wobei $\alpha(1), \alpha(2), \dots$ eine Folge positiver Zahlen sei. Zeigen Sie, dass die Rekordzeiten $L(1), L(2), \dots$ eine Markov-Kette mit Übergangswahrscheinlichkeiten

$$\mathbb{P}[L(n+1) = j \mid L(n) = i] = S(i) \left(\frac{1}{S(j-1)} - \frac{1}{S(j)} \right), \quad 1 \leq i < j,$$

bilden, wobei $S(k) = \alpha(1) + \dots + \alpha(k)$.

Der nächste Satz gibt einen viel einfacheren Algorithmus zur Simulation der Rekordzeiten. Definiere die Abrundungsfunktion (oder die Gauß-Klammer) durch

$$\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}, \quad x \in \mathbb{R}.$$

Satz 8.3.8 (Williams, 1973). Seien U_1, U_2, \dots unabhängig und gleichverteilt auf dem Intervall $[0, 1]$. Definiere $R(1) = 1$ und danach rekursiv $R(n+1) = \lfloor \frac{R(n)}{U_n} \rfloor + 1$ für alle $n \in \mathbb{N}$. 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

$$\mathbb{P}[L(n + 1) = j \mid L(n) = i] = \mathbb{P}[R(n + 1) = j \mid R(n) = i].$$

Die Wahrscheinlichkeit auf der linken Seite ist gleich $\frac{i}{j(j-1)}$ nach Korollar 8.3.5. Wir berechnen die Wahrscheinlichkeit auf der rechten Seite. Es ist

$$\mathbb{P}[R(n + 1) = j \mid R(n) = i] = \mathbb{P}\left[\left\lfloor \frac{R(n)}{U_n} \right\rfloor + 1 = j \mid R(n) = i\right] = \mathbb{P}\left[\left\lfloor \frac{i}{U_n} \right\rfloor + 1 = j \mid R(n) = i\right].$$

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:

$$\begin{aligned} \mathbb{P}[R(n + 1) = j \mid R(n) = i] &= \mathbb{P}\left[\left\lfloor \frac{i}{U_n} \right\rfloor = j - 1\right] = \mathbb{P}\left[\frac{i}{U_n} \in [j - 1, j)\right] \\ &= \mathbb{P}\left[U_n \in \left(\frac{i}{j}, \frac{i}{j - 1}\right]\right] = \frac{i}{j(j - 1)}, \end{aligned}$$

da U_n gleichverteilt auf $[0, 1]$ ist. □

Aufgabe 8.3.9 (Galambos und Seneta, 1975 und Westcott, 1977). Zeigen Sie, dass die Zufallsvariablen

$$T(n) := \left\lfloor \frac{L(n + 1)}{L(n)} \right\rfloor + 1, \quad n \in \mathbb{N},$$

unabhängig und identisch verteilt sind mit $\mathbb{P}[T(n) = k] = \frac{1}{k(k-1)}$, $k \in \{2, 3, \dots\}$.

Satz 8.3.10 (Tata, 1969). Ist $x > 1$ eine natürliche Zahl, so gilt für jedes $n \in \mathbb{N}$:

$$\mathbb{P}\left[\frac{L(n + 1)}{L(n)} > x\right] = \frac{1}{x}.$$

Für eine beliebige reelle Zahl $x > 1$ gilt lediglich

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\frac{L(n + 1)}{L(n)} > x\right] = \frac{1}{x}.$$

Bemerkung 8.3.11. 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, sollte nicht überraschend sein: die Rekordwerte steigen nämlich mit der Zeit und es wird immer schwieriger, neue Rekorde aufzustellen.

BEWEIS. Sei $x > 1$ beliebig reell. Dann gilt wegen des Gesetzes der totalen Wahrscheinlichkeit, dass

$$\begin{aligned}\mathbb{P}[L(n+1) > xL(n)] &= \sum_{i=n}^{\infty} \mathbb{P}[L(n+1) > xi \mid L(n) = i] \cdot \mathbb{P}[L(n) = i] \\ &= \sum_{i=n}^{\infty} \mathbb{P}[L(n+1) > \lfloor xi \rfloor \mid L(n) = i] \cdot \mathbb{P}[L(n) = i],\end{aligned}$$

denn $L(n+1)$ ist ganzzahlig und somit ist $L(n+1) > xi$ genau dann, wenn $L(n+1) > \lfloor xi \rfloor$. Aus Korollar 8.3.5 folgt, dass für alle $i \in \mathbb{N}$ und $k \in \{i, i+1, \dots\}$

$$\mathbb{P}[L(n+1) > k \mid L(n) = i] = \sum_{j=k+1}^{\infty} \frac{i}{j(j-1)} = i \sum_{j=k+1}^{\infty} \left(\frac{1}{j-1} - \frac{1}{j} \right) = \frac{i}{k}.$$

Unter Verwendung dieses Ergebnisses lässt sich der obige Ausdruck zu folgendem vereinfachen:

$$(8.3.2) \quad \mathbb{P}[L(n+1) > xL(n)] = \sum_{i=n}^{\infty} \frac{i}{\lfloor xi \rfloor} \mathbb{P}[L(n) = i].$$

FALL 1. Ist $x \in \{2, 3, \dots\}$ ganzzahlig, so ist $\lfloor xi \rfloor = xi$ und somit erhalten wir

$$\mathbb{P}[L(n+1) > xL(n)] = \sum_{i=n}^{\infty} \frac{i}{\lfloor xi \rfloor} \mathbb{P}[L(n) = i] = \frac{1}{x} \sum_{i=n}^{\infty} \mathbb{P}[L(n) = i] = \frac{1}{x},$$

da $L(n) \geq n$. Das beweist die erste Behauptung des Satzes.

FALL 2. Sei nun $x > 1$ beliebig reell. Aus der Definition der Abrundungsfunktion $\lfloor \cdot \rfloor$ folgt, dass $\lfloor xi \rfloor \leq xi < \lfloor xi \rfloor + 1$. Durch leichte Umformungen folgt:

$$\frac{1}{x} \leq \frac{i}{\lfloor xi \rfloor} < \frac{1}{x} + \frac{1}{\lfloor xi \rfloor x}.$$

Damit kann man die Summe in (8.3.2) wie folgt nach unten abschätzen:

$$\mathbb{P}[L(n+1) > xL(n)] = \sum_{i=n}^{\infty} \frac{i}{\lfloor xi \rfloor} \mathbb{P}[L(n) = i] \geq \frac{1}{x} \sum_{i=n}^{\infty} \mathbb{P}[L(n) = i] = \frac{1}{x}.$$

Außerdem erhält man die folgende Abschätzung nach unten:

$$\mathbb{P}[L(n+1) > xL(n)] = \sum_{i=n}^{\infty} \frac{i}{\lfloor xi \rfloor} \mathbb{P}[L(n) = i] < \sum_{i=n}^{\infty} \left(\frac{1}{x} + \frac{1}{\lfloor xi \rfloor x} \right) \mathbb{P}[L(n) = i] \leq \frac{1}{x} + \frac{1}{x \lfloor xn \rfloor},$$

da $\lfloor xi \rfloor \geq \lfloor xn \rfloor$ für $i \geq n$. Die rechte Seite konvergiert für $n \rightarrow \infty$ gegen $\frac{1}{x}$. Insgesamt folgt mit dem ‘Sandwich-Prinzip’ die zweite Behauptung von Satz 8.3.10. \square

Korollar 8.3.12. Sei U gleichverteilt auf $[0, 1]$, dann gilt

$$\frac{L(n+1)}{L(n)} \xrightarrow[n \rightarrow \infty]{d} \frac{1}{U}.$$

BEWEIS. Die Tailfunktion der Zufallsvariable $\frac{1}{U}$ ist gegeben durch

$$\mathbb{P}\left[\frac{1}{U} > x\right] = \mathbb{P}\left[U < \frac{1}{x}\right] = \frac{1}{x}, \quad x > 1.$$

Somit ist $\frac{1}{U}$ Pareto-verteilt. Die Behauptung folgt nun aus der zweiten Aussage von Satz 8.3.10. \square

Aufgabe 8.3.13 (Shorrock, 1972). Beweisen Sie die folgende multivariate Version des obigen Korollars: Für beliebiges $k \in \mathbb{N}$ gilt

$$\left(\frac{L(n+1)}{L(n)}, \frac{L(n+2)}{L(n+1)}, \dots, \frac{L(n+k)}{L(n+k-1)}\right) \xrightarrow[n \rightarrow \infty]{d} \left(\frac{1}{U_1}, \frac{1}{U_2}, \dots, \frac{1}{U_k}\right),$$

wobei U_1, \dots, U_k unabhängige, auf $[0, 1]$ gleichverteilte Zufallsvariablen sind.

8.4. Zentrale Grenzwertsätze

Der klassische zentrale Grenzwertsatz besagt, dass für u.i.v. Zufallsvariablen X_1, X_2, \dots mit $\mathbb{E}X_1 = 0$ und $\mathbb{E}[X_k^2] = 1$ die Verteilungskonvergenz

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

gilt. Dabei bezeichnen wir mit $\mathcal{N}(0, 1)$ wir die Standardnormalverteilung mit Verteilungsfunktion

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}z^2} dz.$$

Nun zeigen wir, dass auch die Anzahl der Rekorde $N(n)$ nach einer geeigneten Normierung für $n \rightarrow \infty$ gegen die Normalverteilung konvergiert. Die glockenartige Form der Zähldichte von $N(n)$ ist auf Abbildung 2 deutlich zu erkennen.

Satz 8.4.1 (Zentraler Grenzwertsatz für die Anzahl der Rekorde). Es gilt

$$\frac{N(n) - \log n}{\sqrt{\log n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Bemerkung 8.4.2. Wegen der Foata-Korrespondenz (Abschnitt 8.2) gilt ein analoger Satz auch für die Anzahl $Z(n)$ der Zyklen in einer rein zufälligen Permutation von $\{1, \dots, n\}$, nämlich

$$\frac{Z(n) - \log n}{\sqrt{\log n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

In dieser Form wurde der Satz von Goncharov im Jahr 1948 bewiesen.

Der Beweis von Satz 8.4.1 basiert auf einer Verallgemeinerung des zentralen Grenzwertsatzes, die wir hier ohne Beweis angeben. Im Unterschied zum klassischen Grenzwertsatz, in dem es um Folgen von u.i.v. Zufallsvariablen geht, betrachten wir hier ein sogenanntes **Dreiecksschema** von Zufallsvariablen:

$$\begin{array}{c} Z_{1,1}, \\ Z_{1,2}, Z_{2,2}, \\ \dots\dots\dots \\ Z_{1,n}, Z_{2,n}, \dots, Z_{n,n}, \\ \dots\dots\dots \end{array}$$

Satz 8.4.3 (Zentraler Grenzwertsatz von Lindeberg). Es seien die folgenden Bedingungen erfüllt:

- (a) Für jedes $n \in \mathbb{N}$ sind die Zufallsvariablen $Z_{1,n}, Z_{2,n}, \dots, Z_{n,n}$, welche die n -te Zeile des Schemas bilden, unabhängig;
- (b) $\mathbb{E}Z_{k,n} = 0$ und $\mathbb{E}[Z_{k,n}^2] < \infty$ für $k = 1, \dots, n$;
- (c) $\sum_{k=1}^n \mathbb{E}[Z_{k,n}^2] \rightarrow 1$ für $n \rightarrow \infty$;
- (d) Für jedes $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} [Z_{k,n}^2 \mathbb{1}_{\{|Z_{k,n}| > \varepsilon\}}] = 0 \quad \text{(Lindeberg-Bedingung)}.$$

Dann gilt:

$$Z_{1,n} + \dots + Z_{n,n} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Es sei bemerkt, dass die identische Verteiltheit der Zufallsvariablen im zentralen Grenzwertsatz von Lindeberg nicht vorausgesetzt wird. Ebenso wird nicht vorausgesetzt, dass verschiedene Zeilen unabhängig voneinander sind.

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 mit $\xi_k \sim \text{Bern}(\frac{1}{k})$ sind. Definiere

$$Z_{k,n} = \frac{\xi_k - \frac{1}{k}}{\sqrt{\log n}}, \quad k = 1, \dots, n, \quad n \geq 2.$$

Die Zufallsvariablen $Z_{1,n}, \dots, Z_{n,n}$ sind für jedes n unabhängig. Definitionsgemäß gilt $\mathbb{E}Z_{k,n} = 0$ und

$$\sum_{k=1}^n \mathbb{E}[Z_{k,n}^2] = \frac{1}{\log n} \sum_{k=1}^n \text{Var} \xi_k = \frac{1}{\log n} \left(\sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{k^2} \right) \xrightarrow[n \rightarrow \infty]{} 1.$$

Um zu zeigen, dass die Lindeberg-Bedingung gilt, sei ein $\varepsilon > 0$ vorgegeben. Da ξ_k nur die Werte 0 und 1 annimmt, gilt die Abschätzung

$$|Z_{k,n}| = \left| \frac{\xi_k - 1/k}{\sqrt{\log n}} \right| \leq \frac{1}{\sqrt{\log n}} < \varepsilon, \quad k = 1, \dots, n$$

für alle $n > e^{1/\varepsilon^2}$. Somit kann die Bedingung $|Z_{k,n}| > \varepsilon$ nicht erfüllt werden und es gilt sogar $\sum_{k=1}^n \mathbb{E} [Z_{k,n}^2 \mathbb{1}_{\{|Z_{k,n}| > \varepsilon\}}] = 0$ für alle hinreichend große n . Der Satz von Lindeberg liefert nun

$$\frac{N(n) - \sum_{k=1}^n \frac{1}{k}}{\sqrt{\log n}} = \sum_{k=1}^n Z_{k,n} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Die Zentrierungsfolge $\sum_{k=1}^n \frac{1}{k}$ kann nach dem Slutsky-Lemma (???) durch $\log n$ ersetzt werden, denn $\sum_{k=1}^n \frac{1}{k} - \log n = o(\sqrt{\log n})$. \square

Aufgabe 8.4.4. Zeigen Sie, dass für die Anzahl der Rekorde auch das Gesetz der großen Zahlen gilt:

$$\lim_{n \rightarrow \infty} \frac{N(n)}{\log n} = 1 \quad \text{f.s.}$$

Hinweis: ???

