



ulm university universität  
**uulm**

**Universität Ulm**  
**Fakultät für Mathematik und**  
**Wirtschaftswissenschaften**

## **Extremwerttheorie**

**Skript**

Jun.-Prof. Dr. Zakhar Kabluchko

Wintersemester 2011/12

L<sup>A</sup>T<sub>E</sub>X-Version von Benjamin Tempel

# Inhaltsverzeichnis

<b>1</b>	<b>Extremwertverteilungen</b>	<b>1</b>
<b>2</b>	<b>Max-Anziehungsbereiche</b>	<b>7</b>
2.1	Max-Anziehungsbereich der Fréchet-Verteilung $\Phi_\alpha$ . . . . .	7
2.2	Max-Anziehungsbereich der Weibull-Verteilung $\Psi_\alpha$ . . . . .	10
2.3	Max-Anziehungsbereich der Gumbel-Verteilung $G$ . . . . .	12
<b>3</b>	<b>Satz von Fisher–Tippett</b>	<b>16</b>
<b>4</b>	<b>Statistik der Extremwertverteilungen</b>	<b>22</b>
4.1	Statistik der GEV-Verteilungen . . . . .	22
4.2	Statistik der GP-Verteilungen . . . . .	27
<b>5</b>	<b>Ordnungsstatistiken</b>	<b>30</b>
5.1	Extreme Ordnungsstatistiken . . . . .	30
5.2	Allgemeine Eigenschaften der Ordnungsstatistiken . . . . .	33
5.3	Darstellungen der Ordnungsstatistiken . . . . .	35
<b>6</b>	<b>Rekorde</b>	<b>39</b>
6.1	Satz von Rényi . . . . .	39
6.2	Rekordzeiten . . . . .	40
6.3	Anzahl der Rekorde . . . . .	43
6.4	Zentrale Grenzwertsätze für Rekorde . . . . .	45
<b>7</b>	<b>Poisson-Punktprozesse</b>	<b>51</b>
7.1	Poisson-Verteilung . . . . .	51
7.2	Beispiel zu Poisson-Prozessen . . . . .	51
7.3	Definition der Poisson-Prozesse . . . . .	52
7.3.1	Zählmaße . . . . .	52
7.3.2	Punktprozesse . . . . .	53
7.3.3	Poisson-Punktprozesse . . . . .	54
7.4	Eigenschaften der Poisson-Punktprozesse . . . . .	55
7.4.1	Superpositionssatz . . . . .	55
7.4.2	Abbildungssatz . . . . .	56
7.4.3	Laplacefunktionale . . . . .	57

7.5	Markierungen von Poisson-Punktprozessen . . . . .	59
<b>8</b>	<b>Konvergenz von Punktprozessen</b>	<b>62</b>
8.1	Definition der Konvergenz von Punktprozessen . . . . .	62
8.2	Konvergenz der Überschreitungszeitpunkte gegen einen Poisson-Punktprozess	63
8.3	Konvergenz der oberen Ordnungsstatistiken gegen einen Poisson-Punktprozess	65
8.4	Konvergenz der Rekordzeiten gegen einen Poisson-Punktprozess . . . . .	69
8.5	Extremwertprozess . . . . .	71
<b>9</b>	<b>Verschiedenes</b>	<b>72</b>
9.1	Wiederkehrperiode . . . . .	72
9.2	Gumbel-Überschreitungsmethode . . . . .	73

# 1 Extremwertverteilungen

Seien  $X_1, X_2, \dots$  unabhängige und identisch verteilte (u.i.v.) Zufallsvariablen. Mit  $M_n$  bezeichnen wir das **Maximum** von  $X_1, \dots, X_n$ :

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

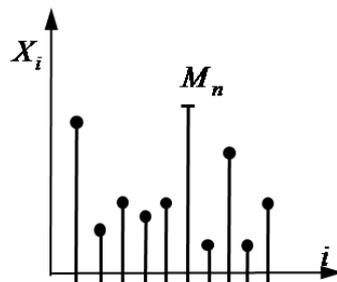


Abbildung 1.1: Veranschaulichung von  $M_n$

Die **Verteilungsfunktion**  $F$  und die **Tailfunktion**  $\bar{F}$  von  $X_i$  sind definiert durch

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

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

**Theorem 1.1.** Die Verteilungsfunktion von  $M_n$  ist gegeben durch

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

*Beweis.* Aus der Definition des Maximums  $M_n$  folgt

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

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

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

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

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

$$x^* = \sup\{t \in \mathbb{R} | F(t) < 1\}.$$

Es kann passieren, dass  $x^*$  den Wert  $+\infty$  annimmt.

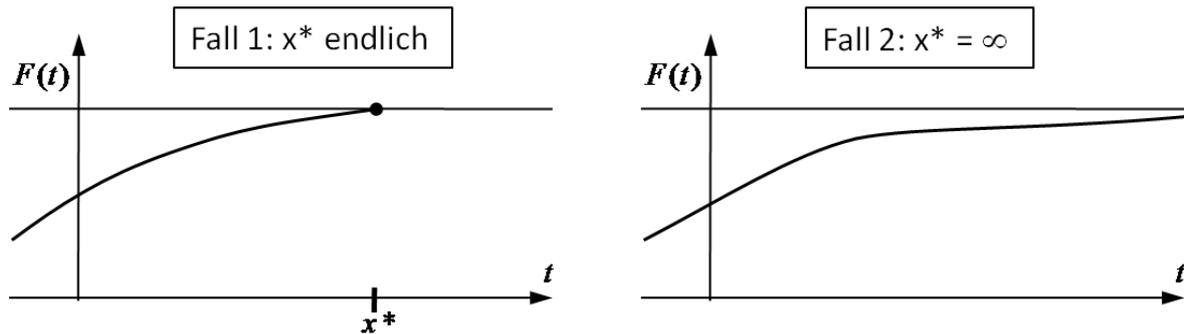


Abbildung 1.2: Veranschaulichung von  $x^*$

**Theorem 1.2.** Für  $n \rightarrow \infty$  konvergiert die Zufallsvariable  $M_n$  in Wahrscheinlichkeit gegen den Wert  $x^*$ .

*Beweis.* Sei zuerst  $x^*$  endlich. Für jedes  $\varepsilon > 0$  gilt  $F(x^* - \varepsilon) < 1$ , wobei die Ungleichung strikt ist. Aus Satz 1.1 folgt, dass

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

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

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

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

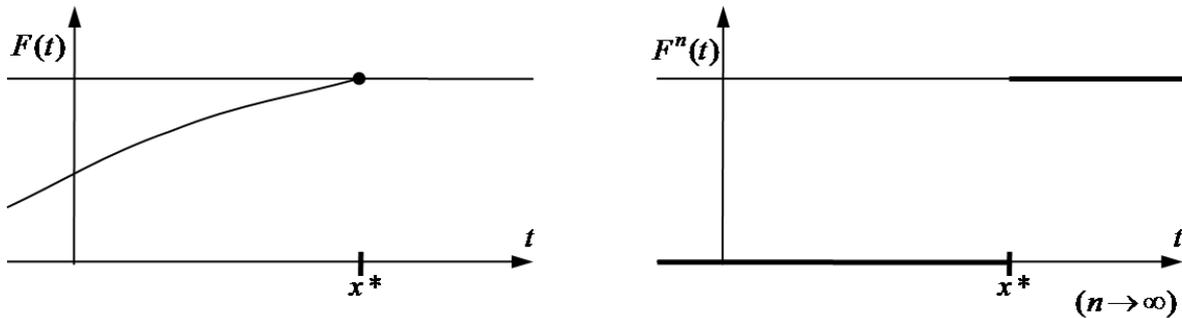
Sei nun  $x^* = +\infty$ . Wir zeigen, dass für jedes  $A \in \mathbb{R}$ , das beliebig groß sein kann,  $\lim_{n \rightarrow \infty} \mathbb{P}[M_n < A] = 0$  gilt. Aus  $x^* = \infty$  folgt, dass  $F(A) < 1$ , wobei die Ungleichung strikt ist. Mit Satz 1.1 folgt

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

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

**Aufgabe 1.3.** Zeigen Sie, dass  $M_n$  sogar fast sicher gegen  $x^*$  konvergiert.

Satz 1.2 gibt nur eine sehr grobe Information darüber, wie die Verteilungsfunktion von  $M_n$  für  $n \rightarrow \infty$  aussieht:



Zum Vergleich besagt das schwache Gesetz der großen Zahlen, dass für u.i.v. Zufallsvariablen  $X_1, X_2, \dots$  mit endlichem Erwartungswert gilt

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow \infty]{P} \mathbb{E}X_1.$$

Dies gibt nur eine sehr grobe Vorstellung darüber, wie die Verteilungsfunktion von  $X_1 + \dots + X_n$  aussieht. Eine präzise Aussage über die Verteilungsfunktion von  $X_1 + \dots + X_n$  liefert der zentrale Grenzwertsatz. Er besagt, dass für u.i.v. Zufallsvariablen  $X_1, X_2, \dots$  mit endlichem Erwartungswert  $\mu$  und endlicher Varianz  $\sigma^2 > 0$

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

im Sinne der Verteilungskonvergenz gilt, wobei  $N(0, 1)$  eine standardnormale Verteilung bezeichnet. Wir stellen nun eine ähnliche Frage für das Maximum  $M_n$ .

**Frage 1.4.** Seien  $X_1, X_2, \dots$  u.i.v. Zufallsvariablen mit Verteilungsfunktion  $F$ . Gibt es Folgen von Konstanten  $a_n > 0$ ,  $b_n \in \mathbb{R}$  und eine Verteilungsfunktion  $G$ , sodass für  $n \rightarrow \infty$  die folgende Verteilungskonvergenz gilt:

$$\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} G? \tag{1.1}$$

Die Verteilungskonvergenz in (1.1) bedeutet, dass für alle  $t \in \mathbb{R}$  in denen  $G$  stetig ist, gilt

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \leq t \right] = G(t).$$

**Definition 1.5.** Eine Verteilungsfunktion  $G$ , für die (1.1) bei irgendeiner Wahl der Normierungskonstanten  $a_n, b_n$  und der Verteilungsfunktion  $F$  gilt, heißt **Extremwertverteilung**. Falls (1.1) gilt, sagen wir, dass  $F$  im **Max-Anziehungsbereich** von  $G$  liegt.

**Bemerkung 1.6.** Durch die “falsche” Wahl der Konstanten  $a_n, b_n$  kann man immer erreichen, dass (1.1) mit einer “ausgearteten” Verteilungsfunktion  $G$  der Form

$$G(t) = \begin{cases} 0, & t < t_0 \\ 1, & t \geq t_0 \end{cases} \quad (1.2)$$

gilt. Eine gemäß  $G$  verteilte Zufallsvariable nimmt nur den Wert  $t_0$  mit Wahrscheinlichkeit 1 an. Zum Beispiel kann man  $a_n$  so groß wählen, dass  $M_n/a_n$  gegen 0 in Verteilung konvergiert. Deshalb werden wir den Fall einer Verteilungsfunktion  $G$  der Form (1.2) immer ausschließen. Eine “ausgeartete” Verteilung gilt nicht als Extremwertverteilung.

Mit der obigen Definition ist es nicht klar, ob Extremwertverteilungen überhaupt existieren. Im Folgenden betrachten wir drei Beispiele, die drei verschiedene Extremwertverteilungen (oder sogar Familien von Extremwertverteilungen) liefern.

**Theorem 1.7.** Seien  $X_i$  exponentialverteilt mit Parameter 1, d.h.  $F(t) = 1 - e^{-t}$  für  $t \geq 0$ . Dann gilt

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

Mit anderen Worten konvergiert die Zufallsvariable  $M_n - \log n$  in Verteilung gegen  $G$ .

*Beweis.* Sei  $t \in \mathbb{R}$  beliebig. Mit Satz 1.1 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  $n$  groß genug. Es folgt

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

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

**Definition 1.8.** Die Verteilungsfunktion  $G(t) = e^{-e^{-t}}$  heißt **Gumbel-Verteilung**.

**Bemerkung 1.9.** Gemäß Satz 1.7 liegt die Exponentialverteilung im Max-Anziehungsbereich der Gumbel-Verteilung. Man kann Satz 1.7 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.

**Theorem 1.10.** Seien  $X_i$  Pareto-verteilt mit Parameter  $\alpha > 0$ , d.h.

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

Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{M_n}{n^{1/\alpha}} \leq t \right] = \begin{cases} e^{-\frac{1}{t^\alpha}}, & t > 0 \\ 0, & t \leq 0 \end{cases} =: \Phi_\alpha(t). \quad (1.3)$$

Mit anderen Worten konvergiert die Zufallsvariable  $n^{-\frac{1}{\alpha}} M_n$  in Verteilung gegen  $\Phi_\alpha$ .

*Beweis.* Sei  $t > 0$  beliebig. Mit Satz 1.1 erhalten wir

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

Da die Zufallsvariablen  $X_i$  Pareto-verteilt sind, folgt

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

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

**Definition 1.11.** Die Verteilungsfunktion  $\Phi_\alpha$  in (1.3) heißt **Fréchet-Verteilung** mit Parameter  $\alpha > 0$ .

**Bemerkung 1.12.** Man kann Satz 1.10 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.

**Theorem 1.13.** Die Verteilungsfunktion  $F$  von  $X_i$  habe die Form

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

wobei  $\alpha > 0$  ein Parameter ist. Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{P}[n^{1/\alpha}(M_n - 1) \leq t] = \begin{cases} e^{-(-t)^\alpha}, & t \leq 0 \\ 1, & t \geq 0 \end{cases} =: \Psi_\alpha(t). \quad (1.4)$$

Mit anderen Worten konvergiert die Zufallsvariable  $n^{1/\alpha}(M_n - 1)$  in Verteilung gegen  $\Psi_\alpha$ .

*Beweis.* Sei  $t \leq 0$  beliebig. Mit Satz 1.1 erhalten wir

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

Aus der Form der Verteilungsfunktion  $F$  folgt, dass

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

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

**Definition 1.14.** Die Verteilungsfunktion  $\Psi_\alpha$  in (1.4) heißt **Weibull-Verteilung** mit Parameter  $\alpha > 0$ .

**Bemerkung 1.15.** Man kann Satz 1.13 wie folgt interpretieren: Für großes  $n$  nähert sich das Maximum  $M_n$  dem Wert 1 von unten an. Die Differenz  $M_n - 1$  nimmt sehr kleine negative Werte auf der Skala  $n^{-1/\alpha}$  an. Reskaliert man  $(M_n - 1)$  mit dem Faktor  $n^{1/\alpha}$ , so erhält man approximativ Weibull-verteilte Werte.

**Bemerkung 1.16.** 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.

**Bemerkung 1.17.** Wir haben drei verschiedene Typen von Extremwertverteilungen kennen gelernt: Gumbel, Fréchet und Weibull. Später werden wir zeigen, dass jede Extremwertverteilung sich auf eine dieser drei Verteilungen zurückführen lässt.

## 2 Max-Anziehungsbereiche

Seien  $X_1, X_2, \dots$  unabhängige identisch verteilte Zufallsvariablen mit Verteilungsfunktion  $F$ . Wir sagen, dass  $F$  im Max-Anziehungsbereich einer Verteilungsfunktion  $H$  liegt, falls es reellwertige Folgen  $a_n > 0$  und  $b_n$  gibt mit

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

Im Folgenden werden wir die Max-Anziehungsbereiche der Verteilungen  $\Phi_\alpha, \Psi_\alpha$  und  $G$  beschreiben.

### 2.1 Max-Anziehungsbereich der Fréchet-Verteilung $\Phi_\alpha$

Wir haben bereits gesehen, dass die Verteilung mit der Tailfunktion  $\bar{F}(t) = t^{-\alpha}$ ,  $t \geq 1$ , im Max-Anziehungsbereich von  $\Phi_\alpha$  liegt. Wir werden zeigen, dass der Anziehungsbereich von  $\Phi_\alpha$  aus allen Verteilungen besteht, deren Tailfunktionen sich im gewissen Sinne wie  $t^{-\alpha}$  verhalten. Der entsprechende Begriff ist die reguläre Variation.

**Definition 2.1.** Eine messbare Funktion  $f : (A, \infty) \rightarrow (0, \infty)$  (wobei  $A \in \mathbb{R}$  beliebig) heißt regulär variierend in  $+\infty$  mit Index  $\alpha \in \mathbb{R}$ , falls

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

Bezeichnung:  $f \in RV_\alpha$ .

**Beispiel 2.2.** Die Funktion  $f(t) = ct^\alpha$ , wobei  $c > 0$ , ist regulär variierend mit Index  $\alpha$ , denn

$$\frac{f(\lambda t)}{f(t)} = \frac{c(\lambda t)^\alpha}{ct^\alpha} = \lambda^\alpha \text{ für alle } \lambda > 0.$$

**Beispiel 2.3.** Die Funktion  $f(t) = ct^\alpha(\log t)^\beta$ , mit  $c > 0$ ,  $\beta \in \mathbb{R}$  ist regulär variierend mit Index  $\alpha$ , denn

$$\frac{f(\lambda t)}{f(t)} = \frac{c(\lambda t)^\alpha (\log(\lambda t))^\beta}{ct^\alpha (\log t)^\beta} = \lambda^\alpha \left( \frac{\log t + \log \lambda}{\log t} \right)^\beta \rightarrow \lambda^\alpha$$

für  $t \rightarrow +\infty$  und alle  $\lambda > 0$ .

**Definition 2.4.** Eine Funktion  $f$  ist langsam variierend in  $+\infty$  wenn  $f \in RV_0$ , d.h.

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

**Beispiel 2.5.** Die Funktion  $f(t) = (\log t)^\beta$  mit  $\beta \in \mathbb{R}$  ist langsam variierend.

**Theorem 2.6.** Sei  $f \in RV_\alpha$ . Dann gibt es eine langsam variierende Funktion  $L$ , sodass  $f(t) = t^\alpha L(t)$ .

*Beweis.* Setze  $L(t) = \frac{f(t)}{t^\alpha}$ . Dann muss man nur noch zeigen, dass  $L(t)$  langsam variierend ist:

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

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

**Theorem 2.7.** Eine Verteilungsfunktion  $F$  mit rechtem Endpunkt  $x^*$  liegt im Max-Anziehungsbereich der Fréchet-Verteilung  $\Phi_\alpha$  mit  $\alpha > 0$  genau dann, wenn folgende zwei Bedingungen erfüllt sind:

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

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

Sind die beiden Bedingungen erfüllt, so gilt

$$\frac{\max\{X_1, \dots, X_n\}}{a_n} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha \text{ mit } a_n = F^{\leftarrow} \left( 1 - \frac{1}{n} \right),$$

wobei  $F^{\leftarrow}$  die Quantilfunktion von  $F$  bezeichnet:

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

Im Folgenden beweisen wir die Rückrichtung von Satz 2.7.

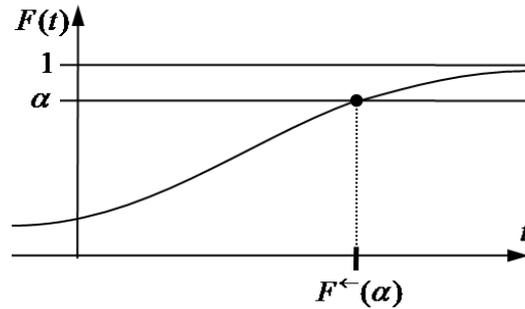


Abbildung 2.1: Veranschaulichung  $F^{\leftarrow}(\alpha)$ .

**Lemma 2.8.** Sei  $x^* = \infty$  und  $\bar{F} \in RV_{-\alpha}$ , dann gilt  $\lim_{n \rightarrow \infty} n\bar{F}(a_n) = 1$ .

