ulm university universität
uulm

Wahrscheinlichkeitstheorie und stochastische Prozesse

Evgeny Spodarev

Vorlesungsskript

18. Dezember 2024

... Everything happens according to Law; Chance is but a name for Law not recognized.

The Kybalion: a study of the hermetic philosophy of ancient Egypt and Greece. Illinois, 1908.

Vorwort

Das vorliegende Vorlesungskript gibt eine maßtheoretische Einführung in die Problemstellungen der Wahrscheinlichkeitstheorie und elementarer stochastischer Prozesse. In der Vorlesung *Elementare Wahrscheinlichkeitsrechnung und Statistik* wurden Grundlagen für diesen Vorlesungskurs bereits geschaffen. Voraussetzungen für das Verständnis sind außerdem übliche Analysis-Vorlesungen sowie Maßtheorie.

Ich danke zahlreichen Kollegen aus dem Institut für Stochastik für die freundliche Unterstützung dieses Vorhabens. Herrn Tobias Brosch, Linus Lach und Artur Bille verdanke ich die schnelle und verantwortungsvolle Umsetzung meiner Vorlesungsnotizen in LaTeX, die sie insbesondere mit vielen schönen Grafiken versehen haben.

Ulm, den 17. April 2020

Inhaltsverzeichnis

Inhaltsverzeichnis **i**

1 Erwartung als Lebesgue–Integral **2**
- 1.1 Lebesguesches Integral 2
- 1.2 Erwartungswert 6
- 1.3 Alternative Darstellung des Erwartungswertes 13
- 1.4 Bedingte Erwartung 15
- 1.5 Reguläre bedingte Erwartung 23
  - 1.5.1 Reguläre bedingte Wahrscheinlichkeiten und Verteilungen 23
  - 1.5.2 Bedingte Dichten 24
  - 1.5.3 Bedingte (Ko)Varianz 27

2 Analytische Hilfsmittel **29**
- 2.1 Komplexwertige Zufallsvariablen 29
- 2.2 Charakteristische Funktionen 30
- 2.3 Erzeugende Funktionen 41

3 Konvergenz von Folgen von Zufallsvariablen **45**
- 3.1 Fast sichere und stochastische Konvergenz 46
- 3.2 L^r–Konvergenz 49
- 3.3 Konvergenz in Verteilung 50
- 3.4 Konvergenz der Funktionale von Zufallsvariablen 59
- 3.5 Gleichgradige Integrierbarkeit 63
- 3.6 Drei–Reihen–Satz von Kolmogorow 69

4 Grenzwertsätze **73**
- 4.1 Gesetze der großen Zahlen 73
  - 4.1.1 Schwaches Gesetz der großen Zahlen 74
  - 4.1.2 Starkes Gesetz der großen Zahlen 75
  - 4.1.3 Anwendung der Gesetze der großen Zahlen 81
- 4.2 Zentraler Grenzwertsatz 82
  - 4.2.1 Klassischer zentraler Grenzwertsatz 83
  - 4.2.2 Grenzwertsatz von Lindeberg 85

4.2.3 Konvergenzgeschwindigkeit im zentralen Grenzwertsatz 94

5 Theorie zufälliger Funktionen **108**
5.1 Zufällige Funktionen 108
5.2 Elementare Beispiele 116
5.3 Regularitätseigenschaften 118
5.4 Differenzierbarkeit von Pfaden 124
5.5 Momente und Kovarianz 125
5.6 Stationarität und Unabhängigkeit 127
5.7 Prozesse mit unabhängigen Zuwächsen 129

6 Zählprozesse **132**
6.1 Erneuerungsprozesse 132
6.2 Inhomogene Poisson-Prozesse 144

7 Wiener-Prozess **155**
7.1 Elementare Eigenschaften 156
7.2 Explizite Konstruktion des Wiener-Prozess 158
7.3 Verteilungs- und Pfadeigenschaften vom Wiener - Prozess . . 167
7.3.1 Donskers Invarianzprinzip 167
7.3.2 Gesetz der großen Zahlen 170
7.3.3 Invarianzeigenschaften 171

Literatur **177**

Index **179**

<

Kapitel 1

Erwartung als Lebesgue–Integral

Wir geben eine Einführung in die Integration nach Lebesgue, um eine allgemeine maßtheoretische Definition der Erwartung zu geben, die für beliebige Zufallsvariablen (im Gegensatz zur Vorlesung Elementare Wahrscheinlichkeitsrechnung und Statistik, Kapitel 3, in dem Erwartungswerte nur für diskrete bzw. absolut stetige Verteilungen definiert waren) gültig bleibt.

1.1 Lebesguesches Integral

Historisch wird das Lebesgue–Integral auf $(\Omega, \mathcal{F}, P)$ durch Fréchet folgendermaßen eingeführt: für eine Zufallsvariable X auf $(\Omega, \mathcal{F}, P)$ gilt

$$\mathbb{E}X = \int_\Omega X(\omega)P(d\omega) = \lim_{\lambda\to 0} \sum_{k=-\infty}^{\infty} k\lambda \cdot P(k\lambda < X(\omega) \leq (k+1)\lambda)\,. \quad (1.1)$$

In dieser Form wurde es auch von Kolmogorov im Jahre 1933 übernommen. Die Erklärung dieser Definition ist einfach: Eine beliebige Zufallsvariable X wird durch eine Folge von diskreten Zufallsvariablen X_λ mit den folgenden Eigenschaften approximiert:

$$X_\lambda(\omega) = k\lambda \text{ , falls } X(\omega) \in (k\lambda, (k+1)\lambda]\,, \quad k \in \mathbb{Z}\,, \lambda \in \mathbb{R}_+\,.$$

Aus der Formel (1.1) folgt

$$\mathbb{E}X_\lambda = \sum_{k=-\infty}^{\infty} k\lambda P(\underbrace{k\lambda < X(\omega) \leq (k+1)\lambda}_{\{X_\lambda = k\lambda\}}).$$

Da $P(\omega \in \Omega : X_\lambda(\omega) \underset{\lambda\to 0}{\longrightarrow} X(\omega)) = 1$ (vgl. Abb. 1.1), kann man erwarten, dass der Grenzwert $\lim_{\lambda\to 0} \mathbb{E}X_\lambda$ unter bestimmten Voraussetzungen existiert. In diesem Fall spricht man vom Erwartungswert $\mathbb{E}X$ von X:

$$\mathbb{E}X = \lim_{\lambda\to 0} \mathbb{E}X_\lambda$$

.

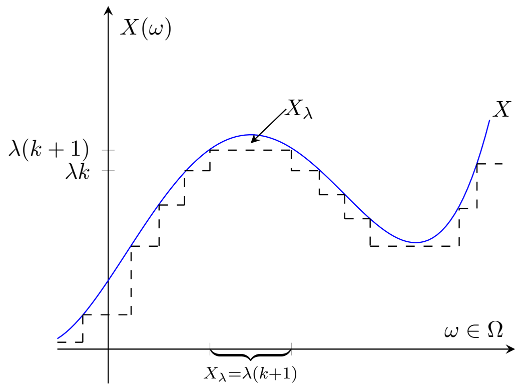


Abbildung 1.1: Definition des Lebesgue–Integrals durch Fréchet

Die Formel (1.1) kann auch folgendermaßen umgeschrieben werden:

$$\int_\Omega X(\omega)P(d\omega) = \lim_{\lambda\to 0} \sum_{k=-\infty}^{\infty} k\lambda\underbrace{[F_X((k+1)\lambda) - F_X(k\lambda)]}_{P(k\lambda<X\leq(k+1)\lambda)},$$

was zur Definition des sogenannten *Lebesgue–Stiltjes–Integrals* $\int_\mathbb{R} xdF_X(x)$ führt. Daher gilt

$$\mathbb{E}X = \int_\Omega X(\omega)P(d\omega) = \int_\mathbb{R} x\, dF_X(x)\,.$$

Wir werden jedoch eine andere Definition des Lebesgue–Integrals benutzen, die in den meisten modernen Büchern über die Wahrscheinlichkeitsrechnung angeführt wird:

Sei $(E, \mathcal{E}, \mu)$ ein messbarer Raum, der aus der Grundmenge E, der σ–Algebra $\mathcal{E}$ der μ–messbaren Teilmengen und des σ-endlichen Maßes μ besteht, wobei ein Maß eine σ–additive nicht–negative Mengenfunktion auf $\mathcal{E}$ ist. Sei $f : E \to \mathbb{R}$ eine $\mathcal{E}$–messbare Abbildung. Wir definieren das Lebesgue–Integral $\int_E f(x)\mu(dx)$ schrittweise:

Definition 1.1.1 Die Abbildung $f : E \to \mathbb{R}_+$ heißt *einfach*, falls

$$f(x) = \sum_{j=1}^{n} x_j I(x \in A_j)\,,$$

wobei $x_j \in \mathbb{R}$, $A_j \in E$ für $j = 1, \dots, n$ und $A_j \cap A_k = \emptyset$, $k \neq j$, $\bigcup_{j=1}^n A_j = E$.

Definition 1.1.2 Das *Lebesgue–Integral* einer einfachen Funktion f bzgl. μ ist

$$\int_E f(x)\mu(dx) = \sum_{j=1}^{n} x_j \mu(A_j) \quad \in \mathbb{R} \cup \{+\infty\}\,.$$

Dieser Wert hängt nicht von der Darstellung von f durch $\sum_{j=1}^{n} x_j I(x \in A_j)$ ab. Das heißt, falls es eine andere messbare Zerlegung $\{B_j\}_{j=1}^{m}$ und Werte $\{y_j\}_{j=1}^{m}$ gibt, sodass

$$f(x) = \sum_{j=1}^{n} x_j I(x \in A_j) = \sum_{j=1}^{m} y_j I(x \in B_j)\,,$$

dann gilt

$$\sum_{i=j}^{n} x_j \mu(A_j) = \sum_{j=1}^{m} y_j \mu(B_j)\,.$$

Übungsaufgabe 1.1.3 Beweisen Sie dies!

Lemma 1.1.4 Sei $f : E \to \mathbb{R}$ eine nicht-negative $\mathcal{E}$–messbare Funktion. Dann existiert eine Folge von einfachen Funktionen $\{f_n\}_{n\in\mathbb{N}}$, $f_n \geq 0,\ n \in \mathbb{N}$, für die gilt

$$f_n(x) \uparrow f(x), \quad x \in E.$$

Beweis Setzen wir $f_n(x) = \sum_{k=1}^{n\cdot 2^n} \frac{k-1}{2^n} I\{\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}\} + nI\{f(x) \geq n\}$ (vgl. Abb. 1.2).

Übungsaufgabe 1.1.5 Zeige, dass $f_n \uparrow f$.

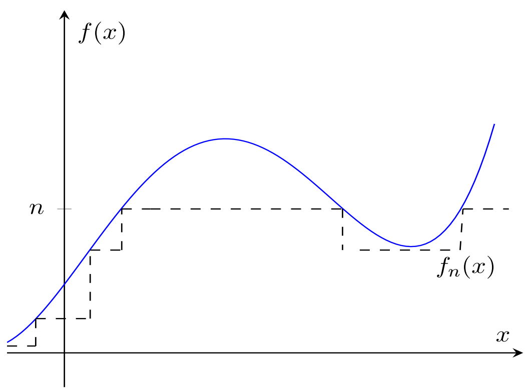


Abbildung 1.2: Approximation durch einfache Funktionen

□

Definition 1.1.6 Sei $f(x) \geq 0$, $x \in E$ eine $\mathcal{E}$–messbare Funktion auf E. Dann heißt

$$\int_E f(x)\mu(dx) = \lim_{n\to\infty} \int_E f_n(x)\mu(dx)$$

das *Lebesgue–Integral* von f. Dieses kann auch unendliche Werte annehmen.

Lemma 1.1.7 Seien $\{f_n\}, g$ nicht–negative einfache Funktionen auf E, $f_n \uparrow f \geq g$. Dann

$$\lim_{n\to\infty} \int_E f_n(x)\mu(dx) \geq \int_E g(x)\mu(dx).$$

Beweis Für alle $\varepsilon > 0$ sei $A_n = \{x \in E \ : \ f_n(x) > g(x) - \varepsilon\}$, dann gilt $A_n \uparrow E$. Insbesondere gilt für alle $x \in E$

$$\begin{aligned} f_n(x) &= f_n(x)I(x \in A_n) + f_n(x)I(x \in A_n^c) \\ &\geq f_n(x)I(x \in A_n) \geq (g(x) - \epsilon)I(x \in A_n). \end{aligned}$$

Aus [20, Satz 4.2.1 3.,4.] folgt

$$\begin{aligned} \int_E f_n(x)\mu(dx) &\geq \int_{A_n} (g(x) - \varepsilon)\mu(dx) = \int_{A_n} g(x)\mu(dx) - \varepsilon\mu(A_n) \\ &= \int_E g(x)\mu(dx) - \int_{A_n^c} g(x)\mu(dx) - \varepsilon\mu(A_n) \\ &\geq \int_E g(x)\mu(dx) - \max_{x\in E} g(x)\mu(A_n^c) - \varepsilon\mu(A_n). \end{aligned}$$

Da $\varepsilon > 0$ beliebig klein sein kann und $\mu(A_n^c) \to 0$, $n \to \infty$, folgt die Aussage. $\square$

Übungsaufgabe 1.1.8 Zeige, dass für zwei unterschiedliche Folgen $f_n \uparrow f$ und $g_n \uparrow f$ gilt

$$\lim_{n\to\infty} \int_E f_n(x)\mu(dx) = \lim_{n\to\infty} \int_E g_n(x)\mu(dx)\,.$$

Das Integral ist also wohldefiniert.

Sei nun $f : E \to \mathbb{R}$ eine beliebige $\mathcal{E}$–messbare Funktion und $x \in E$. Bezeichnen wir mit

$$\begin{aligned} f_+(x) &= \max\{f(x), 0\} && \text{den } \textit{Positivteil von } f \text{ und mit} \\ f_-(x) &= -\min\{f(x), 0\} && \text{den } \textit{Negativteil von } f. \end{aligned} \tag{1.2}$$

Es gilt offensichtlich:

$$\begin{aligned} f_\pm(x) &\geq 0, \\ f(x) &= f_+(x) - f_-(x), \\ |f(x)| &= f_+(x) + f_-(x), \quad x \in E\,. \end{aligned}$$

Definition 1.1.9 Das *Lebesgue–Integral* von f ist durch

$$\int_E f(x)\mu(dx) = \int_E f_+(x)\mu(dx) - \int_E f_-(x)\mu(dx)$$

gegeben, falls $\max\{\int_E f_+(x)\mu(dx), \int_E f_-(x)\mu(dx)\} < \infty$.
Somit gilt $\int_E |f(x)|\mu(dx) < \infty$ und folglich $\int_E f(x)\mu(dx) < \infty$.

Beispiel 1.1.10

1. $E = \mathbb{R}$, $\mu =$ Lebesgue–Maß auf $\mathbb{R}$ (Länge), $\mathcal{E} = \mathcal{B}(\mathbb{R})$. Dann ist

$$\int_E f(x)\,\mu(dx) = \int_{\mathbb{R}} f(x)\,dx$$

das Lebesgue–Integral auf $\mathbb{R}$ im Sinne der mathematischen Analysis.

2. $E = \Omega$ und $\mu = P$ ein Wahrscheinlichkeitsmaß auf Ω, $\mathcal{E} = \mathcal{F}$. Dann heißt

$$\int_\Omega X(\omega)P(d\omega) = \mathbb{E}X$$

Erwartungswert der Zufallsvariablen X. Äquivalent (dies wird später im Satz 1.2.8 bewiesen) dazu kann $\mathbb{E}X$ auch als $\int_{\mathbb{R}} xP_X(dx)$ definiert werden (nach dem Maßtransport–Mechanismus). In diesem Fall ist $E = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, $\mu = P_X$. Da

$$\begin{aligned} P_X(dx) &= P_X((x, x+dx]) = P(x < X \leq x + dx) \\ &= F_X(x+dx) - F_X(x) = dF_X(x)\,, \qquad x \in \mathbb{R}\,, \end{aligned}$$

benutzt man auch die Bezeichnung

$$\mathbb{E}X = \int_{\mathbb{R}} x\,dF_X(x)\,.$$

Es ist das so genannte *Lebesgue–Stiltjes–Integral.*

1.2 Erwartungswert

Somit können wir folgende Definition angeben:

Definition 1.2.1 Sei X eine Zufallsvariable auf $(\Omega, \mathcal{F}, P)$. Sei

$$X(\omega) = X_+(\omega) - X_-(\omega),\ \omega \in \Omega$$

die Zerlegung von X in den Positivteil $X_+(\omega)$ und Negativteil $X_-(\omega)$, $\omega \in \Omega$ (vgl. (1.2)).
Falls

$$\min\left\{\int_\Omega X_+(\omega)P(d\omega), \int_\Omega X_-(\omega)P(d\omega)\right\} < \infty\,,$$

dann wird
$$\int_\Omega X(\omega)P(d\omega)$$
der *Erwartungswert* von X genannt und als $\mathbb{E}X$ bezeichnet. Hier sind die Werte $\mathbb{E}X = \pm\infty$ zulässig. Falls
$$\int_\Omega |X(\omega)|P(d\omega) < \infty$$
(was äquivalent zu $\max\left\{\int_\Omega X_+(\omega)P(d\omega), \int_\Omega X_-(\omega)P(d\omega)\right\} < \infty$ ist), dann heißt die Zufallsvariable *integrierbar*.

Übungsaufgabe 1.2.2 Sei $X \geq 0$ eine Zufallsvariable und sei
$$S_X = \{ \text{ Einfache Zufallsvariablen } Y \ : \ Y \leq X\}.$$
Dann gilt $\mathbb{E}X = \sup\limits_{Y \in S_X} \mathbb{E}Y$.

Die Eigenschaften des so definierten Erwartungswertes fallen mit den Eigenschaften des Lebesgue–Integrals zusammen:

Satz 1.2.3 *(Eigenschaften des Erwartungswertes)*
Seien X und Y Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Dann gilt Folgendes:

1. Falls $X(\omega) = I_A(\omega)$ für ein $A \in \mathcal{F}$, dann ist $\mathbb{E}X = P(A)$.
2. Falls $X(\omega) \geq 0$ für fast alle $\omega \in \Omega$ (man sagt dazu *“fast sicher”* und schreibt *“f.s.”*), dann ist $\mathbb{E}X \geq 0$.
3. *Additivität:* Für beliebige $a, b \in \mathbb{R}$ gilt $\mathbb{E}(aX + bY) = a\mathbb{E}X + b\mathbb{E}Y$, falls X und Y integrierbar sind.
4. *Monotonie:* Falls X und Y integrierbar sind und $X \geq Y$ f.s. ist, dann gilt $\mathbb{E}X \geq \mathbb{E}Y$.
   Falls $0 \leq X \leq Y$ f.s. und Y integrierbar ist, dann ist auch X integrierbar.
5. Für eine integrierbare Zufallsvariable X gilt $|\mathbb{E}X| \leq \mathbb{E}|X|$.
6. Falls X f.s. auf Ω beschränkt ist, dann ist X integrierbar.
7. Falls X und Y unabhängig und integrierbar sind sowie $\mathbb{E}|XY| < \infty$ ist, dann gilt $\mathbb{E}(XY) = \mathbb{E}X \cdot \mathbb{E}Y$.
8. Falls $X \geq 0$ f.s. und $\mathbb{E}X = 0$, dann gilt $X = 0$ f.s.

Beweis

1. Die Behauptung folgt aus der Darstellung $X(\omega) = I_A(\omega) = 1 \cdot I_A(\omega) + 0 \cdot I_{\bar{A}}(\omega)$ und der Definition 1.1.2 des Lebesgue–Integrals für einfache Funktionen.

2. Die Behauptung folgt aus Definition 1.1.6 bis 1.2.1:
$X(\omega) = X_+(\omega) - X_-(\omega)$, wobei $X_-(\omega) = 0$ f.s. Daher gilt

$$\mathbb{E}X = \int_\Omega X_+(\omega)P(d\omega) - \int_\Omega X_-(\omega)P(d\omega) = \lim_{n\to\infty} \int_\Omega X_+^n(\omega)P(d\omega) - 0 \geq 0\,,$$

wobei X_+^n einfache Zufallsvariablen mit den Eigenschaften

$$X_+^n(\omega) \geq 0\,, \omega \in \Omega\,, X_+^n \to X_+$$

sind. Dabei haben wir benutzt, dass aus $X(\omega) = 0$ f.s. $\mathbb{E}X = 0$ folgt.

3. Zunächst soll die Gültigkeit von 3. für einfache Zufallsvariablen X und Y bewiesen werden, dann nach Definition 1.1.6 für alle $X, Y \geq 0$ und schließlich nach Definition 1.1.9 für beliebige X, Y. Wir zeigen es lediglich für einfache X und Y. Der Rest ist eine Übungsaufgabe. Sei

$$X(\omega) = \sum_{j=1}^n x_j I_{A_j}(\omega), \quad Y(\omega) = \sum_{j=1}^m y_j I_{B_j}(\omega)\,.$$

Damit folgt

$$aX + bY = a\sum_{j,k} x_j I_{A_j \cap B_k} + b\sum_{j,k} y_k I_{A_j \cap B_k} = \sum_{j,k}(ax_j + by_k) I_{A_j \cap B_k}$$

mit $\sum_{j,k} = \sum_{j=1}^n \sum_{k=1}^m$ und somit

$$\begin{aligned}\mathbb{E}(aX + bY) &= \sum_{j,k}(ax_j + by_k)P(A_j \cap B_k) \\ &= \sum_{j=1}^n ax_j P(A_j) + \sum_{k=1}^m by_k P(B_k) \\ &= a\sum_{j=1}^n x_j P(A_j) + b\sum_{k=1}^m y_k P(B_k) \\ &= a\mathbb{E}X + b\mathbb{E}Y\,, \quad a, b \in \mathbb{R}\,.\end{aligned}$$

4. Dieser Punkt folgt aus 2. für die Zufallsvariable $X - Y \geq 0$ f.s. Falls $0 \leq X \leq Y$ und Y integrierbar, dann gilt $0 \leq \mathbb{E}X \leq \mathbb{E}Y < \infty$ und somit ist auch X integrierbar.

5. Aussage 5. folgt aus 4., weil $-|X| \leq X \leq |X|$ ist und somit

$$-\mathbb{E}|X| \leq \mathbb{E}X \leq \mathbb{E}|X| < \infty\,.$$

6. Die Behauptung folgt aus 4., da eine Konstante $c > 0$ existiert, sodass $|X| \leq c$ f.s. und somit $\mathbb{E}|X| \leq \mathbb{E}c = c < \infty$.