Nun benutzen wir den obigen Satz um auch einen zentralen Grenzwertsatz für die logarithmierten Rekordzeiten herzuleiten.

Satz 8.4.5 (Zentraler Grenzwertsatz für logarithmierte 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 $n(x) = \lfloor e^{n+x\sqrt{n}} \rfloor$ gilt

$$\mathbb{P} \left[\frac{\log L(n) - n}{\sqrt{n}} \leq x \right] = \mathbb{P}[L(n) \leq e^{n+x\sqrt{n}}] = \mathbb{P}[L(n) \leq n(x)] = \mathbb{P}[N(n(x)) \geq n].$$

Satz 8.4.1 liefert

$$\frac{N(n(x)) - \log n(x)}{\sqrt{\log n(x)}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Es sei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung. Es folgt, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}[N(n(x)) \geq n] = \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{N(n(x)) - \log n(x)}{\sqrt{\log n(x)}} \geq \frac{n - \log n(x)}{\sqrt{\log n(x)}} \right] = 1 - \Phi(-x) = \Phi(x),$$

denn $n(x) = e^{n+x\sqrt{n}}(1 + o(1))$ und somit

$$\lim_{n \rightarrow \infty} \frac{n - \log n(x)}{\sqrt{\log n(x)}} = \lim_{n \rightarrow \infty} \frac{n - (n + x\sqrt{n} + o(1))}{\sqrt{n + x\sqrt{n} + o(1)}} = \lim_{n \rightarrow \infty} \frac{-x\sqrt{n} + o(\sqrt{n})}{\sqrt{n + o(n)}} = -x.$$

Das beweist die Behauptung. \square

Aufgabe 8.4.6. Zeigen Sie, dass für die logarithmierten Rekordzeiten auch das Gesetz der großen Zahlen gilt:

$$\lim_{n \rightarrow \infty} \frac{\log L(n)}{n} = 1 \quad \text{f.s.}$$

Hinweis: Zeigen Sie die entsprechende Aussage für $R(n)$ aus Satz 8.3.8. Zeigen Sie, dass $R(n) \geq 1/(U_1 \dots U_{n-1})$ und finden Sie eine passende Abschätzung nach oben.

Aufgabe 8.4.7. Ist es richtig, dass

$$\lim_{n \rightarrow \infty} \frac{L(n)}{e^n} = 1 \quad \text{f.s.?}$$

Hinweis: Betrachten Sie $L(n+1)/L(n)$. Konvergiert diese Folge fast sicher gegen e ?

Aufgabe 8.4.8 (Neuts, 1967). Zeigen Sie, dass für die Zwischenrekordzeiten $\Delta(n) := L(n) - L(n-1)$ die gleichen Grenzwertsätze wie für $L(n)$ gelten, nämlich

$$\lim_{n \rightarrow \infty} \frac{\log L(n)}{n} = 1 \quad \text{f.s.} \quad \text{und} \quad \frac{\log \Delta(n) - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Hinweis: Korollar 8.3.12.

8.5. Rekordwerte

Seien X_1, X_2, \dots u.i.v. Zufallsvariablen und X eine weitere Zufallsvariable mit der gleichen Verteilung.

Definition 8.5.1. Die **Rekordwerte** sind definiert als

$$X(n) = X_{L(n)} = \max\{X_1, \dots, X_{L(n)}\}.$$

Rekordwerte als Markov-Prozess. Zuerst bestimmen wir die gemeinsame Verteilung von $X(1), \dots, X(n)$.

Satz 8.5.2 (Chandler, 1952). Die Zufallsvariablen X_1, X_2, \dots seien u.i.v. mit Dichte f . Dann ist die gemeinsame Dichte von $(X(1), \dots, X(n))$ durch

$$f(x_1, \dots, x_n) = \frac{f(x_1)f(x_2) \dots f(x_n)}{\bar{F}(x_1)\bar{F}(x_2) \dots \bar{F}(x_{n-1})} \mathbb{1}_{\{x_1 < \dots < x_n < x^*\}}$$

gegeben, wobei x^* den rechten Endpunkt der X_i 's bezeichnet.

BEWEIS. Wir erlauben uns die Freiheit, "unendlich kleine" Intervalle $(x_1, x_1+dx_1), \dots, (x_n, x_n+dx_n)$ zu betrachten und die gesuchte Dichte als

$$f(x_1, \dots, x_n) = \lim_{dx_1, \dots, dx_n \rightarrow 0} \frac{\mathbb{P}[X(1) \in (x_1, x_1+dx_1), \dots, X(n) \in (x_n, x_n+dx_n)]}{dx_1 \dots dx_n}$$

darzustellen. Wegen $X(1) < \dots < X(n) < x^*$ nehmen wir $x_1 < \dots < x_n < x^*$ an. Nun stellen wir das uns interessierende Ereignis

$$(8.5.1) \quad \{X(k) \in (x_k, x_k+dx_k) \text{ für alle } k = 1, \dots, n\}$$

als eine disjunkte Vereinigung der Ereignisse

$$(8.5.2) \quad \{X(k) \in (x_k, x_k + dx_k) \text{ und } L(k) = i(k) \text{ für alle } k = 1, \dots, n\}$$

über alle $1 = i(1) < i(2) < \dots < i(n)$ dar. Definiere $j(k) := i(k) - i(k-1) - 1$, $k \in \{2, \dots, n\}$, so dass $j(2), \dots, j(n) \in \mathbb{N}_0$. Ereignis (8.5.2) tritt genau dann ein, wenn

$$\begin{aligned} X_{i(1)} &\in (x_1, x_1 + dx_1) \text{ und die nächsten } j(2) \text{ Beobachtungen sind kleiner als } X_{i(1)}, \\ X_{i(2)} &\in (x_2, x_2 + dx_2) \text{ und die nächsten } j(3) \text{ Beobachtungen sind kleiner als } X_{i(2)}, \\ &\dots, \\ X_{i(n)} &\in (x_n, x_n + dx_n) \text{ und alle weiteren Beobachtungen sind beliebig.} \end{aligned}$$

Da $X_{i(k)} \in (x_k, x_k + dx_k)$ gefordert wird und dx_k dabei unendlich klein ist, können wir die Bedingung “eine Beobachtung ist kleiner als $X_{i(k)}$ ” durch “eine Beobachtung ist kleiner als x_k ” ersetzen. Uns interessiert also die Wahrscheinlichkeit des Ereignisses

$$\begin{aligned} X_{i(1)} &\in (x_1, x_1 + dx_1) \text{ und die nächsten } j(2) \text{ Beobachtungen sind kleiner als } x_1, \\ X_{i(2)} &\in (x_2, x_2 + dx_2) \text{ und die nächsten } j(3) \text{ Beobachtungen sind kleiner als } x_2, \\ &\dots, \\ X_{i(n)} &\in (x_n, x_n + dx_n). \end{aligned}$$

Wegen Unabhängigkeit ist diese Wahrscheinlichkeit gegeben durch

$$f(x_1)dx_1 F(x_1)^{j(2)} \cdot f(x_2)dx_2 F(x_2)^{j(3)} \cdot \dots \cdot f(x_n)dx_{n-1} F(x_{n-1})^{j(n)} \cdot f(x_n)dx_n.$$

Nach Summation über alle $j(2), \dots, j(n) \in \mathbb{N}_0$ ergibt sich für die Wahrscheinlichkeit des Ereignisses (8.5.1) die Formel

$$\frac{f(x_1)}{\bar{F}(x_1)} dx_1 \cdot \frac{f(x_2)}{\bar{F}(x_2)} dx_2 \cdot \dots \cdot \frac{f(x_{n-1})}{\bar{F}(x_{n-1})} dx_{n-1} \cdot f(x_n) dx_n.$$

Nach Division durch $dx_1 \dots dx_n$ erhalten wir die behauptete Formel für die Dichte $f(x_1, \dots, x_n)$. □

Im nächsten Korollar zeigen wir, dass die Rekordwerte einen Markov-Prozess bilden. Für eine kurze Einführung in Markov-Prozesse mit beliebigen Zustandsraum verweisen wir auf???

Korollar 8.5.3. Die Zufallsvariablen X_1, X_2, \dots seien u.i.v. mit Dichte f . Dann bilden die Rekordwerte $X(1), X(2), \dots$ einen Markov-Prozess mit Zustandsraum $(-\infty, x^*)$ und Übergangsdichten

$$p(x, y) = \begin{cases} f(y)/\bar{F}(x), & \text{falls } x < y < x^*, \\ 0, & \text{sonst.} \end{cases}$$

BEWEIS. Die gemeinsame Dichte von $(X(1), \dots, X(n))$ lässt sich für alle $x_1 < \dots < x_n < x^*$ wie folgt schreiben:

$$f(x_1, \dots, x_n) = f(x_1) \cdot \frac{f(x_2)}{\bar{F}(x_1)} \cdot \dots \cdot \frac{f(x_n)}{\bar{F}(x_{n-1})} = f(x_1)p(x_1, x_2) \dots p(x_{n-1}, x_n).$$

Also bilden die Rekordwerte einen Markov-Prozess mit Anfangsdichte $f(x_1)$ und den oben angegebenen Übergangsdichten $p(x, y)$. \square

Aufgabe 8.5.4 (Chandler, 1952). Beweisen Sie die Formel für die Dichte von $X(n)$:

$$f_{X(n)}(x) = \frac{(-\log \bar{F}(x))^{n-1}}{(n-1)!} f(x), \quad x < x^*.$$

Hinweis: Benutzen Sie Induktion und die Formel für die Übergangsdichte aus Satz 8.5.2.

Aufgabe 8.5.5. Seien X_1, X_2, \dots u.i.v. mit Dichte f . Die Zwischenrekordzeiten sind definiert durch $\Delta(n) := L(n) - L(n-1)$ für $n = 2, 3, \dots$

- (a) Zeigen Sie, dass die gemeinsame Verteilung von $X(1), \dots, X(n)$ und $\Delta(2), \dots, \Delta(n)$ durch die folgende Formel gegeben ist:

$$\begin{aligned} \mathbb{P}[X(1) \in (x_1, x_1 + dx_1), \dots, X(n) \in (x_n, x_n + dx_n), \Delta(2) = \Delta_2, \dots, \Delta(n) = \Delta_n] \\ = f(x_1) \dots f(x_n) F^{\Delta_2}(x_1) \dots F^{\Delta_n}(x_{n-1}) dx_1 \dots dx_n \end{aligned}$$

für alle $x_1 < \dots < x_n$ und $\Delta_2, \dots, \Delta_n \in \mathbb{N}$.

- (b) Zeigen Sie, dass bedingt auf die Rekordwerte $X(1) = x_1, X(2) = x_2, \dots$ die Zwischenrekordzeiten $\Delta(2), \Delta(3), \dots$ unabhängig und geometrisch verteilt sind mit

$$\mathbb{P}[\Delta(n) = \Delta \mid X(1) = x_1, X(2) = x_2, \dots] = \bar{F}(x_{n-1}) F^\Delta(x_{n-1}), \quad \Delta \in \mathbb{N}.$$

Nun möchten wir eine Interpretation von Korollar 8.5.3 vorstellen. Zuerst benötigen wir eine

Zwischenüberlegung: Bedingtes Sampling. Stellen wir uns vor, dass u.i.v. Zufallsvariablen X_1, X_2, \dots erzeugt werden, bis eine dieser Variablen einen vorgegebenen Wert t mit $t < x^*$ überschreitet. Wir definieren also die ‘‘Stopzeit’’

$$N = \min\{n \in \mathbb{N} : X_n > t\}$$

und schauen uns die Zufallsvariable X_N an. Wie ist diese verteilt? Zunächst einmal müssen wir berücksichtigen, dass der Index N zufällig ist, weshalb die Verteilung von X_N ganz anders als die Verteilung von X sein kann. Wir erinnern, dass X eine weitere Zufallsvariable mit der gleichen Verteilungsfunktion F ist. Wir behaupten nun:

Lemma 8.5.6. Die Verteilung von X_N stimmt mit der bedingten Verteilung von X gegeben, dass $X > t$, überein, d.h.

$$\mathbb{P}[X_N \in A] = \mathbb{P}[X \in A \mid X > t].$$

für jede Borel-Menge $A \subset \mathbb{R}$.