*Beweis.* Ist  $F$  streng monoton steigend und stetig, so gilt  $F(a_n) = 1 - \frac{1}{n}$ , denn  $F^{\leftarrow}$  ist dann die inverse Funktion von  $F$ . In diesem Fall ist die Aussage des Lemmas gültig, denn es ist sogar  $n\bar{F}(a_n) = 1$ . Leider gilt das im Fall eines beliebigen  $F$  nicht, wie die folgenden Bilder zeigen:

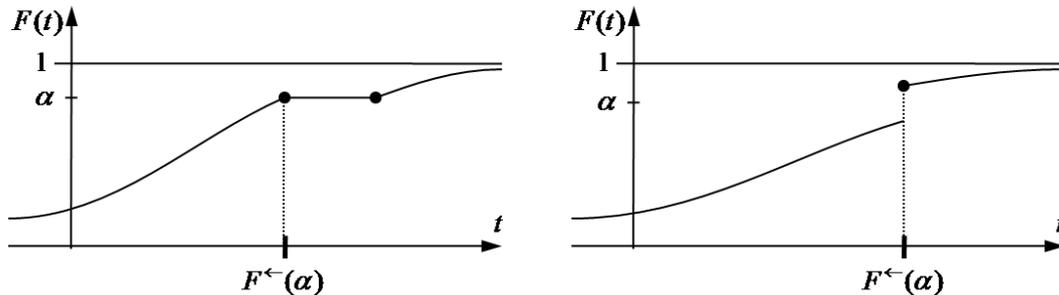


Abbildung 2.2: Problemfälle

Im allgemeinen Fall ist der Beweis etwas komplizierter. Wir zeigen zunächst, dass

$$F(a_n) = F\left(F^{\leftarrow}\left(1 - \frac{1}{n}\right)\right) \geq 1 - \frac{1}{n}. \quad (2.1)$$

Sei  $\varepsilon > 0$ , dann ist  $F(a_n + \varepsilon) \geq 1 - \frac{1}{n}$ . Lassen wir  $\varepsilon$  gegen Null gehen, so gilt  $F(a_n + \varepsilon) \rightarrow F(a_n)$ , weil  $F$  als Verteilungsfunktion rechtsstetig ist. Daraus ergibt sich (2.1). Damit folgt, dass  $\bar{F}(a_n) \leq \frac{1}{n}$ , woraus sich direkt ergibt, dass

$$\limsup_{n \rightarrow \infty} n\bar{F}(a_n) \leq 1. \quad (2.2)$$

Es bleibt also noch zu zeigen, dass

$$\liminf_{n \rightarrow \infty} n\bar{F}(a_n) \geq 1. \quad (2.3)$$

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

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

Damit folgt unmittelbar:

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

da  $\bar{F}$  regulär variierend mit Index  $-\alpha$  ist. Es ergibt sich also, dass  $\liminf_{n \rightarrow \infty} n\bar{F}(a_n) \geq x^\alpha$  für alle  $x \in (0, 1)$ . Wenn man nun  $x$  gegen 1 gehen lässt ergibt sich (2.3). Damit ist insgesamt  $\lim_{n \rightarrow \infty} n\bar{F}(a_n) = 1$  und das Lemma ist bewiesen.  $\square$

*Beweis von Satz 2.7.* Wir beweisen nur die Rückrichtung. Wir nehmen also an, dass  $x^* = \infty$  und  $\bar{F} \in \text{RV}_{-\alpha}$ . Sei  $t > 0$ . Nach Lemma 2.8 gilt  $n\bar{F}(a_n) \rightarrow 1$ . Da  $\bar{F}$  regulär variierend mit Index  $-\alpha$  ist, ergibt sich

$$n\bar{F}(ta_n) = n\bar{F}(a_n) \cdot \frac{\bar{F}(ta_n)}{\bar{F}(a_n)} \rightarrow 1 \cdot t^{-\alpha} = t^{-\alpha}, \quad n \rightarrow \infty.$$

Dadurch folgt:

$$\mathbb{P} \left[ \frac{M_n}{a_n} \leq t \right] = (1 - \bar{F}(ta_n))^n \rightarrow e^{-t^{-\alpha}} = \Phi_\alpha(t), \quad n \rightarrow \infty.$$

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

## 2.2 Max-Anziehungsbereich der Weibull-Verteilung $\Psi_\alpha$

Der Max-Anziehungsbereich der Weibull-Verteilung  $\Psi_\alpha$  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  $\Psi_\alpha$  liegt, muss  $F$  in  $x^*$  im gewissen Sinne regulär variierend sein. Wir geben nun eine präzise Definition.

**Definition 2.9.** Eine messbare Funktion  $f : (0, A) \rightarrow (0, \infty)$  (wobei  $A \in \mathbb{R}$  beliebig) heißt regulär variierend in 0 mit Index  $\alpha \in \mathbb{R}$ , falls

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

Bezeichnung:  $f \in RV_\alpha(0)$ .

**Beispiel 2.10.** Die Funktion  $f(t) = t^\alpha$  ist regulär variierend in 0 mit Index  $\alpha$ .

**Theorem 2.11.** Eine Verteilungsfunktion  $F$  mit rechtem Endpunkt  $x^*$  liegt im Max-Anziehungsbereich der Weibull-Verteilung  $\Psi_\alpha$ , dann und nur dann, wenn folgende zwei Bedingungen erfüllt sind:

1.  $x^* < \infty$
2. Die Funktion  $1 - F(x^* - t)$  ist regulär variierend in 0 mit Index  $\alpha$ , d.h.

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

Sind die Bedingungen erfüllt, so gilt

$$\frac{\max\{X_1, \dots, X_n\} - x^*}{a_n} \xrightarrow[n \rightarrow \infty]{d} \Psi_\alpha \text{ mit } a_n = x^* - F^{\leftarrow}\left(1 - \frac{1}{n}\right).$$

**Beispiel 2.12.** Betrachte die Verteilungsfunktion  $F(t) = 1 - (1 - t)^\alpha$  mit  $t \in [0, 1]$  und  $\alpha > 0$ . Dann ist  $x^* = 1$  und  $1 - F(1 - t) = t^\alpha \in RV_\alpha(0)$ .

*Beweis von Satz 2.11.* Wir beweisen wieder nur die Rückrichtung. Sei  $x^* < \infty$  und  $1 - F(x^* - t) \in RV_\alpha(0)$ . Definiere  $F^*(t) = F(x^* - \frac{1}{t})$ ,  $t > 0$ . Es ist leicht einzusehen, dass  $F^*$  auch eine Verteilungsfunktion ist. Betrachte

$$\lim_{t \rightarrow +\infty} \frac{1 - F^*(\lambda t)}{1 - F^*(t)} = \lim_{t \rightarrow +\infty} \frac{1 - F(x^* - \frac{1}{\lambda t})}{1 - F(x^* - \frac{1}{t})} = \lim_{s \downarrow 0} \frac{1 - F(x^* - \frac{s}{\lambda})}{1 - F(x^* - s)} = \lambda^{-\alpha},$$

wobei wir die Substitution  $t = 1/s$  benutzt haben. Es folgt, dass  $F^* \in RV_{-\alpha}(\infty)$ . Mit Theorem 2.7 folgt nun, dass  $F^*$  im Max-Anziehungsbereich von  $\Phi_\alpha$  liegt und

$$\lim_{n \rightarrow \infty} F^{*n}(a_n^* t) = e^{-1/t^\alpha}, t > 0, \text{ wobei } a_n^* = F^{*\leftarrow}\left(1 - \frac{1}{n}\right).$$

Deshalb gilt für alle  $s = -1/t < 0$

$$\lim_{n \rightarrow \infty} F^n \left( x^* + \frac{s}{a_n^*} \right) = \lim_{n \rightarrow \infty} F^n \left( x^* - \frac{1}{a_n^* t} \right) = \lim_{n \rightarrow \infty} F^{*n}(a_n^* t) = e^{-t^{-\alpha}} = e^{-(-s)^\alpha},$$

was sich auch äquivalent wie folgt ausdrücken lässt:

$$\lim_{n \rightarrow \infty} \mathbb{P}[a_n^*(M_n - x^*) \leq s] = \Psi_\alpha(s), \text{ für alle } s < 0.$$

Es bleibt also nur noch zu zeigen, dass  $a_n = \frac{1}{a_n^*}$ :

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

Substituiert man  $x^* - \frac{1}{t}$  durch  $u$ , so ergibt sich

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

Es ergibt sich also  $a_n = 1/a_n^*$ . □

## 2.3 Max-Anziehungsbereich der Gumbel-Verteilung $G$

**Theorem 2.13.** *Eine Verteilungsfunktion  $F$  mit rechtem Endpunkt  $x^*$  liegt im Max-Anziehungsbereich der Gumbel-Verteilung  $G(t) = e^{-e^{-t}}$  genau dann, wenn es eine positive und messbare Funktion  $g(t)$  gibt mit*

$$\lim_{t \uparrow x^*} \frac{\bar{F}(t + xg(t))}{\bar{F}(t)} = e^{-x} \text{ für alle } x \in \mathbb{R}. \quad (2.4)$$

Ist die Bedingung erfüllt, so gilt

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

mit  $b_n = F^{\leftarrow} \left( 1 - \frac{1}{n} \right)$  und  $a_n = g(a_n)$ .

**Bemerkung 2.14.**  $x^*$  kann endlich oder unendlich sein.

*Beweis.* Es wird hier nur ein Beweis für die Rückrichtung gegeben. Man betrachte dazu

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

denn es gilt  $n\bar{F}(b_n) \rightarrow 1$  (Beweis ähnlich wie in Lemma 2.8). Es folgt

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

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

**Beispiel 2.15.** Wir haben bereits gesehen, dass die Exponentialverteilung mit  $\bar{F}(t) = e^{-t}$ ,  $t > 0$ , im Max-Anziehungsbereich von  $G$  liegt. Man kann nachrechnen, dass Bedingung (2.4) mit  $g(t) = 1$  erfüllt ist.

Im Folgenden werden wir zeigen, dass die Normalverteilung zum Max-Anziehungsbereich der Gumbel-Verteilung gehört. Dazu benötigen wir ein Lemma.

**Definition 2.16.** Zwei Funktionen  $f$  und  $g$  heißen asymptotisch äquivalent (in  $+\infty$ ), falls

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

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

**Lemma 2.17.** Für die Tailfunktion  $\bar{F}$  der Standardnormalverteilung gilt

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

*Beweis.* Es gilt

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

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

**Theorem 2.18.** Die Standardnormalverteilung liegt im Max-Anziehungsbereich der Gumbel-Verteilung  $G$ .

*Beweis.* Wir werden zeigen, dass Bedingung (2.4) des Satzes 2.13 mit  $g(t) = 1/t$  gilt. Mit Lemma 2.17 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  $x \in \mathbb{R}$ . □

Wir werden nun die Konstanten  $a_n$  und  $b_n$  so bestimmen, dass für das Maximum von  $n$  unabhängigen standardnormalverteilten Zufallsvariablen gilt:  $\frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} G$ . Laut Satz 2.13 sollten wir  $b_n$  so wählen, dass  $n\bar{F}(b_n) = 1$ . Da aber die Lösung dieser Gleichung keine explizite Darstellung besitzt, werden wir ein bisschen anders vorgehen. Wir werden  $b_n$  so wählen, dass  $\lim_{n \rightarrow \infty} n\bar{F}(b_n) = 1$ , was genau dann der Fall ist wenn

$$\frac{1}{\sqrt{2\pi}b_n} e^{-b_n^2/2} \sim \frac{1}{n}, \quad n \rightarrow \infty. \quad (2.5)$$

Auf der linken Seite scheint  $e^{-b_n^2/2}$  derjenige Term zu sein, der am schnellsten gegen 0 geht. Wir können also als erste Annäherung  $b_n$  so wählen, dass  $e^{-b_n^2/2} = 1/n$ , d.h.  $b_n = \sqrt{2 \log n}$ . Dann ist aber  $\frac{1}{\sqrt{2\pi}b_n} e^{-b_n^2/2} = \frac{1}{\sqrt{2\pi}\sqrt{2 \log n}} \frac{1}{n}$  – also sind wir noch nicht am Ziel. Wir wählen also  $b_n = \sqrt{2 \log n} + \delta_n$ , wobei  $\delta_n$  noch genauer spezifiziert werden muss. Dann folgt:

$$\frac{1}{\sqrt{2\pi}b_n} e^{-b_n^2/2} = \frac{1}{\sqrt{2\pi}(\sqrt{2 \log n} + \delta_n)} e^{-\log n - \sqrt{2 \log n} \delta_n - \frac{\delta_n^2}{2}}.$$

Wähle  $\delta_n$  so dass  $\sqrt{2\pi}\sqrt{2 \log n} = e^{-\sqrt{2 \log n} \delta_n}$ . Dann folgt durch umformen, dass  $\delta_n = -\frac{1}{2\sqrt{2 \log n}}(\log 4\pi + \log \log n)$ . Es ist schließlich leicht nachzuweisen, dass mit

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

die gewünschte asymptotische Äquivalenz (2.5) gilt. Als  $a_n$  wählt man schließlich

$$a_n = g(b_n) = \frac{1}{b_n} \sim \frac{1}{\sqrt{2 \log n}}.$$

**Theorem 2.19.** *Seien  $X_1, X_2, \dots$  unabhängig und standardnormalverteilt. Dann gilt für alle  $x \in \mathbb{R}$ :*

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

*Beweis.* Übungsaufgabe. □

**Bemerkung 2.20.** Der Satz lässt sich wie folgt interpretieren: das Maximum  $M_n$  nimmt Werte an, die sehr nahe bei  $b_n = \sqrt{2 \log n} - \frac{\log \log n + \log(4\pi)}{2\sqrt{2 \log n}}$  sind. Die Differenz zwischen  $M_n$  und  $b_n$  ist von der Größenordnung  $a_n = 1/\sqrt{2 \log n}$ . Reskaliert man  $M_n - b_n$  mit dem Faktor  $\sqrt{2 \log n}$ , so erhält man eine approximativ Gumbel-verteilte Zufallsvariable.

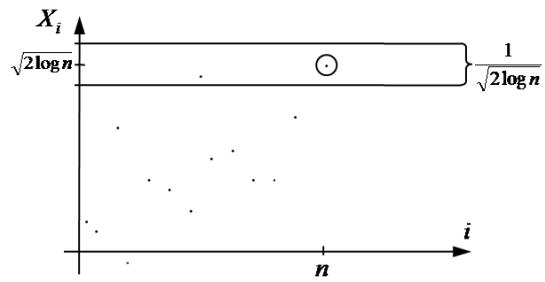


Abbildung 2.3: Maximum liegt mit hoher Wahrscheinlichkeit im dargestellten Bereich

### 3 Satz von Fisher–Tippett

**Theorem 3.1** (Satz von Fisher–Tippett; extremal types theorem). *Eine Verteilung  $G$  ist eine Extremwertverteilung genau dann, wenn es  $c > 0$ ,  $d \in \mathbb{R}$  und  $\gamma \in \mathbb{R}$  gibt mit*

$$G(t) = G_\gamma(ct + d).$$

*Dabei ist  $G_\gamma$  eine Verteilungsfunktion, die durch  $G_\gamma(t) = \exp\left\{-\left(1 + \gamma t\right)^{-\frac{1}{\gamma}}\right\}$  für  $1 + \gamma t > 0$  definiert ist.*

**Bemerkung 3.2.** Wir können die folgenden drei Fälle betrachten.

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

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

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

Im Folgenden werden wir den Satz von Fisher–Tippett beweisen.

**Lemma 3.3** (Khintchine). Seien  $Z_1, Z_2, \dots$  Zufallsvariablen mit Verteilungsfunktionen  $F_1, F_2, \dots$ . Sei  $G$  eine Verteilungsfunktion und  $c_n > 0, d_n \in \mathbb{R}$  Konstanten mit

$$\frac{Z_n - d_n}{c_n} \xrightarrow[n \rightarrow \infty]{d} G.$$

Sei  $\tilde{G}$  eine weitere Verteilungsfunktion und  $\tilde{c}_n > 0, \tilde{d}_n \in \mathbb{R}$  Konstanten mit

$$\frac{Z_n - \tilde{d}_n}{\tilde{c}_n} \xrightarrow[n \rightarrow \infty]{d} \tilde{G}.$$

Seien zudem  $G$  und  $\tilde{G}$  nicht ausgeartet. Dann gibt es Konstanten  $c > 0$  und  $d \in \mathbb{R}$  mit

$$c = \lim_{n \rightarrow \infty} \frac{c_n}{\tilde{c}_n} \text{ bzw. } d = \lim_{n \rightarrow \infty} \frac{\tilde{d}_n - d_n}{c_n}$$

und es gilt:  $\tilde{G}(t) = G(ct + d)$ .

*Beweis.* Ohne Beweis. □

**Definition 3.4.** Eine nicht ausgeartete Verteilungsfunktion  $G$  heißt *max-stabil*, falls es für alle  $n \in \mathbb{N}$  Konstanten  $c_n > 0$  und  $d_n \in \mathbb{R}$  gibt mit

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

**Beispiel 3.5.** Die Gumbel-Verteilung  $G(t) = e^{-e^{-t}}$  ist max-stabil, denn

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

Analog lässt sich zeigen, dass Fréchet- und Weibull-Verteilung  $\Phi_\alpha$  und  $\psi_\alpha$  max-stabil sind.

**Proposition 3.6.** Eine nicht ausgeartete Verteilungsfunktion  $G$  ist max-stabil dann und nur dann, wenn es eine Folge  $F_1, F_2, \dots$  von Verteilungsfunktionen und Konstanten  $c_n > 0, d_n \in \mathbb{R}$  gibt, sodass für alle  $k \in \mathbb{N}$  gilt:

$$F_n(c_n k t + d_n k) \xrightarrow[n \rightarrow \infty]{d} G^{1/k}(t). \quad (3.1)$$

*Beweis.* Zuerst soll (3.1) gelten, d.h.  $F_n(c_n k t + d_n k) \xrightarrow[n \rightarrow \infty]{d} G^{1/k}(t)$ . Mit  $k = 1$  folgt  $F_n(c_n t + d_n) \xrightarrow[n \rightarrow \infty]{d} G(t)$ . Mit Lemma 3.3 folgt, dass es  $\alpha_k > 0$  und  $\beta_k \in \mathbb{R}$  gibt mit  $G^{1/k} =$

$G(\alpha_k t + \beta_k)$ . Daraus folgt schließlich, dass  $G(t) = G^k(\alpha_k t + \beta_k)$  und somit  $G$  max-stabil ist.

Sei nun  $G$  als max-stabil vorgegeben. Wir zeigen, dass dann (3.1) gelten muss. Da  $G$  max-stabil ist, gibt es  $c_k > 0, d_k \in \mathbb{R}$  mit  $G^k(c_k t + d_k) = G(t)$ . Setze nun  $F_n = G^n$ . Dann gilt:

$$F_n(c_{nk}t + d_{nk}) = G^n(c_{nk}t + d_{nk}) = (G^{nk}(c_{nk}t + d_{nk}))^{1/k} = G^{1/k}(t).$$

Somit gilt (3.1). □

**Proposition 3.7.** *Eine Verteilungsfunktion  $G$  ist max-stabil dann und nur dann, wenn  $G$  eine Extremwertverteilung ist.*

*Beweis.* Sei  $G$  max-stabil. Dann gibt es  $c_n > 0, d_n \in \mathbb{R}$ , so dass gilt  $G^n(c_n t + d_n) = G(t)$ . Es gilt also:

$$G^n(c_n t + d_n) \xrightarrow[n \rightarrow \infty]{d} G(t),$$

weshalb  $G$  eine Extremwertverteilung ist.

Sei  $G$  nun eine Extremwertverteilung. Dann gibt es eine Verteilungsfunktion  $F$  und  $c_n > 0, d_n \in \mathbb{R}$ , so dass gilt:

$$F^n(c_n t + d_n) \xrightarrow[n \rightarrow \infty]{d} G(t)$$

und damit für alle  $k \in \mathbb{N}$

$$F^{nk}(c_{nk}t + d_{nk}) \xrightarrow[n \rightarrow \infty]{d} G(t) \text{ bzw. } F^n(c_{nk}t + d_{nk}) \xrightarrow[n \rightarrow \infty]{d} G^{1/k}(t).$$

Mit Proposition 3.6 folgt schließlich, dass  $G$  max-stabil ist. □

**Proposition 3.8.** *Ist  $G$  max-stabil, so gibt es Funktionen  $c(s) > 0, d(s) \in \mathbb{R}$ , so dass für alle positiven reellen Zahlen  $s$  gilt:*

$$G^s(c(s)t + d(s)) = G(t). \tag{3.2}$$

**Bemerkung 3.9.** Die Gleichung (3.2) ist deshalb stärker als die Gleichung aus Definition 3.4, weil  $s$  keine ganze Zahl sein muss. Für natürliche Zahlen  $s$  ist also nichts zu zeigen.

*Beweis.* Wir benötigen die folgende Notation:  $[t] = \max\{n \in \mathbb{Z} | n \leq t\}$ .

Sei  $G$  max-stabil. Dann gibt es  $c_n > 0, d_n \in \mathbb{R}$ , sodass für alle  $n \in \mathbb{N}$  folgendes gilt:  $G^n(c_n t + d_n) = G(t)$ . Für positive  $s$  folgt daraus  $G^{[ns]}(c_{[ns]}t + d_{[ns]}) = G(t)$ , womit sich direkt ergibt, dass

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

Gleichzeitig gilt aber auch  $G^n(c_n t + d_n) \xrightarrow[n \rightarrow \infty]{d} G(t)$ . Mit Lemma 3.3 folgt also, dass es Konstanten  $c(s) > 0, d(s) \in \mathbb{R}$  gibt mit

$$c(s) = \lim_{n \rightarrow \infty} \frac{c_{[ns]}}{c_n} \quad \text{und} \quad d(s) = \lim_{n \rightarrow \infty} \frac{d_n - d_{[ns]}}{c_n}$$

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

**Theorem 3.10.** *Eine max-stabile Verteilung  $G$  ist vom gleichen Typ wie eine der folgenden Verteilungen: Gumbel  $\Lambda$ , Fréchet  $\Phi_\alpha$  mit  $\alpha > 0$  oder Weibull  $\Psi_\alpha$  mit  $\alpha > 0$ .*

**Bemerkung 3.11.** Die Aussage des Satzes zusammen mit Proposition 3.7 impliziert den Satz von Fisher–Tippett.

*Beweis.* Nachzuweisen, dass die Gumbel, Fréchet und Weibull Verteilungen max-stabil sind, sei dem Leser als Übungsaufgabe überlassen.