7. Zeigen wir es für einfache Zufallsvariablen X und Y. Der allgemeine Fall wird später behandelt. Falls $X = \sum_i x_i I_{A_i}, \quad Y = \sum_j y_j I_{B_j}$, dann gilt

$$\begin{aligned}
\mathbb{E}(XY) &= \mathbb{E}(\sum_j x_j I_{A_j})(\sum_k y_k I_{B_k}) \\
&= \mathbb{E}(\sum_{i,j} x_j y_k I_{A_j \cap B_k}) = \sum_{i,j} x_j y_k P(A_j \cap B_k) \\
&\underset{\text{X,Y unabh.}}{=} \sum_{i,j} x_j y_k P(A_j) P(B_k) \\
&= \sum_j x_j P(A_j) \cdot \sum_k y_k P(B_k) = \mathbb{E}X \cdot \mathbb{E}Y \,,
\end{aligned}$$

wobei wir benutzt haben, dass

$$\begin{aligned}
P(A_j \cap B_k) &= P(\{X = x_j\} \cap \{Y = y_k\}) = P(X = x_j) \cdot P(Y = y_k) \\
&= P(A_j) \cdot P(B_k) \,, \qquad \text{für alle } j, k \,.
\end{aligned}$$

8. Sei $X \geq 0$ f.s., $\mathbb{E}X = 0$. Nehmen wir an, dass $X \neq 0$ f.s., das heißt $P(X > 0) > 0$. Es gilt $\{X > 0\} = \bigcup_{n=1}^{\infty} \{X > \varepsilon_n\}$, wobei $\varepsilon_n \downarrow 0$, $\varepsilon_n > 0$ für alle n. Die Folge von Ereignissen $\{X > \varepsilon_n\}$ ist monoton wachsend, weil $\varepsilon_n \geq \varepsilon_{n+1}, n \in \mathbb{N}$. Daraus folgt, dass

$$0 < P(X > 0) = \lim_{n \to \infty} P(X > \varepsilon_n)$$

wegen der Stetigkeit von Wahrscheinlichkeitsmaßen und somit existiert ein $n \in \mathbb{N}$ mit $P(X > \varepsilon_n) > 0$. Da $\varepsilon_n > 0$, erhält man

$$\begin{aligned}
\mathbb{E}X \geq \mathbb{E}\left(X I_{\{X > \varepsilon_n\}}\right) &\geq \mathbb{E}\left(\varepsilon_n I_{\{X > \varepsilon_n\}}\right) \\
&= \varepsilon_n \mathbb{E} I_{\{X > \varepsilon_n\}} = \varepsilon_n P(X > \varepsilon_n) > 0 \,.
\end{aligned}$$

Damit ist $\mathbb{E}X > 0$, was im Widerspruch zu $\mathbb{E}X = 0$ steht.

□

Folgende Konvergenzsätze werden üblicherweise in der Integrationstheorie für allgemeine Integrationsräume $(E, \mathcal{E}, \mu)$ bewiesen (vgl. z.B. [19], S. 186-187):

Satz 1.2.4 *(Monotone Konvergenz)*
Seien $\{X_n\}_{n \in \mathbb{N}}$, X, Y Zufallsvariablen auf $(\Omega, \mathcal{F}, P)$.

1. Falls $X_n \geq Y$ f.s., $n \in \mathbb{N}$, $\mathbb{E}Y > -\infty$ und $X_n \uparrow X$, dann gilt

$$\mathbb{E}X_n \uparrow \mathbb{E}X, \quad n \to \infty \,.$$

2. Falls $X_n \leq Y$ f.s., $n \in \mathbb{N}$, $\mathbb{E}Y < \infty$ und $X_n \downarrow X$, dann gilt
$$\mathbb{E}X_n \downarrow \mathbb{E}X, \quad n \to \infty\,.$$

Beweis

1. Sei zunächst $Y \geq 0$. Für jedes $k \in \mathbb{N}$ existiert eine Folge von einfachen Zufallsvariablen $X_k^{(n)}$: $X_k^{(n)} \uparrow X_k$, $n \to \infty$. Für $X^{(n)} := \max_{1\leq k\leq n} X_k^{(n)}$ gilt
$$X^{(n-1)} \leq X^{(n)} \leq \max_{1\leq k\leq n} X_k = X_n.$$
Sei dann $Z = \lim_{n\to\infty} X^{(n)}$. Da $X_k^{(n)} \leq X^{(n)} \leq X_n$, bekommt man $X_k \leq Z \leq X$, $k \in \mathbb{N}$ durch Grenzwertbildung für $n \to \infty$. Da $\lim_{k\to\infty} X_k = X$, folgt daraus $Z = X$. Da Zufallsvariablen $X^{(n)}$ einfach sind und $Z = \lim_{n\to\infty} X^{(n)}$, mit Hilfe von Lemma 1.1.7 und des Satzes 4.1.2, 4. (ElemWR) bekommen wir
$$\mathbb{E}X = \mathbb{E}Z \leq \lim_{n\to\infty} \mathbb{E}X^{(n)} \leq \lim_{n\to\infty} \mathbb{E}X_n.$$
Andererseits, $X_n \leq X_{n+1} \leq X$ und somit $\lim_{n\to\infty} \mathbb{E}X_n \leq \mathbb{E}X$, was $\lim_{n\to\infty} \mathbb{E}X_n = \mathbb{E}X$ bedeutet.
Sei Nun Y beliebig. Falls $\mathbb{E}Y = \infty$, so gilt nach Satz 4.1.2, 4. (ElemWR) $\mathbb{E}X_n = \mathbb{E}X = \infty$. Insbesondere ist die Aussage des Satzes dann trivialerweise gültig. Falls also $\mathbb{E}|Y| < \infty$, dann $0 \leq X_n - Y \uparrow X - Y$ f.s. Wie im Fall $Y \geq 0$ gilt, dass $\mathbb{E}(X_n - Y) \uparrow \mathbb{E}(X - Y)$, was nach dem Satz 4.1.2, 3. (ElemWR) bedeutet, dass $\mathbb{E}X_n - \mathbb{E}Y \uparrow \mathbb{E}X - \mathbb{E}Y$. Unter der Berücksichtigung von $\mathbb{E}|Y| < \infty$ gilt $\mathbb{E}X_n \uparrow \mathbb{E}X$.

2. Setze $\tilde{X}_n = -X_n$, $\tilde{X} = -X$ und $\tilde{Y} = -Y$, dann folgt die Aussage aus 1.

□

Folgerung 1.2.5 Falls $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von nicht–negativen Zufallsvariablen ist, dann gilt
$$\mathrm{E}\sum_{n=1}^{\infty} X_n = \sum_{n=1}^{\infty} \mathrm{E}X_n\,,$$
wobei diese Reihen auch gegen $+\infty$ divergieren können.

Beweis Die Aussage folgt aus Theoremen 1.2.4 und 1.2.3, 3 mit
$$Y_n = \sum_{i=1}^{n} X_i \uparrow Y = \sum_{i=1}^{\infty} X_i\,.$$

□

Lemma 1.2.6 *(Fatou)*
Seien $\{X_n\}$ und Y Zufallsvariablen. Falls

1. $X_n \geq Y$ f.s., $n \in \mathbb{N}$ und $\mathbb{E}Y > -\infty$,
   dann $\mathbb{E} \liminf_{n\to\infty} X_n \leq \liminf_{n\to\infty} \mathbb{E}X_n$.
2. $X_n \leq Y$ f.s., $n \in \mathbb{N}$ und $\mathbb{E}Y < \infty$,
   dann $\limsup_{n\to\infty} \mathbb{E}X_n \leq \mathbb{E} \limsup_{n\to\infty} X_n$.
3. $|X_n| \leq Y$ f.s., $n \in \mathbb{N}$ und $\mathbb{E}Y < \infty$, dann
   $$\mathbb{E} \liminf_{n\to\infty} X_n \leq \liminf_{n\to\infty} \mathbb{E}X_n \leq \limsup_{n\to\infty} \mathbb{E}X_n \leq \mathbb{E} \limsup_{n\to\infty} X_n.$$

Beweis

1. Sei $Z_n = \inf_{m\geq n} X_m$ f.s., damit für alle n
   $$X_n \geq Z_n, \quad Z_n \uparrow \liminf_{n\to\infty} X_n, \quad Z_n \geq Y \quad \text{f.s.}$$
   Nach 1.2.4 1. gilt
   $$\mathbb{E} \liminf_{n\to\infty} X_n = \mathbb{E} \lim_{n\to\infty} Z_n = \lim_{n\to\infty} \mathbb{E}Z_n = \liminf_{n\to\infty} \mathbb{E}Z_n \leq \liminf_{n\to\infty} \mathbb{E}X_n,$$
   damit ist 1. bewiesen.
2. folgt direkt aus 1.
3. folgt aus 1. und 2.

□

Satz 1.2.7 *Satz von Lebesgue über die majorisierte Konvergenz*
Sei $|X_n| \leq Y$ f.s., $n \in \mathbb{N}$. Falls $\mathbb{E}Y < \infty$ und $X_n \underset{n\to\infty}{\longrightarrow} X$ f.s., dann ist X integrierbar und es gilt

$$\mathbb{E}|X_n - X| \underset{n\to\infty}{\longrightarrow} 0 \quad (L^1\text{–Konvergenz})$$
$$\text{und somit} \qquad \mathbb{E}X_n \underset{n\to\infty}{\longrightarrow} \mathbb{E}X\,.$$

Beweis Für die Integrierbarkeit gilt, da
$\liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n = X$ f.s., nach Lemma 1.2.6 3., dass

$$\mathbb{E} \liminf_{n\to\infty} X_n = \liminf_{n\to\infty} \mathbb{E}X_n = \limsup_{n\to\infty} \mathbb{E}X_n = \mathbb{E} \limsup_{n\to\infty} X_n = \mathbb{E}X.$$

Da auch $|X| \leq Y$ f.s., gilt somit $\mathbb{E}|X| \leq \mathbb{E}Y < \infty$. Unter der Verwendung der Dreiecksungleichung und dem obigen Argument folgt die Aussage. □

Satz 1.2.8 *(Maßtransport im Lebesgue–Integral)*
Sei $X = (X_1, \ldots, X_n)$ ein Zufallsvektor auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und sei $g : \mathbb{R}^n \to \mathbb{R}$ eine $\mathcal{B}_{\mathbb{R}^n}$–messbare Funktion. Dann gilt

$$\mathbb{E}g(X) = \int_\Omega g(X(\omega))P(d\omega) = \int_{\mathbb{R}^n} g(x)P_X(dx) \tag{1.3}$$

in dem Sinne, dass wenn eines der beiden Integrale existiert, auch das andere existiert und beide den gleichen Wert annehmen. Dabei benutzen wir die Bezeichnung $dx = dx_1 \ldots dx_n$ für $x = (x_1, \ldots, x_n)$.

Beweis Nehmen wir zunächst an, dass $g(x) = I_A(x)$ ist, $A \in \mathcal{B}_{\mathbb{R}^n}$. Dann gilt

$$\mathbb{E}g(X) = \mathbb{E}I_{\{X\in A\}} = P(X \in A) = P_X(A) = \int_{\mathbb{R}^n} I_A(x)P_X(dx)\,,$$

und (1.3) ist erfüllt.
Per Additivität gilt (1.3) also auch für beliebige einfache Funktionen $g : \mathbb{R}^n \to \mathbb{R}_+$. Aus dem Satz 1.2.4, für monotone Konvergenz des Lebesgue–Integrals folgt die Gültigkeit von (1.3) für beliebige $g \geq 0$. Der allgemeine Fall resultiert aus der Darstellung $g = g_+ - g_-$. □

Folgerung 1.2.9

1. Das Integral $\int_{\mathbb{R}^n} g(x)P_X(dx)$ wird oft als Lebesgue–Stiltjes–Integral $\int_{\mathbb{R}^n} g(x)dF_X(x)$ geschrieben (vgl. S. 3).
   Daher gilt $\mathbb{E}g(X) = \int_{\mathbb{R}^n} g(x)dF_X(x)$.

2. Falls X absolut stetig verteilt ist mit der Dichte f_X, dann gilt

$$\mathbb{E}g(X) = \int_{\mathbb{R}^n} g(x)f_X(x)\,dx\,.$$

3. Falls X diskret verteilt ist mit dem Wertebereich
   $C = \{x_1, x_2, \ldots\} \subset \mathbb{R}^n$, dann gilt

$$\mathbb{E}g(X) = \sum_j g(x_j)P(X = x_j) = \sum_{x\in C} g(x)P(X = x)\,.$$

4. Falls X eine Zufallsvariable ist, dann gelten die Aussagen des Satzes 1.2.8 und der Folgerung 1.2.9, 1) – 3) für $n = 1$. Insbesondere gilt

$$\mathbb{E}X = \int_{\mathbb{R}} xP_X(dx) = \int_{\mathbb{R}} xdF_X(x)$$

   im allgemeinen Fall,

$$\mathbb{E}X = \int_{\mathbb{R}} xf_X(x)\,dx$$

im Fall einer absolut stetig verteilten Zufallsvariablen X mit Dichte f_X und

$$\mathbb{E}X = \sum_{x\in C} xP(X = x)$$

im Fall einer diskret verteilten Zufallsvariablen X mit Zähldichte $P(X = x)$ und abzählbaren Wertebereich $C \subset \mathbb{R}$.

1.3 Alternative Darstellung des Erwartungswertes

Definition 1.3.1 Falls X eine Zufallsvariable mit der Verteilungsfunktion $F_X(x)$ ist, dann heißt $\bar{F}_X(x) = 1 - F_X(x) = P(X > x), \quad x \in \mathbb{R}$, die *Tailfunktion* der Verteilung von X.

Satz 1.3.2 Sei X eine Zufallsvariable mit $\mathbb{E}|X|^n < \infty$ für ein $n \in \mathbb{N}$. Dann gilt

$$\mathbb{E}X^n = n\int_0^{+\infty} x^{n-1}\bar{F}_X(x)\,dx - n\int_{-\infty}^0 x^{n-1}F_X(x)\,dx\,, \tag{1.4}$$

$$\mathbb{E}|X|^n = n\int_0^{+\infty} x^{n-1}\left(\bar{F}_X(x) + F_X(-x)\right)\,dx\,. \tag{1.5}$$

Beweis Beweisen wir die Formel (1.4). Nach Satz 1.2.8 schreibt man

$$\mathbb{E}X^n = \int_{-\infty}^{\infty} x^n P_X(dx) \quad = \quad \int_0^{\infty} x^n P_X(dx) + \int_{-\infty}^0 x^n P_X(dx)\,.$$

Ferner gilt nach dem Satz von Fubini:

$$\begin{aligned}
\int_0^{\infty} x^n P_X(dx) &= \int_0^{\infty}\int_0^{\infty} ny^{n-1}I(y\in(0,x))\,dyP_X(dx) \\
&= n\int_0^{\infty}\int_0^{\infty} y^{n-1}I(0<y<x)P_X(dx)\,dy \\
&= n\int_0^{\infty} y^{n-1}\int_y^{+\infty} P_X(dx)\,dy \\
&= n\int_0^{\infty} y^{n-1}\bar{F}_X(y)\,dy\,,
\end{aligned}$$

$$
\begin{aligned}
\int_{-\infty}^{0} x^n P_X(dx) &= \int_{-\infty}^{0} (-1) \int_{x}^{0} n y^{n-1}\, dy P_X(dx) \\
&= -n \int_{-\infty}^{0} \int_{-\infty}^{0} y^{n-1} I(x < y < 0)\, dy P_X(dx) \\
&= -n \int_{-\infty}^{0} \int_{-\infty}^{0} y^{n-1} I(x < y) P_X(dx)\, dy \\
&= -n \int_{-\infty}^{0} y^{n-1} \cdot \underbrace{\int_{-\infty}^{y} P_X(dx)}_{P(X \le y) = F_X(y)}\, dy \\
&= -n \int_{-\infty}^{0} y^{n-1} F_X(y)\, dy\,.
\end{aligned}
$$

Zusammengefasst haben wir

$$
\mathbb{E}X^n = n \int_0^{+\infty} y^{n-1} \bar{F}_X(y)\, dy - n \int_{-\infty}^{0} y^{n-1} F_X(y)\, dy\,.
$$

Die Formel (1.5) wird analog bewiesen. □

Jetzt ist es an der Zeit die Aussage 7. des Satzes 1.2.3 in folgender allgemeinen Form zu beweisen:

Satz 1.3.3 Seien $X_1, \ldots, X_n$ Zufallsvariablen mit

$$
\mathbb{E}|X_j| < \infty, = 1, \ldots, n\,,\ \mathbb{E}|X_1 \cdot \ldots \cdot X_n| < \infty\,.
$$

Falls $X_1, \ldots, X_n$ unabhängig sind, dann gilt

$$
\mathbb{E}(X_1 \cdot \ldots \cdot X_n) = \mathbb{E}X_1 \cdot \ldots \cdot \mathbb{E}X_n\,.
$$

Beweis Wir benutzen den Transformationssatz 1.2.8 für

$$
X = (X_1, \ldots, X_n),\ \varphi(x_1, \ldots, x_n) = x_1 \cdot \ldots \cdot x_n\,.
$$

$$
\begin{aligned}
\mathbb{E}(X_1 \cdot \ldots \cdot X_n) &= \int_{\mathbb{R}^n} x_1 \cdot \ldots \cdot x_n\, P_X(dx_1 \ldots dx_n) \\
&\underbrace{=}_{\text{[20, Lemma 3.5.2]}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} x_1 \cdot \ldots \cdot x_n\, P_{X_1}(dx_1) \cdot \ldots \cdot P_{X_n}(dx_n) \\
&= \prod_{i=1}^{n} \int_{\mathbb{R}} x_i P_{X_i}(dx_i) = \prod_{i=1}^{n} \mathbb{E}X_i\,.
\end{aligned}
$$

□

1.4 Bedingte Erwartung

Es sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsaum, und $B \in \mathcal{F}$: $P(B) > 0$. Sei

$$P(\cdot|B) = P(\cdot \cap B)/P(B)$$

das *Wahrscheinlichkeitsmaß unter der Bedingung* B (vgl. [20, Übungsaufgabe 2.3.6]). Betrachte die Erwartung bzgl. des Maßes $P(\cdot|B)$: für eine Zufallsvariable X

$$\begin{aligned} \mathrm{E}(X|B) &= \int_\Omega X(\omega) P(d\omega|B) \\ &= \frac{1}{P(B)} \int_B X(\omega) P(d\omega) \\ &= \frac{\mathrm{E}(X I_B)}{P(B)}. \end{aligned}$$

Definition 1.4.1 Die *bedingte Erwartung* einer Zufallsvariablen X unter der Bedingung $B \in \mathcal{F}$, $P(B) > 0$ wird definiert als

$$\mathrm{E}(X|B) = \frac{\mathrm{E}(X I_B)}{P(B)}.$$

Wie definiert man nun $\mathrm{E}(X|B)$, wenn $P(B) = 0$?

Seien X und Y zwei Zufallsvariablen, wobei Y eine absolut stetige Verteilung besitzt. Dann folgt $P(Y = y) = 0 \quad \forall y \in \mathbb{R}$. Deshalb kann die bedingte Wahrscheinlichkeit $P(X \in B \,|\, Y = y)$ auf dem gewöhnlichen Wege

$$P(X \in B|Y = y) = \frac{P(X \in B,\, Y = y)}{P(Y = y)}$$

nicht definiert werden. Aus der Praxis ist aber eine Reihe von Fragestellungen bekannt (z.B. Bayessche Analyse), in denen Wahrscheinlichkeiten $P(X \in B|Y = y)$ ausgewertet werden müssen. Deswegen werden wir eine neue Definition der bedingten Wahrscheinlichkeit geben, die solche Situationen berücksichtigt. Diese Definition erfolgt durch die Definition der bedingten Erwartung.
Schema:

1. Es wird die bedingte Erwartung von der Zufallsvariablen X bzgl. der σ-Algebra $\mathcal{B}$ als Zufallsvariable $\mathbb{E}(X|\mathcal{B})$ eingeführt, wobei $\mathcal{B}$ eine Teil-σ-Algebra von $\mathcal{F}$ und $(\Omega, \mathcal{F}, P)$ der Wahrscheinlichkeitsraum ist.

2. Die bedingte Erwartung von X unter der Bedingung Y wird als $\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma_Y)$ eingeführt, wobei σ_Y die von Y erzeugte σ-Algebra ist.

3. $P(X \in B|Y = y)$ wird als Zufallsvariable $\mathbb{E}(I(X \in B)|Y)$ auf der Menge
$\{\omega \in \Omega : Y(\omega) = y\}$ eingeführt.

Gehen wir nun dieses Schema im Detail durch:

1. Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $\mathcal{B}$ eine Teil-σ-Algebra von $\mathcal{F}$, d.h. $\mathcal{B} \subseteq \mathcal{F}$.

Definition 1.4.2 Der *bedingte Erwartungswert* einer Zufallsvariablen X definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ bezüglich einer σ-Algebra $\mathcal{B} \subseteq \mathcal{F}$ ist in dem Fall $\mathbb{E}|X| < \infty$ als eine $\mathcal{B}$-messbare Zufallsvariable Y definiert, die die Eigenschaft

$$\int_B Y(\omega)\, P(d\omega) = \int_B X(\omega)\, P(d\omega)\,, \qquad \forall\, B \in \mathcal{B}$$

besitzt. Dabei wird die Bezeichnung $Y = \mathbb{E}(X|\mathcal{B})$ verwendet.

Warum existiert diese Zufallsvariable Y?

- Zerlegen wir X in den positiven X_+ und negativen X_- Anteil $X = X_+ - X_-$ und beweisen die Existenz von $\mathbb{E}(X_\pm|\mathcal{B})$. Danach setzen wir $\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(X_+|\mathcal{B}) - \mathbb{E}(X_-|\mathcal{B})$.
- Somit genügt es zu zeigen, dass der Erwartungswert $\mathbb{E}(X|\mathcal{B})$ einer nicht negativen Zufallsvariablen $X \geq 0$ fast sicher existiert.
- Sei $Q(B) = \int_B X(\omega)\, P(d\omega)$. Man kann zeigen, dass $Q(\cdot)$ ein Maß auf $(\Omega, \mathcal{F})$ ist. Dabei folgt aus $P(B) = 0$ die Gleichheit $Q(B) = 0$ für $B \in \mathcal{B}_\mathbb{R}$ (bzw. $B \in \mathcal{B}$). Somit ist Q absolut stetig bzgl. P. Weiter existiert nach dem folgenden Satz eine Dichte $Y(\omega)$, die messbar bzgl. $\mathcal{B}$ ist und für die

$$Q(B) = \int_B Y(\omega)\, P(d\omega) \quad \Longrightarrow \quad Y(\omega) = \mathbb{E}(X|\mathcal{B})$$

gilt.

Um den eben angesprochenen Satz formulieren zu können, wird die Definition von signierten Maßen benötigt.

Definition 1.4.3 Ein *signiertes Maß* λ auf dem Messraum $(\Omega, \mathcal{B})$ ist $\lambda = \lambda_1 - \lambda_2$, wobei λ_1, λ_2 Maße auf $(\Omega, \mathcal{B})$ sind, sodass mindestens eins davon endlich ist.

Satz 1.4.4 (Satz von Radon-Nikodým (1930)) Seien μ ein σ–endliches und λ ein signiertes Maß auf $(\Omega, \mathcal{B})$. Sei λ absolut stetig bzgl. μ. Dann existiert eine Dichte $f : \Omega \to \bar{\mathbb{R}}$ ($\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$) messbar bezüglich $\mathcal{B}$ mit

$$\lambda(A) = \int_A f(\omega)\mu(d\omega), \quad A \in \mathcal{B}.$$

Dabei ist f eindeutig bis auf eine Menge mit μ–Maß 0. Falls λ ein Maß ist, so gilt $f \geq 0$.

Bemerkung 1.4.5 Aus der obigen Beweisskizze wird ersichtlich, dass $Y(\omega) = \mathbb{E}(X|\mathcal{B})$ nur P-fast sicher definiert ist. Somit kann man mehrere Versionen von $Y(\omega)$ angeben, die sich auf einer Menge der Wahrscheinlichkeit 0 unterscheiden.

Satz 1.4.6 (Eigenschaften des bedingten Erwartungswertes) Seien X und Y Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit der Eigenschaft $\mathbb{E}|X| < \infty$, $\mathbb{E}|Y| < \infty$ und $\mathbb{E}|XY| < \infty$ (dies kann noch ein wenig abgeschwächt werden, ist hier allerdings ausreichend). Seien $\mathcal{B}$, $\mathcal{B}_1$ und $\mathcal{B}_2$ Teil-σ-Algebren von $\mathcal{F}$. Es gelten folgende Eigenschaften (im fast sicheren Sinne):

(a) $\mathbb{E}(X|\{\emptyset, \Omega\}) = \mathbb{E}X$, $\mathbb{E}(X|\mathcal{F}) = X$.