BEWEIS. Indem wir den Wahrscheinlichkeitsraum in die disjunkten Ereignisse $\{N = 1\}, \{N = 2\}, \dots$ zerlegen, erhalten wir, dass

$$\begin{aligned} \mathbb{P}[X_N \in A] &= \sum_{n=1}^{\infty} \mathbb{P}[X_N \in A, N = n] = \sum_{n=1}^{\infty} \mathbb{P}[X_1 \leq t, \dots, X_{n-1} \leq t, X_n \in A, X_n > t] \\ &= \sum_{n=1}^{\infty} \mathbb{P}[X \leq t]^{n-1} \mathbb{P}[X \in A, X > t] = \frac{\mathbb{P}[X \in A, X > t]}{1 - \mathbb{P}[X \leq t]} = \mathbb{P}[X \in A | X > t], \end{aligned}$$

wobei wir die Formel $\sum_{n=1}^{\infty} q^{n-1} = 1/(1 - q)$, $|q| < 1$, für die Summe einer geometrischen Reihe benutzt haben. \square

INTERPRETATION VON KOROLLAR 8.5.3. Stellen wir uns vor, dass uns die ganze Entwicklung des Prozesses X_1, X_2, \dots bis zur n -ten Rekordzeit $L(n)$ bekannt ist. Insbesondere seien die Werte $L(n) = l$ und $X(n) = X_l = x$ bekannt. Nun beobachten wir die Werte X_{l+1}, X_{l+2}, \dots und warten auf den ersten Wert, der x übersteigt. Dieser Wert ist dann der nächste Rekordwert $X(n+1)$. Nach der obigen Zwischenüberlegung hat $X(n+1)$ die gleiche Verteilung, wie X_1 gegeben, dass $X_1 > x$, also

$$\mathbb{P}[X(n+1) \in A | \text{Vorgeschichte bis zum Zeitpunkt } L(n)] = \mathbb{P}[X_1 \in A | X_1 > x].$$

Da diese Verteilung nur vom aktuellen Wert $X(n) = x$ abhängt, bilden $X(1), X(2), \dots$ eine Markov-Kette mit den oben angegebenen Übergangskernen.

Bemerkung 8.5.7 (Simulation von Zufallsvariablen). Eine leichte Abwandlung der obigen Überlegung kann man verwenden, um Zufallsvariablen mit einer gegebenen Verteilung am Rechner zu simulieren. Möchte man z.B. einen auf einer gegebenen Menge $G \subset \mathbb{R}^2$ gleichverteilten Punkt simulieren, so erzeugt man u.i.v. gleichverteilte Punkte auf einem hinreichend großen Quadrat $Q \supset G$, bis einer dieser Punkte in G fällt. Dieser Punkt ist dann gleichverteilt auf G (Übungsaufgabe).

Bemerkung 8.5.8. Im Fall, wenn die Zufallsvariablen X_1, X_2, \dots u.i.v. mit *beliebiger* Verteilung sind, bilden die Rekordwerte $X(1), X(2), \dots$ immer noch einen Markov-Prozess auf dem Zustandsraum $(-\infty, x^*]$. Dieser hat allerdings nunmehr keine Übergangsdichten, sondern wird durch den folgenden Übergangskern charakterisiert:

$$\mathbb{P}[X(n+1) \in A | X(n) = x] = \mathbb{P}[X \in A | X > x],$$

für $x < x^*$ und jede Borel-Menge $A \subset (-\infty, x^*]$. Im Fall, wenn $\mathbb{P}[X_1 = x^*] \neq 0$ muss man besonders vorsichtig sein, denn dann wird der Wert x^* f.s. von einem der X_n 's angenommen, wonach keine weiteren Rekorde aufgestellt werden können. In diesem Fall bricht die Serie der Rekordwerte $X(1), X(2), \dots$ nach endlicher Zeit ab. Wir müssen den Zustand x^* als einen "Grabzustand" des Markov-Prozesses auffassen. Sobald dieser Zustand erreicht wird, "stirbt" der Prozess.

Rekordwerte exponentialverteilter Zufallsvariablen. 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 besonders einfache und schöne Darstellung durch den homogenen Poisson-Prozess.

Satz 8.5.9 (Tata, 1969). Seien Z_1, Z_2, \dots unabhängig und exponentialverteilt mit Parameter 1. Dann gilt:

$$(Z(1), Z(2), \dots, Z(n)) \stackrel{d}{=} (E_1, E_1 + E_2, \dots, E_1 + \dots + E_n),$$

wobei E_1, E_2, \dots ebenfalls unabhängig und exponentialverteilt mit Parameter 1 sind.

Bemerkung 8.5.10. Die Folge der Rekordwerte $Z(1), Z(2), \dots$ bildet also einen Poisson-Prozess mit Intensität 1. Die Zufallsvariable $Z(n)$ ist gammaverteilt mit Parametern $(n, 1)$ und Dichte

$$f_{Z(n)}(t) = \frac{t^{n-1}}{(n-1)!} e^{-t} \mathbb{1}_{\{t>0\}}.$$

BEWEIS VON SATZ 8.5.9. Für die Tailfunktion und die Dichte der Exponentialverteilung mit Parameter 1 gilt $\bar{F}(x) = f(x) = e^{-x}$ für $x > 0$. Gemäß Satz 8.5.2 ist die gemeinsame Dichte von $(Z(1), \dots, Z(n))$ gegeben durch

$$f(x_1, \dots, x_n) = e^{-x_n} \mathbb{1}_{\{0 < x_1 < \dots < x_n\}}.$$

Die gleiche Formel haben wir im Beweis von Satz 7.4.2 für die Dichte des Zufallsvektors $(E_1, E_1 + E_2, \dots, E_1 + \dots + E_n)$ hergeleitet. \square

INTERPRETATION VON SATZ 8.5.9. Im Mittelpunkt steht schon wieder die Gedächtnislosigkeit der Exponentialverteilung. Ist nämlich $Z \sim \text{Exp}(1)$, so ist für jedes $t > 0$ die bedingte Verteilung von $Z - t$ gegeben, dass $Z > t$, ebenfalls eine Exponentialverteilung mit Parameter 1. Nun werden wir induktiv begründen, warum die Zufallsvariablen $Z(1), Z(2) - Z(1), Z(3) - Z(2), \dots$ standard exponentialverteilt und unabhängig sind.

Z(1): Nach Voraussetzung ist $Z(1) = Z_1 \sim \text{Exp}(1)$.

Z(2) - Z(1): Wir halten nun $Z_1 = a_1$ fest und warten auf den zweiten Rekordwert $Z(2)$. Nach der obigen Zwischenüberlegung hat der Exzess $Z(2) - a_1$ die gleiche Verteilung wie $Z - a_1$ gegeben, dass $Z > a_1$, also (wegen der Gedächtnislosigkeit!) wieder eine Exponentialverteilung mit Parameter 1.

Z(3) - Z(2): Nun halten wir die ganze Geschichte des Prozesses bis zur zweiten Rekordzeit $L(2)$ fest. Insbesondere sei $Z(2) = a_2$ gegeben. Wir warten nun auf den dritten Rekordwert $Z(3)$. Nach der Zwischenüberlegung hat der Exzess $Z(3) - a_2$ die gleiche Verteilung, wie $Z - a_2$ gegeben, dass $Z > a_2$, also wieder die Exponentialverteilung mit Parameter 1. Dabei hängt die bedingte Verteilung des Exzesses nicht von der Vorgeschichte des Prozesses ab, also ist die Zufallsvariable $Z(3) - Z(2)$ unabhängig von $Z(1)$ und $Z(2)$.

Somit sind die Zuwächse $Z(1), Z(2) - Z(1), Z(3) - Z(2), \dots$ standard exponentialverteilt und unabhängig.

Korollar 8.5.11. 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 8.5.9 wissen wir, dass $X(n) \stackrel{d}{=} E_1 + \dots + E_n$, wobei E_1, E_2, \dots unabhängig und standard exponentialverteilt sind. Also ist $\mathbb{E}E_i = \text{Var } E_i = 1$ und der klassische zentrale Grenzwertsatz ergibt, dass

$$\frac{E_1 + \dots + E_n - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

Daraus folgt die Behauptung. □

Bemerkung 8.5.12. In Satz 1.2.1 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 der Index $L(n)$ zufällig ist, völlig verändert.

Allgemeine Darstellung durch den Poisson-Prozess. Nun formulieren wir eine allgemeine Darstellung für Rekordwerte beliebiger u.i.v. stetig verteilter Zufallsvariablen X_1, X_2, \dots . Die Hauptidee besteht dabei darin, dass solche Zufallsvariablen als eine geeignete monoton steigende Transformation der standardexponentialverteilten Zufallsvariablen Z_1, Z_2, \dots dargestellt werden können. Entscheidend ist die Beobachtung, dass monoton steigende Transformationen Rekordwerte in Rekordwerte überführen.

Satz 8.5.13. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion F . Dann gilt

$$(X(1), X(2), \dots, X(n)) \stackrel{d}{=} (H(E_1), H(E_1 + E_2), \dots, H(E_1 + \dots + E_n)),$$

wobei $H(x) = F^{\leftarrow}(1 - e^{-x})$, $x > 0$.

BEWEIS. Seien Z_1, Z_2, \dots unabhängig und standard exponentialverteilt. Dann sind die Zufallsvariablen $1 - e^{-Z_i}$ gleichverteilt auf $[0, 1]$. Somit sind die Zufallsvariablen $H(Z_i)$ genauso verteilt wie die X_i 's:

$$(X_1, X_2, \dots) \stackrel{d}{=} (H(Z_1), H(Z_2), \dots).$$

Da die Funktion H strikt monoton steigend ist, stimmen die Rekordwerte der Folge $H(Z_1), H(Z_2), \dots$ mit $H(Z(1)), H(Z(2)), \dots$ überein und wir erhalten

$$(X(1), X(2), \dots) \stackrel{d}{=} (H(Z(1)), H(Z(2)), \dots).$$

Nun können wir auf $(Z(1), Z(2), \dots)$ die Tata-Darstellung aus Satz 8.5.9 anwenden, was zur gewünschten Behauptung führt. □

Aufgabe 8.5.14 (Verteilung von $X(n)$). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion F . Es sei $X(n)$ der n -te Rekordwert.

(a) 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.

(b) Zeigen Sie, dass

$$Q_n(s) = \frac{1}{(n-1)!} \int_0^{-\log(1-s)} t^{n-1} e^{-t} dt.$$

Hinweis: Teil (a) gilt auch für exponentialverteilte Zufallsvariablen.

Aufgabe 8.5.15 (Zwischenrekordzeiten, Neuts, 1967). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion F . Betrachte die Zwischenrekordzeiten $\Delta(n) = L(n) - L(n-1)$, $n = 2, 3, \dots$

(a) Zeigen Sie, dass für alle $k \in \{0, 1, \dots\}$,

$$\mathbb{P}[\Delta(n+1) > k] = \int_0^\infty \frac{x^{n-1}}{(n-1)!} e^{-x} (1 - e^{-x})^k dx.$$

Hinweis: Zeigen Sie, dass die Verteilung von $\Delta(n+1)$ nicht von der Wahl von F abhängt und wählen Sie dann $X_i \sim \text{Exp}(1)$.

(b) Für zwei Zufallsvariablen ξ und η sagen wir, dass ξ stochastisch größer als η ist, falls $\mathbb{P}[\xi > u] \geq \mathbb{P}[\eta > u]$ für alle $u \in \mathbb{R}$ ist. Zeigen Sie: $\Delta(n+1)$ ist stochastisch größer als $\Delta(n)$. (D.h. die Abstände zwischen den Rekorden sind "stochastisch steigend").

Hinweis: $\mathbb{P}[\Delta(n+1) > k] = \mathbb{E}[(1 - e^{E_1 + \dots + E_n})^k]$.

Aufgabe 8.5.16 (k -te Rekordwerte, Dziubdziela und Kopocinsky (1976)). Seien Z_1, Z_2, \dots unabhängige, mit Parameter 1 exponentialverteilte Zufallsvariablen. Sei $k \in \mathbb{N}$ fest. Definiere: ??? Zeigen Sie, dass

$$(X(1, k), X(2, k), \dots, X(n, k)) \stackrel{d}{=} \left(\frac{E_1}{k}, \frac{E_1 + E_2}{k}, \dots, \frac{E_1 + \dots + E_n}{k} \right),$$

wobei E_1, E_2, \dots unabhängig unabhängig und exponentialverteilt mit Parameter 1 sind.

Hinweis: ??? Schwierige Aufgabe.

8.6. Grenzwertverteilungen der Rekordwerte

Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion F und Rekordwerten $X(1), X(2), \dots$. Wir stellen uns die Frage, welche nichtdegenerierte Verteilungen als Grenzwerte von

$$\frac{X(n) - A(n)}{B(n)}$$

für $n \rightarrow \infty$ und geeignete deterministische Folgen $A(n) \in \mathbb{R}$ und $B(n) > 0$ entstehen können. Für die Maxima M_n anstelle der Rekordwerte $X(n)$ wurde die entsprechende Frage durch den Satz von Fisher-Tippett-Gnedenko vollständig beantwortet. Auf den ersten Blick sieht $X(n) = M_{L(n)}$ wie "ein Spezialfall" von M_n aus, allerdings ist der Index $L(n)$ zufällig. Der

nächste Satz birgt eine Überraschung: Die Klasse der Grenzwertverteilungen für Rekordwerte ist anders als die für Maxima.

Satz 8.6.1 (Resnick, 197???) . Gibt es Normierungskonstanten $A(n) \in \mathbb{R}$ und $B(n) > 0$, derart dass

$$\frac{X(n) - A(n)}{B(n)} \xrightarrow[n \rightarrow \infty]{d} G$$

für eine nichtdegenerierte Verteilungsfunktion G , so ist G vom gleichen Typ wie die Verteilungsfunktion einer der folgenden Zufallsvariablen:

- (a) $\xi \sim \mathcal{N}(0, 1)$ (Standardnormalverteilung);
- (b) $e^{\sigma\xi}$ für ein $\sigma > 0$ (Lognormalverteilung);
- (c) $-e^{\sigma\xi}$ für ein $\sigma > 0$ (das Negative der Lognormalverteilung).