Im folgenden sei  $G$  also max-stabil. Wir beweisen, dass  $G$  dann vom gleichen Typ ist wie eine der drei genannten Verteilungen. Nach Proposition 3.8 gibt es im Falle der max-Stabilität von  $G$  Konstanten  $c(s) > 0, d(s) \in \mathbb{R}$  mit  $G^s(c(s)t + d(s)) = G(t)$  für alle positiven reellen Zahlen  $s$ . Diese Gleichung lässt sich durch Anwendung von  $-\log$  auf beiden Seiten folgendermaßen umformen:

$$-s \log G(c(s)t + d(s)) = -\log G(t).$$

Nochmaliges Anwenden von  $-\log$  auf beiden Seiten führt zu:

$$-\log(-\log G(c(s)t + d(s))) - \log s = -\log(-\log G(t)).$$

Mit  $\psi(t) = -\log(-\log G(t))$  folgt

$$\psi(c(s)t + d(s)) - \log s = \psi(t).$$

Da  $G$  eine Verteilungsfunktion ist, folgt, dass  $\psi$  monoton steigend ist. Außerdem gilt:

$$\lim_{t \rightarrow -\infty} \psi(t) = -\log(-\log 0) = -\infty \quad \text{und} \quad \lim_{t \rightarrow \infty} \psi(t) = -\log(-\log 1) = \infty$$

Definiere  $\mathcal{U}(y) = \psi^\leftarrow(y) = \inf\{t \in \mathbb{R} \mid \psi(t) \geq y\}$  als die verallgemeinerte Inverse von  $\psi$ . Dann ist

$$\mathcal{U}(y) = \inf\{t \in \mathbb{R} \mid \psi(t) \geq y\} = \inf\{t \in \mathbb{R} \mid \psi(c(s)t + d(s)) - \log s \geq y\}$$

mit  $p = c(s)t + d(s)$  und entsprechend  $t = \frac{p-d(s)}{c(s)}$  folgt:

$$\mathcal{U}(y) = \inf \left\{ \frac{p - d(s)}{c(s)} \in \mathbb{R} \mid \psi(p) \geq y + \log s \right\} = \frac{\mathcal{U}(y + \log s) - d(s)}{c(s)} \text{ mit } y \in \mathbb{R}, s > 0.$$

Setzt man  $y = 0$  so gilt  $\mathcal{U}(0) = \frac{\mathcal{U}(\log s) - d(s)}{c(s)}$  und damit:

$$\mathcal{U}(y) - \mathcal{U}(0) = \frac{\mathcal{U}(y + \log s) - \mathcal{U}(\log s)}{c(s)}.$$

Ersetzt man schließlich  $\log s$  durch  $z$  dann gilt:

$$\mathcal{U}(y) - \mathcal{U}(0) = \frac{\mathcal{U}(y + z) - \mathcal{U}(z)}{c(e^z)}.$$

Mit dieser Gleichung lässt sich nun die Behauptung beweisen. Dafür unterscheiden wir die folgenden zwei Fälle

1.  $c(e^z) \equiv 1$  und
2.  $c(e^z) \neq 1$  für mindestens ein  $z$ .

Im ersten Fall  $c(e^z) \equiv 1$  gilt:  $\mathcal{U}(y) - \mathcal{U}(0) = \mathcal{U}(y + z) - \mathcal{U}(z)$ . Definiert man  $\tilde{\mathcal{U}}(y) = \mathcal{U}(y) - \mathcal{U}(0)$  folgt  $\tilde{\mathcal{U}}(y + z) - \tilde{\mathcal{U}}(z) = \tilde{\mathcal{U}}(y)$  und damit ist  $\tilde{\mathcal{U}}(y + z) = \tilde{\mathcal{U}}(y) + \tilde{\mathcal{U}}(z)$ . Bei dieser Gleichung handelt es sich um eine Hamel-Gleichung, was zusammen mit dem Umstand, dass  $\tilde{\mathcal{U}}$  monoton steigend ist impliziert, dass  $\tilde{\mathcal{U}}(y) = \rho \cdot y$  für ein  $\rho > 0$  gilt. Es folgt  $\mathcal{U}(y) = \rho \cdot y + b$  mit  $b = \mathcal{U}(0)$ . Insgesamt ergibt sich also, dass in diesem Fall  $\psi(t) = \frac{t-b}{\rho}$  und daher

$$G(t) = e^{-e^{-\psi(t)}} = e^{-e^{-\frac{t-b}{\rho}}}$$

vom gleichen Typ wie die Gumbel-Verteilung ist.

Im zweiten Fall gibt es ein  $z$  für das gilt  $c(e^z) \neq 1$ . Sei nun wieder  $\tilde{\mathcal{U}}(y) = \mathcal{U}(y) - \mathcal{U}(0)$ . Dann ist

$$\tilde{\mathcal{U}}(y + z) - \tilde{\mathcal{U}}(z) = c(e^z) - \tilde{\mathcal{U}}(y).$$

Vertauscht man hier  $y$  und  $z$  erhält man

$$\tilde{\mathcal{U}}(y + z) - \tilde{\mathcal{U}}(y) = c(e^y) - \tilde{\mathcal{U}}(z).$$

Subtrahiert man die beiden obigen Gleichungen erhält man

$$\tilde{\mathcal{U}}(y) = \tilde{\mathcal{U}}(z) \frac{1 - c(e^y)}{1 - c(e^z)}, \tag{3.3}$$

wobei hier wegen  $c(e^z) \neq 1$  keine Division durch Null stattfinden kann. Setzt man nun

### 3 Satz von Fisher–Tippett

---

$\tilde{\mathcal{U}}(y) = f(z)(1 - c(e^y))$  mit  $f(z) = \frac{\tilde{\mathcal{U}}(z)}{1 - c(e^z)}$  in (3.3) ein, erhält man:

$$\tilde{\mathcal{U}}(y + z) = f(z)(1 - c(e^z)) + c(e^z)f(z)(1 - c(e^y)) = f(z) - f(z)c(e^z)c(e^y)$$

Hieraus ergibt sich schließlich:

$$c(e^{x+y}) = c(e^x) \cdot c(e^y)$$

Hierbei handelt es sich wieder um eine Hamel-Gleichung, die für  $\rho \in \mathbb{R}$  durch  $c(s) = s^\rho$  gelöst wird. Insgesamt ergibt sich damit

$$\mathcal{U}(y) = \mathcal{U}(0) + f(z)(1 - e^{\rho y}) = f \cdot (1 - e^{\rho y}) = \alpha + \beta e^{\rho y},$$

weil  $f(z)$  konstant ist. Daher ist  $\psi(z) = \frac{1}{\rho} \log\left(\frac{z - \alpha}{\beta}\right)$  für  $z > \alpha$ , falls  $\beta > 0$  bzw. für  $z < \alpha$ , falls  $\beta < 0$ . Dann ist

$$G(t) = e^{-e^{-\Phi(t)}} = \exp\left\{-\left(\frac{z - \alpha}{\beta}\right)^{-1/\rho}\right\}.$$

Man kann beobachten, dass im Fall  $z > \alpha$  (falls  $\beta > 0$ )  $G$  vom gleichen Typ wie die Fréchet-Verteilung und im Fall  $z < \alpha$  (falls  $\beta < 0$ )  $G$  vom gleichen Typ wie die Weibull-Verteilung ist.  $\square$

# 4 Statistik der Extremwertverteilungen

In diesem Kapitel beschäftigen wir uns mit statistischen Anwendungen der Extremwerttheorie. Wir werden zwei verschiedene Zugänge zur Modellierung von Extremwerten betrachten. Der erste Zugang basiert auf den 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).

## 4.1 Statistik der GEV-Verteilungen

Wir haben bisher gesehen, dass Extremwertverteilungen folgende Form haben:

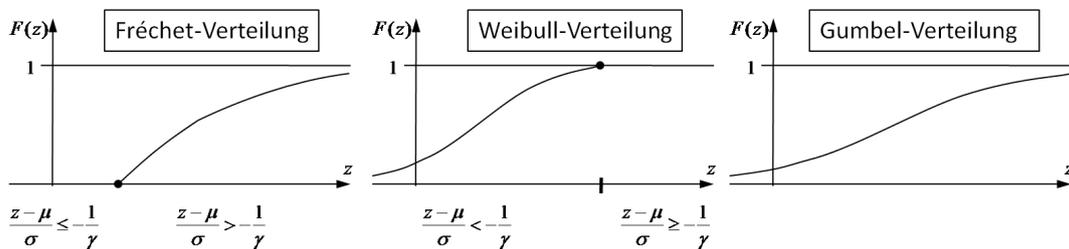
$$G_{\gamma,\mu,\sigma}(z) = \exp \left\{ - \left( 1 + \gamma \frac{z - \mu}{\sigma} \right)^{-\frac{1}{\gamma}} \right\} \text{ für } 1 + \gamma \frac{z - \mu}{\sigma} > 0.$$

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  $\gamma$  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 (GEV) Verteilungen genannt.

**Beispiel 4.1** (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  $X_i = \max_{j=1, \dots, 365} X_{ij}$  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 u.i.v. Zufallsvariablen mit einer GEV-Verteilung  $G_{(\gamma, \mu, \sigma)}$  mit Parametervektor  $\theta = (\gamma, \mu, \sigma) \in \mathbb{R}^2 \times \mathbb{R}_+$ .

**Bemerkung 4.2.** 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  $\mu$  und  $\sigma$  enthalten sind.

**Bemerkung 4.3.** 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.

Unser Problem besteht nun darin, den Parametervektor  $\theta$  zu schätzen. Wir werden die Maximum-Likelihood (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  $\gamma \neq 0$  gilt:

$$f_\theta(z) = f_{(\gamma, \mu, \sigma)}(z) = \begin{cases} \frac{1}{\sigma} (1 + \gamma \frac{z-\mu}{\sigma})^{-\frac{1}{\gamma}-1} \exp\{- (1 + \gamma \frac{z-\mu}{\sigma})^{-\frac{1}{\gamma}}\}, & 1 + \gamma \frac{z-\mu}{\sigma} > 0, \\ 0, & \text{sonst} \end{cases}$$

und für  $\gamma = 0$  gilt:

$$f_\theta(z) = f_{(0, \mu, \sigma)} = \frac{1}{\sigma} e^{-\frac{z-\mu}{\sigma}} \exp\left\{-e^{-\frac{z-\mu}{\sigma}}\right\}, \text{ für } z \in \mathbb{R}.$$

Mit Hilfe der Dichten kann man die Log-Likelihoodfunktion

$$l(\theta) := l(\theta|x_1, \dots, x_n) := \sum_{i=1}^n \log f_\theta(x_i)$$

bestimmen. Für  $\gamma \neq 0$  gilt:

$$l(\theta) = -n \log \sigma - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^n \log \left(1 + \gamma \frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^n \left(1 + \gamma \frac{x_i - \mu}{\sigma}\right)^{-\frac{1}{\gamma}}$$

falls  $1 + \gamma \frac{x_i - \mu}{\sigma} > 0$  für alle  $i = 1, \dots, n$  und  $l(\theta) = -\infty$  sonst. Für  $\gamma = 0$  gilt:

$$l(\theta) = -n \log \sigma - \sum_{i=1}^n \frac{x_i - \mu}{\sigma} - \sum_{i=1}^n e^{-\frac{x_i - \mu}{\sigma}}.$$

Mit der log-Likelihoodfunktion lässt sich der Maximum-Likelihood Schätzer

$$\hat{\theta} = (\hat{\gamma}, \hat{\mu}, \hat{\sigma}) = \operatorname{argmax} l(\gamma, \mu, \sigma)$$

herleiten. Hier kann  $\hat{\theta}$  nicht analytisch bestimmt werden, sondern muss numerisch ermittelt werden.

Nachdem der Parameter  $\theta$  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  $p$ -Quantil des jährlichen Maximums zu schätzen. Wir schätzen  $Z_p$  indem wir die Gleichung

$$G_{(\hat{\gamma}, \hat{\mu}, \hat{\sigma})}(\hat{Z}_p) = 1 - p$$

lösen (falls es mehrere Lösungen gibt, betrachten wir die kleinste):

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

Für  $\hat{\gamma} < 0$  (im Fall der Weibull-Verteilung) besitzt die Verteilung  $G_{(\hat{\gamma}, \hat{\mu}, \hat{\sigma})}$  einen endlichen rechten Endpunkt, der übrigens per Definition  $Z_0$  ist. In diesem Fall gehen wir davon aus, dass es einen absolut höchsten Wasserstand gibt, der niemals überschritten wird. Der Schätzer für  $Z_0$  ist dann gegeben durch:

$$\hat{Z}_0 = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\gamma}}.$$

Nachdem nun das Problem gelöst wurde, stellt sich die Frage, wie wir die Lösung verifizieren können. Wie können wir überprüfen, ob die Daten  $X_1, \dots, X_n$  durch die Verteilungsfunktion  $\hat{G} = G_{(\hat{\gamma}, \hat{\mu}, \hat{\sigma})}$  tatsächlich gut beschrieben werden? Zur Verifikation des Modells gibt es mehrere Methoden, die wir im Folgenden betrachten.

Ordnen wir die Stichprobe  $X_1, \dots, X_n$  monoton aufsteigend an, so erhalten wir die Ordnungsstatistiken

$$X_{(1)} \leq \dots \leq X_{(n)}.$$

**Definition 4.4** (Probability Plot). *Der PP-Plot ist die Menge*

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

Trifft die Annahme, dass die  $X_i$  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 4.1).

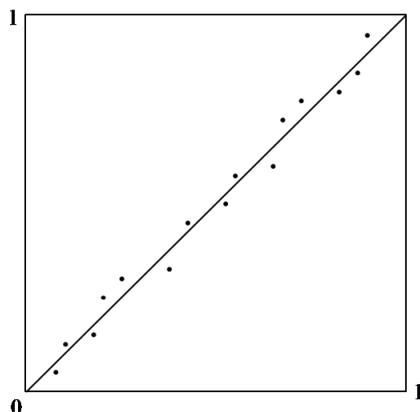


Abbildung 4.1: PP-Plot

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  $\hat{G}^{\leftarrow}(q)$  der Verteilung  $\hat{G}$  ist definiert als (die kleinste) Lösung  $z$  der Gleichung

$$\hat{G}(z) = q.$$

**Definition 4.5** (Quantil Plot). *Der QQ-Plot ist die Menge*

$$\left\{ \left( \hat{G}^{\leftarrow} \left( \frac{i}{n+1} \right), X_{(i)} \right), i = 1, \dots, n \right\} \in \mathbb{R}^2.$$

Beim QQ-Plot werden auf der  $x$ -Achse die  $\frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}$ -Quantile der Verteilung  $\hat{G}$  abgetragen und auf der  $y$ -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 Grafik 4.2 auf der Winkelhalbierenden liegen.

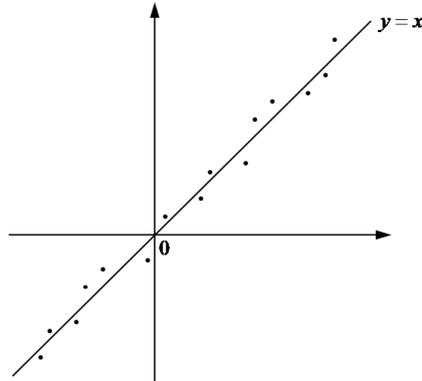


Abbildung 4.2: QQ-Plot

Es kann vorkommen, dass die Daten  $X_1, \dots, X_n$  einen Trend aufweisen (z.B. werden die jährlichen Maxima höher). Wir betrachten nun ein Modell, das der Nichtstationarität der Daten Rechnung trägt. Es seien  $X_1, \dots, X_n$  unabhängig aber nicht identisch verteilt 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  $\gamma$ , einem konstanten Skalenparameter  $\sigma$  und einem linearen Trend, der im Lageparameter  $\mu$  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 4.6.** Der Ansatz kann verallgemeinert werden, ohne dass sich das Modell grundsätzlich ändert. So ist es zum Beispiel problemlos möglich, einen exponentiellen Trend zu modellieren:

$$\gamma(i) = \gamma, \quad \sigma(i) = \sigma, \quad \mu(i) = e^{\beta_0 + \beta_1 i}.$$

## 4.2 Statistik der GP-Verteilungen

Die oben beschriebene Methode basiert auf der Betrachtung von Block-Maxima (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  $X$ . Wie ist der sogenannte Exzess  $X - u$  verteilt, gegeben dass  $X > u$ ? Dabei geht  $u \rightarrow \infty$ . Wir betrachten drei Beispiele.

**Beispiel 4.7.** 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  $\lambda$ . Dies ist die Gedächtnislosigkeit der Exponentialverteilung.

**Beispiel 4.8.** Sei  $X$  aus dem Max-Anziehungsbereich der Fréchet-Verteilung  $\Phi_\alpha$ ,  $\alpha > 0$ . 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 4.9.** Sei  $X$  standardnormalverteilt. Folgende Relation kann mit der Regel von L'Hospital bewiesen werden:

$$\mathbb{P}[X > u] \sim \frac{1}{\sqrt{2\pi}u} e^{-u^2/2} \text{ für } u \rightarrow \infty.$$

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. Die Verteilungsfunktion

$$P_{\gamma,\sigma}(t) = 1 - \left(1 + \frac{\gamma t}{\sigma}\right)^{-\frac{1}{\gamma}}$$

heißt die verallgemeinerte Pareto-Verteilung (GPD, Generalized Pareto Distribution) mit Index  $\gamma \in \mathbb{R}$  und Skalenparameter  $\sigma > 0$ . In dieser Formel ist  $t > 0$ , falls  $\gamma > 0$ , und  $t \in [0, -\frac{\sigma}{\gamma}]$ , falls  $\gamma < 0$ . 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.

**Theorem 4.10** (Pickands–Balkema–de Haan). *Sei  $X$  eine Zufallsvariable mit Verteilungsfunktion  $F$ , die im rechten Endpunkt  $x^*$  stetig ist. Dann liegt  $F$  im Max-Anziehungsbereich von  $G_\gamma$  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  $G_\gamma$  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 unabhängige und identisch verteilte Zufallsvariablen, 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

$$f_{\gamma,\sigma}(t) = \frac{1}{\sigma} \left(1 + \frac{\gamma t}{\sigma}\right)^{-\frac{1}{\gamma}-1}.$$

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  $(\hat{\gamma}, \hat{\sigma}) = \operatorname{argmax}_{\gamma, \sigma} l(\gamma, \sigma)$  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

$$Z_p = u + \frac{\sigma}{\gamma} \left[ \left( \frac{\mathbb{P}[X > u]}{p} \right)^{\gamma} - 1 \right].$$

Dabei haben wir Schätzer für  $\sigma$  und  $\gamma$  bereits hergeleitet. Die Wahrscheinlichkeit  $\mathbb{P}[X > u]$  kann durch  $\frac{1}{n} \sum_{i=1}^n 1_{X_i > u}$  geschätzt werden.

# 5 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  $X_{n:n} = \max\{X_1, \dots, X_n\}$ . Wir benutzen 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.

## 5.1 Extreme Ordnungsstatistiken

Wir haben die möglichen Grenzwertverteilungen von  $M_n^{(1)}$  für  $n \rightarrow \infty$  bereits beschrieben. Im Folgenden beschreiben wir die Grenzwertverteilungen von  $M_n^{(k)}$ , wobei  $k \in \mathbb{N}$  fest und  $n \rightarrow \infty$ .