(b) Falls $X \leq Y$, dann gilt ebenso $\mathbb{E}(X|\mathcal{B}) \leq \mathbb{E}(Y|\mathcal{B})$.

(c) Es gilt $\mathbb{E}(aX + bY|\mathcal{B}) = a\mathbb{E}(X|\mathcal{B}) + b\mathbb{E}(Y|\mathcal{B})$ für alle $a, b \in \mathbb{R}$.

(d) $\mathbb{E}(c|\mathcal{B}) = c$ für $c =$*const.*

(e) Es gilt $\mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1) = \mathbb{E}(X|\mathcal{B}_1)$ und $\mathbb{E}(\mathbb{E}(X|\mathcal{B}_1)|\mathcal{B}_2) = \mathbb{E}(X|\mathcal{B}_1)$, falls $\mathcal{B}_1 \subseteq \mathcal{B}_2$.

(f) Falls X unabhängig von $\mathcal{B}$ ist (d.h., die σ-Algebren $\sigma_X = X^{-1}(\mathcal{B}_\mathbb{R})$ und $\mathcal{B}$ sind unabhängig), dann gilt $\mathbb{E}(X|\mathcal{B}) = \mathbb{E}X$.

(g) $\mathbb{E}(\mathbb{E}(X|\mathcal{B})) = \mathbb{E}X$.

(h) $\mathbb{E}(XY|\mathcal{B}) = X\mathbb{E}(Y|\mathcal{B})$, falls X $\mathcal{B}$-messbar ist.

Beweis

(a) Folgt direkt aus Definition 1.4.2.

(b) Da aus $X \leq Y$ f.s. $XI_A \leq YI_A$ f.s. folgt, $A \in \mathcal{B}$, bekommt man aus der Monotonie– Eigenschaft des Erwartungswertes (vgl. [20, Satz 4.1.2, 4]) für alle $A \in \mathcal{B}$, dass

$$\begin{aligned}\int_A \mathbb{E}(X|\mathcal{B})\, P(d\omega) &= \int_A X(\omega)P(d\omega) = \mathbb{E}(XI_A) \leq \mathbb{E}(YI_A) \\ &= \int_A Y(\omega)P(d\omega) = \int_A \mathbb{E}(Y|\mathcal{B})\, P(d\omega).\end{aligned}$$

Daher $\mathbb{E}(X|\mathcal{B}) \leq \mathbb{E}(Y|\mathcal{B})$ nach Definition 1.4.2.

(c) Aus der Linearität des Erwartungswertes (vgl. [20, Satz 4.1.2, 3.])

gilt für alle $A \in \mathcal{B}$

$$\begin{aligned}
&\int_A (aX(\omega) + bY(\omega))\ P(d\omega) \\
&= a\int_A X(\omega)P(d\omega) + b\int_A Y(\omega)P(d\omega) \\
&= a\int_A \mathbb{E}(X|\mathcal{B})P(d\omega) + b\int_A \mathbb{E}(Y|\mathcal{B})P(d\omega) \\
&= \int_A (a\mathbb{E}(X|\mathcal{B}) + b\mathbb{E}(Y|\mathcal{B}))\ P(d\omega).
\end{aligned}$$

(d) Da $c = const$ $\mathcal{B}$–messbar ist, folgt das Ergebnis ebenso direkt aus Definition 1.4.2.

(e) Sei $A \in \mathcal{B}_1$ und $B \in \mathcal{B}_2$. Da $\mathcal{B}_1 \subseteq \mathcal{B}_2$, gilt $A \in \mathcal{B}_2$ und

$$\begin{aligned}
&\int_A \mathbb{E}(X|\mathcal{B}_1)\, P(d\omega) = \int_A X(\omega)P(d\omega) \\
&= \int_A \mathbb{E}(X|\mathcal{B}_2)\, P(d\omega) = \int_A \mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1)P(d\omega),
\end{aligned}$$

daher $\mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1) = \mathbb{E}(X|\mathcal{B}_1)$.
Nach Definition 1.4.2 gilt für $\mathbb{E}(\mathbb{E}(X|\mathcal{B}_1)|\mathcal{B}_2)$

$$\int_B \mathbb{E}(X|\mathcal{B}_1)\, P(d\omega) = \int_B \mathbb{E}(\mathbb{E}(X|\mathcal{B}_1)|\mathcal{B}_2)\, P(d\omega).$$

$\mathbb{E}(X|\mathcal{B}_1)$ ist $\mathcal{B}_1$-messbar, und da $\mathcal{B}_1 \subseteq \mathcal{B}_2$, auch $\mathcal{B}_2$-messbar. Also gilt $\mathbb{E}(X|\mathcal{B}_1) = \mathbb{E}(\mathbb{E}(X|\mathcal{B}_1)|\mathcal{B}_2)$.

(f) Da $\mathbb{E}X$ $\mathcal{B}$–messbar ist, sollten wir zeigen, dass für alle $A \in \mathcal{B}$

$$\int_A \mathbb{E}X\, P(d\omega) = \int_A X(\omega)P(d\omega),$$

was gilt, da $\mathbb{E}(XI_A) = \mathbb{E}X \cdot \mathbb{E}I_A = \mathbb{E}X \cdot P(A)$ wegen der Unabhängigkeit von X und I_A.

(g) Folgt direkt aus der Turmeigenschaft (e) mit $\mathcal{B}_1 = \{\emptyset, \Omega\}$, $\mathcal{B}_2 = \mathcal{B}$ und aus (a).

(h) Algebraische Induktion.
- Sei $X = I_B$ für $B \in \mathcal{B}$. Dann gilt für jedes $A \in \mathcal{B}$

$$\begin{aligned}
&\int_A X(\omega)Y(\omega)P(d\omega) = \int_{A\cap B} Y(\omega)P(d\omega) \\
&= \int_{A\cap B} \mathbb{E}(Y|\mathcal{B})P(d\omega) = \int_A I_B\mathbb{E}(Y|\mathcal{B})P(d\omega) \\
&= \int_A X(\omega)\mathbb{E}(Y|\mathcal{B})P(d\omega),
\end{aligned}$$

  daher gilt (h) für Indikatoren.

- Aus der Linearität der Bedingten Erwartung folgt (h) für einfache $\mathcal{B}$ -messbare Zufallsvariablen X.
- Sie nun X eine beliebige $\mathcal{B}$–messbare Zufallsvariable mit $\mathbb{E}|X| < \infty$. Es existiert eine Folge von einfachen $\mathcal{B}$–messbaren Zufallsvariablen X_n mit $|X_n| \leq |X|$, $X_n \to X$ f.s. für $n \to \infty$. Nach dem Satz von Lebesgue (dominierte Konvergenz für bedingte Erwartung) gilt:

  $$\mathbb{E}(XY|\mathcal{B}) = \mathbb{E}(\lim_{n\to\infty} X_n Y|\mathcal{B}) = \lim_{n\to\infty} X_n \mathbb{E}(Y|\mathcal{B}) = X\mathbb{E}(Y|\mathcal{B}),$$

  da $X_n Y \to XY$ für $n \to \infty$ und $|X_n Y| \leq |XY|$ f.s.

□

Bemerkung 1.4.7 Wie bereits im obigen Beweis zu sehen ist, behält die bedingte Erwartung folgende Eigenschaften des Lebesgue–Integrals bei:

- Satz über die monotone Konvergenz
- Lemma von Fatou
- Satz von Lebesgue über die dominierte Konvergenz
- Ungleichungen von Markow, Jensen, Cauchy–Bunjakowski–Schwarz
- Dreiecksungleichung: $|\mathbb{E}(X|\mathcal{B})| \leq \mathbb{E}(|X||\mathcal{B})$ f.s.

Beweise wiederholen im Wesentlichen dieselben Ideen wie in den Originalsätzen. Sie können in vielen Büchern wie z.B. von A. Borovkov, A. Klenke, A. Shiryaev, nachgeschlagen werden.

In [20, Bemerkung 4.3.3, 2)] wurde gezeigt, dass für eine Zufallsvariable X mit $\mathbb{E}X^2 < \infty$ der Erwartungswert $\mathbb{E}X$ die Funktion $\mathbb{E}(X-a)^2$ bzgl. $a \in \mathbb{R}$ minimiert. Eine ähnliche Eigenschaft gilt für bedingte Erwartungen:

Satz 1.4.8 Sei $X \in L^2(\Omega, \mathcal{F}, P)$, und sei $\mathcal{B}$ eine Teil-σ-Algebra von $\mathcal{F}$. Dann ist $\mathbb{E}(X|\mathcal{B})$ die orthogonale Projektion von X auf den Unterraum $L^2(\Omega, \mathcal{B}, P) \subset L^2(\Omega, \mathcal{F}, P)$, Vergleich Abbildung 1.3:

$$\mathbb{E}(X|\mathcal{B}) = \operatorname*{arg\,min}_{Y \in L^2(\Omega,\mathcal{B},P)} \mathbb{E}(X-Y)^2.$$

Beweis Nach Satz 1.4.6, (g) gilt $\mathbb{E}(X-Y)^2 = \mathbb{E}\left(\mathbb{E}((X-Y)^2|\mathcal{B})\right)$. Da Y $\mathcal{B}$–messbar ist, folgt mit 1.4.6 (h)

$$\begin{aligned}
&\mathbb{E}((X-Y)^2|\mathcal{B}) = \mathbb{E}((X-\mathbb{E}(X|\mathcal{B}) + \mathbb{E}(X|\mathcal{B}) - Y)^2|\mathcal{B}) \\
&= \mathbb{E}((X-\mathbb{E}(X|\mathcal{B}))^2|\mathcal{B}) + 2\mathbb{E}((X-\mathbb{E}(X|\mathcal{B}))(\mathbb{E}(X|\mathcal{B}) - Y)|\mathcal{B}) \\
&+ \mathbb{E}((\mathbb{E}(X|\mathcal{B}) - Y)^2|\mathcal{B}) \\
&= \mathbb{E}((X-\mathbb{E}(X|\mathcal{B}))^2|\mathcal{B}) + (\mathbb{E}(X|\mathcal{B}) - Y)^2.
\end{aligned}$$

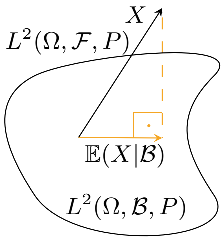


Abbildung 1.3: Bedingter Erwartungswert als Orthogonalprojektion

Nun gilt, dass der Term $(\mathbb{E}(X|\mathcal{B}) - Y)^2$ minimiert wird, wenn $Y = \mathbb{E}(X|\mathcal{B})$. Dabei haben wir genutzt, dass nach 1.4.6 (h)

$$\begin{aligned}&\mathbb{E}((X - \mathbb{E}(X|\mathcal{B}))(\mathbb{E}(X|\mathcal{B}) - Y)|\mathcal{B})\\ &\qquad = (\mathbb{E}(X|\mathcal{B}) - Y)\underbrace{\mathbb{E}(X - \mathbb{E}(X|\mathcal{B})|\mathcal{B})}_{=0} = 0.\end{aligned}$$

□

Beispiel 1.4.9 Sei $\mathcal{B} = \sigma(\{A_1, \ldots, A_n\})$, wobei $\{A_1, \ldots, A_n\}$ eine messbare Zerlegung des Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$ ist, d.h. $\bigcup_{i=1}^n A_j = \Omega$, $A_j \cap A_k = \emptyset$, $i \neq j$, $P(A_j) > 0$, $i = 1, \ldots, n$. Was ist $\mathbb{E}(X|\mathcal{B})$? Da $\mathbb{E}(X|\mathcal{B})$ $\mathcal{B}$-messbar ist, können wir die allgemeine Form der Funktionen ausnutzen, die messbar bezüglich einer endlich erzeugten σ-Algebra $\mathcal{B} = \sigma(\{A_1, \ldots, A_n\})$ sind: $\mathbb{E}(X(\omega)|\mathcal{B}) = \sum_{i=1}^n k_j I(\omega \in A_j)$. Diese Form wird im folgenden Lemma gezeigt.

Lemma 1.4.10 Für eine endlich erzeugte σ-Algebra $\mathcal{B} = \sigma(\{A_1, \ldots, A_n\})$ wie in Beispiel 1.4.9 ist eine $\mathcal{B}$-messbare Zufallsvariable Y einfach, d.h.

$$Y(\omega) = \sum_{j=1}^n k_j I(\omega \in A_j).$$

Beweis Zeige, dass für alle $j = 1, \ldots, n$ die Zufallsvariable Y auf A_j konstant ist, d.h., es existieren $k_j \in \mathbb{R}$ mit $A_j \subseteq \{\omega \in \Omega : Y(\omega) = k_j\}$ für alle $j = 1, \ldots, n$. Setze

$$k_j = \sup\{c \in \mathbb{R} : A_j \cap \{\omega \in \Omega : Y(\omega) < c\} = \emptyset\},$$

dann folgt aus der Definition des Supremums

$$A_j \cap \{\omega \in \Omega : Y(\omega) < k_j\} = A_j \cap \bigcup_{c \in \mathbb{Q}, c < k_j} \{\omega \in \Omega : Y(\omega) < c\} = \emptyset.$$

Für $c > k_j$ gilt nun $A_j \cap \{\omega \in \Omega : Y(\omega) < c\} \neq \emptyset$, und

$$\{\omega \in \Omega : Y(\omega) < c\} \subset \bigcup_{l=1}^{n} A_l.$$

Wegen der $\mathcal{B}$–Messbarkeit von Y und der Tatsache, dass die A_l paarweise disjunkt sind, gilt

$$A_j \cap \{\omega \in \Omega : Y(\omega) < c\} = A_j.$$

Für alle $c > k_j$ folgt daher $A_j \cap \{\omega \in \Omega : Y(\omega) \geq c\} = \emptyset$ und auch

$$A_j \cap \{\omega \in \Omega : Y(\omega) > k_j\} = A_j \cap \bigcup_{c\in\mathbb{Q}, c>k_j} \{\omega \in \Omega : Y(\omega) \geq c\} = \emptyset.$$

Also gilt $A_j \cap \{\omega \in \Omega : Y(\omega) \neq k_j\} = \emptyset$, und $A_j \subseteq \{\omega \in \Omega : Y(\omega) = k_j\}$. □

Berechnen wir k_j: Aus der Definition 1.4.2 folgt für $B = A_j$

$$\begin{aligned}
\int_B \mathbb{E}(X|\mathcal{B})\, P(d\omega) &= \int_{A_j} \sum_{l=1}^{n} k_l \cdot I(\omega \in A_l)\, P(d\omega) = k_j \cdot P(A_j) \\
&= \int_B X\, P(d\omega) = \int_{A_j} X\, P(d\omega) = \mathbb{E}(X \cdot I_{A_j}) \\
\Longrightarrow k_j &= \frac{\mathbb{E}(X \cdot I_{A_j})}{P(A_j)} = \mathbb{E}(X|A_j)\,, \quad j = 1, \ldots, n\,. \\
\Longrightarrow \mathbb{E}(X|\mathcal{B})(\omega) &= \frac{\mathbb{E}(X \cdot I_{A_j})}{P(A_j)}\,, \quad \text{falls } \omega \in A_j\,, \quad j = 1, \ldots, n\,,
\end{aligned}$$

vergleiche die Definition 1.4.1.

2. *Bedingte Erwartung bzgl. einer Zufallsvariablen Y:*

   Definition 1.4.11 Seien X und Y zwei Zufallsvariablen definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Der *bedingte Erwartungswert von X unter der Bedingung Y* wird als $\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma_Y)$ eingeführt, wobei σ_Y die von Y erzeugte σ-Algebra ist: $\sigma_Y = Y^{-1}(\mathcal{B}_\mathbb{R})$.

   Lemma 1.4.12 Es existiert eine Borel-messbare Funktion $g : \mathbb{R} \to \mathbb{R}$, für die gilt, dass $\mathbb{E}(X|Y) = g(Y)$ f.s.

   Beweis Das Maß

$$Q(B) = \int_{Y^{-1}(B)} X(\omega) P(d\omega), \quad B \in \mathcal{B}_\mathbb{R}$$

ist ein signiertes Maß auf $\mathcal{B}_\mathbb{R}$ und absolut stetig bzgl. P_Y. Nach dem Satz von Radon–Nikodým existiert eine $\mathcal{B}_\mathbb{R}$-messbare Funktion $g : \mathbb{R} \to \mathbb{R}$, welche eine Dichte von Q bezüglich P_Y ist. Nach der Definition des bedingten Erwartungswert und dem Maßtransport im Lebesgue–Integral gilt für alle $B \in \mathcal{B}_\mathbb{R}$, dass

$$\int\limits_{Y^{-1}(B)} X(\omega)P(d\omega) = \int\limits_{B} g(x)P_Y(dx) = \int\limits_{Y^{-1}(B)} g(Y(\omega))P(d\omega).$$

Die Funktion $g(Y)$ ist σ_Y–messbar, und alle Elemente $A \in \sigma_Y$ haben die Form $Y^{-1}(B)$, $B \in \mathcal{B}_\mathbb{R}$, also folgt $\mathbb{E}(X|Y) = g(Y)$ f. s.. $\square$

Daher wird die Schreibweise $\mathbb{E}(X|Y = y)$ als $g(y)$ verstanden: $\mathbb{E}(X|Y = y) = g(y)$ oder $\mathbb{E}(X|Y = y)$ ist der Wert von $\mathbb{E}(X|Y)$ auf der Menge $\{\omega \in \Omega : Y(\omega) = y\}$.

3. *Bedingte Wahrscheinlichkeit bzgl. einer σ-Algebra bzw. einer Zufallsvariable:*

Definition 1.4.13 Die *bedingte Wahrscheinlichkeit* von $A \in \mathcal{F}$ unter der Bedingung $\mathcal{B}$ ist gegeben durch $P(A|\mathcal{B}) = \mathbb{E}(I_A|\mathcal{B})$ fast sicher. Analog dazu definieren wir $P(A|Y) = \mathbb{E}(I_A|Y)$ für eine Zufallsvariable Y.

Wir wissen, dass die bedingte Erwartung nur f.s.-eindeutig bestimmt ist. Verschiedene Varianten der bedingten Erwartung, die sich auf einer Menge $B \in \mathcal{F}$ mit $P(B) = 0$ unterscheiden, nennen wir *Versionen von einander.*

Bemerkung 1.4.14 $P(\cdot|\mathcal{B})(\omega)$ ist i.A. kein Maß!
Die soeben definierte Familie von Zufallsvariablen $P(\cdot|\mathcal{B})(\omega)$ erfüllen (für jedes $\omega \in \Omega$ wie f.s.) nicht die Eigenschaften eines Wahrscheinlichkeitsmaßes: Es gilt zwar

$$0 \le P(A|\mathcal{B}) \le 1 \quad \text{f.s. für alle } A \in \mathcal{F},$$

aber die Menge derjenigen $\omega \in \Omega$, für die die Eigenschaft der σ-Additivität

$$P\left(\bigcup_{j=1}^{\infty} A_j \middle| \mathcal{B}\right)(\omega) = \sum_{i=j}^{\infty} P(A_j|\mathcal{B})(\omega) \tag{1.6}$$

für paarweise disjunkte $\{A_j\} \subset \mathcal{F}$ gilt, hängt evtl. von der Folge $\{A_j\}$ sowie von der Version $P(\cdot|\mathcal{B})$ ab. Es bedeutet, dass es i.A. keine Version von $P(\cdot|\mathcal{B})$ existieren muss, so dass die Eigenschaft (1.6) für *alle* $\omega \in M^C$ gilt, wobei $M \in \mathcal{F}$ die Ausnahme–Menge ist mit $P(M) = 0$.

Es wäre jedoch von Vorteil, wenn wir *eine* Version von $P(\cdot|\mathcal{B})$ einführen könnten, die die Eigenschaften eines Wahrscheinlichkeitsmaßes für *alle* $\omega \in \Omega$ hätte. Dann könnte man z.B. gewöhnlich nach Lebesgue integrieren bzgl. des neuen Wahrscheinlichkeitsmaßes, vgl. Satz 1.5.2 unten.

Dies führt uns zum Begriff der *regulären* bedingten Wahrscheinlichkeit.

1.5 Reguläre bedingte Erwartung

1.5.1 Reguläre bedingte Wahrscheinlichkeiten und Verteilungen

Definition 1.5.1 Eine Abbildung $P_{\mathcal{B}} : \mathcal{F} \times \Omega \to [0,1]$ heisst *reguläre bedingte Wahrscheinlichkeit* bzgl. der σ–Algebra $\mathcal{B} \subset \mathcal{F}$, falls

- $P_{\mathcal{B}}(\cdot,\omega)$ ist ein Wahrscheinlichkeitsmaß auf $\mathcal{F}$ für jedes $\omega \in \Omega$.
- $P_{\mathcal{B}}(B,\omega)$ ist eine Version von $P(B|\mathcal{B})(\omega)$ für jedes $B \in \mathcal{F}$.

Satz 1.5.2 Sei $X : \Omega \to \mathbb{R}$ eine integrierbare Zufallsvariable, und sei $P_{\mathcal{B}}$ die reguläre bedingte Wahrscheinlichkeit bzgl. der σ–Algebra $\mathcal{B} \subset \mathcal{F}$. Dann gilt

$$\mathbb{E}(X|\mathcal{B}) = \int_\Omega X(\omega)\, P_{\mathcal{B}}(d\omega). \tag{1.7}$$

Beweis Zunächst prüfen wir die Behauptung für $X(\omega) = I_B(\omega)$, $B \in \mathcal{F}$:

$$\mathbb{E}(I_B|\mathcal{B}) = P_{\mathcal{B}}(B),$$

was aus Definition 1.5.1 ersichtlich wird. Der Rest des Beweises folgt der algebraischen Induktion bei der Konstruktion des Lebesgue–Integrals (vgl. Abschnitt 1.1): Man approximiert nichtnegative X durch eine Folge einfacher Zufallsvariablen und benutzt schliesslich die Zerlegung $X = X_+ - X_-$. $\square$

Die *reguläre bedingte Verteilung* $P_{X|\mathcal{B}}(A)$, $A \in \mathcal{T}$, eines Zufallselements $X : \Omega \to T$ mit Werten in einem topologischen Raum $(T, \mathcal{T})$ bekommt man, indem man die Definition 1.5.1 auf die Urbilder $X^{-1}(A)$, $A \in \mathcal{T}$ anwendet:

$$P_{X|\mathcal{B}}(A) := P(\{X \in A\}|\mathcal{B}), \quad A \in \mathcal{T}.$$

Für eine Zufallsvariable $X : \Omega \to \mathbb{R}$ wird die zufällige Funktion

$$F_{X|\mathcal{B}}(x) := P_{X|\mathcal{B}}((-\infty, x]), \quad x \in \mathbb{R}$$

die *reguläre bedingte Verteilungsfunktion von X unter der Bedingungs-σ-Algebra $\mathcal{B} \subset \mathcal{F}$* genannt.

Weiterhin definieren wir Messräume, die zu einer Borelschen Teilmenge von $\mathbb{R}$ äquivalent sind. So sind ihre Eigenschaften vergleichbar mit denen von $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$:

Definition 1.5.3 Ein Messraum (T, τ) heißt *Borelsch, falls eine Bijektion* $\varphi : (T, \tau) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ *existiert mit den Eigenschaften*

- $\varphi(T) \in \mathcal{B}_{\mathbb{R}}$,
- φ *ist* $\tau|\mathcal{B}_{\mathbb{R}}$*–messbar,*
- φ^{-1} *ist* $\mathcal{B}_{\mathbb{R}}|\tau$*–messbar.*

Es kann gezeigt werden, dass alle vollständigen separablen metrischen Messräume (darunter $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$, $d \in \mathbb{N} \cup \{\infty\}$) Borelsch sind.