BEWEIS. SCHRITT 1: TRANSFORMATION. Seien E_1, E_2, \dots u.i.v. standard exponentialverteilte Zufallsvariablen. Gemäß Satz 8.5.13 können wir $X(n)$ als eine Transformation von $S_n := E_1 + \dots + E_n$ darstellen, nämlich

$$X(n) \stackrel{d}{=} F^{\leftarrow}(1 - e^{-S_n}).$$

Es folgt

$$\begin{aligned} \mathbb{P} \left[\frac{X(n) - A(n)}{B(n)} \leq x \right] &= \mathbb{P}[X(n) \leq A(n) + B(n)x] = \mathbb{P}[F^{\leftarrow}(1 - e^{-S_n}) \leq A(n) + B(n)x] \\ &= \mathbb{P}[S_n \leq T(A(n) + B(n)x)], \end{aligned}$$

wobei

$$T(y) := -\log(1 - F(y)) \in [0, \infty].$$

Indem wir den zentralen Grenzwertsatz im Hinterkopf behalten, schreiben wir das wie folgt um:

$$(8.6.1) \quad \mathbb{P} \left[\frac{X(n) - A(n)}{B(n)} \leq x \right] = \mathbb{P} \left[\frac{S_n - n}{\sqrt{n}} \leq \frac{T(A(n) + B(n)x) - n}{\sqrt{n}} \right].$$

Die Voraussetzung des Satzes lautet also

$$(8.6.2) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{S_n - n}{\sqrt{n}} \leq \frac{T(A(n) + B(n)x) - n}{\sqrt{n}} \right] = G(x)$$

für alle Stetigkeitspunkte x von G . Im Folgenden betrachten wir nur solche x . Dabei behauptet der zentrale Grenzwertsatz, dass

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{S_n - n}{\sqrt{n}} \leq t_n \right] = \Phi(t)$$

für jede Folge t_1, t_2, \dots mit $t := \lim_{n \rightarrow \infty} t_n \in [-\infty, +\infty]$. Dabei definieren wir sinngemäß $\Phi(-\infty) := 0$ und $\Phi(+\infty) := 1$.

SCHRITT 1: EXISTENZ VON $g(x)$. Wir behaupten nun, dass der Grenzwert

$$(8.6.3) \quad g(x) := \lim_{n \rightarrow \infty} \frac{T(A(n) + B(n)x) - n}{\sqrt{n}} \in [-\infty, +\infty]$$

existiert. Wir beweisen das durch Widerspruch. Dazu nehmen wir an, dass die Folge in (8.6.3) zwei verschiedene Häufungspunkte, g_1 und g_2 in $[-\infty, \infty]$, besitzt. Indem wir auf die entsprechenden Teilfolgen den zentralen Grenzwertsatz anwenden, erhalten wir, dass auch die Folge in (8.6.2) zwei Häufungspunkte $\Phi(g_1)$ und $\Phi(g_2)$ besitzt. Wegen der strikten Monotonie von Φ sind diese Häufungspunkte verschieden, was im Widerspruch zu (8.6.2) steht. Also existiert der Grenzwert in (8.6.3). Wegen des zentralen Grenzwertsatzes und (8.6.2) gilt dann $G(x) = \Phi(g(x))$.

Sei $x \in \mathbb{R}$ ein Wert mit $g(x) \neq \pm\infty$. Dann folgt aus (8.6.3), dass

$$T(A(n) + B(n)x) = n + g(x)\sqrt{n} + o(\sqrt{n}) = n \left(1 + \frac{g(x) + o(1)}{\sqrt{n}} \right), \quad n \rightarrow \infty.$$

Daraus folgt mit der Taylor-Entwicklung $\sqrt{1+z} = 1 + z/2 + o(z)$ für $z \rightarrow 0$, dass

$$\sqrt{T(A(n) + B(n)x)} = \sqrt{n} \left(1 + \frac{g(x) + o(1)}{\sqrt{n}} \right)^{1/2} = \sqrt{n} + \frac{g(x) + o(1)}{2}, \quad n \rightarrow \infty.$$

SCHRITT 3: EXISTENZ VON $g(x)$. Betrachte nun die Tailfunktion

$$\bar{G}(x) = e^{-\sqrt{T(x)}} = e^{-\sqrt{-\log(1-F(x))}}.$$

Es folgt

$$\bar{G}(A(n) + B(n)x) = e^{-\sqrt{T(A(n)+B(n)x)}} \sim e^{-\sqrt{n} - \frac{g(x)}{2}}, \quad n \rightarrow \infty.$$

Indem wir nun $n = \lfloor \log^2 m \rfloor$ mit $m \in \mathbb{N}$ einsetzen und $A^*(m) := A(\lfloor \log^2 m \rfloor)$, $B^*(m) := B(\lfloor \log^2 m \rfloor)$ definieren, erhalten wir

$$\bar{G}(A^*(m) + B^*(m)x) \sim \frac{1}{m} e^{-\frac{g(x)}{2}}, \quad m \rightarrow \infty.$$

Der Leser möge dabei nachrechnen, dass $e^{-\sqrt{\lfloor \log^2 m \rfloor}} \sim 1/m$. Daraus folgt schließlich

$$G^m(A^*(m) + B^*(m)x) = (1 - \bar{G}(A^*(m) + B^*(m)x))^m \xrightarrow{m \rightarrow \infty} e^{-e^{-\frac{g(x)}{2}}}.$$

Sind nun V_1, V_2, \dots u.i.v. Zufallsvariablen mit Verteilungsfunktion G , so gilt

$$\frac{\max\{V_1, \dots, V_n\} - A^*(m)}{B^*(m)} \xrightarrow{m \rightarrow \infty} e^{-e^{-\frac{g(x)}{2}}}.$$

Die Grenzwertfunktion $G(x) = \Phi(g(x))$ ist nicht degeneriert genau dann, wenn g an mindestens einer Stelle endlich ist und dabei nicht konstant ist. Die Verteilungsfunktion ist nicht degeneriert, also können wir den Satz von Fisher-Tippett-Gnedenko anwenden. \square

Beispiel 8.6.2 (Lognormalverteilung als Grenzwert). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit Werten in $(1, \infty)$ und Verteilungsfunktion

$$F(x) = 1 - e^{-\alpha^2 \log^2 x}, \quad x \geq 1,$$

wobei $\alpha > 0$ ein Parameter sei. Wir behaupten, dass

$$\frac{X(n)}{e^{\sqrt{n}/\alpha}} \xrightarrow{n \rightarrow \infty} e^{\xi/(2\alpha)} \quad \text{mit } \xi \sim \mathcal{N}(0, 1).$$

BEWEIS. Für die angegebene Verteilung gilt $T(y) = \alpha^2 \log^2 y$ für $y \geq 1$. Mit $A(n) = 0$ und $B(n) = e^{\sqrt{n}/\alpha}$ ergibt sich

$$\frac{T(A(n) + B(n)x) - n}{\sqrt{n}} = \frac{\alpha^2 (\sqrt{n}/\alpha + \log x)^2 - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{} 2\alpha \log x$$

für alle $x > 0$. Mit (8.6.1) und dem zentralen Grenzwertsatz erhalten wir, dass

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{X(n)}{e^{\sqrt{n}/\alpha}} \leq x \right] = \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{S_n - n}{\sqrt{n}} \leq 2\alpha \log x + o(1) \right] = \Phi(2\alpha \log x) = \mathbb{P}[e^{\xi/(2\alpha)} \leq x]$$

für alle $x > 0$, wobei $\xi \sim \mathcal{N}(0, 1)$. □

Beispiel 8.6.3 (Das Negative der Lognormalverteilung als Grenzwert). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit Werten in $(-1, 0)$ und Verteilungsfunktion

$$F(x) = 1 - e^{\alpha^2 \log^2(-x)}, \quad x \in (-1, 0),$$

wobei $\alpha > 0$ ein Parameter sei. Dann gilt

$$e^{\sqrt{n}/\alpha} X(n) \xrightarrow[n \rightarrow \infty]{d} -e^{\xi/(2\alpha)} \quad \text{mit } \xi \sim \mathcal{N}(0, 1).$$

BEWEIS. Übungsaufgabe. □

Aufgabe 8.6.4. Seien X_1, X_2, \dots unabhängig und standardnormalverteilt. Zeigen Sie, dass

$$X(n) - \sqrt{2n} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1/2).$$

Hinweis: Zeigen Sie mit Hilfe von Lemma 5.4.1, dass $T(y) = y^2/2 + o(y)$ für $y \rightarrow +\infty$ und benutzen Sie dann (8.6.1).

Aufgabe 8.6.5. Seien X_1, X_2, \dots u.i.v. mit Verteilungsfunktion $F(x) = 1 - x^{-\alpha}$, $x \geq 1$, wobei $\alpha > 0$. Zeigen Sie: Es gibt keine Normierungsfolgen $A(n) \in \mathbb{R}$ und $B(n) > 0$ derart, dass $(X(n) - A(n))/B(n)$ gegen eine nichtdegenerierte Verteilung konvergiert.

8.7. Rekorde für diskrete Zufallsvariablen

In den vorangegangenen Abschnitten haben wir meistens Zufallsvariablen mit stetiger Verteilungsfunktion betrachtet. Eine ähnliche Theorie lässt sich auch für diskrete Zufallsvariablen entwickeln, wobei der diskrete Fall oft sogar eleganter ist. Wir formulieren die entsprechenden Resultate von Shorrock als eine Serie von Übungsaufgaben.

Seien X, X_1, X_2, \dots u.i.v. diskrete Zufallsvariablen mit Werten in $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. Die strikten Rekordzeiten $L(1), L(2), \dots$ seien gegeben durch $L(1) = 1$ und

$$L(n+1) = \min \{k \geq L(n) + 1 : X_k > X_{L(n)}\} \quad \text{für } n \in \mathbb{N}.$$

Es sei weiterhin $X(n) = X_{L(n)}$ der n -te strikte Rekordwert. Dabei nehmen wir an, dass der Wertebereich der X_n 's nach oben unbeschränkt ist, d.h. $\mathbb{P}[X > k] > 0$ für alle $k \geq 0$. Das garantiert, dass immer neue Rekorde aufgestellt werden können und die Reihe der Rekordzeiten niemals abbricht.

Aufgabe 8.7.1 (Maxima bilden eine Markov-Kette). Zeigen Sie, dass M_1, M_2, \dots eine Markov-Kette auf \mathbb{N}_0 mit Übergangswahrscheinlichkeiten

$$p_{ij} = \begin{cases} p_j, & \text{falls } j > i, \\ p_1 + \dots + p_i, & \text{falls } j = i, \\ 0, & \text{falls } j < i, \end{cases}$$

bilden. Dabei ist $M_n = \max\{X_1, \dots, X_n\}$, wie immer.

Aufgabe 8.7.2 (Verteilung der Rekordwerte). Zeigen Sie für alle $n \in \mathbb{N}$ und ganzzahlige $0 \leq i_1 < i_2 < \dots < i_n$, dass

$$\mathbb{P}[X(1) = x_1, \dots, X(n) = x_n] = \mathbb{P}[X = x_n] \prod_{k=1}^{n-1} \frac{\mathbb{P}[X = x_k]}{\mathbb{P}[X > x_k]}.$$

Hinweis: Bestimmen Sie zuerst $\mathbb{P}[L(1) = i_1, X(1) = x_1, \dots, L(n) = i_n, X(n) = x_n]$ und summieren Sie dann über alle $1 = i_1 < i_2 < \dots < i_n$.

Aufgabe 8.7.3 (Rekordwerte bilden eine Markov-Kette). Zeigen Sie, dass $X(1), X(2), \dots$ eine Markov-Kette mit Übergangswahrscheinlichkeiten

$$\mathbb{P}[X(n+1) = j | X(n) = i] = \begin{cases} \mathbb{P}[X = j] / \mathbb{P}[X > i], & \text{falls } 0 \leq i < j, \\ 0, & \text{sonst} \end{cases}$$

bilden.

Aufgabe 8.7.4 (Indikatoren der Rekordwerte). Definiere die Rekordwertindikatoren η_0, η_1, \dots durch

$$\eta_k := \begin{cases} 1, & \text{falls } X(n) = k \text{ für ein } n \in \mathbb{N}, \\ 0, & \text{sonst.} \end{cases}$$

Zeigen Sie: η_0, η_1, \dots sind unabhängig und $\mathbb{P}[\eta_k = 1] = \mathbb{P}[X = k] / \mathbb{P}[X \geq k]$.

Hinweis: Nevzorov??? Es gibt einen analogen Satz für stetige Zufallsvariablen; siehe Aufgabe 9.6.11.

Aufgabe 8.7.5 (Diskrete Version der Tata-Darstellung). Seien nun X_1, X_2, \dots unabhängig und geometrisch verteilt mit $\mathbb{P}[X_1 = k] = p(1-p)^{k-1}$, $k \in \mathbb{N}$, wobei $p \in (0, 1)$. Beweisen Sie die folgende diskrete Version der Tata-Darstellung:

$$(X(1), X(2), \dots, X(n)) \stackrel{d}{=} (G_1, G_1 + G_2, \dots, G_1 + \dots + G_n),$$

wobei G_1, G_2, \dots unabhängig sind und die gleiche geometrische Verteilung wie die X_i 's besitzen.

Kommentare

Dieser Abschnitt steht unter starkem Einfluss des exzellenten Buches von Nevzorov [20]. Eine umfassende Übersicht über die historische Entwicklung des Gebiets mit Verweisen auf die Originalliteratur findet sich in [20, Appendix 1]. [4], [22], [1], [23].