**Theorem 5.1.** Sei  $u_n \in \mathbb{R}$  eine Folge mit  $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau$ , wobei  $0 < \tau < \infty$ . Dann gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \sum_{i=1}^n 1_{X_i > u_n} = k \right] = e^{-\tau} \frac{\tau^k}{k!}, \quad k = 0, 1, \dots$$

*Beweis.* Wir bezeichnen mit  $S_n := \sum_{i=1}^n 1_{X_i > u_n}$  die Anzahl der Beobachtungen, die oberhalb von  $u_n$  liegen. Es gilt  $S_n \sim \text{Bin}(n, \bar{F}(u_n))$ . Nach Voraussetzung gilt  $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau$ . Da dies gilt, darf der Poissongrenzwertsatz angewendet werden:

$$S_n \xrightarrow[n \rightarrow \infty]{d} \text{Poi}(\tau) \text{ für } n \rightarrow \infty.$$

Mit anderen Worten,  $\lim_{n \rightarrow \infty} \mathbb{P}[S_n = k] = e^{-\tau} \frac{\tau^k}{k!}$ , für alle  $k = 0, 1, \dots$  □

**Theorem 5.2.** *Es gebe Normierungsfolgen  $a_n$  und  $b_n$ , so dass*

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

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  $S_n = \sum_{i=1}^n 1_{X_i > u_n}$  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 Theorem 5.1 folgt:

$$\lim_{n \rightarrow \infty} \sum_{s=0}^{k-1} \mathbb{P}[S_n = s] = \sum_{s=0}^{k-1} e^{-\tau} \frac{\tau^s}{s!} \quad (5.1)$$

mit  $\tau = \lim_{n \rightarrow \infty} n \bar{F}(u_n)$ . Es bleibt also noch die Bestimmung von  $\tau$ . Nach Voraussetzung gilt:

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

Deshalb gilt für geeignete reelle Folgen  $a_n, b_n$

$$F^n(a_n + b_n x) \xrightarrow[n \rightarrow \infty]{} G(x).$$

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 folgt schließlich  $\lim_{n \rightarrow \infty} n F(u_n) = -\log G(x)$ . Deshalb gilt  $\tau = -\log G(x)$ . Einsetzen von  $\tau$  in (5.1) liefert die Behauptung.  $\square$

Im Folgenden beschreiben wir die Grenzwertverteilung des Vektors  $(M_n^{(1)}, \dots, M_n^{(r)})$ , mit  $r$  fest und  $n \rightarrow \infty$ . Betrachte die Schwellenwerte  $u_n^{(1)} \geq u_n^{(2)} \geq \dots \geq u_n^{(r)}$  mit

$$\lim_{n \rightarrow \infty} n \bar{F}(u_n^{(1)}) = \tau_1, \dots, \lim_{n \rightarrow \infty} n \bar{F}(u_n^{(r)}) = \tau_r, \text{ mit } 0 < \tau_1 \leq \dots \leq \tau_r < \infty.$$

Sei  $S_n^{(k)} := \sum_{i=1}^n 1_{X_i > u_n^{(k)}}$  die Anzahl der Überschreitungen des Schwellenwerts  $u_n^{(k)}$ .

**Theorem 5.3.** Für alle  $k_1, k_2, \dots, k_r \in \{0, 1, 2, \dots\}$  gilt

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

*Beweis.* Definiere  $p_{n,k} = \bar{F}(u_n^{(k)})$ . Setzt man

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

dann gilt

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

Damit kann der Beweis faktorweise durchgeführt werden. Für den ersten Faktor gilt

$$\binom{n}{k_1} p_{n,1}^{k_1} = \frac{n(n-1) \dots (n-k_1+1)}{k_1!} \cdot \bar{F}(u_n^{(1)})^{k_1} \sim \frac{n^{k_1} \bar{F}(u_n^{(1)})^{k_1}}{k_1!} \rightarrow \frac{\tau_1^{k_1}}{k_1!}.$$

Ähnlich gilt für den zweiten Faktor

$$\binom{n-k_1}{k_2} (p_{n,2} - p_{n,1})^{k_2} \sim \frac{n^{k_2} (\bar{F}(u_n^{(2)}) - \bar{F}(u_n^{(1)}))^{k_2}}{k_2!} \rightarrow \frac{(\tau_2 - \tau_1)^{k_2}}{k_2!},$$

was sich nun analog für sämtliche Faktoren, außer dem letzten, fortführen lässt. Für den letzten Faktor gilt schließlich:

$$(1 - p_{n,r})^{n-k_1-\dots-k_r} \rightarrow e^{-\tau_r}, \text{ für } n \rightarrow \infty,$$

denn es gilt:

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

□

Nun können wir die gemeinsame Verteilung von  $(M_n^{(1)}, \dots, M_n^{(r)})$  beschreiben. Es gilt

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

Der Grenzwert der rechten Seite kann mit dem obigen Satz berechnet werden, was zu einer langen Formel führt. Im nächsten Satz betrachten wir den Spezialfall  $r = 2$ , d.h. wir

beschreiben die gemeinsame Grenzwertverteilung der zwei größten Werte der Stichprobe.

**Theorem 5.4.** *Es gebe Normierungsfolgen  $a_n$  und  $b_n$ , so dass*

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

wobei  $G$  eine Extremwertverteilung ist. Dann gilt für alle  $x_1 > x_2$ :

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{M_n^{(1)} - a_n}{b_n} \leq x_1, \frac{M_n^{(2)} - a_n}{b_n} \leq x_2 \right] = G(x_2)(\log G(x_1) - \log G(x_2) + 1),$$

falls  $G(x_2) \neq 0$  und 0 sonst.

*Beweis.* Setze  $u_n^{(1)} := a_n + b_n x_1$  bzw.  $u_n^{(2)} := a_n + b_n x_2$ . 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  $\lim_{n \rightarrow \infty} n \bar{F}(u_n^{(1)}) = -\log G(x_1) =: \tau_1$  bzw.  $\lim_{n \rightarrow \infty} n \bar{F}(u_n^{(2)}) = -\log G(x_2) =: \tau_2$ . Es gilt:

$$\begin{aligned} \mathbb{P}[M_n^{(1)} \leq u_n^{(1)}, M_n^{(2)} \leq u_n^{(2)}] &= \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} \leq 1] \\ &= \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 0] + \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 1]. \end{aligned}$$

Mit Theorem 5.3 folgt dann

$$\lim_{n \rightarrow \infty} (\mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 0] + \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 1]) = e^{-\tau_2} + (\tau_2 - \tau_1)e^{-\tau_2}.$$

□

## 5.2 Allgemeine Eigenschaften der Ordnungsstatistiken

**Theorem 5.5.** *Für alle  $1 \leq k \leq n$  und  $t \in \mathbb{R}$  gilt*

$$\mathbb{P}[X_{k:n} \leq t] = \sum_{m=k}^n \binom{n}{m} F(t)^m \bar{F}(t)^{n-m}.$$

*Beweis.* Setze  $L := \sum_{i=1}^n 1_{X_i \leq t}$ , dann gilt  $L \sim \text{Bin}(n, F(t))$  und somit

$$\mathbb{P}[X_{k:n} \leq t] = \mathbb{P}[L \geq k] = \sum_{m=k}^n \mathbb{P}[L = m] = \sum_{m=k}^n \binom{n}{m} F(t)^m \bar{F}(t)^{n-m}.$$

□

Eine Zufallsvariable  $X$  heißt gleichverteilt auf  $[0, 1]$ , falls die Dichte von  $X$  gegeben ist durch  $f_X(t) = 1_{[0,1]}(t)$ .

**Beispiel 5.6.** Seien  $X_1, \dots, X_n$  unabhängig und gleichverteilt auf  $[0, 1]$ . Dann ist die Dichte von  $X_{k:n}$  gegeben durch

$$f_{X_{k:n}}(t) = \begin{cases} k \binom{n}{k} t^{k-1} (1-t)^{n-k}, & t \in [0, 1] \\ 0, & \text{sonst.} \end{cases}$$

Das heißt,  $X_{k:n}$  hat eine Beta-Verteilung  $\text{Beta}(k, n - k + 1)$ .

**Aufgabe 5.7.** Zeigen Sie, dass  $\mathbb{E}X_{k:n} = \frac{k}{n+1}$ .

**Theorem 5.8.** Seien  $X_1, \dots, X_n$  unabhängig und identisch verteilt mit Dichte  $f$ . Dann gilt für die gemeinsame Dichte von  $X_{1:n}, \dots, X_{n:n}$ :

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

*Beweisidee.* Es gibt  $n!$  Anordnungen von  $X_1, \dots, X_n$ . □

Auch wenn die Zufallsvariablen  $X_1, \dots, X_n$  unabhängig sind, sind die Ordnungsstatistiken  $X_{1:n}, \dots, X_{n:n}$  im Allgemeinen nicht unabhängig. Es gilt jedoch eine schwächere Eigenschaft, die sogenannte Markov-Eigenschaft.

**Theorem 5.9** (Bedingte Unabhängigkeit von Ordnungsstatistiken). Seien  $X_1, \dots, X_n$  unabhängige und identisch verteilte Zufallsvariablen mit Dichte  $f$ . Die bedingte Dichte  $f(u|x_1, \dots, x_n)$  von  $X_{k+1:n}$  gegeben, dass  $X_{1:n} = x_1, \dots, X_{k:n} = x_k$ , stimmt mit der bedingten Dichte  $f(u|x_k)$  von  $X_{k+1:n}$  gegeben, dass  $X_{k:n} = x_k$ , überein.

**Bemerkung 5.10.** Angenommen, die ersten  $k$  Ordnungsstatistiken der Stichprobe  $X_1, \dots, X_n$  sind bekannt:  $X_{1:n} = x_1, \dots, X_{k:n} = x_k$ . Wo liegt nun die nächste Ordnungsstatistik  $X_{k+1:n}$ ? Der obige Satz behauptet, dass für die Beantwortung dieser Frage nur der Wert  $X_{k:n} = x_k$  relevant ist. Die Werte der vorherigen Ordnungsstatistiken  $x_1, \dots, x_{k-1}$  tauchen in der bedingten Verteilung von  $X_{k+1:n}$  nicht auf.

*Beweis.* Sei  $f_{1,\dots,k:n}$  die Dichte von  $X_{1:n}, \dots, X_{k:n}$ . Dann gilt

$$f(u|x_1, \dots, x_k) = \frac{f_{1,\dots,k+1:n}(u, x_1, \dots, x_k)}{f_{1,\dots,k:n}(x_1, \dots, x_k)}. \quad (5.2)$$

Sei  $f_{k,k+1:n}$  die Dichte von  $(X_{k:n}, X_{k+1:n})$  und  $f_{k:n}$  die Dichte von  $X_{k:n}$ . Dann gilt

$$f(u|x_k) = \frac{f_{k,k+1:n}(x_k, u)}{f_{k:n}(x_k)}, \quad (5.3)$$

Es gilt:

$$f_{1,\dots,k:n}(x_1, \dots, x_k) = f(x_1) \cdot \dots \cdot f(x_k) \cdot (1 - F(x_k))^{n-k} \cdot n \cdot (n-1) \cdot \dots \cdot (n-k+1),$$

falls  $x_1 < \dots < x_k$  und 0 sonst. Analog,

$$f_{1,\dots,k+1:n}(x_1, \dots, x_{k+1}) = f(x_1) \cdot \dots \cdot f(x_{k+1}) \cdot (1 - F(x_{k+1}))^{n-k-1} \cdot n \cdot (n-1) \cdot \dots \cdot (n-k),$$

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 (5.2) bzw. (5.3) ergibt

$$(5.2) = (5.3) = \frac{(n-k)(1 - F(u))^{n-k-1} f(u)}{(1 - F(x_k))^{n-k}}.$$

□

**Aufgabe 5.11.** Seien  $X_1, \dots, X_n$  unabhängig und exponentialverteilt mit Parameter  $\lambda > 0$ . Welche Verteilung besitzt  $X_{k+1:n} - x_k$  gegeben, dass  $X_{1:n} = x_1, \dots, X_{k:n} = x_k$ ?

### 5.3 Darstellungen der Ordnungsstatistiken

Eine Zufallsvariable  $Z$  heißt exponentialverteilt mit Parameter  $\lambda > 0$ , falls  $\mathbb{P}[Z > t] = e^{-\lambda t}$  für  $t \geq 0$ . Die Ordnungsstatistiken der exponentialverteilten Zufallsvariablen besitzen eine besonders schöne Darstellung.

**Theorem 5.12.** Seien  $Z_1, \dots, Z_n$  unabhängig und exponentialverteilt mit Parameter  $\lambda = 1$ . Dann gilt

$$(Z_{1:n}, \dots, Z_{n:n}) \stackrel{d}{=} \left( \frac{\nu_1}{n}, \frac{\nu_1}{n} + \frac{\nu_2}{n-1}, \dots, \frac{\nu_1}{n} + \frac{\nu_2}{n-1} + \dots + \frac{\nu_n}{1} \right),$$

wobei  $\nu_1, \dots, \nu_n$  unabhängig und exponentialverteilt mit Parameter  $\lambda = 1$  sind.

**Bemerkung 5.13.** Die Abstände  $Z_{1:n}, Z_{2:n} - Z_{1:n}, \dots, Z_{n:n} - Z_{n-1:n}$  sind somit unabhängig und (nicht identisch) exponentialverteilt:

$$Z_{k:n} - Z_{k-1:n} \stackrel{d}{=} \frac{\nu_k}{n-k+1} \sim \text{Exp}(n-k+1).$$

**Aufgabe 5.14.** Berechnen Sie mit dem obigen Satz  $\mathbb{E}Z_{k:n}$ ,  $\text{Var} Z_{k:n}$  und  $\text{Cov}(Z_{k:n}, Z_{l:n})$ .

*Beweis.* Die Dichte von  $(Z_{1:n}, \dots, Z_{n:n})$  ist gegeben durch

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

Wegen der Unabhängigkeit ist die Dichte von  $(\nu_1, \dots, \nu_n)$  gegeben durch

$$g(y_1, \dots, y_n) = \begin{cases} e^{-y_1} \cdot \dots \cdot e^{-y_n}, & y_i > 0 \text{ für alle } i \in \{1, \dots, n\}, \\ 0, & \text{sonst.} \end{cases}$$

Wir betrachten die lineare Transformation

$$T : (y_1, \dots, y_n) \mapsto \left( \frac{y_1}{n}, \frac{y_1}{n} + \frac{y_2}{n-1}, \dots, \frac{y_1}{n} + \frac{y_2}{n-1} + \dots + \frac{y_n}{1} \right) =: (z_1, z_2, \dots, z_n).$$

Die Jacobi-Determinante dieser Transformation ist

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

Mit der Transformationsformel können wir die Dichte von  $T(\nu_1, \dots, \nu_n)$  berechnen:

$$h(z_1, \dots, z_n) = n!g(y_1, \dots, y_n) = \begin{cases} n!e^{-(y_1+\dots+y_n)} = n!e^{-(z_1+\dots+z_n)}, & 0 < z_1 < \dots < z_n, \\ 0, & \text{sonst} \end{cases}$$

Bei obiger Transformation wurde die Dichte von  $(\nu_1, \dots, \nu_n)$  durch die Jacobi-Determinante geteilt, um zur Dichte von  $T(\nu_1, \dots, \nu_n)$  zu gelangen. Man sieht, dass die Dichte  $f$  von  $(Z_{1:n}, \dots, Z_{n:n})$  und die Dichte  $h$  von  $T(\nu_1, \dots, \nu_n)$  übereinstimmen, was die Behauptung impliziert.  $\square$

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

**Theorem 5.15.** *Seien  $U_1, \dots, U_n$  unabhängig und gleichverteilt auf  $[0, 1]$ . Dann gilt:*

$$(U_{1:n}, \dots, U_{n:n}) \stackrel{d}{=} \left( \frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}} \right),$$

wobei  $S_k = \nu_1 + \dots + \nu_k$  und  $\nu_1, \dots, \nu_{n+1}$  unabhängige, mit Parameter 1 exponentialverteilte Zufallsvariablen sind.

**Bemerkung 5.16.** Die Folge  $S_1 < S_2 < \dots$  ist der Poisson-Prozess auf  $(0, \infty)$  mit Intensität 1. Intuitiv kann man die Aussage des Satzes so verstehen, dass die ersten  $n$  Punkte  $S_1, \dots, S_n$  des Poisson-Prozesses auf dem Intervall  $[0, S_{n+1}]$  gleichverteilt sind.

*Beweis.* Die Dichte von  $(U_{1:n}, \dots, U_{n:n})$  ist gegeben durch:

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

Die Dichte von  $(\nu_1, \dots, \nu_{n+1})$  ist wegen der Unabhängigkeit gegeben durch:

$$g(y_1, \dots, y_{n+1}) = \begin{cases} e^{-y_1} \cdot \dots \cdot e^{-y_{n+1}}, & y_i > 0 \text{ für alle } 1 \leq i \leq n+1, \\ 0, & \text{sonst.} \end{cases}$$

Wir betrachten die lineare Transformation

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

Die Jacobi-Determinante von  $T$  ist

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

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

$$h(z_1, \dots, z_{n+1}) = \begin{cases} e^{-(y_1 + \dots + y_{n+1})} = e^{-z_{n+1}} & , 0 < z_1 < z_2 < \dots < z_{n+1}, \\ 0, & \text{sonst.} \end{cases}$$

Nun betrachten wir eine weitere Transformation

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

Die Jacobi-Determinante von  $S$  ist

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

Die Dichte von  $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}}, S_{n+1})$  kann nun mit der Transformationsformel wie folgt berechnet werden:

$$h(w_1, \dots, w_{n+1}) = \begin{cases} e^{-z_{n+1}} z_{n+1}^n = e^{-w_{n+1}} w_{n+1}^n, & 0 < w_1 < w_2 < \dots < w_n < 1, w_{n+1} > 0, \\ 0, & \text{sonst.} \end{cases}$$

Die Dichte von  $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$  ist eine Marginaldichte von  $h$ :

$$r(w_1, \dots, w_n) = \begin{cases} \int_{\mathbb{R}} h(w_1, \dots, w_{n+1}) dw_{n+1} = n!, & 0 < w_1 < w_2 < \dots < w_n < 1, \\ 0, & \text{sonst.} \end{cases}$$

Man sieht, dass die Dichte  $f$  von  $(U_{1:n}, \dots, U_{n:n})$  und die Dichte  $r$  von  $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$  übereinstimmen, woraus die Behauptung folgt.  $\square$

# 6 Rekorde

Seien  $X_1, X_2, \dots$  unabhängige und identisch verteilte Zufallsvariablen mit stetiger Verteilungsfunktion  $F$ . Wir setzen  $M_n = \max\{X_1, \dots, X_n\}$ . Seien  $\xi_1, \xi_2, \dots$  Zufallsvariablen mit  $\xi_1 = 1$  und

$$\xi_n = 1_{X_n > M_{n-1}}, \quad n = 2, 3, \dots$$

Das Ereignis  $\{\xi_n = 1\}$  tritt genau dann ein, wenn der Wert  $X_n$  größer als alle vorherigen Werte  $X_1, \dots, X_{n-1}$  ist. Somit ist  $\xi_n$  die Indikatorvariable des Ereignisses, dass zum Zeitpunkt  $n$  ein neuer Rekord aufgestellt wird.

## 6.1 Satz von Rényi

**Theorem 6.1** (Rényi). *Es gilt  $\mathbb{P}[\xi_n = 1] = \frac{1}{n}$ . Außerdem sind die Zufallsvariablen  $\xi_1, \dots, \xi_n, M_n$  unabhängig.*

*Beweis.* Aus der Stetigkeit von  $F$  folgt, dass  $\mathbb{P}[X_i = X_j] = 0$  für  $i \neq j$ . 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. Es folgt, dass  $\mathbb{P}[\xi_n = 1] = \frac{1}{n}$ .

Wir zeigen die Unabhängigkeit. Seien nun  $1 \leq \alpha(1) < \alpha(2) < \dots < \alpha(s) \leq n$  und  $x \in \mathbb{R}$ . Wir beweisen, dass

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

Es sei zuerst  $s = 1$ . Schreibe  $M_{k,l} = \max\{X_k, \dots, X_l\}$  mit  $k \leq l$ . Wir zeigen, dass

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

Dies kann man wie folgt mit Hilfe der Identitäten

$$\mathbb{P}[X_{\alpha(1)} = u] = F(du), \quad \mathbb{P}[M_{\alpha(1)-1} < u] = (F(u))^{\alpha(1)-1}$$

umschreiben:

$$\mathbb{P}[M_{\alpha(1)-1} < X_{\alpha(1)} < x](F(x))^{n-\alpha(1)} = (F(x))^{n-\alpha(1)} \cdot \int_{-\infty}^x F(du)(F(u))^{\alpha(1)-1}.$$

Das Integral lässt sich einfach zu folgendem Ausdruck integrieren:

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

was die Aussage im Fall  $s = 1$  beweist. Sei nun  $s \in \mathbb{N}$ . Es gilt

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

Dies lässt sich ähnlich wie im Fall  $s = 1$  schreiben als:

$$(F(x))^{n-\alpha(s)} \int_{-\infty}^x F(du_1)(F(u_1))^{\alpha(1)-1} \dots \int_{u_{s-1}}^x F(du_s)(F(u_s))^{\alpha(s)-\alpha(s-1)-1}.$$

Als Übungsaufgabe bleibt zu zeigen, dass sich obiges Integral zu Folgendem errechnen lässt:

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

□

**Bemerkung 6.2.** Es sei  $N(n) = \sum_{k=1}^n \xi_k$  die Anzahl der Rekorde im Intervall  $1, \dots, n$ . Nach dem Satz von Rényi gilt

$$\mathbb{E}N(n) = \sum_{k=1}^n \mathbb{E}\xi_k = \sum_{k=1}^n \frac{1}{k} \sim \log n, \quad n \rightarrow \infty.$$

Es werden also relativ wenige Rekorde erwartet.