Reguläre bedingte Verteilungen existieren auf Borelschen Messräumen, wie der folgende Satz zeigt:

Satz 1.5.4 Sei $X : \Omega \to T$ ein Zufallselement mit den Werten im Borelschen Messraum $(T, \mathcal{T})$, und sei $\mathcal{B}$ eine Teil–σ–Algebra von $\mathcal{F}$. Dann existiert die reguläre bedingte Wahrscheinlichkeit $P(X \in \cdot|\mathcal{B})$.

Den Beweis dieses Satzes kann man z.B. in den Büchern [19, Theorem II.7.4], [15, Theorem 8.28, 8.36] nachlesen.

1.5.2 Bedingte Dichten

Sei (X, Y) ein bivariater Zufallsvektor mit reellwertigen Koordinaten, der absolut stetig verteilt ist mit gemeinsamer Dichte $f_{X,Y}$. Wie kann der bedingte Erwartungswert $\mathbb{E}(X|Y = y)$ berechnet werden? Dafür ist der Begriff der bedingten Dichte oft günstig:

Definition 1.5.5 Die Funktion

$$f_{X|Y}(x|y) := \frac{f_{X,Y}(x, y)}{f_Y(y)} I_{\{f_Y(y) \neq 0\}}, \quad x, y \in \mathbb{R}$$

heisst die *bedingte Dichte von X gegeben Y.*

Es ist offensichtlich, dass sich f_Y durch $f_{X,Y}$ errechnen lässt als

$$f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dx.$$

Übungsaufgabe 1.5.6 (Bayessche Formel für bedingte Dichten) Sei (X, Y) ein absolut stetig verteilter Zufallsvektor mit gemeinsamer Dichte $f_{X,Y}$ sowie Randdichten $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dy$ und $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dx$. Zeigen Sie die folgende Verbindung zwischen den bedingten Dichten $f_{X|Y}$ und $f_{Y|X}$:

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y) f_Y(y)}{f_X(x)} I_{\{f_X(x) \neq 0\}}, \quad x, y \in \mathbb{R}.$$

Zeigen wir die folgende praktische Formel, die auch zur Berechnung von $\mathbb{E}(X|Y = y)$ verwendet werden kann:

Satz 1.5.7 Es sei (X, Y) ein absolut stetig verteilter Zufallsvektor mit gemeinsamer Dichte $f_{X,Y}$, und sei $g : \mathbb{R}^2 \to \mathbb{R}$ eine beliebige messbare Funktion mit $\mathbb{E}|g(X, Y)| < +\infty$. Dann gilt folgende Darstellung:

$$\mathbb{E}\left(g(X, Y)|Y = y\right) = \int\limits_{\mathbb{R}} g(x, y) f_{X|Y}(x|y)\, dx \tag{1.8}$$

für P_Y–fast alle $y \in \mathbb{R}$.

Beweis Bezeichnen wir die rechte Seite der Relation (1.8) mit $h(y)$. Zu zeigen ist, dass $\mathbb{E}\left(g(X, Y)|Y\right) = h(Y)$ f.s. Per Definition der bedingten Erwartung genügt es zu zeigen, dass

$$\int\limits_B h(Y(\omega))\, P(d\omega) = \int\limits_B g(X(\omega), Y(\omega))\, P(d\omega)$$

für beliebige $B \in \sigma_Y$, also, $B = Y^{-1}(A)$ für Borelsche Mengen $A \subset \mathbb{R}$. Einerseits gilt

$$\int\limits_B g(X(\omega), Y(\omega))\, P(d\omega) = \int\limits_A \int\limits_{\mathbb{R}} g(x, y) f_{X,Y}(x, y)\, dx\, dy\,. \tag{1.9}$$

Andererseits haben wir mit Hilfe des Maßtransports und der Definition 1.5.5, dass

$$\begin{aligned}\int\limits_B h(Y(\omega))\, P(d\omega) = \int\limits_A h(y)\, P_Y(dy) &= \int\limits_A \int\limits_{\mathbb{R}} g(x, y) f_{X|Y}(x|y)\, dx\, f_Y(y)\, dy \\ &= \int\limits_A \int\limits_{\mathbb{R}} g(x, y) f_{X,Y}(x, y)\, dx\, dy\,.\end{aligned}$$

Der letzte Ausdruck und die rechte Seite von (1.9) sind identisch. $\square$

Bemerkung 1.5.8 Den wichtigen Spezialfall der Formel (1.7) für $\mathcal{B} = \sigma_Y$ liefert Satz 1.5.7 mit $g(X, Y) = X$:

$$\mathbb{E}(X|Y = y) = \int_{\mathbb{R}} x f_{X|Y}(x|y) dx$$

Insbesondere gilt für $g(x, y) = I(x \in A)$, $A \in \mathcal{B}_{\mathbb{R}}$, dass

$$P(X \in A|Y = y) = \int\limits_A f_{X|Y}(x|y)\, dx \quad P_Y\text{–f.s.}$$

Dies ist ein explizites Beispiel für die Konstruktion des regulären bedingten Wahrscheinlichkeitsmaßes, dessen Existenz im letzten Abschnitt behauptet wurde.

Schliesslich sei ein explizites Beispiel der Berechnung von regulären Wahrscheinlichkeiten zu diskutieren:

Beispiel 1.5.9 (**Verteilungen von Zufallsvektoren mit sphärischen Konturen**) Auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, sei $R > 0$ eine f.s. positive Zufallsvariable, und sei $\eta = (\eta_1, \ldots, \eta_n) \sim \mathcal{U}(S^{n-1})$, $n \geq 2$, ein stochastisch unabhängiger zufälliger Einheitsvektor gleichverteilt auf der Einheitsshpäre S^{n-1}. Man sagt, dass die Verteilung des n-dimensionalen Zufallsvektors $X = (X_1, \ldots, X_n)^\top := R\eta$ *sphärische Konturen* besitzt. Falls z.B. $R^2 \sim \chi^2_n$ eine χ^2–Verteilung mit n Freiheitsgraden hat, so ist $X \sim N(0, I_n)$ n–variat Gaußsch verteilt. Führen wir folgende Bezeichnungen ein: $R_k =: \sum_{j=1}^k X_j^2$, $\bar{R}_{n-k} :=: \sum_{j=k+1}^n X_j^2$, $k = 1, \ldots, n$, so dass $R^2 = R_n$. Laut [5, Lemma 2], [6], [18, Lemma 2.5] gibt es folgende markante Verteilungseigenschaft:

$$(R_k, \bar{R}_{n-k}) \stackrel{d}{=} R^2(B_k, 1 - B_k),$$

wobei $B_k \sim \mathrm{Beta}(k/2, (n-k)/2)$ eine Beta–verteilte Zufallsvariable ist, die stochastisch unabhängig von R ist. Wir wollen nun den Spezialfall dieser Relation beweisen, und zwar für $n = 2$ und stochastisch unabhängige Zufallsvariablen $X_1, X_2 \sim N(0, 1)$ werden wir zeigen, dass X_1^2/R^2 bedingt durch R^2 eine $\mathrm{Beta}(1/2, 1/2)$–Verteilung hat. Etwas formeller soll folgende Gleichheit der bedingten Dichte

$$f_{X_1^2/R^2|R^2}(x|y) = f_{B_1}(x), \qquad x \in (0, 1),\ y > 0 \tag{1.10}$$

gezeigt werden. Dazu soll laut Definition 1.5.5 die gemeinsame Dichte von $(X_1^2/R^2, R^2)$ mit Hilfe des Dichtetransformationssatzes (vgl. [20, Theorem 3.6.2]) aus der Dichte $f_{(X_1^2, X_2^2)}$ errechnet werden. Es gilt $X_j^2 \sim \Gamma(1/2, 1/2)$ mit Dichte

$$f_{X_j^2}(x) = \frac{e^{-x/2} I(x > 0)}{\sqrt{2\pi x}}.$$

Wegen der stochastischen Unabhängigkeit von X_j^2 bekommt man

$$f_{(X_1^2, X_2^2)}(x, y) = \frac{e^{-(x+y)/2} I(x, y > 0)}{2\pi\sqrt{xy}}.$$

Sei die Transformation $\varphi(x, y) = \left(\frac{x}{x+y}, x + y\right)$, $x, y > 0$, mit $\varphi(X_1^2, X_2^2) = (X_1^2/R^2, R^2)$. Ihre Inverse schreibt sich als $\varphi^{-1}(x, y) = (xy, (1-x)y)$, $x, y > 0$ mit Jacobi–Determinante

$$J = det \begin{pmatrix} y & x \\ -y & 1-x \end{pmatrix} = y.$$

Dann gilt

$$
\begin{aligned}
f_{(X_1^2/R^2,R^2)}(x,y) = f_{(X_1^2,X_2^2)}(\varphi^{-1}(x,y))|J| &= \frac{I(x\in(0,1))}{\pi\sqrt{x(1-x)}}\cdot\frac{1}{2}e^{-y/2}I(y>0)\\
&= f_{B_1}(x)\cdot f_{R^2}(y), \qquad x\in(0,1),\ y>0,
\end{aligned}
$$

wobei $R^2 \sim Exp(1/2)$. Damit ist die Relation (1.10) bewiesen.

Das obige Beispiel kann auch mit der multivariaten Dirichlet–Verteilung in Verbindung gebracht werden. Definieren wir die n–variate *Dirichlet–Verteilung* mit Parametern $(p_0, p_1, \ldots, p_n)$ als die absolut stetige Verteilung eines Zufallsvektors $Y = (Y_1, \ldots, Y_n)^\top$ mit gemeinsamer Dichte

$$
f_Y(x_1,\ldots,x_n) = \frac{\Gamma\left(\sum_{j=0}^n p_j\right)}{\prod_{j=0}^n \Gamma(p_j)}\left(1-\sum_{j=1}^n x_j\right)^{p_0-1}\prod_{j=1}^n x_j^{p_j-1}, \quad 0\le x_j, \sum_{j=1}^n x_j \le 1,
$$

für $p_0, p_1, \ldots, p_n > 0$. Es ist ratsam, diese Verteilung als ein multivariates Analogon der Beta–Verteilung aufzufassen, da für jede Randverteilung von Y_j gilt $Y_j \sim \text{Beta}\left(p_j, \sum_{l\neq j} p_l\right)$, $j = 1, \ldots, n$. Es kann gezeigt werden, dass $Y_j \stackrel{d}{=} Z_j/\sum_{l=0}^n Z_l$, $j = 1, \ldots, n$, wobei $Z_l \sim \Gamma(1/2, p_l)$, $l = 0, \ldots, n$ stochastisch unabhängige Zufallsvariablen sind, vgl. [6]. Im Beispiel 1.5.9 wählt man $n = 1$, $Z_0 = X_2^2$, $Z_1 = X_1^2$, $p_0 = p_1 = 1/2$ mit $Y_1 = X_1^2/R^2$.

1.5.3 Bedingte (Ko)Varianz

Ähnlich wie im unbedingten Fall, können höhere bzw. gemischte bedingte Momente eingeführt weden. Wir tun es hier für bedingte Momente zweiter Ordnung:

Definition 1.5.10 Sei $\mathcal{B} \subset \mathcal{F}$ eine Teil–σ-Algebra.

1. Seien $X, Y : \Omega \to \mathbb{R}$ integrierbare Zufallsvariablen mit $\mathbb{E}|XY| < \infty$. Ihre ***bedingte Kovarianz unter Bedingung*** $\mathcal{B}$ ist gleich

$$
\text{Cov}\ (X,Y|\mathcal{B}) := \mathbb{E}\left[(X-\mathbb{E}(X|\mathcal{B}))\left(Y-\mathbb{E}(Y|\mathcal{B})\right)|\mathcal{B}\right].
$$

2. Sei $Z : \Omega \to \mathbb{R}$ eine Zufallsvariable mit $\mathbb{E}Z^2 < \infty$. Ihre ***bedingte Varianz unter Bedingung*** $\mathcal{B}$ ist gleich

$$
\text{Var}\ (Z|\mathcal{B}) := \mathbb{E}\left[(Z-\mathbb{E}(Z|\mathcal{B}))^2\,|\mathcal{B}\right].
$$

Offensichtlich gilt Var $(Z|\mathcal{B}) =$ Cov $(Z, Z|\mathcal{B})$ wie im Fall der unbedingten (Ko)Varianzen.

Lemma 1.5.11 Sei $\mathcal{B} \subset \mathcal{F}$ eine Teil–σ-Algebra, und seien $X, Y, Z : \Omega \to \mathbb{R}$ Zufallsvariablen wie in Definition 1.5.10. Für sie gilt

$$\text{Cov }(X, Y) = \mathbb{E}\,[\text{Cov }(X, Y|\mathcal{B})] + \text{Cov }\,[\mathbb{E}(X|\mathcal{B}), \mathbb{E}(Y|\mathcal{B})] \tag{1.11}$$

bzw.

$$\text{Var } Z = \mathbb{E}\,[\text{Var }(Z|\mathcal{B})] + \text{Var }\,[\mathbb{E}(Z|\mathcal{B})]\ . \tag{1.12}$$

Beweis Die Formel (1.12) folgt aus (1.11) für $X = Y = Z$. Um nun die Relation (1.11) zu beweisen, verwenden wir die nahrhafte Null, indem wir $\mathbb{E}(X|\mathcal{B}), \mathbb{E}(Y|\mathcal{B})$ im Ausdruck für Cov (X, Y) aufaddieren und gleichzeitig abziehen, sowie Theorem 1.4.6 1g:

$$\begin{aligned}
\text{Cov }(X, Y) =\ & \mathbb{E}\,(X - \mathbb{E}X)\,(Y - \mathbb{E}Y) \\
=\ & \mathbb{E}\,\mathbb{E}\,[(X - \mathbb{E}(X|\mathcal{B}) + \mathbb{E}(X|\mathcal{B}) - \mathbb{E}X)\,(Y - \mathbb{E}(Y|\mathcal{B}) + \mathbb{E}(Y|\mathcal{B}) - \mathbb{E}Y)\,|\mathcal{B}] \\
=\ & \mathbb{E}\,\mathbb{E}\,[(X - \mathbb{E}(X|\mathcal{B}))\,(Y - \mathbb{E}(Y|\mathcal{B}))\,|\mathcal{B}] + \mathbb{E}\,\mathbb{E}\,[(X - \mathbb{E}(X|\mathcal{B}))\,(\mathbb{E}(Y|\mathcal{B}) - \mathbb{E}Y)\,|\mathcal{B}] \\
& + \mathbb{E}\,\mathbb{E}\,[(\mathbb{E}(X|\mathcal{B}) - \mathbb{E}X)\,(Y - \mathbb{E}(Y|\mathcal{B}))\,|\mathcal{B}] + \mathbb{E}\,\mathbb{E}\,[(\mathbb{E}(X|\mathcal{B}) - \mathbb{E}X)\,(\mathbb{E}(Y|\mathcal{B}) - \mathbb{E}Y)\,|\mathcal{B}] \\
=\ & \mathbb{E}\,[\text{Cov }(X, Y|\mathcal{B})] + \mathbb{E}\,(\mathbb{E}(X|\mathcal{B}) - \mathbb{E}(X|\mathcal{B}))\,(\mathbb{E}(Y|\mathcal{B}) - \mathbb{E}Y) \\
& + \mathbb{E}\,(\mathbb{E}(X|\mathcal{B}) - \mathbb{E}X)\,(\mathbb{E}(Y|\mathcal{B}) - \mathbb{E}(Y|\mathcal{B})) + \text{Cov }\,[\mathbb{E}(X|\mathcal{B}), \mathbb{E}(Y|\mathcal{B})] \\
=\ & \mathbb{E}\,[\text{Cov }(X, Y|\mathcal{B})] + \text{Cov }\,[\mathbb{E}(X|\mathcal{B}), \mathbb{E}(Y|\mathcal{B})]\ .
\end{aligned}$$

In den letzten drei Zeilen wurde vor allem Theorem 1.4.6 1h benutzt, weil $\mathbb{E}(X|\mathcal{B})$ und $\mathbb{E}(Y|\mathcal{B})$ $\mathcal{B}$–messbare Zufallsvariablen sind. Demnach entfallen zwei Terme dort, die Null sind. □

Kapitel 2

Analytische Hilfsmittel der Wahrscheinlichkeitsrechnung

Um Gesetze der großen Zahlen und weitere Grenzwertsätze der Wahrscheinlichkeitsrechnung beweisen zu können, brauchen wir einige analytische Werkzeuge aus der Analysis. Sie werden in diesem Kapitel eingeführt und ausführlich behandelt.

2.1 Komplexwertige Zufallsvariablen

Um die charakteristischen Funktionen einführen zu können, brauchen wir den Begriff der komplexwertigen Zufallsvariablen. Sei $\mathbb{C}$ die Menge aller komplexen Zahlen und $\mathcal{G}$ die minimale σ–Algebra, die alle Teilmengen
$\{z \in \mathbb{C} : z = x + iy,\, a_1 < x \le b_1,\, a_2 < y \le b_2\}$,
$a_1, b_1, a_2, b_2 \in \mathbb{R}$, $a_1 < b_1$, $a_2 < b_2$ beinhaltet. Sei $B_1(0) = \{z \in \mathbb{C} : |z| \le 1\}$ der komplexe Einheitskreis.

Definition 2.1.1 Ein Zufallselement

$$X : (\Omega, \mathcal{F}, P) \to (\mathbb{C}, \mathcal{G})$$

heißt *komplexwertige Zufallsvariable.*
Offensichtlich kann eine komplexwertige Zufallsvariable X als

$$X(\omega) = X_1(\omega) + iX_2(\omega), \;\; \omega \in \Omega$$

dargestellt werden, wobei $X_1(\omega) = \mathrm{Re}\,(X(\omega))$, $X_2(\omega) = \mathrm{Im}\,(X(\omega))$ zwei reellwertige Zufallvariablen sind. Durch diese Darstellung werden viele wichtige Begriffe, die für reellwertige Zufallsvariablen eingeführt wurden, direkt auf komplexwertige Zufallsvariablen übertragbar. Zum Beispiel wird der *Erwartungswert einer komplexwertigen Zufallsvariablen* X als $\mathrm{E}X = \mathrm{E}X_1 + i\mathrm{E}X_2$ eingeführt. Dabei gilt

$$|X| = \sqrt{|X_1|^2 + |X_2|^2} \text{ und } \mathrm{E}|X|^2 = \mathrm{E}|X_1|^2 + \mathrm{E}|X_2|^2\,.$$

Weiterhin sind komplexwertige Zufallsvariablen $X = X_1 + iX_2$ und $Y = Y_1 + iY_2$ *stochastisch unabhängig*, falls die σ–Algebren, die von den Zufallsvektoren (X_1, X_2) und (Y_1, Y_2) erzeugt werden, stochastisch unabhängig sind. Die *komplex Konjugierte* von $X = X_1+iX_2$ führt man als $\overline{X} = X_1-iX_2$ ein.

Übungsaufgabe 2.1.2 Zeigen Sie, dass

1. für eine komplexwertige Zufallsvariable X auch $|\mathrm{E}X| \leq \mathrm{E}|X|$ gilt.

2. für unabhängige komplexwertige Zufallsvariablen X und Y die Aussage des Satzes 1.3.3 gilt: $\mathrm{E}(XY) = \mathrm{E}X \cdot \mathrm{E}Y$.

2.2 Charakteristische Funktionen

Definition 2.2.1

1. Sei $X : \Omega \to \mathbb{R}$ eine reellwertige Zufallsvariable. Die *charakteristische Funktion* $\varphi_X : \mathbb{R} \to \mathbb{C}$ von X wird durch
$$\varphi_X(t) := \mathrm{E}e^{itX} = \int_{\mathbb{R}} e^{itx}\, dF_X(x), \quad t \in \mathbb{R}$$
eingeführt.

2. Sei $X : \Omega \to \mathbb{R}^n$ ein reellwertiger Zufallsvektor. Die *charakteristische Funktion* φ_X von X wird als $\varphi_X : \mathbb{R}^n \to \mathbb{C}$ durch
$$\varphi_X(t) := \mathrm{E}e^{i\langle t,X\rangle} = \int_{\mathbb{R}^n} e^{i\langle t,X\rangle}\, dF_X(x), \quad t \in \mathbb{R}^n$$
eingeführt. Hierbei ist $\langle\cdot\,,\cdot\rangle$ das euklidische Skalarprodukt in $\mathbb{R}^n$.

Wir werden im Folgenden die Eigenschaften von charakteristischen Funktionen für Zufallsvariablen formulieren, obwohl sehr viele von ihnen auch leicht auf charakteristische Funktionen von Zufallsvektoren übertragen werden können.

Satz 2.2.2 Sei $X : \Omega \to \mathbb{R}$ eine beliebige Zufallsvariable mit charakteristischer Funktion φ_X.

1. Es gilt
$$|\varphi_X(t)| \leq \varphi_X(0) = 1\,,\ t \in \mathbb{R}\,,$$
also $\varphi_X : \mathbb{R} \to B_1(0) \subseteq \mathbb{C}$.

2. $\varphi_X(-t) = \overline{\varphi_X}(t)$ für alle $t \in \mathbb{R}$.

3. $\varphi_{aX+b}(t) = e^{itb} \cdot \varphi_X(at)$ für alle $a, b \in \mathbb{R}$ und $t \in \mathbb{R}$.

4. φ_X ist gleichmäßig stetig auf $\mathbb{R}$.

5. Falls $X_1, \ldots, X_n$ stochastisch unabhängige reellwertige Zufallsvariablen auf $(\Omega, \mathcal{F}, P)$ mit charakteristischen Funktionen φ_{X_i} sind und $S_n = \sum_{j=1}^n X_j$, dann gilt
$$\varphi_{S_n}(t) = \prod_{j=1}^n \varphi_{X_j}(t), \quad t \in \mathbb{R}.$$

Beweis

1. Es gilt $e^{itX} \in \partial B_1(0) = \{z \in \mathbb{C} : |z| = 1\}$ f.s. für komplexwertige Zufallsvariablen. Daher folgt nach Übungsaufgabe 2.1.2
$$|\varphi_X(t)| = |\mathrm{E}e^{itX}| \leq \mathrm{E}|e^{itX}| = \mathrm{E}(1) = 1.$$
Somit existiert der Erwartungswert in der Definition 2.1.1 immer, und die charakteristische Funktion einer beliebigen Zufallsvariablen existiert für alle $t \in \mathbb{R}$. Die Eigenschaft $\varphi_X(0) = 1$ ist offensichtlich.

2. Wir schreiben mit Hilfe der Polardarstellung komplexer Zahlen:
$$\begin{aligned}\varphi_X(-t) &= \mathrm{E}e^{-itX} = \mathrm{E}(\cos(-tX) + i\sin(-tX))\\ &= \mathrm{E}(\cos(tX) - i\sin(tX)) = \mathrm{E}\cos(tX) - i\mathrm{E}(\sin(tX))\\ &= \overline{(\mathrm{E}\cos(tX) + i\sin(tX))} = \overline{\mathrm{E}e^{itX}}\\ &= \overline{\varphi_X}(t), \quad t \in \mathbb{R}.\end{aligned}$$

3. Für alle $t \in \mathbb{R}$ und $a, b \in \mathbb{R}$ haben wir
$$\begin{aligned}\varphi_{aX+b}(t) &= \mathrm{E}e^{it(aX+b)} = \underbrace{e^{itb}}_{=const}\, \mathrm{E}\left(e^{iatX}\right)\\ &= e^{itb} \cdot \varphi_X(at).\end{aligned}$$

4. Zeigen wir, dass: $|\varphi_X(t+h) - \varphi_X(t)| \underset{h \to 0}{\longrightarrow} 0$ unabhängig von $t \in \mathbb{R}$ ist. In der Tat gilt
$$\begin{aligned}|\varphi_X(t+h) - \varphi_X(t)| &= |\mathrm{E}e^{i(t+h)X} - \mathrm{E}e^{itX}| = \left|\mathrm{E}\left(e^{itX}\left(e^{ihX} - 1\right)\right)\right|\\ &\leq \mathrm{E}\Bigg(\underbrace{\left|e^{itX}\right|}_{=1} \cdot \left|e^{ihX} - 1\right|\Bigg) = \mathrm{E}\left|e^{ihX} - 1\right| \underset{h \to 0}{\longrightarrow} 0\end{aligned}$$
nach dem Satz 1.2.7 über die majorisierte Konvergenz, weil
$$\left|e^{ihX} - 1\right| \leq \left|e^{ihX}\right| + 1 = 2 \text{ für alle } h \in \mathbb{R}.$$

5. Es gilt

$$\varphi_Y(t) = \mathrm{E}e^{it\sum_{j=1}^n X_j} = \mathrm{E}\left(\prod_{j=1}^n e^{itX_j}\right) = \prod_{j=1}^n \mathrm{E}e^{itX_j}$$
$$= \prod_{j=1}^n \varphi_{X_j}(t)$$

für alle $t \in \mathbb{R}$, weil aus der Unabhängigkeit von $X_1, \ldots, X_n$ die Unabhängigkeit von $e^{itX_1}, \ldots, e^{itX_n}$ folgt (vgl. Übungsaufgabe 2.1.2).