Anhang: Markov-Prozesse

9.1. Markov-Ketten mit diskretem Zustandsraum

Ein Exkurs über Markov-Ketten. Man stelle sich ein System vor, das seinen Zustand in diskreter Zeit ändern kann. Die Menge aller möglichen Zustände sei mit E bezeichnet und höchstens abzählbar. Befindet sich das System gerade in Zustand i , so sei die Wahrscheinlichkeit, dass es im nächsten Zeitpunkt zum Zustand j wechselt, bezeichnet mit p_{ij} . Sinngemäß muss dann $p_{ij} \geq 0$ und $\sum_{j \in E} p_{ij} = 1$ für jedes $i \in E$ erfüllt sein. Wir gehen davon aus, dass die folgende Markov-Eigenschaft erfüllt ist: Die Auswahl des nächsten Zustandes j geschieht unabhängig davon, wie der aktuelle Zustand i erreicht wurde. Um die Verteilung der Markov-Kette komplett zu beschreiben, müssen wir noch die Anfangsbedingung angeben: die Wahrscheinlichkeit, dass sich das System zum Zeitpunkt 1 im Zustand i befindet, sei mit α_i bezeichnet, wobei $\alpha_i \geq 0$ und $\sum_{i \in E} \alpha_i = 1$. Dann ist die Wahrscheinlichkeit, dass das System in den Zeitpunkten $1, \dots, n$ die Zustände $i(1), \dots, i(n) \in E$ durchläuft, gegeben durch

$$(9.1.1) \quad \mathbb{P}[\eta(1) = i(1), \eta(2) = i(2), \dots, \eta(n) = i(n)] = \alpha_{i(1)} p_{i(1)i(2)} \cdots p_{i(n-1)i(n)}.$$

Dabei ist $\eta(k)$ die Zufallsvariable mit Werten in E , die den Zustand angibt, in welchem sich die Markov-Kette zum Zeitpunkt k befindet. Wir werden (9.1.1) als Definition einer Markov-Kette verwenden.

Definition 9.1.1. Eine Folge $\eta(1), \eta(2), \dots$ von Zufallsvariablen mit Werten in E heißt **Markov-Kette**, wenn ihre endlich-dimensionalen Verteilungen durch (9.1.1) gegeben sind.

Nun können wir die **Markov-Eigenschaft** herleiten.

Proposition 9.1.2. Für beliebige Zustände $i(1), i(2), \dots, i(n) \in E$ und $j \in E$ gilt

$$\begin{aligned} \mathbb{P}[\eta(n+1) = j \mid \eta(1) = i(1), \eta(2) = i(2), \dots, \eta(n) = i(n)] \\ = \mathbb{P}[\eta(n+1) = j \mid \eta(n) = i(n)] \end{aligned}$$

vorausgesetzt, dass $\mathbb{P}[\eta(1) = i(1), \eta(2) = i(2), \dots, \eta(n) = i(n)] \neq 0$.

INTERPRETATION. Die bedingte Verteilung von $\eta(n+1)$ gegeben $\eta(n)$ ist die gleiche, wie die bedingte Verteilung von $\eta(n)$ gegeben $\eta(1), \eta(2), \dots, \eta(n)$. Gegeben die Gegenwart (also $\eta(n) = i(n)$) ist die Zukunft (also $\eta(n+1)$) unabhängig von der Vergangenheit (also $\eta(1), \dots, \eta(n-1)$).

Außerdem gilt für beliebige zwei Zustände $i, j \in E$

$$\mathbb{P}[\eta(n+1) = j \mid \eta(n) = i] = p_{ij}$$

vorausgesetzt $\mathbb{P}[\eta(n) = i] \neq 0$.

9.2. Markov-Ketten mit allgemeinem Zustandsraum

Wir stellen uns ein System vor, das zu diskreten Zeitpunkten $1, 2, \dots$ seinen Zustand zufällig ändern kann. Die Menge der möglichen Zustände sei dabei ein Intervall E auf der reellen Achse¹ Man kann sich z.B. ein zufällig herumspringendes Teilchen auf der reellen Geraden vorstellen. Dabei gehen wir davon aus, dass das Teilchen gedächtnislos im folgenden Sinne ist: gegeben die aktuelle Position $\eta(n)$ des Teilchens zum Zeitpunkt n , hängt seine zukünftige Bewegung zwar von $\eta(n)$ ab, nicht aber von $\eta(1), \dots, \eta(n-1)$. In diesem Fall wird das System durch die Übergangskerne

die künftige Bewegung des Teilchens na Befindet sich das System zu einem Zeitpunkt n in Zustand x , so wird der Zustand zum Zeitpunkt

$$\mathbb{P}[\eta(1) \in (y_1, y_1 + dy_1), \dots, \eta(n) \in (y_n, y_n + dy_n)] = \alpha(dy_1)K(y_1, dy_2) \dots K(y_{n-1}, dy_n).$$

Oft haben die Übergangskerne eine Dichte

$$K(x, A) = \int_A$$

Beispiel 9.2.1 (Irrfahrt). Seien Z_1, Z_2, \dots

Beispiel 9.2.2 (Multiplikative Irrfahrt). Seien Z_1, Z_2, \dots

Beispiel 9.2.3 (Maxima). Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit Verteilungsfunktion F . Dann bilden die Maxima M_1, M_2, \dots einen Markov-Prozess auf $(-\infty, x^*]$ mit Übergangskern, der wie folgt charakterisiert wird:

$$K(x, (-\infty, y]) = \begin{cases} F(y), & \text{falls } y \in [x, x^*], \\ 0, & \text{falls } y < x. \end{cases}$$

9.3. Markov-Prozesse mit stetiger Zeit

¹Allgemeiner kann man einen beliebigen Messraum (E, \mathcal{E}) als die Menge der Zustände zulassen.

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

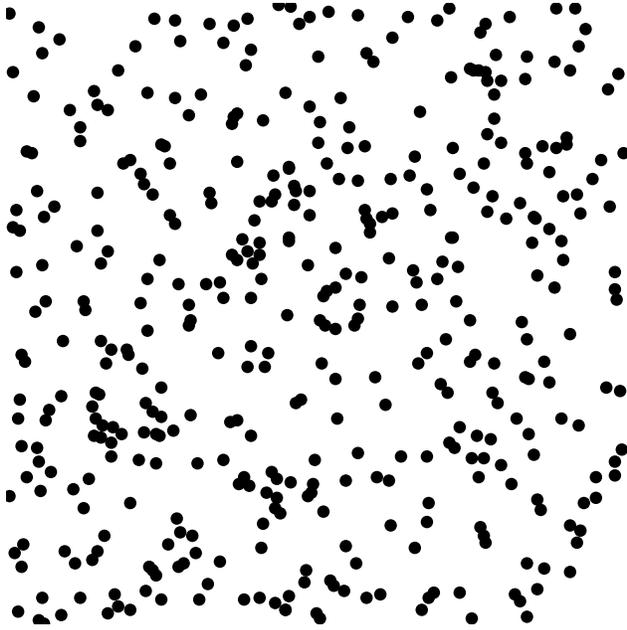


ABBILDUNG 1. Homogener Poisson-Punktprozess.

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.

Zum Beispiel können wir versuchen, ein stochastisches Modell für die Verteilung der Sterne im Himmel zu finden. Ist $A \subset \mathbb{R}^2$ eine Borel-Menge (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 $\mathbb{N}_0 \cup \{\infty\} = \{0, 1, 2, \dots, \infty\}$ auf. Wir gehen davon

aus, dass die folgenden natürlichen Eigenschaften gelten:

- (1) Für disjunkte Borel-Mengen 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 $\varepsilon \approx 0$ einen Stern gibt, ist $\approx \lambda\varepsilon$.

Der Koeffizient $\lambda > 0$ beschreibt dabei die Intensität der Sterne im Himmel.

Wie ist nun die Anzahl der Sterne $\Pi(A)$ in einer beliebigen Borel-Menge A verteilt? Um dies herauszufinden, unterteilen wir A in kleine Gebiete mit Fläche ε . Die Anzahl dieser Gebiete ist $\approx \text{Fläche}(A)/\varepsilon$. Aus unseren Voraussetzungen und dem Poisson-Grenzwertsatz folgt, dass

$$\Pi(A) \approx \text{Bin}\left(\frac{\text{Fläche}(A)}{\varepsilon}, \lambda\varepsilon\right) \approx \text{Poi}(\lambda \cdot \text{Fläche}(A)) \text{ für } \varepsilon \downarrow 0.$$

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 von Zählmaßen und Punktprozessen

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 kein 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 den Euklid'schen Raum \mathbb{R}^d oder allgemeiner 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 Borel-Funktion $f : E \rightarrow \mathbb{R}$ heißt **lokal integrierbar**, falls $\int_K |f(t)| dt < \infty$ für alle kompakten Mengen $K \subset E$.

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 .

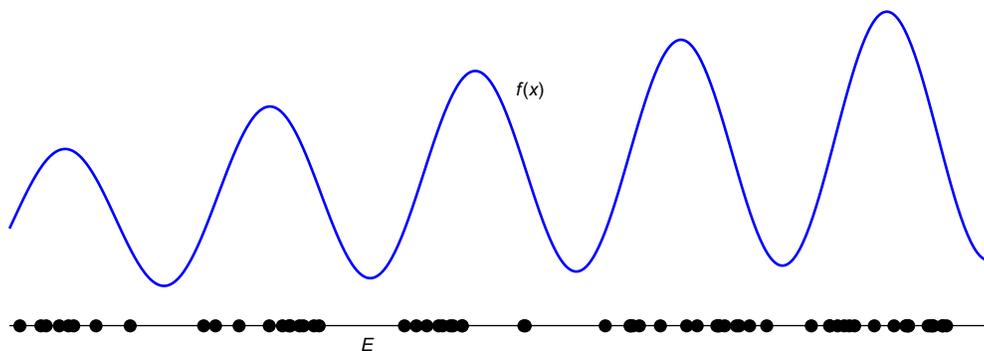


ABBILDUNG 2. Poisson-Prozess auf \mathbb{R} mit Intensität f .

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.

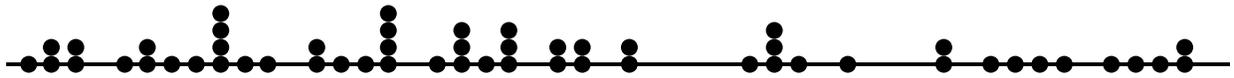


ABBILDUNG 3. Poisson-Prozess auf \mathbb{Z} mit Intensitätsmaß $\mu = \sum_{n \in \mathbb{Z}} \delta_n$.

Beispiel 9.4.6 (Homogener Poisson-Punktprozess auf \mathbb{R}_+). 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).



ABBILDUNG 4. Poisson-Prozess auf \mathbb{R}_+ mit Intensität 1.

Beispiel 9.4.7 (Homogener Poisson-Punktprozess auf \mathbb{R}). 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 .



ABBILDUNG 5. Poisson-Prozess auf \mathbb{R} mit Intensität 1.

9.5. Superpositionssatz

Eine wichtige Eigenschaft der Poisson-Verteilung ist ihre Faltungsstabilität. 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. \square

Aufgabe 9.5.4 (α -stabile Verteilungen). Seien $0 < P_1 < P_2 < \dots$ die Punkte eines homogenen Poisson Punktprozesses auf $(0, \infty)$ mit Intensität 1. Für $\alpha \in (0, 1)$ definiere $S = \sum_{n=1}^{\infty} \frac{1}{P_n^{1/\alpha}}$.

- (a) Zeigen Sie, dass diese Reihe mit Wahrscheinlichkeit 1 konvergiert.
Hinweis: P_n kann in der Form $\nu_1 + \dots + \nu_n$ dargestellt werden, wobei ν_i unabhängig und $\text{Exp}(1)$ -verteilt sind.
- (b) Seien S_1, \dots, S_k unabhängige identisch verteilte Zufallsvariablen mit der gleichen Verteilungsfunktion wie S . Zeigen Sie, dass die Zufallsvariable

$$\frac{S_1 + \dots + S_k}{k^{1/\alpha}}$$

die gleiche Verteilung hat wie S .

Hinweis: Die Zufallsvariable S_i kann mit Hilfe des Poisson-Punktprozesses $\Pi_i = \sum_{n=1}^{\infty} P_{n,i}$ dargestellt werden.

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. \square

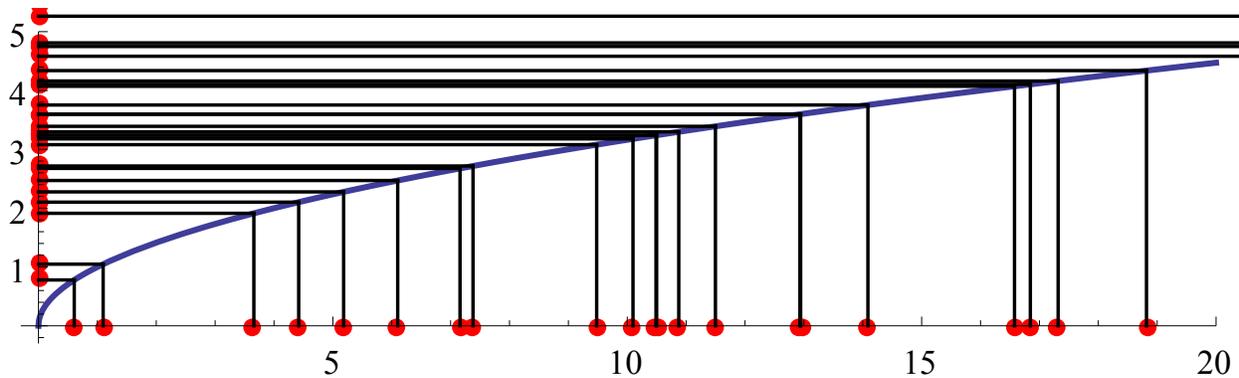


ABBILDUNG 6. Auf der horizontalen Achse sind die Punkte P_1, P_2, \dots aufgetragen, die einen homogenen Poisson-Punktprozess bilden. Diese werden auf die Punkte $\sqrt{P_1}, \sqrt{P_2}, \dots$ abgebildet.