## 6.2 Rekordzeiten

Wir definieren nun die Rekordzeiten  $L(1) < L(2) < \dots$  durch:  $L(1) = 1$ ,  $L(2) = \min\{j > 1 : \xi_j = 1\}$  und allgemein

$$L(n+1) = \min\{j > L(n) : \xi_j = 1\}, \quad n = 2, 3, \dots$$

Somit ist  $L(n)$  der Zeitpunkt, zu dem der  $n$ -te Rekord aufgestellt wird. Im nächsten Satz beschreiben wir die gemeinsame Verteilung des Vektors  $(L(1), \dots, L(n))$ .

**Theorem 6.3.** Für  $1 = j(1) < j(2) < \dots < j(n)$  gilt

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

*Beweis.* Es gilt:

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

Wegen der sich aus Satz 6.1 ergebenden Unabhängigkeit kann man dies in folgenden Ausdruck umschreiben:

$$\mathbb{P}[\xi_2 = 0] \cdot \dots \cdot \mathbb{P}[\xi_{j(n)} = 0] \cdot \frac{\mathbb{P}[\xi_{j(2)} = 1]}{\mathbb{P}[\xi_{j(2)} = 0]} \cdot \dots \cdot \frac{\mathbb{P}[\xi_{j(n)} = 1]}{\mathbb{P}[\xi_{j(n)} = 0]},$$

was sich ebenfalls, wegen des Satzes von Rényi, wie folgt darstellen lässt:

$$\left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \dots \left(1 - \frac{1}{j(n)}\right) \frac{1/j(2)}{1 - 1/j(2)} \dots \frac{1/j(n)}{1 - 1/j(n)}.$$

Durch geschicktes Umformen lässt sich das wie folgt darstellen:

$$\frac{2-1}{2} \frac{3-1}{3} \dots \frac{j(n)-1}{j(n)} \frac{1}{j(2)-1} \dots \frac{1}{j(n)-1} = \frac{1}{j(n)} \frac{1}{j(2)-1} \dots \frac{1}{j(n)-1},$$

wobei sich die Gleichheit ergibt, da die ersten  $j(n) - 1$  Faktoren ein Teleskopprodukt bilden.  $\square$

**Bemerkung 6.4.** Die Verteilung der Rekordzeiten  $L(1), L(2), \dots$  ist (abgesehen von der Stetigkeitsannahme an die Verteilungsfunktion  $F$ ) unabhängig von der Verteilung der Zufallsvariablen  $X_1, X_2, \dots$

**Bemerkung 6.5.** Wir können nun die Verteilung von  $L(n)$  ausrechnen. Es gilt

$$\mathbb{P}[L(n) = k] = \sum_{1=j(1) < j(2) < \dots < j(n)=k} \mathbb{P}[L(1) = j(1), \dots, L(n) = j(n)].$$

Wegen Satz 6.3 erhalten wir

$$\mathbb{P}[L(n) = k] = \sum_{1=j(1) < j(2) < \dots < j(n)=k} \frac{1}{k(j(2) - 1) \cdot \dots \cdot (j(n) - 1)}.$$

Wir werden später eine explizite Formel für die Verteilung von  $L(n)$  herleiten.

Im nächsten Satz werden wir zeigen, dass die Folge der Rekordzeiten  $L(1), L(2), \dots$  eine sogenannte Markov-Kette bildet. 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  $L(n+1)$ ? Es stellt 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 irrelevant. Diese Eigenschaft wird als Markov-Eigenschaft bezeichnet.

**Theorem 6.6.** *Für alle  $1 = i(1) < i(2) < \dots < i(n) = i < j$  gilt die Markov-Eigenschaft der Rekordzeiten:*

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

Außerdem gilt:

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

*Beweis.* Wir berechnen nur die Wahrscheinlichkeit  $\mathbb{P}[L(n+1) = j | L(n) = i]$ , für die gilt:

$$\mathbb{P}[L(n+1) = j | L(n) = i] = \frac{\mathbb{P}[L(n+1) = j, L(n) = i]}{\mathbb{P}[L(n) = i]},$$

was folgendem Ausdruck entspricht:

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

Nach Satz 6.3 lässt sich dies umformen zu:

$$\frac{\sum_{1=i(1)<i(2)<\dots<i(n)=i<j=i(n+1)} \frac{1}{j \cdot (i(2)-1) \cdot \dots \cdot (i(n-1)-1) \cdot (i-1) \cdot (j-1)}}{\sum_{1=i(1)<i(2)<\dots<i(n)=i<j=i(n+1)} \frac{1}{i \cdot (i(2)-1) \cdot \dots \cdot (i(n-1)-1) \cdot (i-1)}} = \frac{i}{j(j-1)}.$$

□

Der obige Satz zeigt, wie man die Folge der Rekordzeiten am Rechner simulieren kann, ohne dafür die Variablen  $X_1, X_2, \dots$  erzeugen zu müssen. Man startet mit  $L(1) = 1$  und geht induktiv vor. Sind die Werte  $L(1), \dots, L(n)$  mit  $L(n) = i$  bekannt, so erzeugt man eine Zufallsvariable auf  $\{i+1, i+2, \dots\}$ , indem man den Wert  $j$  mit Wahrscheinlichkeit  $\frac{j}{i(i-1)}$  auswählt. Dieser Wert ist dann der Wert von  $L(n+1)$ . Danach wiederholt man das Ganze.

Der nächste Satz liefert eine einfache Darstellung der Rekordzeiten. Definiere die Gauß-Klammer durch

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

**Theorem 6.7.** Seien  $U_1, U_2, \dots$  unabhängig und gleichverteilt auf dem Intervall  $[0, 1]$ . Definiere  $R(1) = 1$  und  $R(n+1) = \lfloor \frac{R(n)}{U_n} \rfloor + 1$  für  $n \in \mathbb{N}$ . Dann gilt die Gleichheit der Verteilungen:

$$(L(1), L(2), \dots, L(n)) \stackrel{d}{=} (R(1), R(2), \dots, R(n)).$$

*Beweis.* Wegen der Markov-Eigenschaft reicht es zu zeigen, dass

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

Es ist

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

Die Zufallsvariable  $R(n)$  hängt nur von  $U_1, \dots, U_{n-1}$  ab. Die Ereignisse  $\{\lfloor \frac{i}{U_n} \rfloor + 1 = j\}$  und  $\{R(n) = i\}$  sind also unabhängig und somit vereinfacht sich das Ganze zu folgendem Ausdruck:

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

Da es sich bei  $U_n$  um eine auf  $[0, 1]$  gleichverteilte Zufallsvariable handelt ergibt sich:

$$\mathbb{P}[R(n+1) = j | R(n) = i] = \mathbb{P}\left[U_n \in \left(\frac{i}{j}, \frac{i}{j-1}\right]\right] = \frac{i}{j(j-1)}.$$

Der letzte Term ist gemäß Satz 6.6 gleich der Wahrscheinlichkeit  $\mathbb{P}[L(n+1) = j | L(n) = i]$ , was zu zeigen war.  $\square$

### 6.3 Anzahl der Rekorde

Wir haben schon gezeigt, dass für  $N(n) = \sum_{k=1}^n \xi_k$  (die Anzahl der Rekorde im Intervall  $\{1, \dots, n\}$ ) folgendes gilt:  $\mathbb{E}N(n) = \sum_{k=1}^n \frac{1}{k}$ . Im nächsten Satz berechnen wir die komplette Verteilung von  $N(n)$ . Dabei stoßen wir auf die Stirling-Zahlen  $S_n^k$ , die durch die Gleichung

$$x(x-1)\dots(x-n+1) = \sum_{k=0}^n S_n^k x^k$$

definiert sind.

**Theorem 6.8.** Für die Verteilung der Anzahl der Rekorde gilt

$$\mathbb{P}[N(n) = k] = (-1)^{n+k} \frac{S_n^k}{n!} = \frac{|S_n^k|}{n!}, \quad k = 1, \dots, n.$$

**Bemerkung 6.9.** Setzt man  $x = -1$  in die Definition der Stirling-Zahlen ein, so erhält man  $\sum_{k=1}^n (-1)^{n+k} S_n^k = n!$ . Also summieren sich die Wahrscheinlichkeiten zu 1.

*Beweis.* Die Zufallsvariablen  $\xi_1, \dots, \xi_n$  sind nach dem Satz von Rényi unabhängig mit  $\mathbb{P}[\xi_k = 1] = \frac{1}{k}$  bzw.  $\mathbb{P}[\xi_k = 0] = 1 - \frac{1}{k}$ .

Die erzeugende Funktion einer Zufallsvariable  $Z$  mit Werten in  $\{0, 1, \dots\}$  ist definiert durch

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

Im Folgenden werden wir die erzeugende Funktion von  $N(n)$  angeben. Die erzeugende Funktion von  $\xi_k$  ist

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

Es gilt  $N(n) = \xi_1 + \dots + \xi_n$  und deshalb ist:

$$g_{N(n)}(t) = g_{\xi_1}(t) \dots g_{\xi_n}(t) = \prod_{k=1}^n \frac{k-1+t}{k} = (-1)^n \prod_{k=1}^n \frac{-t-(k-1)}{k}.$$

Mit Hilfe der Definition der Stirling-Zahlen lässt sich dies zu folgendem Ausdruck umformen:

$$g_{N(n)}(t) = \frac{(-1)^n}{n!} \sum_{k=1}^n S_n^k (-t)^k = \sum_{k=1}^n t^k \frac{S_n^k (-1)^{n+m}}{n!}.$$

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

$$\mathbb{P}[N(n) = k] = (-1)^{n+k} \frac{S_n^k}{n!} = \frac{|S_n^k|}{n!},$$

wobei die letzte Gleichheit aus der Nichtnegativität der Wahrscheinlichkeiten folgt.  $\square$

**Aufgabe 6.10.** Zeigen Sie mit dem obigen Satz, dass für die Verteilung der Rekordzeiten gilt

$$P[L(n) = k] = \frac{|S_{n-1}^{k-1}|}{k!}, \quad k = n, n+1, \dots$$

## 6.4 Zentrale Grenzwertsätze für Rekorde

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

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

**Theorem 6.11** (Zentraler Grenzwertsatz für die Anzahl der Rekorde). *Es gilt*

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

Der Beweis basiert auf einer Verallgemeinerung des zentralen Grenzwertsatzes, die wir ohne Beweis angeben.

**Theorem 6.12** (Zentraler Grenzwertsatz von Ljapunow). *Für alle natürlichen Zahlen  $n$  seien  $Z_{n1}, \dots, Z_{nn}$  unabhängige Zufallsvariablen mit  $\mathbb{E}Z_{nk} = 0$ ,  $\sigma_{nk}^2 := \text{Var}Z_{nk} \in (0, \infty)$  für  $k = 1, \dots, n$  und  $\sum_{k=1}^n \sigma_{nk}^2 = 1$ . Außerdem gelte  $\sum_{k=1}^n \mathbb{E}|Z_{nk}|^{2+\delta} \rightarrow 0$  für ein  $\delta > 0$  und  $n \rightarrow \infty$ . Dann gilt:*

$$Z_{n1} + \dots + Z_{nn} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

**Bemerkung 6.13.** Die identische Verteiltheit der Zufallsvariablen wird im zentralen Grenzwertsatz von Ljapunow nicht benötigt und daher nicht vorausgesetzt.

*Beweis von Satz 6.11.* Es ist  $N(n) = \xi_1 + \dots + \xi_n$ , wobei die  $\xi_k$  nach dem Satz von Rényi unabhängig aber nicht identisch verteilt sind. Setze

$$Z_{nk} = \frac{\xi_k - \frac{1}{k}}{\sigma_n}, \quad k = 1, \dots, n,$$

mit

$$\sigma_n^2 = \text{Var}(N(n)) = \sum_{k=1}^n \text{Var}(\xi_k) = \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{k^2} \sim \log n, \quad n \rightarrow \infty,$$

denn  $\sum_{k=1}^n \frac{1}{k} \sim \log n$  und  $\sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$ . Definitionsgemäß gilt

$$\mathbb{E}Z_{nk} = 0, \quad \sum_{k=1}^n \mathbb{E}Z_{nk}^2 = 1.$$

Wir zeigen, dass die Ljapunow-Bedingung mit  $\delta = 3$  gilt. Die Zufallsvariable  $\frac{\xi_k - 1/k}{\sigma_n}$  nimmt nur zwei Werte an, und zwar  $\frac{1/k}{\sigma_n}$  mit Wahrscheinlichkeit  $1 - \frac{1}{k}$  und  $\frac{1-1/k}{\sigma_n}$  mit

Wahrscheinlichkeit  $\frac{1}{k}$ . Es folgt

$$\sum_{k=1}^n \mathbb{E} \left| \frac{\xi_k - \frac{1}{k}}{\sigma_n} \right|^3 \leq \sum_{k=1}^n \left( \frac{1}{k\sigma_n} \right)^3 \cdot 1 + \left( \frac{1}{\sigma_n} \right)^3 \cdot \frac{1}{k} = \frac{1}{\sigma_n^3} \sum_{k=1}^n \left( \frac{1}{k^3} + \frac{1}{k} \right) \rightarrow 0 \text{ für } n \rightarrow \infty,$$

da  $\sigma_n^3 \sim (\log n)^{3/2}$  und  $\sum_{k=1}^n \left( \frac{1}{k^3} + \frac{1}{k} \right) \sim \log n$  für  $n \rightarrow \infty$ . Es sind also alle Voraussetzungen für den zentralen Grenzwertsatz von Ljapunow gegeben und damit folgt:

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

Unter Beachtung des bereits gezeigten Zusammenhangs  $\sigma_n^2 \sim \log n$  führt das zur behaupteten Grenzaussage:

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

(Man kann z.B. das Khintchine-Lemma verwenden). □

Mit Hilfe des obigen Satzes leiten wir einen zentralen Grenzwertsatz für die Rekordzeiten her.

**Theorem 6.14** (Zentraler Grenzwertsatz für Rekordzeiten). *Es gilt*

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

*Beweis.* Sei  $x \in \mathbb{R}$  fest. Mit  $n(x) = e^{n+x\sqrt{n}}$  gilt

$$\mathbb{P} \left[ \frac{\log L(n) - n}{\sqrt{n}} \leq x \right] = \mathbb{P}[L(n) \leq n(x)] = \mathbb{P}[N(n(x)) \geq n].$$

Zur Vereinfachung sei  $n(x) \in \mathbb{Z}$ , dabei verliert der Beweis nicht an Allgemeingültigkeit, da für  $n(x) \notin \mathbb{Z}$  einfach  $[n(x)]$  betrachtet werden kann. Nach Satz 6.11 gilt

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

Es folgt, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}[N(n(x)) \geq n] = \lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{N(n(x)) - \log n(x)}{\sqrt{\log n(x)}} \geq \frac{n - (n + x\sqrt{n})}{\sqrt{n + x\sqrt{n}}} \right] = 1 - \Phi(-x) = \Phi(x),$$

da  $\lim_{n \rightarrow \infty} \frac{n - (n+x\sqrt{n})}{\sqrt{n+x\sqrt{n}}} = -x$ . Dabei bezeichnet  $\Phi(x)$  die Verteilungsfunktion von  $\mathcal{N}(0, 1)$ . □

**Aufgabe 6.15.** Zeigen Sie, dass für die Anzahl der Rekorde auch das Gesetz der großen Zahlen gilt:

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

**Theorem 6.16.** Sei  $x > 1$  eine natürliche Zahl, dann gilt:

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

Sei  $x > 1$  eine beliebige reelle Zahl, dann gilt:

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

Das heißt,  $\frac{L(n+1)}{L(n)}$  konvergiert in Verteilung gegen eine Pareto-verteilte Zufallsvariable mit Verteilungsfunktion  $1 - \frac{1}{x}$ ,  $x > 1$ .

**Bemerkung 6.17.** Man sieht hier noch einmal, dass Rekorde mit  $n \rightarrow \infty$  immer seltener auftreten. Zum Beispiel ist mit Wahrscheinlichkeit  $1/2$  der Abstand zwischen der  $(n+1)$ -ten und der  $n$ -ten Rekordzeit größer als die  $n$ -te Rekordzeit selbst. Mit Wahrscheinlichkeit  $1/3$  ist der Abstand zwischen der  $(n+1)$ -ten und der  $n$ -ten Rekordzeit mindestens doppelt so groß wie die  $n$ -te Rekordzeit selbst, usw. Die Tatsache, dass Rekorde immer seltener auftreten ist ziemlich natürlich: die Rekordwerte steigen nämlich mit der Zeit und es wird immer schwieriger neue Rekorde aufzustellen.

*Beweis.* Sei  $x > 1$ . Dann gilt wegen des Gesetzes der totalen Wahrscheinlichkeit:

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

denn  $L(n+1)$  ist ganzzahlig und somit ist  $L(n+1) > xi$  zu  $L(n+1) > [xi]$  äquivalent. Im Beweis von Satz 6.6 wurde die Wahrscheinlichkeit  $\mathbb{P}[L(n+1) > xi | L(n) = i]$  bereits berechnet. Wir wollen das Ergebnis hier verwenden. So lässt sich obiger Ausdruck zu folgendem vereinfachen:

$$\mathbb{P}[L(n+1) > xL(n)] = \sum_{i=n}^{\infty} \frac{i}{[xi]} \mathbb{P}[L(n) = i]. \quad (6.1)$$

Sei zuerst  $x \in \mathbb{N}$ . Man sieht:  $\sum_{i=n}^{\infty} \frac{i}{[xi]} \mathbb{P}[L(n) = i] = \frac{1}{x} \sum_{i=n}^{\infty} \mathbb{P}[L(n) = i] = \frac{1}{x}$ . Sei nun  $x > 1$  beliebig reell. Aus der Definition der Gauß-Klammer folgt, dass  $[xi] \leq xi < [xi] + 1$ . Durch leichte Umformungen folgt:

$$\frac{1}{x} \leq \frac{i}{[xi]} < \frac{1}{x} + \frac{1}{[xi]x}.$$

Deshalb kann man (6.1) wie folgt nach unten abschätzen:

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

Außerdem kann man daher (6.1) wie folgt nach oben abschätzen:

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

da  $[xi] \geq [xn]$  für  $i \geq n$ . Insgesamt folgt mit dem ‘Sandwich-Prinzip’ die Behauptung.  $\square$

**Aufgabe 6.18.** Zeigen Sie, dass für die Rekordzeiten auch das Gesetz der großen Zahlen gilt:

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

**Aufgabe 6.19.** Ist es richtig, dass

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

Zum Schluss betrachten wir noch die Rekordwerte. Diese sind definiert als

$$X(n) = M_{L(n)} = X_{L(n)}.$$

Im Spezialfall der exponentialverteilten Zufallsvariablen besitzen die Rekordwerte eine schöne Darstellung.

**Theorem 6.20.** Seien  $X_1, X_2, \dots$  unabhängig und exponentialverteilt mit Parameter 1. Dann gilt:

$$(X(1), X(2), \dots, X(n)) \stackrel{d}{=} (\nu_1, \nu_1 + \nu_2, \dots, \nu_1 + \dots + \nu_n),$$

wobei  $\nu_1, \nu_2, \dots$  unabhängig unabhängig und exponentialverteilt mit Parameter 1 sind.

**Bemerkung 6.21.** Mit anderen Worten, die Folge  $X(1), X(2), \dots$  ist ein Poisson-Prozess mit Intensität 1.

*Beweisidee.* Der Beweis basiert auf der Gedächtnislosigkeit der Exponentialverteilung. Ist nämlich  $X \sim \text{Exp}(1)$ , so ist für jedes  $t > 0$  die bedingte Verteilung von  $X - t$  gegeben, dass  $X > t$ , ebenfalls eine Exponentialverteilung mit Parameter 1.

Nach Voraussetzung ist  $X(1) = X_1 \sim \text{Exp}(1)$ . Wir halten nun  $X_1 = a_1$  fest und warten auf den zweiten Rekordwert  $X(2)$ . Der Exzess  $X(2) - a_1$  hat die gleiche Verteilung, wie  $X - a_1$  gegeben, dass  $X > a_1$ , also die Exponentialverteilung mit Parameter 1. Nun halten wir  $X(2) = a_2$  fest und warten auf den dritten Rekordwert  $X(3)$ . Der Exzess  $X(3) - a_2$  hat die gleiche Verteilung, wie  $X - a_2$  gegeben, dass  $X > a_2$ , also wieder die Exponentialverteilung mit Parameter 1, usw.