□

Beispiel 2.2.3

1. Zeigen wir, dass für $X \sim \mathit{Bernoulli}(p)$

$$\varphi_X(t) = pe^{it} + (1-p)\,, \quad t \in \mathbb{R}$$

gilt. In der Tat,

$$\begin{aligned} \varphi_X(t) &= \mathrm{E}e^{itX} = e^{it\cdot 1} \cdot P(X=1) + e^{it\cdot 0} P(X=0) \\ &= e^{it} \cdot p + 1 \cdot (1-p) = pe^{it} + 1 - p\,. \end{aligned}$$

2. Sei $P \sim \text{Poisson}(\lambda)$, mit $\lambda > 0$. Zeigen wir, dass

$$\varphi_X(t) = e^{\lambda(e^{it}-1)}\,, \quad t \in \mathbb{R}.$$

Es gilt:

$$\begin{aligned} \varphi_X(t) &= \mathrm{E}e^{itX} = \sum_{k=0}^{\infty} e^{itk} \cdot P(X=k) = \sum_{k=0}^{\infty} e^{itk} \cdot e^{-\lambda}\frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^{it}\right)^k}{k!} = e^{-\lambda} \cdot e^{\lambda e^{it}} = e^{\lambda\left(e^{it}-1\right)}\,. \end{aligned}$$

3. Sei $X \sim N(\mu, \sigma^2)$. Zeigen wir, dass

$$\varphi_X(t) = e^{it\mu - \frac{t^2\sigma^2}{2}}\,, \quad t \in \mathbb{R}\,.$$

Es gilt $X = \mu + \sigma Y$, wobei $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$. Wegen der Eigenschaft 3) des Satzes 2.2.2 gilt $\varphi_X(t) = e^{it\mu} \cdot \varphi_Y(\sigma t)$. Somit genügt es, $\varphi_Y(t)$ für $Y \sim N(0,1)$ zu berechnen. Zeigen wir, dass $\varphi_Y(t) = e^{-\frac{t^2}{2}}$

ist und somit $\varphi_X(t) = e^{it\mu} \cdot e^{-\frac{\sigma^2 t^2}{2}}$ gilt.

$$\begin{aligned}
\varphi_Y(t) &= \mathrm{E}e^{itY} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} \cdot e^{-\frac{x^2}{2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \sum_{k=0}^{\infty} \frac{(itx)^k}{k!} e^{-\frac{x^2}{2}}\, dx \\
&= \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^n e^{-\frac{x^2}{2}}\, dx}_{=\mathrm{E}Y^k} \\
&\overset{2k=n}{=} \sum_{k=0}^{\infty} \frac{(it)^{2k}}{(2k)!} \cdot \frac{(2k)!}{2^k k!} = \sum_{n=0}^{\infty} \frac{(-t^2)^n}{(2n)!} \cdot \frac{(2n)!}{2^n n!} \\
&= \sum_{n=0}^{\infty} \frac{\left(-\frac{t^2}{2}\right)^n}{n!} = e^{-\frac{t^2}{2}}, \quad t \in \mathbb{R},
\end{aligned}$$

wobei

$$\mathrm{E}Y^k = \begin{cases} 0, & k \text{ ungerade} \\ \frac{k!}{2^{\frac{k}{2}} (\frac{k}{2})!}, & k \text{ gerade} \end{cases} \tag{2.1}$$

Übungsaufgabe 2.2.4

1. Zeigen Sie, dass alle Übergänge (Vertauschung von $\sum$ und $\int$) im letzten Beispiel korrekt sind.
2. Zeigen Sie, dass die Relation (2.1) gilt.

Satz 2.2.5 *(Weitere Eigenschaften von charakteristischen Funktionen)*
Sei X eine Zufallsvariable mit Verteilungsfunktion F_X und charakteristischer Funktion φ_X. Dann gilt Folgendes:

1. φ_X ist genau dann reellwertig, wenn die Verteilung von X symmetrisch ist.
2. Falls $\mathrm{E}|X|^n < \infty$ für ein $n \in \mathbb{N}$, dann existiert die r–te Ableitung $\varphi_X^{(r)}(t)$ von φ_X für alle $r \leq n$, und es gilt

$$\varphi_X^{(r)}(t) = \int_{\mathbb{R}} (ix)^r e^{itx} dF_X(x), \quad t \in \mathbb{R}, \tag{2.2}$$

$$\mathrm{E}X^r = \frac{\varphi_X^{(r)}(0)}{i^r}, \tag{2.3}$$

$$\varphi_X(t) = \sum_{r=0}^{n} \frac{(it)^r}{r!} \mathrm{E}X^r + \frac{(it)^n}{n!} \varepsilon_n(t), \tag{2.4}$$

wobei $|\varepsilon_n(t)| \leq 3\mathbb{E}|X|^n$ und $\varepsilon_n(t) \to 0$ für $t \to 0$. Der Ausdruck (2.4) ist die Taylor-Entwicklung von φ_X um 0.

3. Falls $\varphi^{(2n)}(0)$ existiert und endlich ist, dann gilt
$$\mathrm{E}X^{2n} < \infty\,.$$

4. Falls $\mathrm{E}|X|^n < \infty$ für alle $n \in \mathbb{N}$ und
$$\limsup_{n\to\infty} \frac{\sqrt[n]{\mathrm{E}|X|^n}}{n} = \frac{1}{e\cdot R} < \infty\,, \quad R > 0\,,$$
dann gilt
$$\varphi_X(t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!}\mathrm{E}X^n\,, \quad |t| < R\,.$$

5. Falls $|\varphi_X(t_0)| = 1$ für ein $t_0 \neq 0$, dann ist der Träger C der Verteilung von X (der Wertebereich von X)
$$C = \left\{a + nh : n \in \mathbb{Z},\, h = \frac{2\pi}{t_0}\right\}$$
für ein $a \in \mathbb{R}$.

6. Falls $|\varphi_X(t)| = |\varphi_X(\alpha t)| = 1$ für zwei Punkte t und αt, $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, $t \neq 0$, dann gilt $X \equiv const$ f.s.

7. Falls $|\varphi_X(t)| = 1\,, \quad t \in \mathbb{R}$, dann gilt $X \equiv const$ f.s.

Beweis

1. Die Verteilung von X ist symmetrisch, falls $X \stackrel{d}{=} -X$, also
$$P_X(B) = P_X(-B)\,, \quad B \in \mathcal{B}_{\mathbb{R}}\,,$$
wobei $-B = \{-x : x \in B\}$.

"$\Leftarrow$" Falls P_X symmetrisch ist, dann gilt $\int_{\mathbb{R}} \sin(tx)dF_X(x) = 0$, weil $\sin(tx)$ eine beschränkte ungerade Funktion von $x \in \mathbb{R}$ ist. Somit gilt
$$\varphi_X(t) = \mathrm{E}\cos(tX) + i\,\underbrace{\mathrm{E}(\sin(tX))}_{=0} = \mathrm{E}\cos(tX) \in \mathbb{R}\,, \quad t \in \mathbb{R}\,,$$
und φ_X ist reellwertig.

"$\Rightarrow$" Falls $\varphi_X(t) \in \mathbb{R}\,, t \in \mathbb{R}$, dann folgt aus dem Satz 2.2.2, 2, dass
$$\varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi_X}(t) = \varphi_X(t),\ t \in \mathbb{R}\,.$$
In der Folgerung 2.2.8 wird bewiesen, dass somit $P_X = P_{-X}$ folgt, d.h.
$$P(X \in B) = P(-X \in B) = P(X \in -B),\ B \in \mathcal{B}_{\mathbb{R}}\,,$$
was die Symmetrie der Verteilung von X bedeutet.

2. Falls $\mathrm{E}|X|^n < \infty$, dann folgt aus der Ungleichung von Ljapunow (vgl. [20, Folgerung 4.4.6]) $\mathrm{E}|X|^r < \infty\,, \quad r \leq n\,.$
Beweisen wir die Gültigkeit der Darstellung (2.2) induktiv. Aus dem Beweis des Satzes 2.2.2, 4) folgt

$$\lim_{h\searrow 0} \frac{\varphi_X(t+h) - \varphi_X(t)}{h} = \lim_{h\searrow 0} \mathrm{E}\left(e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right).$$

Da

$$\left|\frac{e^{ihx}-1}{h}\right| \leq |x|\,, \quad x \in \mathbb{R}\,, \quad h \to 0$$

und $\mathrm{E}|X| < \infty$, die majorisierte Konvergenz von Lebesgue (vgl. Satz 1.2.7) ergibt, dass der Grenzwert

$$\begin{aligned}\lim_{h\to 0} \mathrm{E}\left(e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right) &= \mathrm{E}\left(e^{itX}\cdot \lim_{h\to 0}\left(\frac{e^{ihX}-1}{h}\right)\right) = i\mathbb{E}\left(Xe^{itX}\right)\\ &= i\int_{\mathbb{R}} xe^{itx}dF_X(x)\,.\end{aligned}$$

existiert. Somit existiert auch die Ableitung

$$\varphi_X'(t) = i\mathbb{E}\left(Xe^{itX}\right) = i\int_R xe^{itx}dF_X(x)\,.$$

(Basis der Induktion).

Übungsaufgabe 2.2.6 Führen Sie die Induktion zu Ende und zeigen Sie die Gültigkeit von (2.2) für alle $r \leq n$.

Die Formel (2.3) folgt direkt aus (2.2) mit $t = 0$, da

$$\varphi_X^{(r)}(t) = i^r \cdot \mathrm{E}\left(X^r e^{itX}\right).$$

Um die Taylor-Entwicklung (2.4) zu beweisen, benutzen wir die Taylor-Entwicklung für e^{iy}:

$$e^{iy} = \cos y + i\sin y = \sum_{k=0}^{n-1} \frac{(iy)^k}{k!} + \frac{(iy)^n}{n!}(\cos(\Theta_1 y) + i\sin(\Theta_2 y))\,,$$

mit $y, \Theta_1, \Theta_2 \in \mathbb{R}\,, \quad |\Theta_1| \leq 1,\, |\Theta_2| \leq 1$. Somit gilt für $y = tX$

$$e^{itX} = \sum_{k=0}^{n-1} \frac{(itX)^k}{k!} + \frac{(itX)^n}{n!}(\cos(\Theta_1 tX) + i\sin(\Theta_2 tX))\,,$$

wobei Θ_1 und Θ_2 hier Zufallsvariablen sind. Somit

$$\varphi_X(t) = \mathrm{E}\left(e^{itX}\right) = \sum_{k=0}^{n-1} \frac{(it)^k}{k!}\mathrm{E}X^k + \frac{(it)^n}{n!}(\mathrm{E}X^n + \varepsilon_n(t))\,,$$

wobei
$$\varepsilon_n(t) = \mathrm{E}\left[X^n\left(\cos(\Theta_1 tX) + i\sin(\Theta_2 tX) - 1\right)\right] .$$
Daraus folgt, dass $|\varepsilon_n(t)| \leq 3\mathbb{E}|X|^n$. Der Satz von Lebesgue über die majorisierte Konvergenz ergibt $\varepsilon_n(t) \to 0$, falls $t \to 0$.

3. Ohne Beweis (vgl. Beweis in [19], S. 280–281).

4. Ebenfalls ohne Beweis (vgl. Beweis in [19], S. 280–281).

5. Falls $|\varphi_X(t_0)| = 1$, $t_0 \neq 0$, dann existiert ein $a \in \mathbb{R}$ mit der Eigenschaft $\varphi_X(t_0) = e^{it_0 a}$. Daraus folgt
$$e^{it_0 a} = \varphi_X(t_0) = \int_{\mathbb{R}} e^{it_0 x}\, dF_X(x),$$
$$\mathbb{R} \ni 1 = \int_{\mathbb{R}} e^{it_0(x-a)}\, dF_X(x) = \int_{\mathbb{R}} \cos(t_0(x-a))\, dF_X(x),$$
oder anders geschrieben
$$\int_{\mathbb{R}} (1 - \cos(t_0(x-a)))\, dF_X(x) = \mathrm{E}(\underbrace{1 - \cos(t_0(X-a))}_{=Y}) = 0\,,$$
wobei $Y \geq 0$ f.s. Aus der Eigenschaft 8) des Satzes 1.2.3 folgt $Y \equiv 0$ f.s. und somit $\cos(t_0(X-a)) \equiv 1$ f.s., was darauf hindeutet, dass
$$t_0(X-a) = 2\pi n, \quad n \in \mathbb{Z} \Longrightarrow X = a + n\frac{2\pi}{t_0}\,, \quad n \in \mathbb{Z}$$
f.s. und 5) ist bewiesen für $h = \frac{2\pi}{t_0}$.

6. Falls $|\varphi_X(t)| = |\varphi_X(\alpha t)| = 1$, dann folgt aus 5), dass
$$X = a + \frac{2\pi}{t}n = b + \frac{2\pi}{\alpha t}m\,, \quad n, m \in \mathbb{Z} \text{ f.s. für } a, b \in \mathbb{R}$$
Falls $X \neq const$, dann existieren unterschiedliche $n_1, n_2, m_1, m_2 \in \mathbb{Z}$: $n_1 \neq n_2$, $m_1 \neq m_2$ und
$$a + \frac{2\pi}{t}n_1 = b + \frac{2\pi}{\alpha t}m_1, \quad a + \frac{2\pi}{t}n_2 = b + \frac{2\pi}{\alpha t}m_2$$
(D.h. es gibt mindestens zwei gleiche Werte in den Mengen $\{a + \frac{2\pi}{t}n,\, n \in \mathbb{Z}\}$ und $\{b + \frac{2\pi}{\alpha t}m,\, m \in \mathbb{Z}\}$); folglich,
$$\frac{2\pi}{t}(n_1 - n_2) = \frac{2\pi}{\alpha t}(m_1 - m_2)\,,$$
woraufhin $\alpha \in \mathbb{Q}$ ist, was den Bedingungen des Satzes widerspricht. Daher ist $X \equiv const$ f.s.

7. Folgt aus 6), da $|\varphi_X(t)|= 1$ für alle $t \in \mathbb{R}$ impliziert, dass t, $\alpha t \in \mathbb{R}$ existieren, mit $\alpha \notin \mathbb{Q}$, so dass $|\varphi_X(t)|= |\varphi_X(\alpha t)|= 1$.

$\square$

Satz 2.2.7 *(Umkehrformel):*
Sei X eine Zufallsvariable mit Verteilungsfunktion F_X und charakteristischer Funktion φ_X .

1. Für beliebige Punkte $a, b \in \mathbb{R} :\ a < b$ und F_X stetig in a, b gilt

$$F_X(b) - F_X(a) = \lim_{x\to\infty} \frac{1}{2\pi} \int_{-x}^{x} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt\,.$$

2. Falls $\int_{\mathbb{R}} |\varphi_X(t)|\, dt < \infty$, dann ist X absolut stetig verteilt mit Dichte

$$f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_X(t)\, dt\,.$$

Um die Hauptidee des Beweises und die intuitive Bedeutung der Umkehrformel heuristisch zu vermitteln, betrachten wir zunächst den Fall einer absolut stetig verteilten Zufallsvariablen X mit Dichte f_X. Dann gilt

$$\begin{aligned}
F_X(b) - F_X(a) &= \int_a^b f_X(x)\, dx \\
&\underset{\text{nach 2)}}{=} \frac{1}{2\pi} \int_a^b \int_{\mathbb{R}} e^{-itx} \varphi_X(t)\, dt\, dx \\
&\underset{\text{S. v. Fubini}}{=} \frac{1}{2\pi} \int_{\mathbb{R}} \varphi_X(t) \int_a^b e^{-itx}\, dx\, dt \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi_X(t) \frac{e^{-ita} - e^{-itb}}{it}\, dt\,,
\end{aligned}$$

somit gilt die Umkehrformel in 1).

Beweis

1. Für alle $a, b, c \in \mathbb{R}$, $a < b$, $c > 0$ gelten mit dem Satz von Fubini und $\varphi_X(t) = \int_{\mathbb{R}} e^{itx} dF_X(x)$ die folgenden Relationen

$$\begin{aligned}
&\frac{1}{2\pi} \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt \\
&= \int_{\mathbb{R}} \left(\frac{1}{2\pi} \int_{-c}^{c} \frac{e^{it(x-a)} - e^{it(x-b)}}{it}\, dt \right) dF_X(x) \\
&= \int_{\mathbb{R}} \left(\frac{1}{\pi} \int_0^c \frac{\sin(t(x-a))}{t}\, dt - \frac{1}{\pi} \int_0^c \frac{\sin(t(x-b))}{t}\, dt \right) dF_X(x)\,,
\end{aligned}$$

weil

$$
\begin{aligned}
\int_{-c}^{c} \frac{e^{it(x-a)}}{it}\,dt &= \underbrace{\int_{-c}^{c} \frac{\cos(t(x-a))}{it}\,dt}_{=0,\text{ da Funktion ungerade ist}} + \int_{-c}^{c} \frac{\sin(t(x-a))}{t}\,dt \\
&= \int_{0}^{c} \frac{\sin(t(x-a))}{t}\,dt + \underbrace{\int_{-c}^{0} \frac{(-1)^3 \sin(-t(x-a))}{-t} d(-t)}_{\text{Substituiere } y=-t} - \\
&\quad - i\int_{0}^{c} \frac{\cos(t(x-a))}{t} - i\underbrace{\int_{-c}^{0} \frac{(-1)^2 \cos(-t(x-a))}{-t}\,d(-t)}_{\text{Substituiere } y=-t} \\
&= 2\int_{0}^{c} \frac{\sin(t(x-a))}{t} - i\underbrace{\left(\int_{0}^{c} - \int_{0}^{c}\right) \frac{\cos(t(x-a))}{t}\,dt}_{=0} \\
&= 2\int_{0}^{c} \frac{\sin(t(x-a))}{t}\,dt\,.
\end{aligned}
$$

Dasselbe gilt für

$$\int_{-c}^{c} \frac{e^{it(x-b)}}{it}\,dt\,.$$

Aus der Analysis weiß man, dass

$$\lim_{c\to\pm\infty} \int_{0}^{c} \frac{\sin(x)}{x}\,dx = \pm\,\frac{\pi}{2}\,, \tag{2.5}$$

somit ist

$$g_{x,a}(c) = \frac{1}{\pi}\int_{0}^{c} \frac{\sin(t(x-a))}{t}\,dt$$

beschränkt für alle $x, a \in \mathbb{R}$ (als Funktion $g_{x,a} : [0,\infty) \to \mathbb{R}$). Es folgt aus dem Satz von Lebesgue über die majorisierte Konvergenz (vgl.

Satz 1.2.7), dass

$$
\begin{aligned}
\lim_{c\to\infty} \frac{1}{2\pi}\int_{-c}^{c} \frac{e^{-ita}-e^{-itb}}{it}\varphi_X(t)\,dt &= \lim_{c\to\infty}\int_{\mathbb{R}} (g_{x,a}(c)-g_{x,b}(c))\,dF_X(x) \\
&= \int_{\mathbb{R}} \lim_{c\to\infty} (g_{x,a}(c)-g_{x,b}(c))\,dF_X(x) \\
&\underset{(2.5)}{=} \int_{\mathbb{R}} \frac{1}{\pi}\cdot\frac{\pi}{2}(\text{sgn }(x-a)-\text{sgn }(x-b))\,dF_X(x) \\
&\underset{a<b}{=} \underbrace{\frac{1}{2}\int_{-\infty}^{a}(-1-(-1))\,dF_X(x)}_{=0} + \frac{1}{2}\underbrace{\int_{\{a\}}(0-(-1))dF_X(x)}_{=0,\text{ da } F_X \text{ stetig in } a \text{ und } b \text{ ist}} + \\
&\quad + \frac{1}{2}\int_{a}^{b}(1-(-1))\,dF_X(x) + \frac{1}{2}\underbrace{\int_{\{b\}}(1-0)dF_X(x)}_{=0,\text{ da } F_X \text{ stetig in } a \text{ und } b \text{ ist}} + \\
&\quad + \underbrace{\frac{1}{2}\int_{b}^{+\infty}(1-1)\,dF_X(x)}_{=0} = \int_a^b dF_X(x) \\
&= F_X(b)-F_X(a)\,.
\end{aligned}
$$

2. Sei $\int_{\mathbb{R}} |\varphi_X(t)|\,dt < \infty$. Definieren wir $f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi_X(t)\,dt$. Es folgt aus dem Satz von Lebesgue über die majorisierte Konvergenz (vgl. Satz 1.2.7), dass $f(x)$ eine stetige Funktion und somit integrierbar auf $[a,b]$ ist. Es gilt

$$
\begin{aligned}
\int_a^b f(x)\,dx &= \int_a^b \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi_X(t)\,dt\,dx \\
&\underset{\text{Fubini}}{=} \frac{1}{2\pi}\int_{\mathbb{R}}\varphi_X(t)\int_a^b e^{-itx}\,dx\,dt \\
&= \lim_{c\to\infty}\frac{1}{2\pi}\int_{-c}^{c}\varphi_X(t)\frac{e^{-ita}-e^{-itb}}{it}\,dt \\
&\underset{1)}{=} F_X(b)-F_X(a)
\end{aligned}
$$

für alle $a<b$, die Stetigkeitspunkte von F_X sind. Daraus folgt

$$
F_X(b) = \int_{-\infty}^{b} f(y)\,dy \quad \forall b\in\mathbb{R} \implies f_X(x) = f(x) \quad \text{für fast alle } x\in\mathbb{R},
$$

da F_X eine Dichte f_X besitzt.

□

Folgerung 2.2.8 *(Eindeutigkeitssatz)*
Die charakteristische Funktion φ_X einer Zufallsvariablen X bestimmt ihre Verteilung P_X eindeutig.