Beispiel 9.6.8 (Siehe Abbildung 6). 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 (Übungsaufgabe). Sei μ das Lebesgue-Maß auf $[0, \infty)$, dann gilt

$$(T\mu)([0, y]) = \mu(T^{-1}([0, y])) = \mu([0, y^2]) = y^2.$$

Somit hat $T\mu$ die Dichte $f(y) = 2y$, $x > 0$, und die Intensität des Poisson-Punktprozesses mit den Punkten $\sqrt{P_1}, \sqrt{P_2}, \dots$ ist $2y$.

Aufgabe 9.6.9. Sei $\Pi = \sum_{i=1}^{\infty} \delta_{Z_i}$ ein homogener Poisson-Punktprozess auf \mathbb{R}^d mit Intensität λ . Zeigen Sie, dass Π stationär ist: Für jedes $u \in \mathbb{R}^d$ bilden die verschobenen Punkte $Z_i + u$, $i \in \mathbb{N}$, ebenfalls einen homogenen Poisson-Punktprozess mit Intensität λ .

Aufgabe 9.6.10. Die Positionen der Sterne im Universum seien mit einem homogenen Poisson-Punktprozess Π auf \mathbb{R}^3 modelliert, dessen Intensität gleich $\lambda > 0$ sei. Für $x \in \mathbb{R}^3$ sei $\|x\|$ der Abstand von x zum Koordinatenursprung. Die Punkte von Π (also die Positionen der Sterne) seien mit X_1, X_2, \dots bezeichnet, wobei die Nummerierung so gewählt sei, dass $\|X_1\| < \|X_2\| < \dots$

- Bestimmen Sie die Verteilungsfunktion und den Erwartungswert von $\|X_1\|$.
- Zeigen Sie, dass die Punkte $\|X_i\|$ einen Poisson-Punktprozess auf \mathbb{R} bilden und geben Sie die Dichte seines Intensitätsmaßes an.

Aufgabe 9.6.11. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit stetiger Verteilungsfunktion F . Zeigen Sie: Die Rekordwerte $X(1), X(2), \dots$ bilden einen Poisson-Punktprozess auf (x_*, x^*) mit Intensitätsmaß

$$\mu(a, b] = \log(\bar{F}(a)/\bar{F}(b)), \quad x_* < a \leq b < x^*.$$

Hinweis: Für exponentialverteilte Zufallsvariablen folgt das aus der Tata-Darstellung. Im Allgemeinen Fall benutzen Sie Satz 8.5.13. Für eine diskrete Version dieser Aufgabe siehe Aufgabe ??.

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 $f : E \rightarrow [0, \infty)$.

Definition 9.7.2. Sei Π ein Punktprozess auf E . Für $f \in B_+(E)$ definiere die **lineare Statistik** als die Zufallsvariable

$$S_f = \sum_{x \in \Pi} f(x) = \int_E f d\Pi$$

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 gilt für alle Borel-Funktionen $f : E \rightarrow [0, \infty)$, dass

$$(9.7.1) \quad \mathbb{E}[e^{-S_f}] = \exp \left\{ - \int_E (1 - e^{-f(x)}) \mu(dx) \right\}.$$

BEWEIS. SCHRITT 1 (Indikatorfunktionen). Sei zuerst $f(x) = c \mathbb{1}_A(x)$, mit $c \geq 0$, $A \subset E$ Borel. Dann ist $S_f = c\Pi(A)$. Es gilt $\Pi(A) \sim \text{Poi}(\mu(A))$. Es folgt

$$\mathbb{E}[e^{-S_f}] = \mathbb{E}[e^{-c\Pi(A)}] = \sum_{k=0}^{\infty} e^{-\mu(A)} \frac{\mu(A)^k}{k!} e^{-ck} = e^{-\mu(A)} \sum_{k=0}^{\infty} \frac{(\mu(A)e^{-c})^k}{k!} = e^{-\mu(A)(1-e^{-c})}.$$

Auf der anderen Seite gilt für $f(x) = c \mathbb{1}_A(x)$

$$\exp \left\{ - \int_E (1 - e^{-f(x)}) \mu(dx) \right\} = \exp \{ -\mu(A) \cdot (1 - e^{-c}) \}.$$

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 $c_1, \dots, c_n \geq 0$. Dann gilt:

$$\mathbb{E}[e^{-S_f}] = \mathbb{E}[e^{-c_1\Pi(A_1)} \cdot \dots \cdot e^{-c_n\Pi(A_n)}] = \prod_{i=1}^n \mathbb{E}[e^{-c_i\Pi(A_i)}],$$

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 (Approximation). 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-Formel). 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. \square

Aufgabe 9.7.6. Beweisen Sie die folgende Ergänzung zur Campbell-Formel: Für alle Borel-Funktionen $f : E \rightarrow [0, \infty)$,

$$\mathbb{E}[S_f^2] = \int_E f^2(x) \mu(dx) + \left(\int_E f(x) \mu(dx) \right)^2.$$

Aufgabe 9.7.7. In einem zweidimensionalen kreisförmigen Land vom Radius $R > 0$ seien die Städte X_1, \dots, X_N durch einen homogenen Poisson-Punktprozess mit Intensität $\lambda > 0$ modelliert. Die Hauptstadt des Landes liege dabei im Mittelpunkt des Kreises. Es sollen nun alle Städte auf direktem Weg mit der Hauptstadt verbunden werden. Bestimmen Sie

die erwartete Streckenlänge aller Straßen zusammen, d.h. bestimmen Sie

$$\mathbb{E} \left[\sum_{i=1}^N \|X_i\| \right],$$

wobei mit $\|x\|$ der Abstand des Punktes x zum Ursprung bezeichnet sei.

Im nächsten Satz zeigen wir, dass man einen Poisson-Punktprozess an seinem Laplace-Funktional erkennen kann.

Satz 9.7.8. 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. □

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 7. 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 vom Färbungsvorgang) 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.

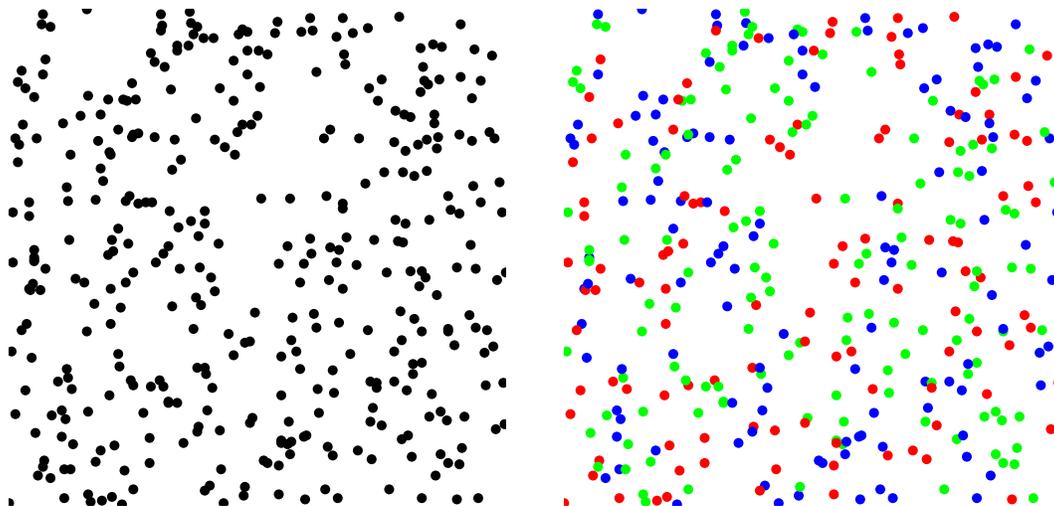


ABBILDUNG 7. Färbungssatz.

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) für jede Borel-Menge $B \subset E_2$ ist $x \mapsto \lambda_x(B)$ eine Borel-Funktion auf E_1 .

Beispiel 9.8.1. Die Menge E_2 aller möglichen Farben sei 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 $p_1, \dots, p_n : E_1 \rightarrow [0, 1]$ Borel-Funktionen mit $\sum_{i=1}^n p_i(x) = 1$ 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.¹

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. Im markierten Punktprozess Π^* sind sowohl die Positionen der Punkte als auch deren Marken gespeichert. 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

$$\mu^*(C) = \iint_{(x,m) \in C} \mu(dx) \lambda_x(dm), \text{ für } C \subset E_1 \times E_2 \text{ Borel.}$$

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 $f : E_1 \times E_2 \rightarrow \mathbb{R}$ eine nichtnegative Borel-Funktion. Wir betrachten die lineare Statistik

$$S^* = \sum_{(x,m) \in \Pi^*} f(x, m).$$

Unser Ziel ist es, das Laplace-Funktional $\mathbb{E}[e^{-S^*}]$ zu bestimmen. Dazu wollen wir zuerst den folgenden bedingten Erwartungswert berechnen:

$$\mathbb{E}[e^{-S^*} | \Pi] = \mathbb{E} \left[\prod_{x \in \Pi} e^{-f(x, m_x)} \middle| \Pi \right] = \prod_{x \in \Pi} \mathbb{E}[e^{-f(x, m_x)} | \Pi] = \prod_{x \in \Pi} \int_{E_2} e^{-f(x, m)} \lambda_x(dm).$$

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 ergibt sich

$$\mathbb{E}[e^{-S^*}] = \mathbb{E}[\mathbb{E}[e^{-S^*} | \Pi]] = \mathbb{E} \left[\exp \left\{ - \sum_{x \in \Pi} f^*(x) \right\} \right]$$

¹Hätte der Punktprozess Π multiple Punkte, so müssten wir für jedes $x \in E_1$ einen Vorrat aus unendlich vielen unabhängigen Marken erzeugen müssen, um damit im Notfall alle an der Stelle x auftretenden Atome von Π versorgen zu können. Da das die Konstruktion verkomplizieren würde, nehmen wir an, dass Π keine multiplen Punkte hat.

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:

$$\mathbb{E}[e^{-S^*}] = \exp \left\{ - \int_{E_1} \int_{E_2} (1 - e^{-f(x,m)}) \lambda_x(dm) \mu(dx) \right\}.$$

Mit $\mu^*(dx, dm) := \mu(dx) \lambda_x(dm)$ und mit Satz 9.7.8 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.

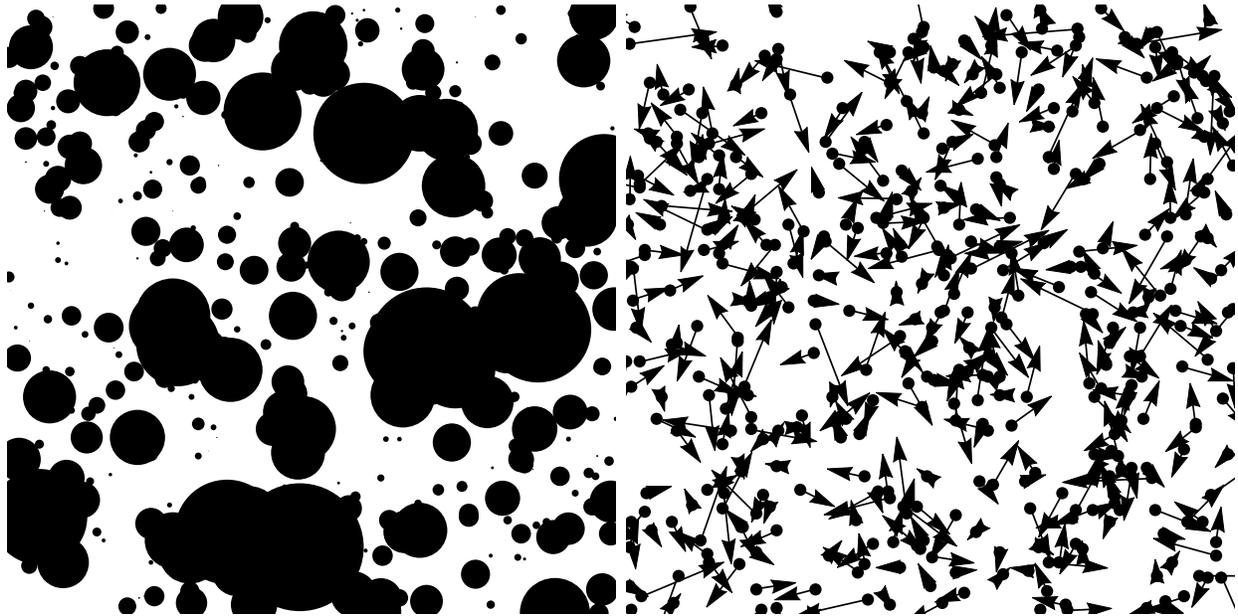


ABBILDUNG 8. Links: Boole'sches Modell (Aufgabe 9.8.5). Rechts: Ideales Gas (Aufgabe 9.8.6).

Aufgabe 9.8.5 (Boole'sches Modell (Abbildung 8)). Es sei $\Pi = \sum_{i=1}^{\infty} \delta_{Z_i}$ ein homogener Poisson-Punktprozess auf \mathbb{R}^2 mit Intensität 1. Seien R_1, R_2, \dots unabhängige identisch verteilte nichtnegative Zufallsvariablen mit Verteilungsfunktion F , die auch von Π unabhängig

sind. Betrachte die zufällige Menge

$$B = \bigcup_{i=1}^{\infty} B_{R_i}(Z_i) \subset \mathbb{R}^2,$$

wobei $B_{R_i}(Z_i) = \{z \in \mathbb{R}^2 : \|z - Z_i\| \leq R_i\}$ ein Kreis um Z_i mit Radius R_i ist. Bestimmen Sie $\mathbb{P}[0 \in B]$. Bestimmen Sie den Erwartungswert von des Lebesgue-Maßes der zufälligen Menge $[0, 1]^2 \cap B$.

Hinweis: Benutzen Sie den Markierungssatz.

Aufgabe 9.8.6 (Ideales Gas, Abbildung 8).