Anschaulich kann man die Verteilung der Rekordwerte  $X(1), X(2), \dots$  wie folgt erklären: Die Zuwächse im Bild 6.1 sind immer standard exponentialverteilt und unabhängig voneinander.

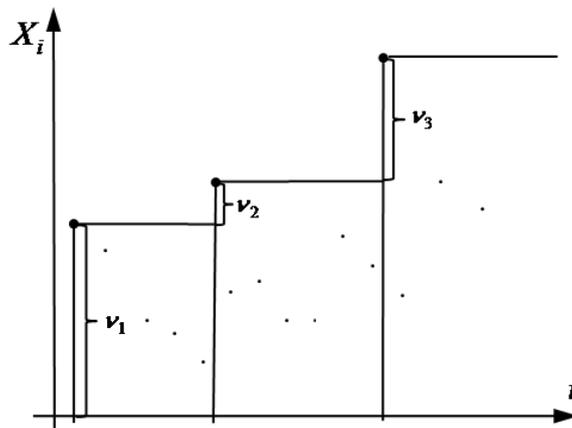


Abbildung 6.1: Darstellung der Zuwächse  $X(i) - X(i - 1)$

□

**Korollar 6.22.** *Unter Voraussetzungen des obigen Satzes gilt:*

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

*Beweis.* Aus Satz 6.20 wissen wir, dass  $X(n) \stackrel{d}{=} \nu_1 + \dots + \nu_n$ . Dabei ist  $\mathbb{E}\nu_i = \text{Var } \nu_i = 1$ . Der zentrale Grenzwertsatz ergibt, dass

$$\frac{\nu_1 + \dots + \nu_n - n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1) \text{ für } n \rightarrow \infty.$$

Daraus folgt die Behauptung.

□

**Bemerkung 6.23.** Im Satz 1.7 haben wir nachgewiesen, dass für standard exponentialverteilte Zufallsvariablen  $M_n - \log n$  gegen die Gumbel-Verteilung konvergiert. Unter passenden Normierungen gilt also  $M_n \rightarrow G(t)$  und gleichzeitig  $M_{L(n)} \rightarrow \mathcal{N}(0, 1)$ . Die Grenzverteilung wird also offenbar durch den Umstand, dass  $L(n)$  zufällig ist völlig verändert.

# 7 Poisson-Punktprozesse

Poisson-Punktprozesse sind natürliche Modelle für zufällige Konfigurationen von Punkten im Raum. Wie der Name sagt, spielt die Poisson-Verteilung eine entscheidende Rolle. Wir werden also mit der Definition der Poisson-Verteilung anfangen.

## 7.1 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.

**Theorem 7.1.** *Es sei  $p_n$  eine Folge mit  $\lim_{n \rightarrow \infty} np_n = \lambda$ , wobei  $\lambda > 0$ . 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 7.2.** *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.

## 7.2 Beispiel zu Poisson-Prozessen

Wir betrachten nun wieder eine sehr große Zahl von unabhängigen Experimenten mit sehr kleinen Erfolgswahrscheinlichkeiten. Diesmal stellen wir uns aber vor, dass jedes

Experiment außerdem 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.

Wir werden versuchen, ein stochastisches Modell für die Verteilung der Sterne im Himmel zu finden. Ist  $A$  ein Gebiet (ein Teil des Himmels), so bezeichnen wir mit  $\pi(A)$  die Anzahl der Sterne in  $A$ . Diese Anzahl fassen wir als eine Zufallsvariable mit Werten in  $\mathbb{N}_0$  auf. Wir gehen davon aus, dass folgende Eigenschaft gilt:

*Für disjunkte Gebiete  $A_1, \dots, A_n$  sind die Zufallsvariablen  $\pi(A_1), \dots, \pi(A_n)$  unabhängig.*

Außerdem machen wir folgende Annahme. Ist  $Q$  ein kleines Gebiet mit Fläche  $\epsilon \approx 0$ , so gilt

$$\mathbb{P}[\text{“Es gibt einen Stern in } Q\text{”}] \approx \lambda\epsilon, \quad \epsilon \approx 0.$$

Das heißt, die Wahrscheinlichkeit, einen Stern in einem sehr kleinen Gebiet zu finden, ist proportional zum Flächeninhalt dieses Gebiets. Der Koeffizient  $\lambda > 0$  beschreibt dabei die Intensität der Sterne im Himmel.

Wie ist nun  $\pi(A)$  für ein beliebiges Gebiet  $A$  verteilt? Dazu unterteilen wir  $A$  in kleine Gebiete mit Fläche  $\epsilon$ . Die Anzahl dieser Gebiete ist  $\approx \text{Fläche}(A)/\epsilon$ . Aus unseren Voraussetzungen und dem Poisson-Grenzwertsatz folgt, dass

$$\pi(A) \approx \text{Bin}\left(\frac{\text{Fläche}(A)}{\epsilon}, \lambda\epsilon\right) \approx \text{Poi}(\lambda \cdot \text{Fläche}(A)) \text{ für } \epsilon \downarrow 0.$$

Es gilt also:

*Für jedes Gebiet  $A$  ist  $\pi(A)$  Poisson-verteilt mit Parameter  $\lambda \cdot \text{Fläche}(A)$ .*

Eine zufällige Konfiguration von Punkten im Raum, die die beiden oben genannten Eigenschaften hat, bezeichnen wir als einen Poisson-Punktprozess.

## 7.3 Definition der Poisson-Prozesse

Wir werden nun eine mathematische Definition der Poisson-Punktprozesse geben.

### 7.3.1 Zählmaße

Die erste Frage ist, wie man eine “Punktekonfiguration” definiert. Wir bezeichnen mit  $\mathcal{B}^d$  die  $\sigma$ -Algebra der Borel-Mengen in  $\mathbb{R}^d$  und mit  $\mathcal{B}_0^d$  die Familie der beschränkten Borel-Mengen.

**Definition 7.3.** *Ein Maß  $\mu$  auf  $\mathbb{R}^d$  heißt Zählmaß, falls für alle  $A \in \mathcal{B}_0^d$ ,  $\mu(A)$  eine nicht-negative ganze Zahl ist.*

**Beispiel 7.4.** Sei  $x \in \mathbb{R}^d$ . Das Dirac-Maß  $\delta_x$  ist definiert durch

$$\delta_x(A) = \begin{cases} 1, & x \in A, \\ 0, & \text{sonst.} \end{cases}$$

Offensichtlich ist  $\delta_x$  ein Zählmaß. Man kann sich  $\delta_x$  als eine Punktekonfiguration vorstellen, die aus einem Punkt  $x$  besteht. Etwas allgemeiner, jede endliche oder abzählbare Summe  $\mu = \sum_{i=1}^n \delta_{x_i}$ , wobei  $n \in \mathbb{N} \cup \{\infty\}$  und  $x_1, x_2, \dots \in \mathbb{R}^d$ , ist ein Zählmaß, wenn man zusätzlich fordert, dass die Folge  $x_1, x_2, \dots$  keine Häufungspunkte besitzt.

**Beispiel 7.5.** Kein Zählmaß hingegen ist:  $\mu = \sum_{n=1}^{\infty} \delta_{1/n}$ , denn hier ist  $\mu([0, 1]) = \infty$ .

Umgekehrt kann man zeigen, dass jedes Zählmaß  $\mu$  sich als

$$\mu = \sum_{i=1}^n \delta_{x_i}$$

darstellen lässt, wobei  $n \in \mathbb{N} \cup \{\infty\}$  und  $x_1, x_2, \dots$  eine endliche oder abzählbar unendliche Folge von Punkten in  $\mathbb{R}^d$  ist, die keine Häufungspunkte besitzt. Zählmaße sind somit (deterministische) Punktekonfigurationen ohne Häufungspunkte.

### 7.3.2 Punktprozesse

Wir wollen nun definieren, was eine *zufällige* Punktekonfiguration ohne Häufungspunkte (ein zufälliges Zählmaß) ist. Solche zufälligen Punktekonfigurationen heißen Punktprozesse. Wir bezeichnen mit  $\mathbb{M}$  die Menge aller Zählmaße.

**Definition 7.6.** Sei  $\mathcal{M} \subset 2^{\mathbb{M}}$  definiert als die  $\sigma$ -Algebra erzeugt von Mengen der Form  $\{\mu \in \mathbb{M} : \mu(A_i) = k_i, i = 1, \dots, n\}$ , wobei  $n \in \mathbb{N}$ ,  $k_1, \dots, k_n \in \mathbb{N} \cup \{0\}$  und  $A_1, \dots, A_n \in \mathcal{B}^d$ .

**Definition 7.7.** Ein Punktprozess ist eine messbare Abbildung von einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathbb{P})$  nach  $(\mathbb{M}, \mathcal{M})$ .

Bezeichnet  $\pi$  einen Punktprozess, so sei  $\pi(A)$  definiert als die Anzahl der Punkte in einer Menge  $A \in \mathcal{B}^d$ . Somit ist  $\pi(A)$  eine Zufallsvariable mit Werten in  $\mathbb{N}_0$ .

**Beispiel 7.8.** 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 Binomialpunktprozess genannt, denn die Anzahl der Punkte in einer Menge  $A \in \mathcal{B}^d$  ist binomialverteilt:  $\pi(A) \sim \text{Bin}(N, \mathbb{P}[X_1 \in A])$ .

### 7.3.3 Poisson-Punktprozesse

**Definition 7.9.** Sei  $\mu$  ein Maß auf  $\mathbb{R}^d$  mit  $\mu(A) < \infty$  für alle  $A \in \mathcal{B}_0^d$  (Radon-Maß). Ein Punktprozess  $\pi$  auf  $\mathbb{R}^d$  heißt Poisson-Punktprozess mit Intensitätsmaß  $\mu$ , falls folgende zwei Bedingungen gelten:

1. Für alle  $A \in \mathcal{B}^d$ :  $\pi(A) \sim \text{Poi}(\mu(A))$ .
2. Für alle disjunkten Mengen  $A_1, \dots, A_n \in \mathcal{B}^d$  sind die Zufallsvariablen  $\pi(A_1), \dots, \pi(A_n)$  unabhängig.

Wir benutzen die Schreibweise  $\pi \sim \text{PPP}(\mu)$ .

**Bemerkung 7.10.** 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.

**Bemerkung 7.11.** Eine messbare Funktion  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  heißt lokal integrierbar, falls  $\int_B |f(t)| dt < \infty$  für jede Menge  $B \in \mathcal{B}_0^d$ . Ist nun  $f \geq 0$  eine lokal-integrierbare Funktion, so kann man ein Radon-Maß  $\mu$  mit  $\mu(B) = \int_B f(t) dt$  definieren. Die Funktion  $f$  heißt die Dichte von  $\mu$  und wir schreiben  $\mu(dt) = f(t) dt$ . 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  $\pi$ .

**Beispiel 7.12.** 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*.

**Beispiel 7.13.** Seien  $E_1, E_2, \dots$  unabhängige und identisch verteilte Zufallsvariablen mit  $\mathbb{P}[E_i > t] = e^{-\lambda t}$ ,  $t > 0$ . Dann ist  $\sum_{n=1}^{\infty} \delta_{E_1 + \dots + E_n}$  ein Poisson-Punktprozess mit Intensität  $\lambda 1_{(0, \infty)}$ . Um nun den eindimensionalen homogenen Poisson-Punktprozess mit Intensität  $\lambda$  zu konstruieren, muss man auf die gleiche Weise unabhängig einen Poisson-Punktprozess auf  $(-\infty, 0)$  erzeugen und die Vereinigung der beiden Punktprozesse nehmen.

## 7.4 Eigenschaften der Poisson-Punktprozesse

### 7.4.1 Superpositionssatz

Die Poisson-Verteilung ist faltungsstabil: sind  $X_1 \sim \text{Poi}(\lambda_1), \dots, X_n \sim \text{Poi}(\lambda_n)$  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.

**Theorem 7.14.** 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 7.15.** Wenn  $\sum_{i=1}^{\infty} \lambda_i = \infty$ , dann gilt  $S = \infty$  fast sicher.

*Beweis.* 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

$$\{S_1 \leq r\} \supseteq \{S_2 \leq r\} \supseteq \dots \text{ und } \bigcap_{i=1}^{\infty} \{S_i \leq r\} = \{S \leq r\}$$

Deshalb gilt:

$$\mathbb{P}[S \leq r] = \lim_{n \rightarrow \infty} \mathbb{P}[S_n \leq r] = \lim_{n \rightarrow \infty} \sum_{k=0}^r e^{-\sigma_n} \frac{\sigma_n^k}{k!},$$

da  $S_n \sim \text{Poi}(\sigma_n)$ . Man kann hier zwei Fälle unterscheiden:

**Fall 1:**  $\sigma_n \rightarrow \sigma$ ,  $\sigma < \infty$ . 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:  $\mathbb{P}[S \leq r] = 0$  und dann ist  $S = \infty$  fast sicher. □

Die Superposition  $\pi = \pi_1 + \pi_2 + \dots$  der Punktprozesse  $\pi_1, \pi_2, \dots$  ist die Vereinigung aller Punkte dieser Punktprozesse. Hierbei ist  $\pi$  nicht immer ein Zählmaß. Im nächsten Satz beweisen wir, dass die Superposition von unabhängigen Poisson-Punktprozessen wieder ein Poisson-Punktprozess ist.

**Theorem 7.16.** Seien  $\pi_i \sim \text{PPP}(\mu_i)$ ,  $i \in \mathbb{N}$ , unabhängige Poisson-Punktprozesse. 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  $A \in \mathcal{B}^d$ . Es gilt  $\pi_i(A) \sim \text{Poi}(\mu_i(A))$ , weil  $\pi_i \sim \text{PPP}(\mu_i)$ . Mit Theorem 7.14 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 gezeigt.

Seien nun  $A_1, \dots, A_n \in \mathcal{B}^d$  disjunkt. 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  $\pi_1, \pi_2, \dots$  unabhängig. Es folgt, dass

$$\pi(A_1) = \sum_{i=1}^{\infty} \pi_i(A_1), \dots, \pi(A_n) = \sum_{i=1}^{\infty} \pi_i(A_n) \text{ unabhängig sind.}$$

Daher ist auch die zweite Eigenschaft aus der Definition nachgewiesen, woraus die Behauptung folgt.  $\square$

**Bemerkung 7.17.** Aus  $A \in \mathcal{B}_0^d$  folgt, dass  $\pi(A) \sim \text{Poi}(\mu(A)) < \infty$ , woraus folgt, dass  $\pi$  ein Punktprozess ist.

## 7.4.2 Abbildungssatz

**Definition 7.18.** Sei  $T : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$  eine Borel-Abbildung. Sei  $\nu$  ein Maß auf  $\mathbb{R}^{d_1}$ . Das Bild von  $\nu$  ist ein Maß  $T\nu$  auf  $\mathbb{R}^{d_2}$  mit

$$(T\nu)(A) = \nu(T^{-1}(A)), \quad A \in \mathcal{B}^{d_2}.$$

**Beispiel 7.19.** Sei  $\delta$  ein Zählmaß, mit  $\delta = \sum_i \delta_{x_i}$ . Dann ist  $T\delta = \sum_i \delta_{Tx_i}$ .

**Beispiel 7.20.** Wir konstruieren ein Radon-Maß  $\nu$  und eine Abbildung  $T$  mit der Eigenschaft, dass  $T\nu$  kein Radon-Maß ist. Sei  $T : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $T(x, y) = x$  und sei  $\nu$  das Lebesgue-Maß auf  $\mathbb{R}^2$ . Dann folgt  $T\nu((a, b)) = \infty$ , für alle  $a < b$ . Somit ist  $T\nu$  kein Radon-Maß.

**Theorem 7.21.** Sei  $T : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$  eine Borel-Abbildung und  $\pi \sim \text{PPP}(\mu)$ . Sei  $T\mu$  ein Radon-Maß auf  $\mathbb{R}^{d_2}$ . Dann gilt  $T\pi \sim \text{PPP}(T\mu)$ .

*Beweis.* Wir überprüfen, ob die Bedingungen aus Definition 7.9 erfüllt sind. Sei  $A \in \mathcal{B}^{d_2}$ . Es gilt:

$$(T\pi)(A) = \pi(T^{-1}(A)) \sim \text{Poi}(\mu(T^{-1}(A))) = \text{Poi}((T\mu)(A)).$$

Für  $A \in \mathcal{B}_0^{d_2}$  folgt außerdem, dass  $(T\pi)(A) \sim \text{Poi}((T\mu)(A)) < \infty$  fast sicher, weshalb  $T\pi$  ein Punktprozess ist.

Seien nun  $A_1, \dots, A_n \in \mathcal{B}^{d_2}$  disjunkt, dann sind auch die Urbilder  $T^{-1}(A_1), \dots, T^{-1}(A_n)$  disjunkt. Es folgt, dass die Zufallsvariablen  $\pi(T^{-1}(A_1)), \dots, \pi(T^{-1}(A_n))$  unabhängig sind, da  $\pi \sim \text{PPP}(\mu)$ . Deshalb sind die Zufallsvariablen  $(T\pi)(A_1), \dots, (T\pi)(A_n)$  unabhängig.  $\square$

### 7.4.3 Laplacefunktionale

**Definition 7.22.** Es sei  $B(\mathbb{R}^d)$  die Menge aller Borel-Funktionen  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ . Es sei  $B_+(\mathbb{R}^d)$  die Menge aller  $f \in B(\mathbb{R}^d)$  mit  $f \geq 0$ .

**Definition 7.23.** Sei  $\pi$  ein Punktprozess auf  $\mathbb{R}^d$ . Für  $f \in B_+(\mathbb{R}^d)$  definiere die Zufallsvariable  $S_f = \sum_{x \in \pi} f(x)$ . Dann heißt

$$\psi_\pi : B(\mathbb{R}^d) \rightarrow [0, \infty) \text{ mit } \psi_\pi(f) = \mathbb{E}[e^{-S_f}]$$

das Laplace-Funktional von  $\pi$ .

Im nächsten Satz berechnen wir das Laplace-Funktional eines Poisson-Prozesses.

**Theorem 7.24.** Sei  $\pi$  ein Poisson-Punktprozess auf  $\mathbb{R}^d$  mit Intensität  $\mu$ , dann ist

$$\mathbb{E}[e^{-S_f}] = \exp \left\{ - \int_{\mathbb{R}^d} (1 - e^{-f(x)}) d\mu(x) \right\}. \quad (7.1)$$

*Beweis. Schritt 1.* (Indikatorfunktionen) Sei zuerst  $f(x) = c \cdot 1_A(x)$ , mit  $c > 0$ ,  $A \in \mathcal{B}_0^d$ . Dann ist  $S_f = c \cdot \pi(A)$ . Es gilt  $\pi(A) \sim \text{Poi}(\mu(A))$ . Es folgt

$$\mathbb{E}[e^{-S_f}] = \mathbb{E}[e^{-c\pi(A)}] = \sum_{k=0}^{\infty} e^{-\mu(A)} \frac{\mu(A)^k}{k!} e^{-ck} = e^{-\mu(A)} \sum_{k=0}^{\infty} \frac{(\mu(A)e^{-c})^k}{k!} = e^{-\mu(A)(1-e^{-c})}.$$

Somit gilt die Behauptung für  $f(x) = c \cdot 1_A(x)$ .

*Schritt 2.* (Einfache Funktionen) Sei nun  $f(x) = c_1 \cdot 1_{A_1}(x) + \dots + c_n \cdot 1_{A_n}(x)$  mit  $A_1, \dots, A_n \in \mathcal{B}_0^d$ , disjunkt, und  $c_1, \dots, c_n > 0$ . Dann gilt:

$$\mathbb{E}[e^{-S_f}] = \mathbb{E}[e^{-c_1\pi(A_1)} \dots e^{-c_n\pi(A_n)}] = \prod_{i=1}^n \mathbb{E}[e^{-c_i\pi(A_i)}],$$

da  $\pi$  ein Poisson-Punktprozess ist und daher die  $\pi(A_i)$  unabhängig sind. Da wir die Gültigkeit der Behauptung für diese Art von Funktionen bereits im ersten Schritt gezeigt haben, folgt:

$$\prod_{i=1}^n \mathbb{E}[e^{-c_i \pi(A_i)}] = \exp \left\{ - \sum_{i=1}^n \mu(A_i) (1 - e^{-c_i}) \right\} = \exp \left\{ - \int_{\mathbb{R}^d} (1 - e^{-f(x)}) d\mu(x) \right\}.$$

*Schritt 3.* Sei nun  $f \geq 0$  eine beliebige Borel-Funktion. Dann gibt es einfache Funktionen  $f_i$ , die punktweise von unten gegen  $f$  konvergieren. Da wir die Richtigkeit der Behauptung für einfache Funktionen bereits bewiesen haben, folgt mit der Monotonen Konvergenz, dass sie auch für  $f$  gilt.  $\square$

**Bemerkung 7.25.** Die Formel (7.1) gilt auch in folgender leicht allgemeiner Form:

$$\mathbb{E}[e^{-\theta S_f}] = \exp \left\{ - \int_{\mathbb{R}^d} (1 - e^{-\theta f(x)}) d\mu(x) \right\}, \quad \theta \geq 0.$$

Die Richtigkeit dieser Behauptung lässt sich nachweisen, indem man  $\theta f(x)$  in (7.1) einsetzt.