Beweis Zu zeigen ist: Falls X und Y zwei Zufallsvariablen mit charakterisitschen Funktionen φ_X und φ_Y sind, so dass $\varphi_X = \varphi_Y$, dann gilt $P_X = P_Y$. Falls $\varphi_X = \varphi_Y$, dann folgt aus dem Satz 2.2.7 $F_X(b) - F_X(a) = F_Y(b) - F_Y(a)$ für alle Stetigkeitspunkte a, b von F_X und F_Y.
Da jede Verteilungsfunktion höchstens abzählbar viele Sprungstellen haben kann (vgl. [20, Bemerkung 3.2.6]), existiert eine Folge $\{a_n\} \subset \mathbb{R} :\ a_n \to -\infty$, wobei a_n die Stetigkeitspunkte von F_X und F_Y sind. Somit gilt

$$F_X(b) = F_X(b) - \underbrace{\lim_{n\to\infty} F_X(a_n)}_{=0} = F_Y(b) - \underbrace{\lim_{n\to\infty} F_Y(a_n)}_{=0} = F_Y(b)\,,$$

für alle Stetigkeitspunkte $b \in \mathbb{R}$ von F_X und F_Y. Somit gilt auch $F_X = F_Y$, weil sie rechtsseitig stetig sind. Aus [20, Satz 3.2.8] ergibt sich die Gleichung $P_X = P_Y$. $\square$

Wir geben jetzt (ohne Beweis) folgende wichtige Charakterisierungsaussagen über die charakteristische Funktion an:

Satz 2.2.9

1. *Satz von Bochner–Kchintschin:*
   Sei $\varphi : \mathbb{R} \to \mathbb{C}$ eine stetige Funktion mit $\varphi(0) = 1$. Sie ist eine charakteristische Funktion einer Verteilung P_X genau dann, wenn sie positiv semidefinit ist, d.h. für alle $n \in \mathbb{N}$, $t_1, \ldots, t_n \in \mathbb{R}$ und $z_1, \ldots, z_n \in \mathbb{C}$, gilt
   $$\sum_{i,j=1}^{n} \varphi(t_i - t_j) z_i \bar{z}_j \geq 0\,.$$

2. *Satz von Pólya:*
   Sei $\varphi : \mathbb{R} \to \mathbb{R}_+$ eine stetige gerade Funktion mit
   $$\varphi(0) = 1,\ \lim_{t\to\infty} \varphi(t) = 0\,,$$
   die konvex auf $\mathbb{R}_+ = [0, \infty)$ ist. Dann ist φ eine charakteristische Funktion einer Verteilung.

3. *Satz von Marcinkiewicz:*
   Falls eine charakteristische Funktion φ_X die Form $\varphi_X(t) = e^{P(t)}$ hat, wobei $P(t)$ ein Polynom des Grades n ist, dann gilt $n \leq 2$.

Übungsaufgabe 2.2.10 Beweisen Sie die Notwendigkeit der positiven Semidefinitheit im Satz von Bochner–Kchintchin.

Bemerkung 2.2.11 Der Satz von Pólya gibt eine bequeme Methode zur Konstruktion von charakteristischen Funktionen an. So sind z.B.

$$\varphi(t) = e^{-|t|} \tag{2.6}$$

und

$$\varphi(t) = \begin{cases} 1 - |t|, & |t| \leq 1 \\ 0, & |t| > 1 \end{cases} \tag{2.7}$$

gültige charakteristische Funktionen (vgl. Abb. 2.1).

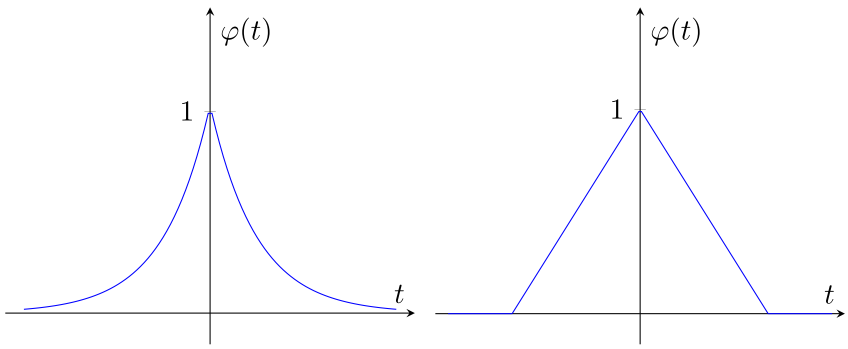


Abbildung 2.1: Grafik der charakteristischen Funktionen (2.6) (links) und (2.7) (rechts)

Andererseits ist nach dem Satz von Marcinkiewicz sofort klar, dass z.B. $\varphi(t) = e^{-t^3}$ oder $\varphi(t) = e^{-t^4}$ keine charakteristischen Funktionen sind.

Übungsaufgabe 2.2.12

1. Zeigen Sie mit Hilfe der charakteristischen Funktionen die Faltungsstabilität der Normal– und Poissonverteilung.

2. Sind die Funktionen

   $$\varphi_1(t) = 1+t,\ \varphi_2(t) = \frac{1}{1+t},\ \varphi_3(t) = \sin t,\ \varphi_4(t) = \cos t,\ \varphi_5(t) = \frac{1}{1+t^2}$$

   charakteristische Funktionen? Falls ja, welche Verteilungen stehen dahinter?

2.3 Erzeugende Funktionen

In diesem Abschnitt betrachten wir Zufallsvariablen X mit Wertebereich $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, d.h. $P(X \in \mathbb{N}_0) = 1$. Die charakteristische Funktion einer solchen Zufallsvariable X mit Zähldichte $\{p_k\}_{k \in \mathbb{N}_0}$, $p_k = P(X = k)$, $k \in \mathbb{N}_0$ sieht folgendermaßen aus:

$$\varphi_X(t) = \mathrm{E}e^{itX} = \sum_{k \in \mathbb{N}_0} e^{itk} p_k = \sum_{k \in \mathbb{N}_0} \left(e^{it}\right)^k p_k = \sum_{k \in \mathbb{N}_0} z^k p_k\,,$$

wobei $z = e^{it} \in B_1(0)$. Daraus folgt, dass die Verteilung von X eindeutig durch das Verhalten der Funktion $g_X(z) = \sum_{k\in\mathbb{N}_0} z^k p_k$, $z \in \mathbb{C}$, $|z| \leq 1$ auf dem Kreis $\{z \in \mathbb{C} : |z| = 1\}$ definiert ist. Die Funktion g_X hat einen Namen: *die erzeugende Funktion von* X.

Definition 2.3.1 Sei X eine diskrete Zufallsvariable mit dem Wertebereich in $\mathbb{N}_0$. Dann heißt

$$g_X(z) = \mathrm{E} z^X, \ |z| \leq 1, \ z \in \mathbb{C}$$

erzeugende Funktion von X.

Alle Eigenschaften der charakteristischen Funktion übertragen sich auf erzeugende Funktionen durch die offensichtliche Substitution der Variablen $z = e^{it}$: $\varphi_X(t) = g_X(e^{it})$, $t \in \mathbb{R}$.

Satz 2.3.2 *(Eigenschaften von erzeugenden Funktionen)*
Sei X eine diskrete Zufallsvariable mit dem Wertebereich $\mathbb{N}_0$ und Zähldichte $\{p_k\}_{k=0}^{\infty}$. Für die erzeugende Funktion $g_X(z) = \sum_{k=0}^{\infty} p_k z^k$, $|z| \leq 1$ gelten folgende Eigenschaften:

1. $g_X(z)$ ist analytisch in $\{z \in \mathbb{C} : |z| < 1\}$.

2. Die Verteilung von X ist eindeutig durch g_X bestimmt:

$$p_k = \frac{g_X^{(k)}(0)}{k!}, \quad k \in \mathbb{Z}_+ .$$

3. (faktorielle Momente)

$$\left.\frac{d^k g_X(z)}{dz^k}\right|_{z=1} = g_X^{(k)}(1) = \mathrm{E}\left(X(X-1) \cdot \ldots \cdot (X-k+1)\right), \ k \in \mathbb{Z}_+ .$$

4. Falls X und Y zwei unabhängige Zufallsvariablen mit dem Wertebereich $\mathbb{Z}$ und erzeugenden Funktion g_X, g_Y sind, dann gilt

$$g_{X+Y}(z) = g_X(z) \cdot g_Y(z), \quad |z| \leq 1 .$$

Übungsaufgabe 2.3.3 Beweisen Sie diesen Satz!

Folgerung 2.3.4 Für Zufallsvariablen X wie im Satz 2.3.2 gilt

$$\mathrm{E}X = g'_X(1), \quad \mathrm{Var}\ X = g''_X(1) + g'_X(1) - \left(g'_X(1)\right)^2 .$$

Beispiel 2.3.5

1. *Binomialverteilung:* Sei $X \sim Bin(n,p)$. Zeigen wir, dass $g_X(z) = (pz + 1 - p)^n$ ist. In der Tat gilt für $z \in \mathbb{C}$

$$\begin{aligned} g_X(z) &= \mathrm{E}z^X = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} z^k = \sum_{k=0}^{n} \binom{n}{k} (zp)^k (1-p)^{n-k} \\ &= (pz+1-p)^n\,, \\ g_X''(z) &= n(n-1)(pz+1-p)^{n-2}p^2\,. \end{aligned}$$

   Dann haben wir nach Satz 2.3.2, 3)

$$\begin{aligned} \mathrm{E}X &= g_X'(1) = \Big(n(pz+1-p)^{n-1}\cdot p\Big)\Big|_{z=1} = np\,, \\ \mathrm{Var}\ X &= g_X''(1) + np - (np)^2 = n(n-1)p^2 + np - (np)^2 = np(1-p)\,. \end{aligned}$$

2. *Poissonverteilung:* Sei $X \sim \mathrm{Poisson}(\lambda)$, $\lambda > 0$. Zeigen wir, dass

$$g_X(z) = e^{-\lambda(1-z)}$$

   für $z \in \mathbb{C}$. Wahrlich, man schreibt

$$\begin{aligned} g_X(z) &= \mathrm{E}z^X = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}z^k = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda z)^k}{k!} = e^{-\lambda}\cdot e^{\lambda z} \\ &= e^{-\lambda(1-z)}\,, \\ g_X''(z) &= \lambda^2 e^{-\lambda(1-z)}\,, \end{aligned}$$

   Somit ergibt sich wieder mit Satz 2.3.2, 3) für den Erwartungswert und die Varianz

$$\begin{aligned} \mathrm{E}X &= g_X'(1) = \Big(\lambda e^{-\lambda(1-z)}\Big)\Big|_{z=1} = \lambda\,, \\ \mathrm{Var}\ X &= g_X''(1) + \lambda - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda\,. \end{aligned}$$

Bemerkung 2.3.6

1. Oft wird g_X auf $[-1,1]$ definiert.

2. Die Definition von g_X als $\mathbb{E}z^X$ kann auch auf Zufallsvariablen $X : \Omega \to \mathbb{Z}$ ausgeweitet werden. Dabei muss auch der Fall $z = 0$ aus dem Definitionsbereich ausgeschlossen werden (vgl. Abbildung 2.3.6.2), also:

$$\{z \in \mathbb{C} \,:\, \delta \le |z| \le 1\},\ \delta > 0.$$

3. Weitere Funktionen, die mit φ_X und g_X verwandt sind:

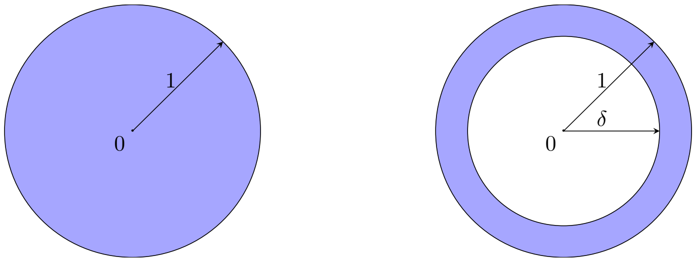


Abbildung 2.2: Definitionsbereich einer erzeugenden Funktion einer Zufallsvariablen X mit Wertebereich (a) $\mathbb{N}$ oder (b) $\mathbb{Z}$

(a) Momenterzeugende Funktion:
$M_X(t) = \mathrm{E}(e^{tX})$, die für $t \in \mathbb{R}$ definiert ist, für die dieser Ausdruck endlich ist. Zum Beispiel ist $M_X(t)$ von $X \geq 0$ wohl definiert zumindest für alle $t \leq 0$.

(b) Kumulantenerzeugende Funktion:
$C_X(t) = \log M_X(t)$, $t \in \mathbb{R}$.

(c) Laplace-Transformierte:
$\hat{l}_X(s) = \mathrm{E}(e^{-sX})$, $s \in \mathbb{C} : \operatorname{Re} s \geq 0$ für fast sicher nicht negative Zufallsvariablen.

Kapitel 3

Konvergenz von Folgen von Zufallsvariablen

Um die Grenzwertsätze und Näherungsformel des Kapitels 4 beweisen zu können, soll definiert werden, in welchem Sinne die Konvergenz von Zufallsvariablen zu verstehen ist. Damit wollen wir uns jetzt befassen:

Definition

Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Sei X eine weitere Zufallsvariable auf $(\Omega, \mathcal{F}, P)$. Man sagt, die Folge $\{X_n\}$ konvergiert gegen X für $n \to \infty$

1. *fast sicher*, falls

$$P(\{\omega \in \Omega : X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega)\}) = 1\,.$$

   Notation: $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$.

2. *stochastisch*, falls für alle $\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0\,.$$

   Notation: $X_n \xrightarrow[n\to\infty]{P} X$.

3. *in* L^r, $r \geq 1$, falls

$$X, X_1, X_2, \ldots \in L^r(\Omega, \mathcal{F}, P) \text{ und } \mathbb{E}|X_n - X|^r \underset{n\to\infty}{\longrightarrow} 0\,.$$

   Notation: $X_n \xrightarrow[n\to\infty]{L^r} X$.
   Einen Spezialfall bildet die L^2–Konvergenz, die man auch *Konvergenz im quadratischen Mittel* nennt.

4. *in Verteilung*, falls
$$F_{X_n}(x) \underset{n\to\infty}{\longrightarrow} F_X(x)$$
für jeden Stetigkeitspunkt x von F_X, wobei F_{X_n} und F_X die Verteilungsfunktionen von X_n und X sind.
Notation: $X_n \xrightarrow[n\to\infty]{d} X$.

5. *schwach*, falls
$$\mathbb{E}h(X_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}h(X)$$
für jede stetige beschränkte Funktion $h : \mathbb{R} \to \mathbb{R}$.
Notation: $X_n \xrightarrow[n\to\infty]{w} X$.

In den folgenden Sätzen werden wir zeigen können, dass die soeben eingeführten Konvergenzarten wie folgt miteinander verbunden sind:

$$\begin{tikzcd}
\boxed{\xrightarrow{L^r}} \arrow[r, Rightarrow] & \boxed{\xrightarrow{P}} \arrow[r, Rightarrow, shift left] & \boxed{\xrightarrow{d}} \arrow[l, Rightarrow, shift left, "X=const"] \arrow[r, Leftrightarrow] & \boxed{\xrightarrow{w}} \\
& \boxed{\xrightarrow{\text{f.s.}}} \arrow[u, Rightarrow] & &
\end{tikzcd} \tag{3.1}$$

3.1 Fast sichere und stochastische Konvergenz

Satz 3.1.1 *(Äquivalente Formulierung der fast sicheren Konvergenz)*
Es gilt $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$ genau dann, wenn $\sup_{k\ge n} |X_k - X| \xrightarrow[n\to\infty]{P} 0\,.$

Beweis

"$\Rightarrow$" Für jedes $\varepsilon > 0$ sei

$$A_\varepsilon = \limsup_{n\to\infty}\{|X_n - X| > \varepsilon\} = \bigcap_n \bigcup_{k\ge n} \{|X_k - X| > \varepsilon\}\,, \tag{3.2}$$

d.h. für alle $n \in \mathbb{N}$ existiert ein $k \ge n :\ |X_k - X| > \varepsilon$. Falls $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$, dann folgt daraus $P(A_\varepsilon) = 0\,,\quad \varepsilon > 0$. Wegen der Stetigkeit von Wahrscheinlichkeitsmaßen gilt außerdem, dass

$$\begin{aligned} 0 = P(A_\varepsilon) &= \lim_{n\to\infty} P\left(\bigcup_{k\ge n} \{|X_k - X| > \varepsilon\}\right) \\ &= \lim_{n\to\infty} P(\sup_{k\ge n} |X_k - X| > \varepsilon)\,, \end{aligned} \tag{3.3}$$

somit
$$\sup_{k\ge n} |X_k - X| \xrightarrow[n\to\infty]{P} 0\,.$$

"$\Leftarrow$" Falls $\sup_{k\geq n}|X_k - X| \xrightarrow{P} 0$, dann gilt $P(A_\varepsilon) = 0$ (siehe die Gleichungskette in (3.3)). Hieraus folgt $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$, weil

$$P\left(\{\omega\in\Omega : X_n(\omega) \underset{n\to\infty}{\not\longrightarrow} X(\omega)\}\right) = P\left(\bigcup_{m=1}^{\infty} A_{\frac{1}{m}}\right) \leq \sum_{m=1}^{\infty} P(A_{\frac{1}{m}}) = 0\,;$$

Dabei gilt die erste Gleichung, weil für $\omega \in \{\omega\in\Omega : X_n(\omega) \underset{n\to\infty}{\not\longrightarrow} X(\omega)\}$ ein $m\in\mathbb{N}$ existiert, sodass es für jedes $n\in\mathbb{N}$ ein $k\geq n$ gibt mit $|X_k(\omega) - X(\omega)| > \frac{1}{m}$.

□

Folgerung 3.1.2

1. Aus $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$ folgt $X_n \xrightarrow[n\to\infty]{P} X$.

2. Falls $\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty\,, \quad \varepsilon > 0$, dann gilt $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$.

Beweis

1. Aus $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$ folgt $\sup_{k\geq n}|X_k - X| \xrightarrow[n\to\infty]{P} 0$ und somit auch $|X_n - X| \xrightarrow[n\to\infty]{P} 0$, also $X_n \xrightarrow[n\to\infty]{P} X$.

2. Sei $A_n = \{|X_n - X| > \varepsilon\}$. Da $\sum_{n=1}^{\infty} P(A_n) < \infty$, ergibt sich nach dem Lemma von Borel-Cantelli (vgl. [20, Lemma 2.2.16])

$$0 = P(\limsup_{n\to\infty} A_n) = P(A_\varepsilon)\,, \quad \varepsilon > 0\,,$$

wobei das Ereignis A_ε in (3.2) eingeführt wurde. Der Rest des Beweises verläuft genauso wie im 2.Teil des Beweises von Satz 3.1.1.

□

Bemerkung 3.1.3

1. Es gilt $X_n \xrightarrow[n\to\infty]{P} X$ genau dann, wenn

$$P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0 \quad \forall\varepsilon > 0\,. \tag{3.4}$$

Nach der Folgerung 3.1.2, 2) ist eine schnelle Konvergenz in (3.4) (z.B. $P(|X_n - X| > \varepsilon) \underset{n\to\infty}{\sim} \frac{1}{n^{1+\delta}}$, $\delta > 0$, da $\sum_{k=1}^{\infty} P(|X_n - X| > \varepsilon < \infty)$) hinreichend für $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$.

2. Aus $X_n \xrightarrow[n\to\infty]{P} X$ folgt nicht $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$, wie dieses Beispiel zeigt:
   Sei $X \equiv 0$ und $\{X_n\}_{n\in\mathbb{N}}$, $X_n \sim \mathit{Bernoulli}(p_n)$ eine Folge von unabhängigen Zufallsvariablen, wobei $p_n \xrightarrow[n\to\infty]{} 0$ und $\sum_{n=1}^{\infty} p_n = \infty$ (z.B. $p_n = \frac{1}{n}$, $n \in \mathbb{N}$). Dann gilt
$$P(|X_n - 0| > \varepsilon) = P(|X_n| > \varepsilon) = P(X_n = 1) = p_n \xrightarrow[n\to\infty]{} 0\,,$$
   also $X_n \xrightarrow[n\to\infty]{P} 0$. Dennoch kann X_n nicht gegen 0 fast sicher konvergieren, denn aus dem Lemma von Borel-Cantelli (vgl. [20, Lemma 2.2.16]) folgt $P(\limsup_{n\to\infty}\{X_n = 1\}) = 1$, weil $\sum_{k=1}^{\infty} p_k = \infty$.

Satz 3.1.4 $X_n \xrightarrow{P} X$ genau dann, wenn für alle Teilfolgen $\{X_{n_i}\}$ von $\{X_n\}$ Teilfolgen $\{X_{n_{i_j}}\} \subset \{X_{n_i}\}$ existieren, sodass $X_{n_{i_j}} \xrightarrow[j\to\infty]{\text{f.s.}} X$.

Beweis

"$\Rightarrow$" Falls $X_n \xrightarrow[n\to\infty]{P} X$, dann gilt $X_{n_i} \xrightarrow[i\to\infty]{P} X$ auch für jede Teilfolge $\{X_{n_i}\} \subset \{X_n\}$. Um die Schreibweise zu vereinfachen, setzen wir $\{X_{n_i}\} = \{X_n\}$. Wir müssen zeigen, dass eine Teilfolge $\{X_{n_j}\}$ von $\{X_n\}$ existiert mit $X_{n_j} \xrightarrow{\text{f.s.}} X$. Da $P(|X_n - X| > \delta) \xrightarrow[n\to\infty]{} 0$ für alle $\delta > 0$, können wir für jedes $n \in \mathbb{N}$ und $\delta = 2^{-n}$ Zahlen n_j so wählen, dass $P(|X_{n_j} - X| > 2^{-n_j}) \leq 2^{-n_j}$. Dann gilt für alle $\varepsilon > 0$, dass
$$\begin{aligned}\sum_{n_j} P(|X_{n_j} - X| > \varepsilon) &\leq \sum_{n_j:\,2^{-n_j} > \varepsilon} P(|X_{n_j} - X| > \varepsilon) + \sum_{n_j:\,2^{-n_j} \leq \varepsilon} P(|X_{n_j} - X| > 2^{-n_j}) \\ &\leq \sum_{n_j:\,2^{-n_j} > \varepsilon} P(|X_{n_j} - X| > \varepsilon) + \sum_{n=1}^{\infty} 2^{-n} < \infty\,.\end{aligned}$$
Folgerung 3.1.2, 2) ergibt $X_{n_j} \xrightarrow{\text{f.s.}} X$.

"$\Leftarrow$" Nehmen wir an, dass $X_n \underset{n\to\infty}{\overset{P}{\not\longrightarrow}} X$. Dann existieren $\varepsilon > 0$, $\delta > 0$ und eine Teilfolge $\{X_{n_i}\}$ von $\{X_n\}$ mit $P(|X_{n_i} - X| > \varepsilon) > \delta$ für alle $i \in \mathbb{N}$. Nach der Folgerung 3.1.2, 1) kann somit keine Teilfolge $\{X_{n_{i_j}}\} \subset \{X_{n_i}\}$ existieren, für die $X_{n_{i_j}} \xrightarrow{\text{f.s.}} X$ gilt.

□

Bemerkung 3.1.5 Die Stochastische Konvergenz ist metrisierbar mit Metrik
$$d(x,y) = \mathbb{E}\left(\frac{|X-Y|}{1+|X-Y|}\right). \tag{3.5}$$

Für diese Metrik gilt

$$X_n \underset{n\to\infty}{\overset{P}{\to}} X \iff d(X_n, X) \to 0 \text{ für } n \to \infty.$$

Somit wird der topologische Raum der Zufallsvariablen mit Topologie der stochastischen Konvergenz zu einem metrischen Raum, in dem
$d(X,Y) = 0 \iff X = Y\ f.s.$ Dieser Raum ist insbesondere vollständig, d.h., jede Cauchy-Folge der Zufallsvariablen $\{X_n\}$ hat eine Zufallsvariable X als Grenzwert:

$$X_n \underset{n\to\infty}{\overset{P}{\longrightarrow}} X \iff \forall \varepsilon, \delta > 0,\ \exists N : \forall n, m > N\ P(|X_n - X_m| > \varepsilon) < \delta.$$

Übungsaufgabe 3.1.6 Sei d definiert wie (3.5). Zeige, dass d eine Metrik ist und für eine Folge von Zufallsvariablen $\{X_n\}$ und eine Zufallsvariable X gilt

$$X_n \underset{n\to\infty}{\overset{P}{\longrightarrow}} X \iff \forall \varepsilon, \delta > 0,\ \exists N : \forall n, m > N\ P(|X_n - X_m| > \varepsilon) < \delta.$$

3.2 L^r–Konvergenz

Satz 3.2.1 *(L^r–Konvergenz und stochastische Konvergenz)*

1. Sei $X_n \in L^r$, $n \in \mathbb{N}$, $X \in L^r$, $r \geq 1$ und $X_n \underset{n\to\infty}{\overset{L^r}{\longrightarrow}} X$. Dann gilt auch $X_n \overset{L^s}{\longrightarrow} X,$ für alle $s < r$.
2. Sei $X_n \in L^r, n \in \mathbb{N}, X \in L^r$ und $X_n \overset{L^r}{\longrightarrow} X, r \geq 1$. Dann gilt $X_n \overset{P}{\longrightarrow} X$.