9.9. Simulation

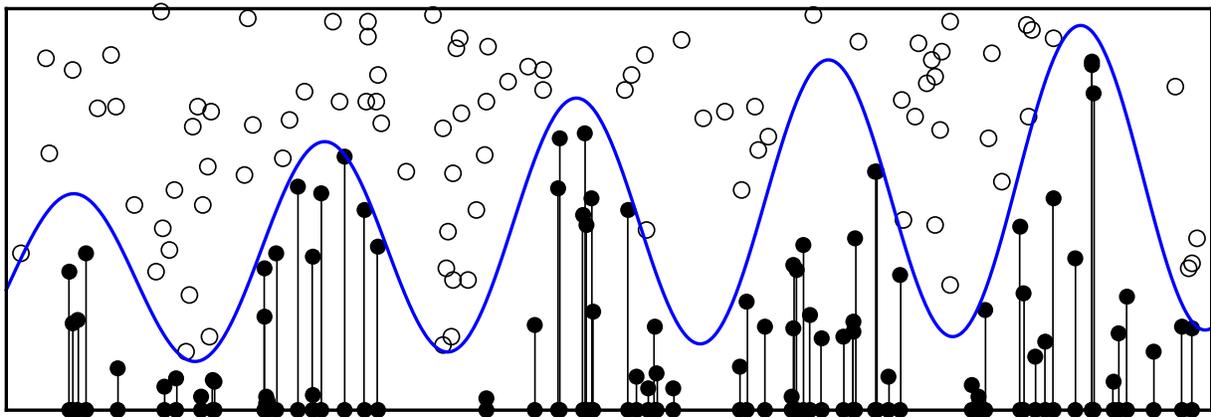


ABBILDUNG 9. Simulation eines Poisson-Punktprozesses mit Intensität f .

9.10. Poisson-Linienprozess

Kommentare

[15], [17], [19].

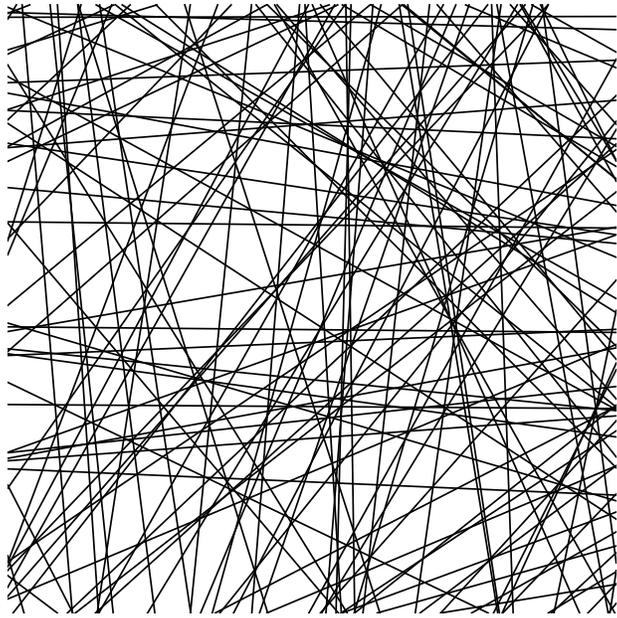


ABBILDUNG 10. Poisson-Linienprozess.

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 $\lim_{n \rightarrow \infty} \psi_{\Pi_n}(f) = \psi_{\Pi}(f)$, 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 Familie $\varepsilon_{ni}, i \in \mathbb{Z}$, 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 eine Funktion $f \in C_c^+(\mathbb{R})$ 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. Laut Satz 10.2.2 reicht es zu zeigen, dass

$$\lim_{n \rightarrow \infty} \psi_n(f) = \psi(f).$$

Wegen der Unabhängigkeit der Familie $\varepsilon_{ni}, i \in \mathbb{Z}$, gilt

$$\psi_n(f) = \mathbb{E} \left[\exp \left\{ - \sum_{i \in \mathbb{Z}} \varepsilon_{ni} f \left(\frac{i}{n} \right) \right\} \right] = \mathbb{E} \prod_{i \in \mathbb{Z}} e^{-\varepsilon_{ni} f(\frac{i}{n})} = \prod_{i \in \mathbb{Z}} \mathbb{E} e^{-\varepsilon_{ni} f(\frac{i}{n})} = \prod_{i \in \mathbb{Z}} (1 + p_n (e^{-f(\frac{i}{n})} - 1)).$$

Durch Logarithmieren ergibt sich

$$\log \psi_n(f) = \sum_{i \in \mathbb{Z}} \log(1 + p_n (e^{-f(\frac{i}{n})} - 1)).$$

Die Entwicklung $\log(1 + x) = x + o(x)$ (für $x \rightarrow 0$) legt nahe, diese Formel als

$$\log \psi_n(f) = \sum_{i \in \mathbb{Z}} p_n (e^{-f(\frac{i}{n})} - 1) + R_n,$$

umzuschreiben, wobei

$$R_n := \sum_{i \in \mathbb{Z}} \left(\log(1 + p_n (e^{-f(\frac{i}{n})} - 1)) - p_n (e^{-f(\frac{i}{n})} - 1) \right)$$

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

$$|R_n| \leq \sum_{i \in \mathbb{Z}} \left| \log(1 + p_n(e^{-f(\frac{i}{n})} - 1)) - p_n(e^{-f(\frac{i}{n})} - 1) \right| \leq \frac{1}{2} \sum_{i \in \mathbb{Z}} p_n^2 (e^{-f(\frac{i}{n})} - 1)^2,$$

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.

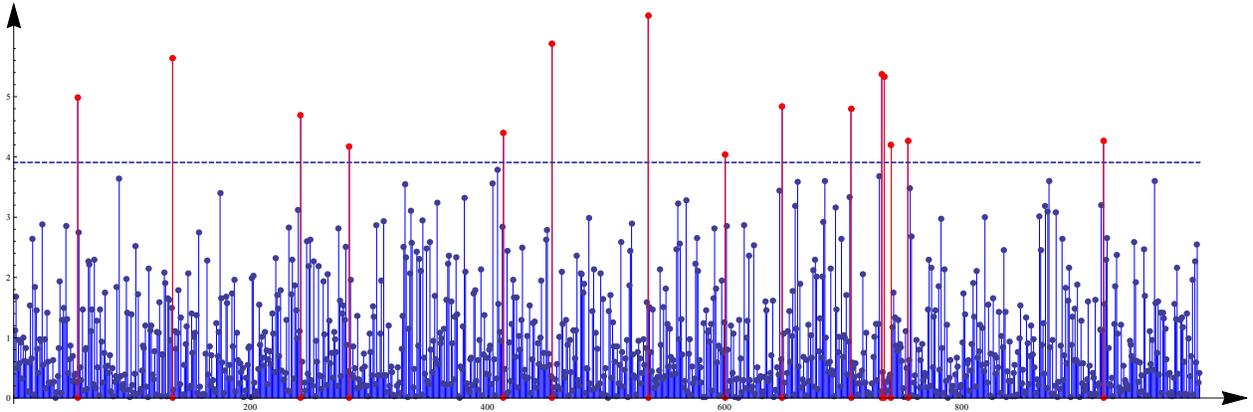


ABBILDUNG 1. Überschreitungen eines hohen Schwellenwerts in einer Folge von u.i.v. Zufallsvariablen

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:

$$\psi_n(f) = \prod_{i=1}^n \mathbb{E} e^{-f(Y_{ni})} = (\mathbb{E} e^{-f(Y_{n1})})^n = \left(\int_E e^{-f(t)} \mu_n(dt) \right)^n = \left(1 - \frac{\int_E (1 - e^{-f(t)}) n\mu_n(dt)}{n} \right)^n.$$

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 $a_n \in \mathbb{R}$ und $b_n > 0$ 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} :

$$\Pi_n := \sum_{i=1}^n \delta_{\frac{X_i - a_n}{b_n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(e^{-t} dt).$$

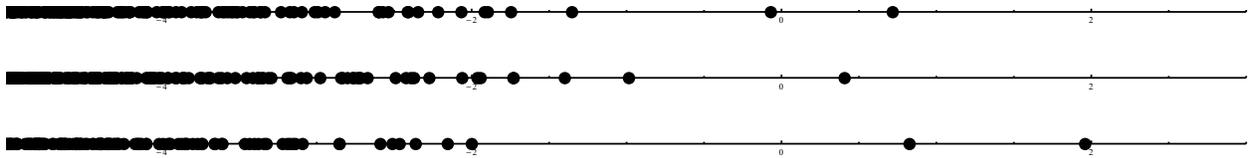


ABBILDUNG 2. Drei unabhängige Realisierungen des Poisson-Punktprozesses mit Intensität e^{-t} auf \mathbb{R} . Das Bild zeigt das Intervall $[-5, 3]$.

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 $\Pi((x, \infty)) \sim \text{Poi}(e^{-x})$.

BEWEISIDEE VON SATZ 10.5.1. Wir wissen, dass die Konvergenz in (10.5.1) genau dann gilt, wenn

$$\lim_{n \rightarrow \infty} n\mathbb{P}[X_1 > a_n + b_n t] = e^{-t}$$

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 10.5.3. 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 1.2.1

$$\max\{X_1, \dots, X_n\} - \log n \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

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 10.5.4. 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

$$\sum_{i=1}^n \delta_{\sqrt{2 \log n} (X_i - a_n)} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(e^{-t} dt).$$

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 10.5.5. Unter Voraussetzungen von Satz 10.5.1 gilt für alle $r \in \mathbb{N}$

$$(10.5.2) \quad \left(\frac{M_n^{(1)} - a_n}{b_n}, \dots, \frac{M_n^{(r)} - a_n}{b_n} \right) \xrightarrow[n \rightarrow \infty]{d} (Y_1, \dots, Y_r),$$

wobei $Y_1 > Y_2 > \dots$ die Punkte des Poisson-Punktprozesses Π mit Intensität e^{-t} auf \mathbb{R} seien. Die gemeinsame Dichte von (Y_1, \dots, Y_r) ist gegeben durch

$$f_{Y_1, \dots, Y_r}(y_1, \dots, y_r) = e^{-e^{-y_r}} e^{-(y_1 + \dots + y_r)} \mathbb{1}_{\{y_1 > y_2 > \dots > y_r\}}.$$

BEWEIS. ???

Wir bestimmen die gemeinsame Dichte von (Y_1, \dots, Y_r) an der Stelle $(y_1, \dots, y_r) \in \mathbb{R}^r$. Wir nehmen an, dass $y_1 > \dots > y_r$, denn sonst verschwindet die Dichte wegen $Y_1 > \dots > Y_r$. Wir betrachten unendlich kleine Intervalle $(y_1, y_1 + dy_1), \dots, (y_r, y_r + dy_r)$ und bestimmen die Wahrscheinlichkeit, dass Y_1, \dots, Y_r in diesen Intervallen liegen. Damit das der Fall ist, muss Folgendes eintreten:

- In jedem der Intervalle $(y_i, y_i + dy_i)$ befindet sich ein Punkt des Poisson-Punktprozesses. Das hat (infinitesimale) Wahrscheinlichkeit $e^{-y_1} dy_1 \cdot \dots \cdot e^{-y_r} dy_r$.
- In den Intervallen $(y_1 + dy_1, \infty), (y_2 + dy_2, y_1), \dots, (y_r + dy_r, y_{r-1})$ liegen keine Punkte des Poisson-Punktprozesses. Ganz allgemein ist die Wahrscheinlichkeit, dass in einem Intervall (a, b) keine Punkte liegen, gegeben durch

$$\mathbb{P}[\Pi((a, b)) = 0] = e^{-\int_a^b e^{-t} dt} = e^{-e^{-a} - e^{-b}}.$$

Somit konvergiert die Wahrscheinlichkeit des uns interessierenden Ereignisses für $dy_1, \dots, dy_r \rightarrow 0$ gegen

$$e^{-e^{-y_1}} \cdot e^{-e^{-y_2} - e^{-y_1}} \cdot \dots \cdot e^{-e^{-y_r} + e^{-y_{r-1}}} = e^{-e^{-y_r}}.$$

Für die Dichte von (Y_1, \dots, Y_r) erhalten wir die Formel

$$\begin{aligned} f_{Y_1, \dots, Y_r}(y_1, \dots, y_r) &= \lim_{dy_1, \dots, dy_r \rightarrow 0} \frac{\mathbb{P}[y_1 < Y_1 < y_1 + dy_1, \dots, y_r < Y_r < y_r + dy_r]}{dy_1 \dots dy_r} \\ &= e^{-e^{-y_r}} e^{-(y_1 + \dots + y_r)}, \end{aligned}$$

was die Behauptung beweist. □

Bemerkung 10.5.6. Wir hatten bereits in Satz (7.2.5) gezeigt, dass eine Grenzwertverteilung in (10.5.3) existiert, konnten diese allerdings nicht identifizieren. Nun haben wir mit Hilfe von Poisson-Punktprozessen eine explizite Formel für die Dichte hergeleitet.

Korollar 10.5.7. Unter Voraussetzungen von Satz 10.5.1 gilt für alle $k \in \mathbb{N}$

$$(10.5.3) \quad \frac{M_n^{(k)} - a_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} Y_k,$$

wobei $Y_1 > Y_2 > \dots$ die Punkte des Poisson-Punktprozesses Π mit Intensität e^{-t} auf \mathbb{R} seien. Die Dichte von Y_k ist gegeben durch

$$f_{Y_k}(y) = e^{-e^{-y}} \frac{e^{-ky}}{(k-1)!}, \quad y \in \mathbb{R}.$$

Aufgabe 10.5.8. Seien $Y_1 > Y_2 > \dots$ die absteigend angeordneten Punkte des Poisson-Punktprozesses mit Intensität e^{-t} auf \mathbb{R} .