**Korollar 7.26** (Campbell). Sei  $\pi \sim PPP(\mu)$  auf  $\mathbb{R}^d$  und  $f \in B_+(\mathbb{R}^d)$ , dann gilt:

$$\mathbb{E}[S_f] = \int_{\mathbb{R}^d} f(x) d\mu(x).$$

*Beweis.* Wegen Satz 7.24 gilt:

$$\mathbb{E}[S_f] = - \frac{d}{d\theta} \log \mathbb{E}[e^{\theta S_f}] \Big|_{\theta=0} = \frac{d}{d\theta} \int_{\mathbb{R}^d} (1 - e^{-\theta f(x)}) d\mu(x) \Big|_{\theta=0}.$$

Durch Vertauschung von Integral und Ableitung lässt sich der Ausdruck wie folgt schreiben und vereinfachen:

$$\mathbb{E}[S_f] = \int_{\mathbb{R}^d} \frac{d}{d\theta} (1 - e^{-\theta f(x)}) d\mu(x) \Big|_{\theta=0} = \int_{\mathbb{R}^d} f(x) d\mu(x).$$

$\square$

**Aufgabe 7.27.** Zeigen Sie:

$$\mathbb{E}[S_f^2] = \int_{\mathbb{R}^d} f^2(x) d\mu(x) + \int_{\mathbb{R}^d} f(x) d\mu(x).$$

Im nächsten Satz zeigen wir, dass man einen Poisson-Punktprozess an seinem Laplace-Funktional erkennen kann.

**Theorem 7.28.** Sei  $\pi$  ein Punktprozess auf  $\mathbb{R}^d$  mit  $\mathbb{E}[e^{-S_f}] = e^{-\int_{\mathbb{R}^d} (1-e^{-f(x)})d\mu(x)}$  für alle Funktionen  $f \in B_+(\mathbb{R}^d)$ , wobei  $\mu$  ein Radon-Maß ist. Dann ist  $\pi \sim PPP(\mu)$ .

*Beweis.* Wir zeigen, dass  $\pi(A) \sim \text{Poi}(\mu(A))$  für alle  $A \in \mathcal{B}^d$ . Sei  $f(x) = \theta 1_A(x)$  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 Intensität  $\mu(A)$  ist.

Seien nun  $A_1, \dots, A_n \in \mathcal{B}^d$  disjunkt. Wir zeigen, dass dann  $\pi(A_1), \dots, \pi(A_n)$  unabhängig sind. Sei

$$f(x) = \theta_1 1_{A_1}(x) + \dots + \theta_n 1_{A_n}(x), \text{ mit } \theta_1, \dots, \theta_n \geq 0.$$

Einsetzen liefert:

$$\mathbb{E}[e^{-\theta_1\pi(A_1)} \dots e^{-\theta_n\pi(A_n)}] = \exp\left\{-\int_{\mathbb{R}^d} (1 - e^{-f(x)})d\mu(x)\right\} = \exp\left\{-\sum_{i=1}^n \mu(A_i)(1 - e^{-\theta_i})\right\}.$$

Weiterhin gilt:

$$\exp\left\{-\sum_{i=1}^n \mu(A_i)(1 - e^{-\theta_i})\right\} = \prod_{i=1}^n e^{-\mu(A_i)(1 - e^{-\theta_i})} = \prod_{i=1}^n \mathbb{E}[e^{-\theta_i\pi(A_i)}],$$

woraus folgt, dass  $\pi(A_1), \dots, \pi(A_n)$  unabhängig sind. □

## 7.5 Markierungen von Poisson-Punktprozessen

Sei  $\pi$  ein Poisson-Punktprozess mit Intensität  $\mu$  auf  $\mathbb{R}^{d_1}$ . Erzeuge für jeden Punkt  $x \in \mathbb{R}^{d_1}$  eine Zufallsvariable  $m_x$  mit Werten in  $\mathbb{R}^{d_2}$ . Wir bezeichnen die Verteilung von  $m_x$  mit  $\lambda_x$  ( $\lambda_x$  ist Wahrscheinlichkeitsmaß auf  $\mathbb{R}^{d_2}$  und darf von  $x$  abhängen). Wir brauchen folgende drei Annahmen:

1.  $m_x, x \in \mathbb{R}^{d_1}$ , sind unabhängige Zufallsvariablen.
2.  $m_x, x \in \mathbb{R}^{d_1}$ , sind unabhängig von  $\pi$ .
3.  $x \mapsto \lambda_x(B)$  ist eine Borel-Funktion für alle  $B \in \mathcal{B}^{d_2}$ .

Wir definieren den markierten Poisson-Punktprozess  $\pi^*$  auf  $\mathbb{R}^{d_1+d_2}$  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  $\pi$  gebildet wird.

**Theorem 7.29.**  $\pi^*$  ist Poisson-Punktprozess mit Intensität  $\mu^*$  auf  $\mathbb{R}^{d_1+d_2}$ , wobei

$$\mu^*(C) = \iint_{(x,m) \in C} d\mu(x) \lambda_x(dm), \text{ für } C \in \mathcal{B}^{d_1+d_2}.$$

*Beweis.* Sei  $f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \rightarrow \mathbb{R}$  mit  $f \geq 0$  eine Borel-Funktion. Wir schreiben  $S^* = \sum_{(x,m) \in \pi^*} f(x, m)$ . Wir wollen zuerst folgenden bedingten Erwartungswert berechnen:

$$\mathbb{E}[e^{-S^*} | \pi] = \prod_{x \in \pi} \mathbb{E}[e^{-f(x, m_x)} | \pi] = \prod_{x \in \pi} \int_{\mathbb{R}^{d_2}} e^{-f(x, m)} d\lambda_x(m).$$

Mit Hilfe einer einfachen Transformation lässt sich obiger Ausdruck wie folgt darstellen:

$$\prod_{x \in \pi} \int_{\mathbb{R}^{d_2}} e^{-f(x, m)} d\lambda_x(m) = \exp \left\{ - \sum_{x \in \pi} - \log \int_{\mathbb{R}^{d_2}} e^{-f(x, m)} d\lambda_x(m) \right\} =: \exp \left\{ - \sum_{x \in \pi} f^*(x) \right\}.$$

Es folgt:

$$\mathbb{E}[e^{-S^*}] = \mathbb{E}_\pi \left[ \exp \left\{ - \sum_{x \in \pi} f^*(x) \right\} \right],$$

was sich mit Hilfe von Satz 7.24 wie folgt umformen lässt:

$$\mathbb{E}_\pi \left[ \exp \left\{ - \sum_{x \in \pi} f^*(x) \right\} \right] = \exp \left\{ - \int_{\mathbb{R}^d} (1 - e^{-f^*(x)}) d\mu(x) \right\}.$$

Einsetzen von  $f^*$  liefert nun folgende Formel:

$$\exp \left\{ - \int_{\mathbb{R}^d} (1 - e^{-f^*(x)}) d\mu(x) \right\} = \exp \left\{ - \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} (1 - e^{-f(x, m)}) d\mu(x) \lambda(dm) \right\},$$

wobei man  $d\mu(x) \lambda(dm) = d\mu^*(x, m)$  setzt. Mit Satz 7.28 folgt schließlich, dass  $\pi^*$  Poisson-Punktprozess mit Intensität  $\mu^*$  ist.  $\square$

**Beispiel 7.30.** Sei  $\pi \sim \text{PPP}(\mu)$ . Färbe die Punkte  $x$  des Poisson-Punktprozesses mit verschiedenen Farben ein, wobei gelten soll, dass  $x$  mit Wahrscheinlichkeit  $p_i$  mit Farbe

$i \in \mathbb{N}$  gefärbt wird und sich die Wahrscheinlichkeiten zu 1 summieren ( $\sum_{i=1}^{\infty} p_i = 1$ ). Alle Punkte werden unabhängig voneinander gefärbt. Sei  $\pi_i$  der Punktprozess der Punkte der Farbe  $i$ . Dann gilt  $\pi_i \sim \text{PPP}(p_i\mu)$  und die Punktprozesse  $\pi_i$ ,  $i \in \mathbb{N}$ , sind unabhängig.

# 8 Konvergenz von Punktprozessen

## 8.1 Definition der Konvergenz von Punktprozessen

Sei  $\pi$  ein Punktprozess und seien  $B_1, \dots, B_k \in \mathcal{B}_0^d$ . 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  $\pi$ .

**Definition 8.1.** Seien  $\pi, \pi_1, \pi_2, \dots$  Punktprozesse auf  $\mathbb{R}^d$ . Wir sagen, dass  $\pi_n$  gegen  $\pi$  in Verteilung konvergiert (Bezeichnung:  $\pi_n \xrightarrow[n \rightarrow \infty]{d} \pi$ ), falls für alle  $B_1, \dots, B_k \in \mathcal{B}_0^d$  mit  $\pi(\partial B_i) = 0$  fast sicher gilt

$$(\pi_n(B_1), \dots, \pi_n(B_k)) \xrightarrow[n \rightarrow \infty]{d} (\pi(B_1), \dots, \pi(B_k)).$$

Mit anderen Worten gilt für alle  $m_1, \dots, m_k \in \mathbb{N}_0$ :

$$\lim_{n \rightarrow \infty} \mathbb{P}[\pi_n(B_1) = m_1, \dots, \pi_n(B_k) = m_k] = \mathbb{P}[\pi(B_1) = m_1, \dots, \pi(B_k) = m_k].$$

Warum in obiger Definition  $\pi(\partial B_i) = 0$  fast sicher gelten muss, soll folgendes Beispiel veranschaulichen.

**Beispiel 8.2.** Sei  $\pi_n = \delta_{1/n}$ ,  $\pi = \delta_0$ . Betrachten wir dazu denn Fall mit  $k = 1$ ,  $B = (0, 2)$ , dann ist  $\pi_n(B) = 1$  und  $\pi(B) = 0$ . Hätten wir die Forderung  $\pi(\partial B_i) = 1$  f.s. weggelassen, so würde  $\pi_n$  nicht gegen  $\pi$  konvergieren. Das wäre sehr unnatürlich.

**Definition 8.3.** Eine Funktion  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  hat kompakten Träger, falls es ein  $R > 0$  gibt mit  $f(t) = 0$  für alle  $t \in \mathbb{R}^d$  mit  $|t| > R$ . Die Menge aller stetigen Funktionen auf  $\mathbb{R}^d$  mit kompaktem Träger sei mit  $C_c(\mathbb{R}^d)$  bezeichnet. Es sei  $C_c^+(\mathbb{R}^d)$  die Menge aller  $f \in C_c(\mathbb{R}^d)$  mit  $f \geq 0$ .

**Theorem 8.4.** Seien  $\pi, \pi_1, \pi_2, \dots$  Punktprozesse auf  $\mathbb{R}^d$ , dann sind die folgenden Aussagen äquivalent:

1.  $\pi_n \xrightarrow[n \rightarrow \infty]{d} \pi$ .
2. für alle  $f \in C_c^+(\mathbb{R}^d)$  gilt  $\sum_{x \in \pi_n} f(x) \xrightarrow[n \rightarrow \infty]{d} \sum_{x \in \pi} f(x)$ .
3. für alle  $f \in C_c^+(\mathbb{R}^d)$  gilt

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} \right] = \mathbb{E} \left[ \exp \left\{ - \sum_{x \in \pi} f(x) \right\} \right].$$

*Beweis.* Ohne Beweis. □

## 8.2 Konvergenz der Überschreitungszeitpunkte gegen einen Poisson-Punktprozess

**Theorem 8.5.** Für alle  $n \in \mathbb{N}$  seien  $\varepsilon_{ni}, i \in \mathbb{Z}$ , unabhängige und identisch verteilte Zufallsvariablen mit  $\mathbb{P}[\varepsilon_{ni} = 1] = p_n$  bzw.  $\mathbb{P}[\varepsilon_{ni} = 0] = 1 - p_n$ , wobei  $\lim_{n \rightarrow \infty} np_n = \lambda > 0$ . Dann gilt:

$$\sum_{i \in \mathbb{Z}} \varepsilon_{ni} \delta_{\frac{i}{n}} \xrightarrow[n \rightarrow \infty]{d} PPP(\lambda dt).$$

*Beweis.* Für  $f \in C_c^+(\mathbb{R})$  definieren wir

$$\phi_n(f) = \mathbb{E} \left[ \exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} \right] \text{ bzw. } \phi(f) = \mathbb{E} \left[ \exp \left\{ - \sum_{x \in \pi} f(x) \right\} \right].$$

Wir zeigen, dass  $\lim_{n \rightarrow \infty} \phi_n(f) = \phi(f)$ . Wegen der Unabhängigkeit von  $\varepsilon_{ni}$  gilt:

$$\phi_n(f) = \mathbb{E} \left[ \exp \left\{ - \sum_{i \in \mathbb{Z}} \varepsilon_{ni} f \left( \frac{i}{n} \right) \right\} \right] = \mathbb{E} \prod_{i \in \mathbb{Z}} e^{-\varepsilon_{ni} f(\frac{i}{n})} = \prod_{i \in \mathbb{Z}} \mathbb{E} e^{-\varepsilon_{ni} f(\frac{i}{n})} = \prod_{i \in \mathbb{Z}} (1 + p_n (e^{-f(\frac{i}{n})} - 1)).$$

Mit der Ungleichung  $|\log(1+x) - x| \leq x^2/2$  ergibt sich

$$\log \phi_n(f) = \sum_{i \in \mathbb{Z}} \log(1 + p_n (e^{-f(\frac{i}{n})} - 1)) = \sum_{i \in \mathbb{Z}} p_n (e^{-f(\frac{i}{n})} - 1) + R_n,$$

wobei wir später zeigen werden, dass  $\lim_{n \rightarrow \infty} R_n = 0$ . Es folgt, dass

$$\lim_{n \rightarrow \infty} \log \phi_n(f) = \lim_{n \rightarrow \infty} (np_n) \cdot \frac{1}{n} \sum_{i \in \mathbb{Z}} (e^{-f(\frac{i}{n})} - 1) = \lambda \int_{\mathbb{R}} (e^{-f(t)} - 1) dt,$$

wobei letzteres einfach die Definition des Riemann-Integrals ist. Zusammenfassend gilt also

$$\lim_{n \rightarrow \infty} \phi_n(f) = \exp \left\{ -\lambda \int_{\mathbb{R}} (1 - e^{-f(t)}) dt \right\},$$

was der Laplace-Transformierten eines Poisson-Punktprozesses mit Intensität  $\lambda$  entspricht. Mit Satz 8.4 folgt

$$\pi_n := \sum_{i \in \mathbb{Z}} \varepsilon_{ni} \delta_{\frac{i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\lambda dt).$$

Der Term  $R_n$  kann wie folgt abgeschätzt werden. Die Funktion  $f$  hat einen kompakten Träger, also verschwindet sie außerhalb eines Intervalls  $[-A, A]$ . Es folgt, dass höchstens  $2An$  Werte von  $f(i/n)$  ungleich 0 sind. Somit gilt

$$|R_n| \leq \sum_{i \in \mathbb{Z}} p_n^2 (e^{-f(\frac{i}{n})} - 1)^2 \leq 2AnMp_n^2 = o(1),$$

wobei  $M$  das Supremum von  $f$  ist. □

**Theorem 8.6.** Seien  $X_i$ ,  $i \in \mathbb{Z}$ , u.i.v. Zufallsvariablen mit Verteilungsfunktion  $F$ . Sei  $u_n$  eine Folge mit  $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \lambda > 0$ . Dann gilt

$$\sum_{i \in \mathbb{Z}: X_i > u_n} \delta_{\frac{i}{n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\lambda).$$

*Beweis.* Sei  $\varepsilon_{ni} := 1_{X_i > u_n}$ . Dann ist

$$\lim_{n \rightarrow \infty} n\mathbb{P}[\varepsilon_{ni} = 1] = \lim_{n \rightarrow \infty} n\bar{F}(u_n) = \lambda.$$

Damit sind die Bedingungen von Satz 8.5 erfüllt und es folgt die Behauptung. □

### 8.3 Konvergenz der oberen Ordnungsstatistiken gegen einen Poisson-Punktprozess

**Theorem 8.7.** Für alle  $n \in \mathbb{N}$  seien  $Y_{n1}, \dots, Y_{nn}$  u.i.v. Zufallsvektoren in  $\mathbb{R}^d$  und sei  $\mu_n$  ein Wahrscheinlichkeitsmaß mit  $\mu_n(B) = \mathbb{P}[Y_{n1} \in B]$  für alle  $B \in \mathcal{B}^d$ . Sei  $\mu$  ein Radon-Maß auf  $\mathbb{R}^d$  mit  $\lim_{n \rightarrow \infty} n\mu_n(B) = \mu(B)$  für alle  $B \in \mathcal{B}_0^d$  mit  $\mu(\partial B) = 0$ . Dann gilt

$$\pi_n := \sum_{i=1}^n \delta_{Y_{ni}} \xrightarrow[n \rightarrow \infty]{d} PPP(\mu) =: \pi.$$

*Beweis.* Sei  $f \in C_c^+(\mathbb{R}^d)$ ,

$$\phi_n(f) := \mathbb{E} \exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} = \mathbb{E} \exp \left\{ - \sum_{i=1}^n f(Y_{ni}) \right\} = \mathbb{E} \prod_{i=1}^n e^{-f(Y_{ni})}.$$

Wegen der Unabhängigkeit der  $Y_{ni}$  kann man diesen Ausdruck wie folgt schreiben:

$$\prod_{i=1}^n \mathbb{E} e^{-f(Y_{ni})} = \prod_{i=1}^n \int_{\mathbb{R}^d} e^{-f(t)} d\mu_n(t) = \prod_{i=1}^n \left( 1 - \frac{\int_{\mathbb{R}^d} (1 - e^{-f(t)}) n d\mu_n(t)}{n} \right).$$

Wegen  $n\mu_n(t) \rightarrow \mu(t)$  für  $n \rightarrow \infty$  folgt schließlich:

$$\left( 1 - \frac{\int_{\mathbb{R}^d} (1 - e^{-f(t)}) n d\mu_n(t)}{n} \right)^n \rightarrow \exp \left\{ - \int_{\mathbb{R}^d} (1 - e^{-f(t)}) d\mu(t) \right\},$$

was die Laplace-Transformierte eines Poisson-Punktprozesses mit Intensität  $\mu$  ist. Mit Satz 8.4 folgt  $\pi_n \xrightarrow[n \rightarrow \infty]{d} PPP(\mu)$ .  $\square$

**Theorem 8.8.** Seien  $X_1, \dots, X_n$  u.i.v. Zufallsvariablen aus dem Max-Anziehungsbereich der Gumbel-Verteilung  $e^{-e^{-x}}$ , d.h. es gibt Folgen  $a_n, b_n$  mit

$$\frac{\max\{X_1, \dots, X_n\} - a_n}{b_n} \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}. \quad (8.1)$$

Dann gilt:

$$\pi_n := \sum_{i=1}^n \delta_{\frac{X_i - a_n}{b_n}} \xrightarrow[n \rightarrow \infty]{d} PPP(e^{-t} dt) =: \pi.$$

**Bemerkung 8.9.** Es gilt:

$$\pi(\mathbb{R}^+) \sim \text{Poi} \left( \int_0^\infty e^{-t} dt \right) = \text{Poi}(1) < \infty \text{ fast sicher}$$

und

$$\pi(\mathbb{R}^-) \sim \text{Poi} \left( \int_{-\infty}^0 e^{-t} dt \right) = \text{Poi}(\infty) = \infty \text{ fast sicher.}$$