Beweis

1. Sei $X_n \overset{L^r}{\longrightarrow} X,$ also $\mathbb{E}|X_n - X|^r \underset{n\to\infty}{\longrightarrow} 0$. Aus der Ljapunow–Ungleichung (vgl. [20, Folgerung 4.6.6]) ergibt sich
$$0 \leq \mathbb{E}|X_n - X|^s \leq (\mathbb{E}|X_n - X|^r)^{\frac{s}{r}} \underset{n\to\infty}{\longrightarrow} 0\,,$$
somit gilt $X_n \underset{n\to\infty}{\overset{L^s}{\longrightarrow}} X$.
2. Aus der Markow–Ungleichung (vgl. [20, Satz 4.6.1]) folgt für alle $\varepsilon > 0$, dass
$$P(|X_n - X| > \varepsilon) = P(|X_n - X|^r > \varepsilon^r) \leq \frac{\mathbb{E}|X_n - X|^r}{\varepsilon^r} \underset{n\to\infty}{\longrightarrow} 0\,,$$
somit gilt $X_n \underset{n\to\infty}{\overset{P}{\longrightarrow}} X$.

□

Bemerkung 3.2.2 Der Satz 3.2.1, 2) behauptet, dass aus der L^r–Konvergenz die stochastische Konvergenz folgt. Zeigen wir, dass es keine "genau dann"–Aussage sein kann:

1. Falls $\mathbb{E}|X_n - X|^r = +\infty\,, n \in \mathbb{N}\,,$ jedoch $X_n \xrightarrow[n\to\infty]{P} X$, so gilt $X_n \underset{n\to\infty}{\overset{L^r}{\not\to}} X$. Zum Beispiel wähle $X_n = 2^{-n}Y$ mit $Y \sim Cauchy(0,1)\,, r = 1$ und $X = 0$.

2. Seien
$$X_n = \begin{cases} c^n, & \text{mit Wkt. } p_n \\ 0, & \text{sonst,} \end{cases} \qquad n \in \mathbb{N}\,.$$
Dabei sei $p_n \xrightarrow[n\to\infty]{} 0$, sodass $c_n^r p_n \underset{n\to\infty}{\not\to} 0$. Dann gilt
$$P(|X_n| > \varepsilon) = P(X_n \neq 0) = p_n \xrightarrow[n\to\infty]{} 0\,, \varepsilon > 0\,,$$
und somit $X_n \xrightarrow[n\to\infty]{P} 0$. Dennoch gilt für alle $r > 0$
$$\mathbb{E}|X_n - 0|^r = \mathbb{E}X_n^r = (c_n)^r \cdot P(X_n = c_n) + 0 \cdot P(X_n = 0) = c_n^r p_n \underset{n\to\infty}{\not\to} 0$$
also $X_n \underset{n\to\infty}{\overset{L^r}{\not\to}} 0$. Wähle zum Beispiel $c_n = e^n, p_n = \frac{1}{n}$.

3.3 Konvergenz in Verteilung

Satz 3.3.1 *(Konvergenz in Verteilung und stochastisch)*
Aus $X_n \xrightarrow[n\to\infty]{P} X$ folgt $X_n \xrightarrow[n\to\infty]{d} X$.

Beweis Falls $X_n \xrightarrow[n\to\infty]{P} X$, dann
$$P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0\,, \varepsilon > 0.$$
Dann
$$\begin{aligned} F_{X_n}(x) &= P(X_n \le x) \\ &= P(X_n \le x,\, |X_n - X| \le \varepsilon) + P(X_n \le x,\, |X_n - X| > \varepsilon) \\ &\le P(X \le x + \varepsilon) + P(|X_n - X| > \varepsilon) \\ &= F_X(x + \varepsilon) + P(|X_n - X| > \varepsilon)\,, \end{aligned}$$
für alle $\varepsilon > 0$, $n \in \mathbb{N}$ und $x \in \mathbb{R}$. Somit gilt
$$\limsup_{n\to\infty} F_{X_n}(x) \le F_X(x + \varepsilon),\ \varepsilon > 0\,,$$

das heißt, $\limsup_{n\to\infty} F_{X_n}(x) \le F_X(x)\,, x \in \mathbb{R}\,,$ weil $F_X(x)$ rechtsseitig stetig ist.
Ähnlich wie oben kann man zeigen, dass

$$F_X(x-\varepsilon) \le F_{X_n}(x) + P(|X_n - X| > \varepsilon)\,, \varepsilon > 0\,,$$

woraus folgt, dass $F_X(x) \le \liminf_{n\to\infty} F_{X_n}(x)$ für jeden Stetigkeitspunkt $x \in \mathbb{R}$ von F_X. Hieraus folgt, dass $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ für jeden Stetigkeitspunkt x von F_X, also $X_n \xrightarrow{d} X$ für $n \to \infty$. □

Bemerkung 3.3.2

1. Die Aussage des Satzes 3.3.1 lässt sich im Allgemeinen nicht umkehren. Dies zeigt folgendes Beispiel:
   Sei $X \sim N(0,1)\,,$ und $X_n = -X,\ n \in \mathbb{N}$. Da die Verteilung von X symmetrisch ist, gilt $X_n = -X \stackrel{d}{=} X,\ n \in \mathbb{N}$ und damit $X_n \xrightarrow[n\to\infty]{d} X\,.$
   Dennoch gilt nicht $X_n \xrightarrow{P} X$, weil

   $$P(|X_n - X| > \varepsilon) = P(|2X| > \varepsilon) = P(|X| > \frac{\varepsilon}{2}) = const \underset{n\to\infty}{\not\longrightarrow} 0\,.$$

2. Falls jedoch $X \equiv const$ ist, dann gilt die Umkehrung des Satzes 3.3.1, vgl. die folgende Aussage.

Satz 3.3.3 Falls $X_n \xrightarrow[n\to\infty]{d} c$ für $c \in \mathbb{R}$, dann gilt $X_n \xrightarrow[n\to\infty]{P} c\,.$

Beweis Falls $X_n \xrightarrow[n\to\infty]{d} X \equiv c$, dann gilt für alle $x \ne c$

$$F_{X_n}(x) \xrightarrow[n\to\infty]{} F_c(x) = \begin{cases} 1, & x \ge c \\ 0, & x < c \end{cases}.$$

Dann haben wir für alle $\varepsilon > 0$ und $n \in \mathbb{N}$

$$\begin{aligned} P(|X_n - c| > \varepsilon) &= P(X_n < c-\varepsilon) + P(X_n > c+\varepsilon) \\ &\le F_{X_n}(c-\varepsilon) + (1 - F_{X_n}(c+\varepsilon)) \\ &\xrightarrow[n\to\infty]{} F_c(c-\varepsilon) + (1 - F_c(c+\varepsilon)) = 0 + 1 - 1 = 0\,, \end{aligned}$$

vgl. die Abbildung 3.1 von $F_c(x)$. Somit gilt $X_n \xrightarrow[n\to\infty]{P} c\,.$ □

Satz 3.3.4 *(Konvergenz in Verteilung und schwache Konvergenz)*
Die Konvergenz in Verteilung und die schwache Konvergenz sind äquivalent.

Beweis Zu zeigen ist $X_n \xrightarrow[n\to\infty]{d} X \iff \mathbb{E}h(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}h(X)$ für beliebige beschränkte stetige Funktionen h.

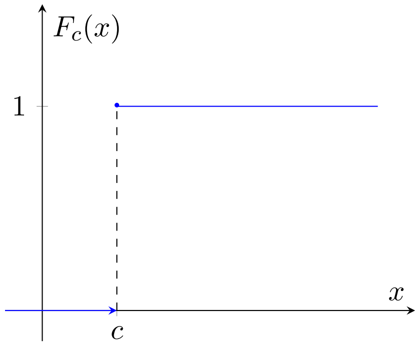


Abbildung 3.1: Verteilungsfunktion F_c

"⇒" Sei $X_n \xrightarrow[n\to\infty]{d} X$, und h eine beschränkte stetige Funktion auf $\mathbb{R}$. Dann sei $b = \sup_{x\in\mathbb{R}} |h(x)| \in (0,\infty)$. Für jedes $\varepsilon > 0$ wählen wir $\nu > 0$ so, dass ν und $-\nu$ Stetigkeitspunkte von F_X sind und $P(|X| > \nu) < \frac{\varepsilon}{b}$. So ein ν existiert, weil die Verteilungsfunktion F_X höchstens abzählbar viele Sprungstellen haben kann und weil $\lim_{x\to\infty} P(|X| > x) = 0$. Aus $F_{X_n}(\pm\nu) \xrightarrow[n\to\infty]{} F_X(\pm\nu)$ folgt außerdem $P(|X_n| > \nu) < \frac{2\varepsilon}{b}$ für hinreichend große $n \in \mathbb{N}$. Da h beschränkt und stetig ist, lässt sich h durch Treppenfunktionen approximieren, d.h.

$$\text{für } \varepsilon, \nu > 0 \quad \text{existiert } \ g(x) = \sum_{j=1}^{k} a_i \cdot I(x_{j-1} < x \leq x_j)$$

für $a_1, \ldots, a_k \in \mathbb{R}$ und Stetigkeitspunkte $-\nu = x_0 < \ldots < x_k = \nu$ von F_X, sodass

$$\sup_{x\in[-\nu,\nu]} |h(x) - g(x)| < \varepsilon\,. \tag{3.6}$$

Wegen $h(x) = h(x) \cdot I(|x| \leq \nu) + h(x) \cdot I(|x| > \nu), \quad x \in [-\nu, \nu]$ gilt

$$|\mathbb{E}h(X_n) - \mathbb{E}h(X)| \leq |\underbrace{\mathbb{E}[h(X_n) \cdot I(|X_n| \leq \nu)] - \mathbb{E}[h(X) \cdot I(|X| \leq \nu)]}_{=:I_1}|$$
$$+ \underbrace{\mathbb{E}[|h(X_n)| \cdot I(|X_n| > \nu)]}_{=:I_2} + \underbrace{\mathbb{E}[|h(X)| \cdot I(|X| > \nu)]}_{=:I_3} = I_1 + I_2 + I_3\,.$$

Es gilt

$$\begin{aligned}
I_2 &\leq b \cdot P(|X_n| > \nu) \leq b \cdot \frac{2\varepsilon}{b} = 2\varepsilon\,,\\
I_3 &\leq b \cdot P(|X| > \nu) \leq b \cdot \frac{\varepsilon}{b} = \varepsilon\,,\\
I_1 &\leq |\mathbb{E}[h(X_n) \cdot I(|X_n| \leq \nu)] - \mathbb{E}g(X_n)| + |\mathbb{E}g(X_n) - \mathbb{E}g(X)|\\
&\quad + |\mathbb{E}g(X) - \mathbb{E}\left(h(X) \cdot I(|X| \leq \nu)\right)|\\
&< \varepsilon + |\mathbb{E}g(X_n) - \mathbb{E}g(X)| + \varepsilon
\end{aligned}$$

wegen (3.6). Da x_i Stetigkeitspunkte von F_X sind und wegen $X_n \xrightarrow[n\to\infty]{d} X$

$$F_{X_n}(x_j) \underset{n\to\infty}{\longrightarrow} F_X(x_j)\,, \quad j = 0, \ldots, k\,,$$

gilt zusätzlich

$$\begin{aligned}
\mathbb{E}g(X_n) &= \sum_{j=1}^{k} a_j (F_{X_n}(x_i) - F_{X_n}(x_{j-1}))\\
&\underset{n\to\infty}{\longrightarrow} \sum_{j=1}^{k} a_j (F_X(x_j) - F_X(x_{j-1})) = \mathbb{E}g(X)\,.
\end{aligned}$$

Somit gilt $|\mathbb{E}g(X_n) - \mathbb{E}g(X)| < \varepsilon$ und $|\mathbb{E}h(X_n) - \mathbb{E}h(X)| < 6\varepsilon$ für hinreichend große n. Damit haben wir gezeigt, dass

$$\mathbb{E}h(X_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}h(X) \text{ und somit } X_n \xrightarrow[n\to\infty]{w} X\,.$$

"$\Leftarrow$" Sei $X_n \xrightarrow[n\to\infty]{w} X$, also

$$\mathbb{E}h(X_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}h(X) \tag{3.7}$$

für jede beschränkte stetige Funktion h auf $\mathbb{R}$. Zu zeigen ist, dass $F_{X_n} \underset{n\to\infty}{\longrightarrow} F_X(x)$ für jeden Stetigkeitspunkt x von F_X.
Fixieren wir so ein $x \in \mathbb{R}$ und $\varepsilon > 0$. Dann existiert eine beschränkte stetige Funktion h auf $\mathbb{R}$, sodass (vgl. Abb. 3.2)

$$I(y \leq x) \leq h(y) \leq I(y \leq x + \varepsilon) \tag{3.8}$$

und somit

$$\begin{aligned}
F_{X_n}(x) &= \mathbb{E}I(X_n \leq x) \leq \mathbb{E}h(X_n)\,,\\
\mathbb{E}h(X) &\leq \mathbb{E}I(X \leq x + \varepsilon) = F_X(x + \varepsilon)\,.
\end{aligned}$$

Daraus folgt

$$\limsup_{n\to\infty} F_{X_n}(x) \leq \lim_{n\to\infty} \mathbb{E}h(X_n) \underset{(3.7)}{=} \mathbb{E}h(X) \leq F_X(x + \varepsilon)\,.$$

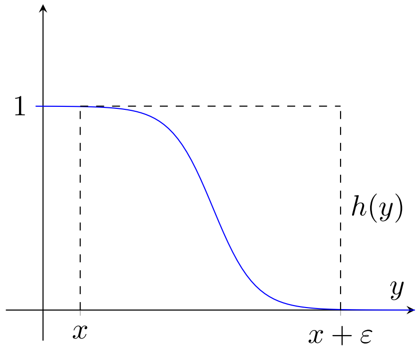


Abbildung 3.2: Grafik von h mit der Eigenschaft (3.8)

Da ε beliebig ist, ergibt sich $\limsup_{n\to\infty} F_{X_n}(x) \le F_X(x)$. Nun, um die Ungleichung $F_X(x) \le \liminf_{n\to\infty} F_{X_n}(x)$ zu zeigen, wählen wir eine stetige beschränkte Funktion h auf $\mathbb{R}$ mit der Eigenschaft (vgl. Abb. 3.3)

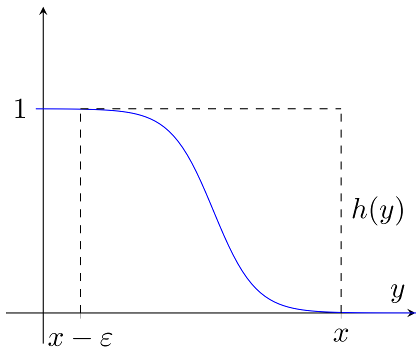


Abbildung 3.3: Grafik von h mit der Eigenschaft (3.9)

$$I(y \le x-\varepsilon) \le h(y) \le I(y \le x) \quad y \in \mathbb{R}\,. \tag{3.9}$$

Dann gilt

$$F_X(x-\varepsilon) = \mathbb{E}I(X \le x-\varepsilon) \le \mathbb{E}h(X) = \lim_{n\to\infty} \mathbb{E}h(X_n) \le \liminf_{n\to\infty} F_{X_n}(x)\,;$$

Für jeden Stetigkeitspunkt $x \in \mathbb{R}$ von F_X bekommt man

$$F_X(x) = \lim_{\varepsilon\to+0} F_X(x-\varepsilon) \le \liminf_{n\to\infty} F_{X_n}(x)\,.$$

Insgesamt gilt für jeden Stetigkeitspunkt $x \in \mathbb{R}$ von F_X

$$F_X(x) \quad \leq \quad \liminf_{n\to\infty} F_{X_n}(x) \leq \limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x)\,,$$

das heißt, $F_{X_n}(x) \to F_X(x)$ für $n \to \infty$. Somit ist bewiesen, dass

$$X_n \xrightarrow{d} X\,.$$

□

Bemerkung 3.3.5 Aus dem Beweis der Hinlänglichkeit in Satz 3.3.4 sieht man leicht, dass die schwache Konvergenz auch folgendermaßen äquivalent eingeführt werden kann: $X_n \xrightarrow{w} X$, falls $\mathbb{E}h(X_n) \to \mathbb{E}h(X)$ für alle Funktionen $h \in C_b^k$ für ein $k \in \mathbb{N}$, wobei die Klasse

$$C_b^k = \{h : \mathbb{R} \to \mathbb{R} \,|\, h \text{ beschränkt und k–mal gleichmäßig stetig diff'bar auf } \mathbb{R}\}\,.$$

Satz 3.3.6 Seien $\{X_n\}_{n\in\mathbb{N}}$, X Zufallsvariablen definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit den charakteristischen Funktionen φ_{X_n}, $n \in \mathbb{N}$ und φ_X. Es gilt $X_n \underset{n\to\infty}{\xrightarrow{d}} X$ genau dann, wenn

$$\varphi_{X_n}(t) \underset{n\to\infty}{\longrightarrow} \varphi_X(t) \quad \forall t \in \mathbb{R}\,.$$

Für den Beweis dieses Satzes benötigen wir die folgenden zwei Hilfssätze:

Lemma 3.3.7 Die Folge von Verteilungsfunktionen $\{F_n\}$ konvergiert schwach gegen eine monoton nicht-fallende Funktion F, falls

$$F_n(x) \underset{n\to\infty}{\longrightarrow} F(x) \quad \forall x \in D,$$

wobei D eine dichte Teilmenge von $\mathbb{R}$ ist.

Beweis Die schwache Konvergenz von F_n gegen F bedeutet per Definition, dass $F_n(x) \underset{n\to\infty}{\longrightarrow} F(x)$ für alle Stetigkeitspunkte x von F. Sei x so ein Punkt. Für alle $x', x'' \in D : x' \leq x \leq x''$ gilt $F_n(x') \leq F_n(x) \leq F_n(x'')$ und somit $\lim_{n\to\infty} F_n(x') \leq \liminf_{n\to\infty} F_n(x) \leq \limsup_{n\to\infty} F_n(x) \leq \lim_{n\to\infty} F_n(x'')$.
Sei $x' \uparrow x$, $x'' \downarrow x$ innerhalb von D. Da x eine Stetigkeitsstelle von F ist, gilt

$$\begin{aligned} F(x) &= \lim_{x'\uparrow x} \lim_{n\to\infty} F_n(x') \leq \liminf_{n\to\infty} F_n(x) \leq \limsup_{n\to\infty} F_n(x) \\ &\leq \lim_{x''\downarrow x} \lim_{n\to\infty} F_n(x'') = F(x), \end{aligned}$$

somit $\lim_{n\to\infty} F_n(x) = F_n(x)$. □

Lemma 3.3.8 (*Helly*)
Für jede Folge von Verteilungsfunktionen $\{F_n\}$ existiert es eine Teilfolge $\{F_{n_k}\} \subset \{F_n\}$, die schwach gegen eine monoton nicht fallende rechtsseitig stetige Funktion F konvergiert.

Beweis Sei $D = \{x_n\}$ eine beliebige abzählbare dichte Teilmenge von $\mathbb{R}$. Die Zahlenfolge $\{F_n(x_1)\}$ ist beschränkt in $\mathbb{R}$, weil $0 \leq F_n(x) \leq 1\,, \quad x \in \mathbb{R}$, deshalb enthält sie eine konvergente Teilfolge $\{F_{1n}(x_1)\}$.
Sei $\lim\limits_{n\to\infty} F_{1n}(x_1) =: F(x_1)$. Betrachte nun $\{F_{1n}(x_2)\}$, und sie besitzt eine konvergente Teilfolge $\{F_{2n}(x_2)\}$ mit Grenzwert $F(x_2)$. Dabei gilt $\lim\limits_{n\to\infty} F_{2n}(x_1) = F(x_1)$. Falls wir diese iterative Konstruktion fortsetzen, gibt es für jedes $k \in \mathbb{N}$ Teilfolgen $\{F_{kn}(x_i)\}$, $i = 1, \ldots, k$, sodass $\lim\limits_{n\to\infty} F_{kn}(x_i) = F(x_i)$. Betrachten wir die diagonale Teilfolge $\{F_{nn}(x)\}$. Für alle $x \in D$ gilt $\lim\limits_{n\to\infty} F_{nn}(x) = F(x)$, wobei $F : D \to [0,1]$ eine monoton nicht-fallende beschränkte Funktion ist (wegen solcher Eigenschaften von F_n für alle n). Per rechtsseitige Stetigkeit lässt sie sich bis auf $F : \mathbb{R} \to [0,1]$ fortsetzen, die monoton nicht-fallend ist. Nach Lemma 3.3.7 konvergiert $\{F_{nn}\}$ gegen F schwach. □

Nun können wir den Satz 3.3.6 beweisen.

Beweis

“$\Rightarrow$”

$$\begin{aligned}\varphi_{X_n}(t) &= \mathbb{E}e^{itX_n} = \mathbb{E}\cos(tX_n) + i \cdot \mathbb{E}\sin(tX_n)\\ &\underset{n\to\infty}{\longrightarrow} \mathbb{E}\cos(tX) + i \cdot \mathbb{E}\sin(tX) = \varphi_X(t)\,,\end{aligned}$$

weil $X_n \xrightarrow{d} X$ nach dem Satz 3.3.4 bedeutet, dass

$$\mathbb{E}g(X_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}g(X)$$

für beliebige beschränkte Funktionen g. Die Funktionen $g(x) = \sin(tx)$ und $g(x) = \cos(tx)$ sind aber beschränkt und stetig für alle $t \in \mathbb{R}$.

“$\Leftarrow$” Zeigen wir Folgendes: Falls $\varphi_{X_n}(t) \underset{n\to\infty}{\longrightarrow} \varphi(t)$ für alle $t \in \mathbb{R}$ und $\varphi(t)$ stetig in t = 0, dann ist φ eine charakteristische Funktion einer Zufallsvariable X mit Verteilungsfunktion F_X, und $X_n \underset{n\to\infty}{\xrightarrow{d}} X$.