(a) Zeigen Sie, dass für jedes $r \in \mathbb{N}$

$$(Y_1, Y_2, \dots, Y_r) \stackrel{d}{=} \left(\gamma + \sum_{k=1}^{\infty} \frac{E_k - 1}{k}, \gamma - 1 + \sum_{k=2}^{\infty} \frac{E_k - 1}{k}, \dots, \gamma - \sum_{k=1}^{r-1} \frac{1}{k} + \sum_{k=r}^{\infty} \frac{E_k - 1}{k} \right),$$

wobei $E_1, E_2, \dots \sim \text{Exp}(1)$ unabhängige Zufallsvariablen sind und γ die Euler-Mascheroni-Konstante bezeichnet.

Hinweis: Es seien $Z_1, Z_2, \dots \sim \text{Exp}(1)$ unabhängig. Auf der einen Seite konvergiert $(Z_{n:n} - \log n, \dots, Z_{n-r+1:n} - \log n)$ gegen (Y_1, \dots, Y_r) in Verteilung. Auf der anderen Seite kann man Satz 7.3.1 benutzen.

(b) Zeigen Sie: $Y_1 - Y_2 \sim \text{Exp}(1), Y_2 - Y_3 \sim \text{Exp}(2), \dots$, und diese Zufallsvariablen sind unabhängig.

(c) Zeigen Sie:

$$\mathbb{E}Y_r = \gamma - \left(1 + \frac{1}{2} + \dots + \frac{1}{r-1} \right).$$

In der nächsten Proposition 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 10.5.9. Es sei $\sum_{n=1}^{\infty} \delta_{P_n}$ ein homogener Poisson-Punktprozess mit Intensität 1 auf $(0, \infty)$. Dann gilt

$$\sum_{n=1}^{\infty} \delta_{-\log P_n} \sim \text{PPP}(e^{-t} dt).$$

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. □

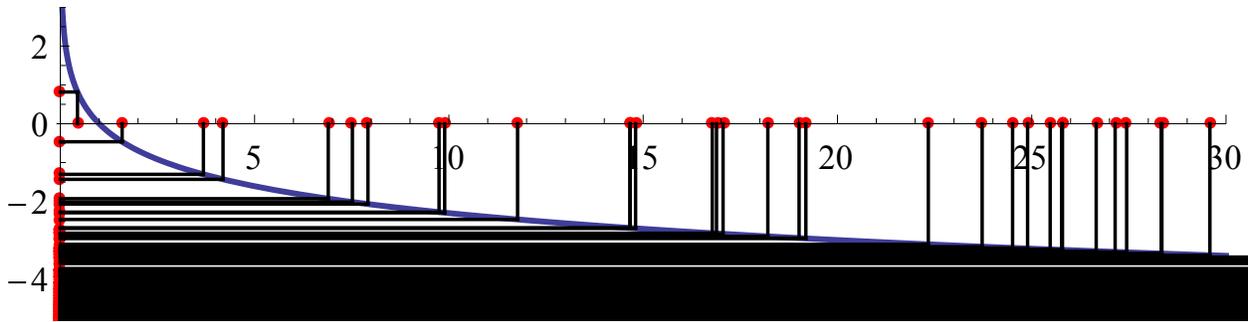


ABBILDUNG 3. Poisson-Punktprozess mit Intensität 1 (horizontale Achse) wird mit der Transformation $T(x) = -\log x$ auf den Poisson-Punktprozess mit Intensität e^{-t} (vertikale Achse) transformiert.

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.

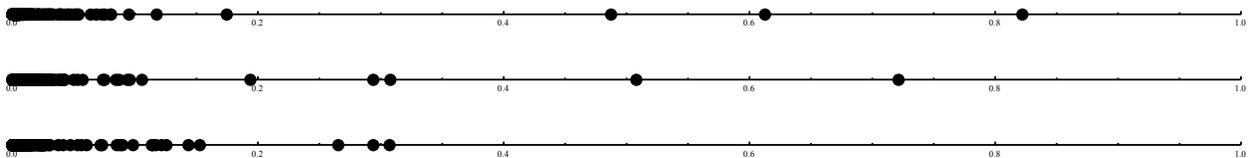


ABBILDUNG 4. Drei unabhängige Realisierungen des Poisson-Punktprozesses mit Intensität $t^{-2}dt$ auf $(0, \infty)$. Das Bild zeigt das Intervall $[0, 1]$.

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 $[t, \infty)$:

$$n\mu_n([t, \infty)) = n\mathbb{P}[Y_{ni} \geq t] = n\mathbb{P}[X_1 \geq b_n t] \xrightarrow{n \rightarrow \infty} \frac{1}{t^\alpha}.$$

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 1.2.5 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

$$\frac{\max\{X_1, \dots, X_n\}}{n} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha$$

mit $\alpha = 1$. Somit folgt aus Satz 10.6.1, dass auf $(0, \infty)$

$$\sum_{\substack{i=1, \dots, n \\ X_i > 0}} \delta_{\frac{X_i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{1}{t^2} dt \right).$$

Aus Symmetriegründen folgt aber auch, dass auf $(-\infty, 0)$

$$\sum_{\substack{i=1, \dots, n \\ X_i < 0}} \delta_{\frac{X_i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{1}{(-t)^2} dt \right).$$

Man kann sogar beide Fälle vereinigen und zeigen, dass auf $E = \mathbb{R} \setminus \{0\}$ gilt

$$\sum_{i=1}^n \delta_{\frac{X_i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left(\frac{1}{|t|^2} dt \right).$$

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. □

Aufgabe 10.7.2. Zeigen Sie: Die Abbildung $x \mapsto e^x$ bildet einen homogenen PPP auf \mathbb{R} mit Intensität 1 auf einen PPP(dt/t) auf $(0, \infty)$ ab.

Aufgabe 10.7.3 (Skalenfreiheit). Zeigen Sie: Ist $\sum_{i=1}^{\infty} \delta_{Y_i}$ ein PPP(dt/t) auf $(0, \infty)$ und $c > 0$ eine Konstante, so ist $\sum_{i=1}^{\infty} \delta_{cY_i}$ ein PPP mit der gleichen Intensität.

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 $a_n \in \mathbb{R}$ und $b_n > 0$ 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

$$\sum_{i=1}^{\infty} \delta_{(U_i, Y_i)} \sim \text{PPP}(ds \times e^{-t} dt).$$

Der obige Satz macht den folgenden “funktionalen Grenzwertsatz” für den Maximumprozess plausibel. Zuerst benötigen wir eine Definition.

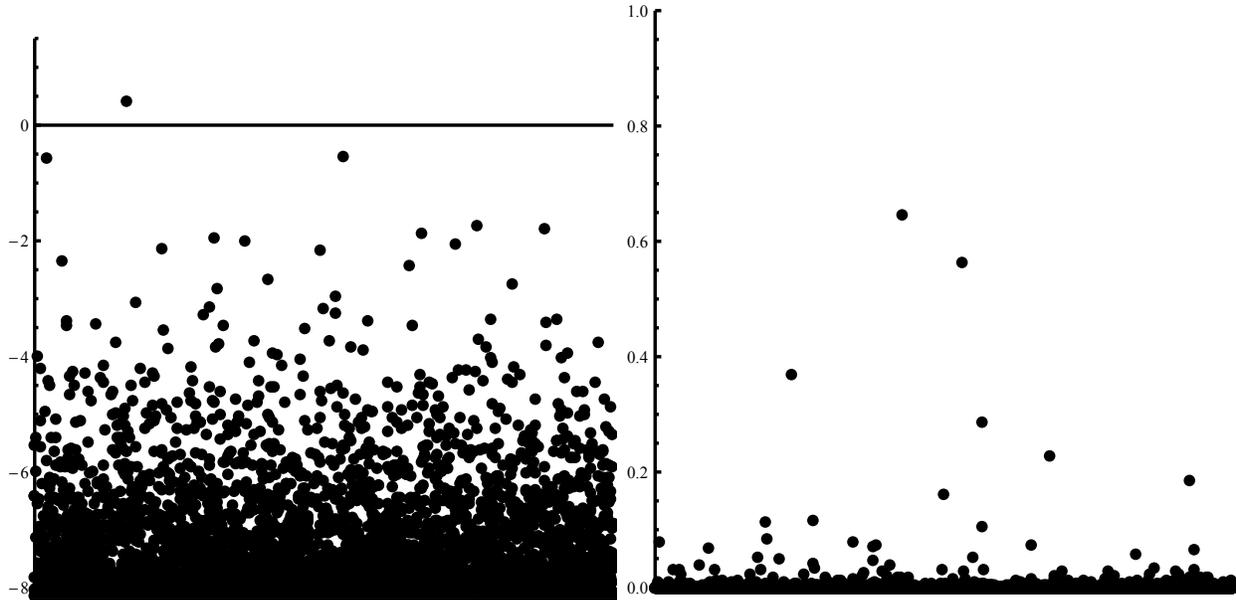


ABBILDUNG 5. Poisson-Punktprozesse aus Sätzen 10.8.1 und 10.8.4.

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

$$\{(M_{[nt]} - a_n)/b_n: t \in [0, 1]\}$$

gegen den Gumbel-Extremwertprozess im Sinne der endlich-dimensionalen Verteilungen.

D.h. für alle $0 < t_1 < \dots < t_k \leq 1$ gilt

$$\left(\frac{M_{[nt_1]} - a_n}{b_n}, \dots, \frac{M_{[nt_k]} - a_n}{b_n} \right) \xrightarrow[n \rightarrow \infty]{d} (Z_{t_1}, \dots, Z_{t_k}).$$

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. □

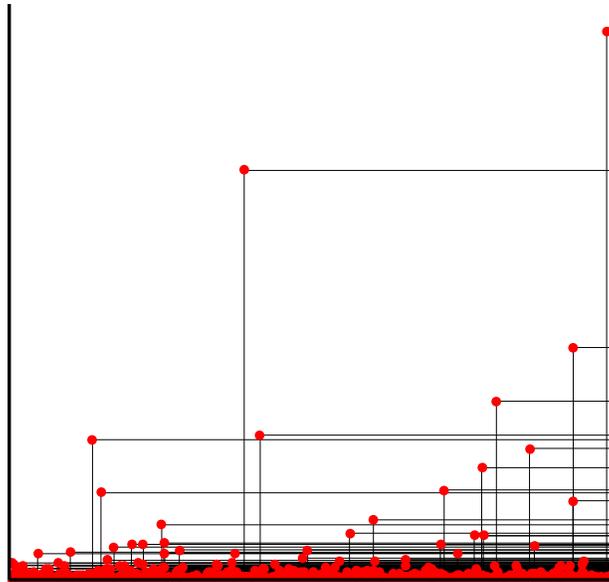


ABBILDUNG 6. Eine Realisierung des Fréchet-Extremwertprozesses aus Satz 10.8.4.

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.

Kommentare

[22].

Literaturverzeichnis

- [1] M. Ahsanullah and V. B. Nevzorov. *Records via probability theory*, volume 6 of *Atlantis Studies in Probability and Statistics*. Atlantis Press, Paris, 2015.
- [2] M. Ahsanullah, V. B. Nevzorov, and M. Shakil. *An introduction to order statistics*, volume 3 of *Atlantis Studies in Probability and Statistics*. Atlantis Press, Paris, 2013.
- [3] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A first course in order statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1992. A Wiley-Interscience Publication.
- [4] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Records*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1998. A Wiley-Interscience Publication.
- [5] A. A. Balkema and L. de Haan. On R. von Mises' condition for the domain of attraction of $\exp(-e^{-x})^1$. *Ann. Math. Statist.*, 43:1352–1354, 1972.
- [6] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers. *Statistics of extremes*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2004. Theory and applications, With contributions from Daniel De Waal and Chris Ferro.
- [7] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.
- [8] S. Coles. *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. Springer-Verlag London, Ltd., London, 2001.
- [9] H. A. David and H. N. Nagaraja. *Order statistics*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2003.
- [10] L. de Haan. *On regular variation and its application to the weak convergence of sample extremes*, volume 32 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1970.
- [11] L. de Haan and A. Ferreira. *Extreme value theory*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006. An introduction.
- [12] J. Galambos. *The asymptotic theory of extreme order statistics*. Robert E. Krieger Publishing Co., Inc., Melbourne, FL, second edition, 1987.
- [13] E. J. Gumbel. *Statistics of extremes*. Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1958 original [Columbia University Press, New York; MR0096342].
- [14] H. Herrlich. *Axiom of choice*, volume 1876 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006.
- [15] J. F. C. Kingman. *Poisson processes*, volume 3 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.
- [16] M. Kuczma. *An introduction to the theory of functional equations and inequalities*. Birkhäuser Verlag, Basel, second edition, 2009. Cauchy's equation and Jensen's inequality, Edited and with a preface by Attila Gilányi.
- [17] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2018.
- [18] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York-Berlin, 1983.
- [19] K. Matthes, J. Kerstan, and J. Mecke. *Infinitely divisible point processes*. John Wiley & Sons, Chichester-New York-Brisbane, 1978. Translated from the German by B. Simon, Wiley Series in Probability and Mathematical Statistics.

- [20] V. B. Nevzorov. *Records: mathematical theory*, volume 194 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2001. Translated from the Russian manuscript by D. M. Chibisov.
- [21] R.-D. Reiss. *Approximate distributions of order statistics*. Springer Series in Statistics. Springer-Verlag, New York, 1989. With applications to nonparametric statistics.
- [22] S. I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.
- [23] M. Q. Shahbaz, M. Ahsanullah, S. Hanif Shahbaz, and B. M. Al-Zahrani. *Ordered random variables: theory and applications*, volume 9 of *Atlantis Studies in Probability and Statistics*. Atlantis Press, Paris, 2016.
- [24] G. R. Shorack and J. A. Wellner. *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.