*Beweis.* Es gilt: (8.1) ist äquivalent zu  $\lim_{n \rightarrow \infty} n \mathbb{P}[X_1 > a_n + tb_n] = e^{-t}$ . 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  $B \in \mathcal{B}^1$ . Dann gilt für Mengen  $B$  der Form  $[t, \infty)$ :

$$n \cdot \mu_n([t, \infty)) = n \cdot \mathbb{P}[Y_{n1} \geq t] = n \cdot \mathbb{P} \left[ \frac{X_1 - a_n}{b_n} \geq t \right] \rightarrow e^{-t}, n \rightarrow \infty,$$

da die Zufallsvariablen im Max-Anziehungsbereich der Gumbelverteilung liegen. Betrachten wir als nächstes Mengen  $B$  der Form  $[t_1, t_2)$ :

$$n \cdot \mu_n([t_1, t_2)) = n \cdot \mu_n([t_1, \infty)) - n \cdot \mu_n([t_2, \infty)) \rightarrow e^{-t_1} - e^{-t_2}, n \rightarrow \infty,$$

aus dem selben Grund wie zuvor. Es folgt:

$$e^{-t_1} - e^{-t_2} = \int_{t_1}^{t_2} e^{-t} dt = \mu([t_1, t_2)).$$

Dieser Beweislogik folgend lässt sich die Konvergenz auch für beliebige Mengen  $B \in \mathcal{B}_0^1$  zeigen, worauf wir hier verzichten wollen. Es folgt:

$$n\mu_n(B) \xrightarrow[n \rightarrow \infty]{} \mu(B) \text{ für alle } B \in \mathcal{B}_0^1.$$

Mit Satz 8.7 folgt die Behauptung. □

**Beispiel 8.10.** Seien  $X_i$  unabhängige und identisch exponentialverteilte Zufallsvariablen mit Parameter 1, d.h.  $\mathbb{P}[X_i > t] = e^{-t}$  für  $t > 0$ . Dann gilt nach Satz 1.7

$$\max\{X_1, \dots, X_n\} - \log n \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}.$$

Mit Satz 8.8 folgt nun

$$\sum_{i=1}^n \delta_{X_i - \log n} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(e^{-t} dt).$$

**Beispiel 8.11.** Seien  $X_i$  unabhängige und identisch standardnormalverteilte Zufallsva-

riablen. Dann haben wir bereits gezeigt, dass gilt:

$$\sqrt{2 \log n}(\max\{X_1, \dots, X_n\} - a_n) \xrightarrow[n \rightarrow \infty]{d} e^{-e^{-x}}$$

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 8.8:

$$\sum_{i=1}^n \delta_{(X_i - a_n)\sqrt{2 \log n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(e^{-t} dt).$$

Im Kapitel Ordnungsstatistiken haben für selbige folgende Notation eingeführt, die wir nachfolgend wieder verwenden wollen:

$$M_n^{(k)} = X_{n-k+1:n}, \quad k = 1, \dots, n.$$

**Korollar 8.12.** *Seien  $e_1, e_2, \dots$  unabhängige und identisch verteilte Zufallsvariablen mit  $\mathbb{P}[e_i > t] = e^{-t}$  für  $t > 0$ . Definiere  $P_k = e_1 + \dots + e_k$ , dann gilt für alle  $k \in \mathbb{N}$*

$$\left( \frac{M_n^{(1)} - a_n}{b_n}, \dots, \frac{M_n^{(k)} - a_n}{b_n} \right) \xrightarrow[n \rightarrow \infty]{d} (-\log P_1, \dots, -\log P_k).$$

Für den Beweis des Satzes wird folgende Eigenschaft des Poisson-Punktprozess als bekannt vorausgesetzt:

$$\sum_{i=1}^{\infty} \delta_{P_i} \sim \text{PPP}(1) \text{ auf } (0, \infty).$$

*Beweis.* Es ist  $\pi \sim \text{PPP}(\mu) =: \sum_{i=1}^{\infty} \delta_{U_i}$  mit  $\mu = e^{-t} dt$  und  $U_1 > U_2 > \dots$ . Aus Satz 8.8 folgt:

$$\left( \frac{M_n^{(1)} - a_n}{b_n}, \dots, \frac{M_n^{(n)} - a_n}{b_n} \right) \xrightarrow[n \rightarrow \infty]{d} (U_1, \dots, U_k).$$

Es ist also noch zu zeigen, dass gilt:  $(U_1, \dots, U_k) \stackrel{d}{=} (-\log P_1, \dots, -\log P_k)$ , was sich auch äquivalent wie folgt formulieren lässt:

$$T(P_1), T(P_2) \dots \sim \text{PPP}(e^{-t} dt), \text{ mit } T(t) = -\log(t).$$

Sei dazu  $\nu$  definiert als das Lebesgue-Maß auf  $(0, \infty)$ , dann ist

$$P_1, P_2, \dots \sim \text{PPP}(\nu).$$

Mit dem Transformationssatz für Poisson-Punktprozesse folgt:

$$TP_1, TP_2, \dots \sim \text{PPP}(T\nu).$$

Um den Beweis abzuschließen, wollen wir nun noch nachweisen, dass  $T\nu$  mit  $\mu$  übereinstimmt:

$$T\nu((a, b)) = \nu T^{-1}(a, b) = \nu((e^{-b}, e^{-a})) = e^{-a} - e^{-b} = \mu((a, b)).$$

Es gilt also  $T\nu = \mu$  und deshalb folgt die Behauptung.  $\square$

**Theorem 8.13.** Seien  $X_1, X_2, \dots$  u.i.v. Zufallsvariablen aus dem Max-Anziehungsbereich der Fréchet-Verteilung  $\Phi_\alpha$ , d.h.

$$\frac{\max\{X_1, \dots, X_n\}}{b_n} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha(t) = e^{-t^{-\alpha}}, \quad t > 0. \quad (8.2)$$

Dann gilt:

$$\sum_{i=1}^n \delta_{\frac{X_i}{b_n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}\left(\frac{\alpha}{t^{\alpha+1}} dt\right) \text{ auf } (0, \infty).$$

*Beweis.* Bedingung (8.2) ist äquivalent zu  $\lim_{n \rightarrow \infty} n\mathbb{P}[X_1 \geq b_n t] = \frac{1}{t^\alpha}$  für alle  $t > 0$ . Wir definieren  $Y_{ni} = \frac{X_i}{b_n} 1_{X_i > 0}$ . Dann ist das Maß  $\mu_n$  auf  $(0, \infty)$  die Verteilung von  $Y_{ni}$ . Betrachten wir zuerst Mengen der Form  $[t, \infty)$ :

$$n\mu_n([t, \infty)) = n\mathbb{P}[Y_{ni} \geq t] = n\mathbb{P}[X_1 \geq tb_n] \rightarrow \frac{1}{t^\alpha}, \quad n \rightarrow \infty,$$

weil die Zufallsvariablen im Max-Anziehungsbereich der Fréchet-Verteilung sind. Betrachten wir nun Mengen der Form  $[t_1, t_2)$  mit  $t_1, t_2 > 0$ :

$$n\mu_n([t_1, t_2)) = n\mu_n([t_1, \infty)) - n\mu_n([t_2, \infty)) \rightarrow \frac{1}{t_1^\alpha} - \frac{1}{t_2^\alpha} = \int_{t_1}^{t_2} \frac{\alpha}{t^{\alpha+1}} dt$$

aus dem selben Grund wie zuvor. Dieser Beweislogik folgend lässt sich die Konvergenz auch für beliebige Mengen  $B \in \mathcal{B}_0^1$  zeigen, worauf wir hier verzichten wollen. Es folgt:

$$\lim_{n \rightarrow \infty} n\mu_n(B) = \int_B \frac{\alpha}{t^{\alpha+1}} dt \text{ für alle } B \in \mathcal{B}_0^1.$$

Damit ist Satz 8.7 anwendbar und die Behauptung ist bewiesen.  $\square$

**Beispiel 8.14.** Seien  $X_i$  u.i.v. Pareto-verteilt mit Parameter  $\alpha$ , d.h.  $\mathbb{P}[X_i > t] = \frac{1}{t^\alpha}$  für alle  $t > 1$ . Wir haben in Satz 1.10 bereits nachgewiesen, dass Pareto-verteilte Zufallsva-

riablen im Max-Anziehungsbereich der Fréchet-Verteilung liegen, d.h.

$$\frac{\max\{X_1, \dots, X_n\}}{n^{1/\alpha}} \xrightarrow[n \rightarrow \infty]{d} \Phi_\alpha.$$

Mit Satz 8.13 gilt dann:

$$\sum_{i=1}^n \delta_{\frac{X_i}{n^{1/\alpha}}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP} \left( \frac{\alpha}{t^{\alpha+1}} dt \right).$$

**Theorem 8.15.** Seien  $X_1, X_2, \dots$  u.i.v. aus dem Max-Anziehungsbereich der Weibull-Verteilung, d.h. der rechte Endpunkt  $x^*$  sei endlich und

$$\frac{\max\{X_1, \dots, X_n\} - x^*}{b_n} \rightarrow \Psi_\alpha(t) = e^{-(-t)^\alpha}, \quad t < 0.$$

Dann gilt:

$$\sum_{i=1}^n \delta_{\frac{X_i - x^*}{b_n}} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(\alpha(-t)^{\alpha-1} dt) \text{ auf } (-\infty, 0).$$

*Beweis.* Ohne Beweis. □

**Beispiel 8.16.** Seien  $X_i$  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 8.15

$$\sum_{i=1}^n \delta_{(X_i - 1)n} \xrightarrow[n \rightarrow \infty]{d} \text{PPP}(dt) \text{ auf } (-\infty, 0).$$

## 8.4 Konvergenz der Rekordzeiten gegen einen Poisson-Punktprozess

Seien  $X_1, X_2, \dots$  u.i.v. mit stetiger Verteilungsfunktion  $F$ . Wie gehabt benutzen wir die Notation  $M_n = \max\{X_1, \dots, X_n\}$  für das Maximum der ersten  $n$  Zufallsvariablen bzw.  $\xi_j = 1_{X_j > M_{j-1}}$ ,  $j = 1, 2, \dots$  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

$$L(n+1) = \min\{j > L(n) : \xi_j = 1\}.$$

**Theorem 8.17.** *Es gilt*

$$\pi_n = \sum_{i=1}^n \delta_{\frac{L(i)}{n}} \xrightarrow[n \rightarrow \infty]{d} PPP(dt/t) \text{ auf } (0, \infty).$$

*Beweis.* Sei  $f \in C_c^+(0, \infty)$ . Es gilt

$$\phi_n(f) := \mathbb{E} \exp \left\{ - \sum_{x \in \pi_n} f(x) \right\} = \mathbb{E} \exp \left\{ - \sum_{i=1}^{\infty} f \left( \frac{L(i)}{n} \right) \right\} = \mathbb{E} \exp \left\{ - \sum_{j=1}^{\infty} \xi_j f \left( \frac{j}{n} \right) \right\},$$

weil  $\xi_j$  immer dann 0 ist, wenn kein neuer Rekord an Stelle  $j$  vorliegt. Man kann obiges auch wie folgt schreiben und wegen der Unabhängigkeit der  $\xi_j$  nach Satz von Rényi umformen:

$$\phi_n(f) = \mathbb{E} \prod_{j=1}^{\infty} \exp \left\{ - \xi_j f \left( \frac{j}{n} \right) \right\} = \prod_{j=1}^{\infty} \mathbb{E} \exp \left\{ - \xi_j f \left( \frac{j}{n} \right) \right\} = \prod_{j=1}^{\infty} \left( 1 - \frac{1}{j} + \frac{1}{j} e^{-f(\frac{j}{n})} \right),$$

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:

$$\log \phi_n(t) = \sum_{j=1}^{\infty} \log \left( 1 + \frac{1}{j} \left( e^{-f(\frac{j}{n})} - 1 \right) \right) = \sum_{j=1}^{\infty} \frac{1}{j} \left( e^{-f(\frac{j}{n})} - 1 \right) + \text{Rest},$$

dabei nutzen wir die Taylorentwicklung von  $\log(1+x) = x + \text{Rest}$  aus, wobei der Rest hier gegen 0 geht für  $n \rightarrow \infty$ . 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 folgt:

$$\frac{1}{n} \sum_{i=1}^{\infty} \frac{n}{j} \left( e^{-f(\frac{j}{n})} - 1 \right) + \text{Rest} \rightarrow \int_0^{\infty} g(t) dt = \int_0^{\infty} (e^{-f(t)} - 1) \frac{dt}{t}.$$

Die Konvergenz von der Reihe zum Integral ergibt sich hier direkt aus der Definition des Riemann-Integrals. Schließlich ist zu beobachten

$$\lim_{n \rightarrow \infty} \phi_n(t) = \exp \left\{ - \int_0^{\infty} (e^{-f(t)} - 1) \frac{dt}{t} \right\},$$

was die Laplace-Transformierte eines Poisson-Punktprozesses mit Intensität  $\frac{dt}{t}$  ist. Es

folgt die Behauptung. □

## 8.5 Extremwertprozess

Seien  $X_1, X_2, \dots$  unabhängig und identisch verteilte Zufallsvariablen mit Verteilungsfunktion  $F$ . Sei  $M_n = \max\{X_1, \dots, X_n\}$ . Es gilt  $M_1 \leq M_2 \leq M_3 \leq \dots$

**Theorem 8.18.** *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:*

$$\begin{aligned} & \mathbb{P}[M_{t_1} \leq x_1, \dots, M_{t_k} \leq x_k] \\ &= F^{t_k - t_{k-1}}(x_k) F^{t_{k-1} - t_{k-2}}(\min\{x_{k-1}, x_k\}) \cdot \dots \cdot F^{t_{k-1} - t_{k-2}}(\min\{x_1, \dots, x_k\}). \end{aligned}$$

**Bemerkung 8.19.** Es gilt die Markov-Eigenschaft:

$$\mathbb{P}[M_{n+1} \leq u | M_1 = m_1, \dots, M_n = m_n] = \mathbb{P}[M_{n+1} \leq u | M_n = m_n].$$

**Theorem 8.20.** *Sei  $\pi = \sum_{k=1}^{\infty} \delta_{(X_k, Y_k)} \sim PPP(\mu)$  auf  $[0, \infty) \times \mathbb{R}$  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*

$$(M_t)_{t \in \mathbb{N}} \stackrel{d}{=} (Z_t)_{t \in \mathbb{N}}.$$

Der obige Satz zeigt, dass man den Maximumsprozess  $(M_t)_{t \in \mathbb{N}}$  in einen sog. Extremwertprozess  $(Z_t)_{t \geq 0}$  in stetiger Zeit einbetten kann.

*Beweis.* Sei  $t_1 < t_2 < \dots < t_k, t_i \in \mathbb{N}$ , 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  $A_j = [t_{j-1}, t_j] \times (\min\{u_k, u_{k-1}, \dots, u_j\}, \infty)$  ist. Obiges lässt sich, da  $\pi$  ein Poisson-Punktprozess ist, zu Folgendem vereinfachen:

$$e^{-\mu(\cup_{j=1}^k A_j)} = \prod_{j=1}^k e^{-\mu(A_j)} = \prod_{j=1}^k e^{(t_j - t_{j-1}) \log F(\min\{u_k, u_{k-1}, \dots, u_j\})} = \mathbb{P}[M_{t_1} \leq u_1, \dots, M_{t_k} \leq u_k].$$

□

# 9 Verschiedenes

## 9.1 Wiederkehrperiode

Seien  $X_1, X_2, \dots$  unabhängige und identisch verteilte Zufallsvariablen und sei  $u \in \mathbb{R}$  ein Schwellenwert. Es sei  $p := \mathbb{P}[X_i > u] \neq 0$ . Wir definieren

$$T_1 := \min\{i \in \mathbb{N} : X_i > u\} \text{ und } T_{n+1} := \min\{i > T_n : X_i > u\}.$$

Man kann sich  $T_1, T_2, \dots$  als Zeitpunkte denken, zu denen ein bestimmtes Ereignis (eine Überflutung) eintritt. Wir definieren außerdem  $W_n = T_n - T_{n-1}$ , wobei  $T_0 = 0$  gesetzt wird.

**Theorem 9.1.** Für alle  $n \in \mathbb{N}$  gilt  $W_n \sim \text{Geo}(p)$ , d.h.  $\mathbb{P}[W_n = k] = p(1-p)^{k-1}$  für alle  $k = 1, 2, \dots$

**Bemerkung 9.2.** Es folgt  $\mathbb{E}W_n = \frac{1}{p}$ . Die Zahl  $\frac{1}{p}$  heißt die Wiederkehrperiode. Es ist der erwartete Abstand zwischen zwei aufeinanderfolgenden Überflutungen.

*Beweis von Satz 9.1.* Wir beweisen den Satz nur für  $n = 1$ :

$$\mathbb{P}[W_1 = k] = \mathbb{P}[X_1 \leq u, \dots, X_{k-1} \leq u, X_k > u] = (1-p)^{k-1}p.$$

□

**Beispiel 9.3.** Per Definition ist die Wahrscheinlichkeit, dass ein 100-Jahresereignis in einem Jahr eintritt, gleich  $p = 1/100$ . Die Wiederkehrperiode ist somit 100 Jahre. Wie groß ist die Wahrscheinlichkeit, dass ein 100-Jahresereignis in den nächsten 100 Jahren nicht eintritt? Diese Wahrscheinlichkeit ist

$$\left(1 - \frac{1}{100}\right)^{100} \approx \frac{1}{e} \approx 0,3679.$$

Außerdem können wir direkt angeben, wieviele 100-Jahresereignisse etwa in 100 Jahren stattfinden. Diese Zahl ist nämlich  $\text{Bin}(100, \frac{1}{100})$ , also approximativ  $\text{Poi}(1)$ -verteilt.

## 9.2 Gumbel-Überschreitungsmethode

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

$$S_{k,r,n} = S = \sum_{i=1}^r 1_{X_{n+i} > M_n^{(k)}}.$$

Somit ist  $S_{k,r,n}$  die Anzahl der Überschreitungen des Schwellenwerts  $M_n^{(k)}$  im Zeitintervall  $n+1, \dots, n+r$ .

**Theorem 9.4.** Die Verteilung von  $S_{k,r,n}$  ist gegeben durch

$$\mathbb{P}[S = j] = \frac{\binom{r+n-k-j}{n-k} \binom{j+k-1}{k-1}}{\binom{r+n}{n}}, \quad j = 0, \dots, r.$$

*Beweis.* Unter Anwendung der totalen Wahrscheinlichkeit gilt:

$$\mathbb{P}[S = j] = \int_{\mathbb{R}} \mathbb{P}[S = j | M_n^{(k)} = u] dF_{n,k}(u).$$

Dabei ist  $F_{n,k}$  die Verteilungsfunktion von  $M_n^{(k)}$ . Es ist zu beachten, dass

$$S | M_n^{(k)} = u \sim \text{Bin}(r, \bar{F}(u))$$

beziehungsweise

$$dF_{n,k}(u) = \frac{n!}{(k-1)!(n-k)!} F^{n-k}(u) \bar{F}^{k-1}(u) dF(u).$$

Es folgt:

$$\mathbb{P}[S = j] = \int_{\mathbb{R}} \binom{r}{j} \bar{F}^j(u) F^{r-j}(u) \frac{n!}{(k-1)!(n-k)!} F^{n-k}(u) \bar{F}^{k-1}(u) dF(u),$$

was sich unter der Substitution  $F(u) = t$  wie folgt schreiben lässt:

$$\mathbb{P}[S = j] = \int_0^1 \binom{r}{j} \frac{n!}{(k-1)!(n-k)!} t^{r-j+n-k} (1-t)^{j+k-1} dt,$$

was sich schließlich unter Beachtung der bekannten Formel für die Beta-Funktion

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = \frac{(\alpha-1)!(\beta-1)!}{(\alpha+\beta-1)!}$$

zu Folgendem umschreiben lässt:

$$\mathbb{P}[S = j] = \binom{r}{j} \frac{n!}{(k-1)!(n-k)!} \frac{(k+j-1)!(r-j+n-k)!}{(r+n)!}.$$

□

**Beispiel 9.5.** Die maximalen jährlichen Wasserstände an einem Fluss seien seit  $n = 100$  Jahren bekannt und mit  $X_1, \dots, X_n$  bezeichnet. Wie hoch muss ein Deich sein, bei dem die Wahrscheinlichkeit einer Überflutung in den nächsten  $r = 12$  Jahren  $P = 0,3$  sein soll? Wir setzen die gesuchte Deichhöhe als  $M_n^{(k)}$  mit einem unbekanntem  $k$  an. Im obigen Satz wurde gezeigt, dass

$$\mathbb{P}[S_{k,r,n} = 0] = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{(r+n) \cdot (r+n-1) \cdot \dots \cdot (r+n-k+1)} \stackrel{!}{=} 0,7.$$

Als Lösung der obigen Gleichung mit  $n = 100$ ,  $r = 12$  ergibt sich  $k = 3$ . Die Deichhöhe kann als die dritte Ordnungsstatistik  $M_n^{(3)}$  gewählt werden. Der Vorteil dieser Methode besteht darin, dass die (unbekannte) Verteilungsfunktion von  $X_i$  keine Rolle spielt (und nicht geschätzt werden muss). Der Nachteil ist, dass die Methode bei einer kleineren Überflutungswahrscheinlichkeit  $P$  nicht funktioniert.