Zur Vereinfachung der Bezeichnung verwenden wir

$$F_n := F_{X_n},\ F := F_X,\ \varphi_n := \varphi_{X_n},\ \varphi := \varphi_X, \quad n \in \mathbb{N}.$$

Nach Lemma 3.3.8 existiert eine Teilfolge $\{F_{n_k}\}_{k\in\mathbb{N}}$ von $\{F_n\}_{n\in\mathbb{N}}$, die schwach gegen eine monoton nicht fallende rechtsseitig stetige Funktion F konvergiert. Da $F : \mathbb{R} \to [0,1]$ ist, sollten wir zeigen, dass F eine Verteilungsfunktion ist, indem wir zeigen, dass

$$\lim_{x\to-\infty} F(x) = 0, \qquad \lim_{x\to+\infty} F(x) = 1. \tag{3.10}$$

Sei Var F die Variation von F, also

$$\text{Var } F = \sup_{\substack{n\in\mathbb{N}\\ t_0<\cdots<t_n}} \sum_{j=1}^{n} |F(t_j) - F(t_{j-1})|.$$

Die Relationen (3.10) sind äquivalent zu Var $F = 1$, da wegen der Monotonie von F gilt

$$\begin{aligned}\text{Var } F &= \sup_{\substack{n\in\mathbb{N}\\ t_0<\cdots<t_n}} \sum_{j=1}^{n} (F(t_j) - F(t_{j-1})) = \sup_{t_0<t_n} (F(t_n) - F(t_0)) \\ &= \sup_{t\in\mathbb{R}} F(t) - \inf_{t\in\mathbb{R}} F(t).\end{aligned}$$

Nehmen wir das Gegenteil an, also $0 < \text{Var } F = \delta < 1$, und führen wir es zu einem Widerspruch. Da $\varphi_n(t) \underset{n\to\infty}{\longrightarrow} \varphi(t), \quad t \in \mathbb{R}$, und $\varphi_n(t) = \mathbb{E}e^{itX_n}$, $n \in \mathbb{N}$ beschränkt auf $\mathbb{R}$ und somit integrierbar auf jedem Segment $[-\tau;\tau]$ sind, ist auch φ integrierbar auf $[-\tau;\tau]$, $\tau > 0$. Wähle ein $\varepsilon < 1 - \delta$. Da $\varphi(0) = \lim_{n\to\infty} \varphi_n(0) = 1$, und φ stetig in 0 ist, existiert für $\varepsilon > 0$ ein $\tau > 0$ mit der Eigenschaft

$$\frac{1}{2\tau}\left|\int_{-\tau}^{\tau} \varphi(t)dt\right| > 1 - \varepsilon/2 > \delta + \varepsilon/2.$$

In der Tat folgt die rechte Ungleichung aus $\varepsilon < 1 - \delta$ durch $\varepsilon/2 + \varepsilon/2 < 1 - \delta$. Die linke Ungleichung folgt aus der Tatsache, dass

$$\left|\int_{-\tau}^{\tau} \varphi(t)dt\right|^2 \geq \left(\int_{-\tau}^{\tau} Re\varphi(t)dt\right)^2$$

und aus der Abschätzung $Re\,\varphi(t) > 1 - \varepsilon/2$ für $t \in \mathbb{R} : |t| \leq \tau$, die aus $|\varphi(t) - 1| < \varepsilon/2$ für $|t| \leq \tau$ folgt.
Zeigen wir nun, dass

$$\frac{1}{2\tau}\left|\int_{-\tau}^{\tau} \varphi(t)dt\right| \leq \delta + \varepsilon/2,$$

was uns zum gewünschten Widerspruch führen wird.

Für alle $k \in \mathbb{N}$ gilt wegen des Satzes von Fubini:

$$\begin{aligned}\left|\int_{-\tau}^{\tau} \varphi_{n_k}(t)dt\right| &= \left|\int_{\mathbb{R}}\int_{-\tau}^{\tau} e^{itx}dtdF_{n_k}(x)\right| \\ &\leq \left|\int_{|x|<a}\int_{-\tau}^{\tau} e^{itx}dtdF_{n_k}(x)\right| + \left|\int_{|x|\geq a}\int_{-\tau}^{\tau} e^{itx}dtdF_{n_k}(x)\right| \\ &\leq \underbrace{\int_{-a}^{a}\left|\int_{-\tau}^{\tau} e^{itx}dt\right|dF_{n_k}(x)}_{=I_1} + \underbrace{\int_{|x|\geq a}\left|\int_{-\tau}^{\tau} e^{itx}dt\right|dF_{n_k}(x)}_{=I_2}\end{aligned}$$

Für das Integral I_1 gilt

$$\begin{aligned} I_1 &\leq \int_{-a}^{a} \int_{-\tau}^{\tau} \underbrace{|e^{itx}|}_{\leq 1} \, dt \, dF_{n_k}(x) \leq \int_{-a}^{a} 2\tau \, dF_{n_k}(x) \\ &= 2\tau(F_{n_k}(a) - F_{n_k}(-a)) \leq 2\tau(F(a) - F(-a) + \varepsilon/4) \\ &< 2\tau(\delta + \varepsilon/4) \end{aligned}$$

für ausreichend große k. Für die Abschätzung des Integrals I_2 bemerken wir, dass

$$\left| \int_{-\tau}^{\tau} e^{itx} dt \right| = \left| \frac{e^{i\tau x} - e^{-i\tau x}}{ix} \right| = \left| \frac{2 \sin \tau x}{x} \right| \leq \frac{2}{|x|}.$$

Das heißt es gilt:

$$I_2 \leq \frac{2}{a} \underbrace{\int_{|x| \geq a} dF_{n_k}(x)}_{\leq 1} \leq 2\tau\varepsilon/4$$

für $a > \frac{4}{\tau\varepsilon}$. Nun erhalten wir insgesamt, dass $\left| \int_{-\tau}^{\tau} \varphi_{n_k}(t) dt \right| \leq 2\tau(\delta + \varepsilon/2)$. Somit haben wir bewiesen, dass für $k \to \infty$

$$\frac{1}{2\tau} \left| \int_{-\tau}^{\tau} \varphi(t) dt \right| \leq \delta + \varepsilon/2,$$

was zum Widerspruch führt.

Warum ist die Grenzwertfunktion F identisch für alle Teilfolgen $\{F_{n_k}\}$? Nehmen wir das Gegenteil an, also

$$\begin{aligned} &\exists \{F_{n_k}\} \subset \{F_n\}, \\ &\{\tilde{F}_{n_k}\} \subset \{F_n\} : \\ &\qquad F_{n_k} \underset{k \to \infty}{\to} F, \tilde{F}_{n_k} \underset{k \to \infty}{\to} \tilde{F}, \, F \neq \tilde{F}. \end{aligned}$$

Dann nach Satz 3.3.6: 1) konvergieren die charakteristischen Funktionen

$$\begin{aligned} \varphi_{n_k}(t) &= \int e^{its} dF_{n_{k_j}}(s) \underset{k \to \infty}{\to} \tilde{\varphi}_1(t), \\ \tilde{\varphi}_{n_k}(t) &= \int e^{its} d\tilde{F}_{n_{k_j}}(s) \underset{k \to \infty}{\to} \tilde{\varphi}_2(t), \quad t \in \mathbb{R}, \end{aligned}$$

wobei $\tilde{\varphi}_1 = \tilde{\varphi}_2 = \varphi$.

Da der Zusammenhang $\tilde{\varphi}_1 \leftrightarrow F$, $\tilde{\varphi}_2 \leftrightarrow \tilde{F}$ eindeutig ist, so soll $F = \tilde{F}$ sein.

Dann bedeutet es die schwache Konvergenz von $\{F_n\}$ zu F.

□

3.4 Konvergenz der Funktionale von Zufallsvariablen

In diesem Abschnitt wird untersucht, ob die oben betrachteten Konvergenzarten bei Addition, Multiplikation und Anwendung stetiger Funktionale auf Folgen von Zufallsvariablen erhalten bleiben.

Satz 3.4.1 *(Addition von Zufallsvariablen)*
Seien $\{X_n\}_{n\in\mathbb{N}}$, X, $\{Y_n\}_{n\in\mathbb{N}}$, Y Zufallsvariablen definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$.

1. Falls $X_n \underset{n\to\infty}{\longrightarrow} X$ und $Y_n \underset{n\to\infty}{\longrightarrow} Y$ (fast sicher, stochastisch oder in L^r, $r \geq 1$), dann konvergiert auch ihre Summe $X_n + Y_n \underset{n\to\infty}{\longrightarrow} X + Y$ im selben Sinne.

2. (*Satz von Slutsky*)
Die Aussage 1. lässt sich auf die Konvergenz in Verteilung nur beschränkt übertragen, es gilt nämlich: $X_n + Y_n \xrightarrow[n\to\infty]{d} X + c$, falls $X_n \xrightarrow[n\to\infty]{d} X$ und $Y_n \xrightarrow[n\to\infty]{d} c \equiv const$.

Beweis

1. *Fast sichere Konvergenz:*
Falls $X_n \xrightarrow{\text{f.s}} X$, $Y_n \xrightarrow{\text{f.s.}} Y$, dann gilt $X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega)$, $Y_n(\omega) \underset{n\to\infty}{\longrightarrow} Y(\omega)$ für fast alle $\omega \in \Omega$ und somit
$$X_n(\omega) + Y_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega) + Y(\omega)$$
für fast alle $\omega \in \Omega$. Daraus folgt $X_n + Y_n \xrightarrow[n\to\infty]{\text{f.s.}} X + Y$.
Stochastische Konvergenz:
Falls $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, dann gilt: Für alle Teilfolgen $\{n_i\}$ von $\mathbb{N}$ existiert eine Teilfolge $\{n_{i_j}\} \subset \{n_i\}$, sodass
$$X_{n_{i_j}} \xrightarrow[j\to\infty]{\text{f.s.}} X,\ Y_{n_{i_j}} \xrightarrow[j\to\infty]{\text{f.s.}} Y$$
nach dem Satz 3.1.4. Aus dem hier vorangegangenen Beweis der fast sicheren Konvergenz folgt sofort $X_{n_{i_j}} + Y_{n_{i_j}} \xrightarrow[j\to\infty]{\text{f.s.}} X + Y$, und somit nach der erneuten Anwendung des Satzes 3.1.4
$$X_n + Y_n \xrightarrow[n\to\infty]{P} X + Y\,.$$
Konvergenz in L^r:
Sei $X_n \xrightarrow{L^r} X$, $Y_n \xrightarrow{L^r} Y$, $r \geq 1$. Das heißt, dass
$$\mathbb{E}|X_n - X|^r \underset{n\to\infty}{\longrightarrow} 0,\ \mathbb{E}|Y_n - Y|^r \underset{n\to\infty}{\longrightarrow} 0\,.$$

Aus der Minkowski–Ungleichung (vgl. [20, Satz 4.6.10]) folgt

$$(\mathbb{E}|X_n + Y_n - (X+Y)|^r)^{\frac{1}{r}} \le (\mathbb{E}|X_n - X|^r)^{\frac{1}{r}} + (\mathbb{E}|Y_n - Y|^r)^{\frac{1}{r}} \underset{n\to\infty}{\longrightarrow} 0\,,$$

somit gilt $X_n + Y_n \xrightarrow{L^r} X + Y$.

2. Falls $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{d} c$, dann gilt $\mathbb{E}h(X_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}h(X)$ für jede Funktion $h \in C_b^0$ (vgl. Bemerkung 3.3.5), $Y_n \xrightarrow{P} c$ (vgl. Satz 3.3.3) und somit

$$P(|Y_n - c| > \delta) \underset{n\to\infty}{\longrightarrow} 0\,, \quad \delta > 0\,.$$

Zu zeigen ist:

$$\mathbb{E}h(X_n + Y_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}h(X + c)\,, \quad h \in C_b^0\,.$$

Da $h : \mathbb{R} \to \mathbb{R}$ beschränkt und gleichmäßig stetig ist, gilt

$$b = \sup_{x\in\mathbb{R}} |h(x)| < \infty$$

und für alle $\varepsilon > 0$ existiert ein $\delta > 0$, sodass für alle $x, y \in \mathbb{R}$ mit der Eigenschaft $|x - y| \le \delta$ gilt, dass $|h(x) - h(y)| < \varepsilon$. Somit gilt

$$\begin{aligned}
&|\mathbb{E}h(X_n + Y_n) - \mathbb{E}h(X + c)| \\
&= |\mathbb{E}\left((h(X_n + Y_n) - h(X_n + c)) \cdot (I(|Y_n - c| \le \delta)\right. \\
&\qquad \left. + I(|Y_n - c| > \delta)) + h(X_n + c) - h(X + c)\right)| \\
&\le \mathbb{E}(\underbrace{|h(X_n + Y_n) - h(X_n + c)|}_{\le\varepsilon} \cdot \underbrace{I\{|Y_n - c| \le \delta\}}_{\le 1}) \\
&\qquad + \mathbb{E}(\underbrace{|h(X_n + Y_n) - h(X_n + c)|}_{\le 2b} \cdot I\{|Y_n - c| > \delta\}) \\
&\qquad + |\mathbb{E}\left(h(X_n + c) - h(X + c)\right)| \\
&\le \varepsilon + 2b \cdot P(|Y_n - c| > \delta) + |\mathbb{E}g(X_n) - \mathbb{E}g(X)|\,, \qquad (3.11)
\end{aligned}$$

wobei $g(\cdot) = h(\cdot + c) \in C_b^0$.

Da $P(|Y_n - c| > \delta) \underset{n\to\infty}{\longrightarrow} 0$ und $\mathbb{E}g(X_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}g(X)$, können die Summanden $2b \cdot P(|Y_n - c| > \delta)$ und $|\mathbb{E}g(X_n) - \mathbb{E}g(X)|$ in der Summe (3.11) für ausreichend große $n \in \mathbb{N}$ beliebig klein werden. Somit gilt $\mathbb{E}h(X_n + Y_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}h(X + c)$, $h \in C_b^0$ und $X_n + Y_n \underset{n\to\infty}{\xrightarrow{d}} X + c$ mit dem Satz 3.3.4 und der Bemerkung 3.3.5.

□

Beispiel 3.4.2 Zeigen wir, dass im Allgemeinen aus $X_n \xrightarrow[n\to\infty]{d} X$, $Y_n \xrightarrow[n\to\infty]{d} Y$ nicht $X_n + Y_n \xrightarrow[n\to\infty]{d} X + Y$ folgen kann.
Seien dazu X_n, $n \in \mathbb{N}$ unabhängige, identisch verteilte Zufallsvariablen,

$$X_n \stackrel{d}{=} X \sim \text{Bernoulli}(1/2)\,.$$

Setzen wir $Y_n = 1 - X_n$. Es gilt $Y_n \stackrel{d}{=} Y \sim$Bernoulli($1/2$), $X_n + Y_n = 1\,, n \in \mathbb{N}$, somit $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{d} Y$ für $n \to \infty$, X und Y stochastisch unabhängig. Andererseits ist $Z = X + Y$ eine diskret verteilte Zufallsvariable mit der Zähldichte

$$P(Z = 0) = P(Z = 2) = \frac{1}{4},\; P(Z = 1) = \frac{1}{2}\,.$$

Somit $1 = X_n + Y_n \underset{n\to\infty}{\overset{d}{\not\longrightarrow}} Z = X + Y$.

Satz 3.4.3 *(Multiplikation von Zufallsvariablen)*
Seien $\{X_n\}_{n\in\mathbb{N}}$, X, $\{Y_n\}_{n\in\mathbb{N}}$ und Y Zufallsvariablen definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$.

1. Falls $X_n \to X$ und $Y_n \to Y$ fast sicher oder stochastisch, dann konvergiert auch ihr Produkt $X_n Y_n \xrightarrow[n\to\infty]{} XY$ im selben Sinne.

2. Falls $X_n \xrightarrow[n\to\infty]{L^{2r}} X$, $Y_n \xrightarrow[n\to\infty]{L^{2r}} Y$, $r \geq 1$, dann gilt

$$X_n Y_n \xrightarrow[n\to\infty]{L^r} XY \text{ für } X_n,\, X,\, Y_n,\, Y \in L^{2r}\,, \quad n \in \mathbb{N}\,.$$

3. *Satz von Slutsky:*
   Falls $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{d} c \equiv const$, dann gilt $X_n Y_n \xrightarrow[n\to\infty]{d} Xc$.

Beweis

1. wird analog zum Satz 3.4.1, 1)bewiesen.

2. Falls $X_n \xrightarrow{L^{2r}} X$, $Y_n \xrightarrow{L^{2r}} Y$, dann gilt

$$\mathbb{E}|X_n - X|^{2r} \xrightarrow[n\to\infty]{} 0 \text{ und } \mathbb{E}|Y_n - Y|^{2r} \xrightarrow[n\to\infty]{} 0\,.$$

Nach dem Satz von Minkowski (vgl. [20, Satz 4.6.10]) gilt

$$\begin{aligned}
(\mathbb{E}|X_nY_n - XY|^r)^{\frac{1}{r}} &= (\mathbb{E}|(X_nY_n - X_nY) + (X_nY - XY)|^r)^{\frac{1}{r}} \\
&\leq (\mathbb{E}|X_n(Y_n - Y)|^r)^{\frac{1}{r}} + (\mathbb{E}|Y(X_n - X)^r)^{\frac{1}{r}} \\
\leq (\underbrace{\mathbb{E}|X_n|^{2r}}_{\xrightarrow[n\to\infty]{} \mathbb{E}|X|^{2r}} &\cdot \underbrace{\mathbb{E}|Y_n - Y|^{2r}}_{\xrightarrow[n\to\infty]{} 0})^{\frac{1}{2r}} + (\mathbb{E}|Y|^{2r} \cdot \underbrace{\mathbb{E}|X_n - X|^r}_{\xrightarrow[n\to\infty]{} 0})^{\frac{1}{2r}} \xrightarrow[n\to\infty]{} 0\,,
\end{aligned}$$

wobei die letzte Ungleichung aus der Ungleichung von Cauchy-Schwarz (vgl. [20, Satz 4.6.4]) folgt. Somit haben wir gezeigt, dass

$$X_n Y_n \xrightarrow[n\to\infty]{L^r} XY\,.$$

3. Sei $h \in C_b^0$. Zu zeigen ist

$$\mathbb{E}h(X_n Y_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}h(Xc)\,,$$

falls $X_n \xrightarrow[n\to\infty]{d} X$ und $Y_n \xrightarrow[n\to\infty]{d} c$. Da h beschränkt und gleichmäßig stetig ist, gilt $b = \sup_{x\in\mathbb{R}} |h(x)| < \infty$ und für alle $\varepsilon > 0$ existiert ein $\delta > 0$, sodass für $x, y \in \mathbb{R}$ mit $|x-y| \le \delta$ gilt, dass $|h(x) - h(y)| \le \varepsilon$. Sei $v > 0$ so gewählt, dass $\pm v$ Stetigkeitsstellen von F_X sind und $P(|X| > v) < \varepsilon$. Dies ist möglich, da $\bar{F}_X(v) \underset{v\to+\infty}{\longrightarrow} 0$, $F_X(-v) \underset{v\to+\infty}{\longrightarrow} 0$, und F_X abzählbar viele Sprungstellen hat. Aus $X_n \xrightarrow[n\to\infty]{d} X$ folgt somit, dass $P(|X_n| > v) < 2\varepsilon$ für hinreichend große $n \in \mathbb{N}$. Ähnlich wie im Beweis des Satzes 3.4.1, 2) gilt

$$\begin{aligned}
|\mathbb{E}h(X_nY_n) - \mathbb{E}h(Xc)| &\le \mathbb{E}\left(|h(X_nY_n) - h(X_n\cdot c)| \cdot I\{|Y_n - c| > \delta/v\}\right) \\
&\quad + \mathbb{E}\left(|h(X_nY_n) - h(X_n\cdot c)| \cdot I\{|Y_n - c| \le \delta/v\} \cdot I\{|X_n| > v\}\right) \\
&\quad + \mathbb{E}(\underbrace{|h(X_nY_n) - h(X_n\cdot c)|}_{<\varepsilon} \cdot I\{|Y_n - c| \le \delta/v\} \cdot I\{|X_n| \le v\}) \\
&\quad + |\mathbb{E}h(cX_n) - \mathbb{E}h(cX)| \\
&\le 2b\cdot \underbrace{P(|Y_n - c| > \delta/v)}_{<\varepsilon} + 2b\cdot \underbrace{P(|X_n| > v)}_{<2\varepsilon} + \varepsilon + |\underbrace{\mathbb{E}g(X_n) - \mathbb{E}g(X)}_{<\varepsilon}|\,,
\end{aligned}$$

wobei wegen

$$Y_n \xrightarrow[n\to\infty]{d} c \quad \Longrightarrow \quad Y_n \xrightarrow[n\to\infty]{P} c$$

die Abschätzung $P(|Y_n - c| > \delta/v) < \varepsilon$ für hinreichend große $n \in \mathbb{N}$ gilt und $g(\cdot) = h(\cdot c) \in C^0$. Aus $X_n \xrightarrow{w} X$ folgt, dass auch $|\mathbb{E}g(X_n) - \mathbb{E}g(X)| < \varepsilon$ für große $n \in \mathbb{N}$. Somit gilt

$$|\mathbb{E}h(X_nY_n) - \mathbb{E}h(X\cdot c)| \le (6b+2)\cdot\varepsilon$$

für alle $\varepsilon > 0$ und für hinreichend große $n \in \mathbb{N}$ und deshalb

$$X_nY_n \xrightarrow[n\to\infty]{d} X\cdot c\,.$$

$\square$

Satz 3.4.4 *(Stetigkeitssatz)*
Seien X_n, $n \in \mathbb{N}$, X beliebige Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und $h : \mathbb{R} \to \mathbb{R}$ eine stetige Funktion.

a) Falls $X_n \underset{n\to\infty}{\longrightarrow} X$ fast sicher, stochastisch oder in Verteilung konvergiert, dann konvergiert auch $h(X_n) \underset{n\to\infty}{\longrightarrow} h(X)$ im selben Sinne.

b) Falls h hölder-stetig mit Hölder-Exponenten $\beta > 0$ ist, $X_n \xrightarrow[n\to\infty]{L^{\beta r}} X$ für $r > 0$, und $\mathbb{E}|h(X_n)|^r$, $E|h(X)|^r < \infty \quad \forall n \in \mathbb{N}$, dann $h(X_n) \xrightarrow[n\to\infty]{L^r} h(X)$.

Beweis a) Falls $X_n \xrightarrow[n\to\infty]{\text{f.s.}} X$ oder $X_n \xrightarrow[n\to\infty]{P} X$, dann wird die Gültigkeit von $h(X_n) \xrightarrow[n\to\infty]{\text{f.s.}} h(X)$ bzw. $h(X_n) \xrightarrow[n\to\infty]{P} h(X)$ genauso wie im Satz 3.4.1, 1) bewiesen.

Falls $X_n \xrightarrow[n\to\infty]{d} X$, dann gilt $\mathbb{E}g(X_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}g(X)$ für beliebige stetige beschränkte Funktionen $g : \mathbb{R} \to \mathbb{R}$. Die Funktion $g \circ h : \mathbb{R} \to \mathbb{R}$ ist ebenfalls beschränkt und stetig, somit gilt

$$\mathbb{E}g(h(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}g(h(X)) \quad \text{und} \quad h(X_n) \xrightarrow[n\to\infty]{d} h(X)\,.$$

b) Falls $h : \mathbb{R} \to \mathbb{R}$ Hölder-stetig ist, so gilt

$$\exists c > 0 : \; |h(x) - h(y)| \le c \cdot (x-y)^\beta \quad \forall x, y \in \mathbb{R} \Rightarrow \mathbb{E}|h(x_n) - h(x)|^r \le c^r.$$

$$\mathbb{E}|X_n - X|^{\beta r} \underset{n\to\infty}{\longrightarrow} 0\,.$$

□

3.5 Gleichgradige Integrierbarkeit

Es seien $X, X_1, X_2, \dots$ Zufallsvariablen auf dem W-Raum $(\Omega, \mathcal{F}, P)$. Wir wissen bereits, dass im Allgemeinen weder aus $X_n \xrightarrow[n\to\infty]{f.s.} X$ noch aus $X_n \xrightarrow[n\to\infty]{P} X$ folgt, dass $X_n \xrightarrow[n\to\infty]{L^1} X$.

Definition 3.5.1 Eine Folge von Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}}$ heißt *gleichgradig integrierbar*, falls

$$\mathbb{E}|X_n| < \infty, \; n \in \mathbb{N}$$

und

$$\sup_{n\in\mathbb{N}} \mathbb{E}(|X_n| I(|X_n| > x)) \underset{x\to\infty}{\longrightarrow} 0.$$

Bemerkung 3.5.2 Seien $\{X_n\}_{n\in\mathbb{N}}$ und $\{Y_n\}_{n\in\mathbb{N}}$ zwei gleichgradig integrierbare Folgen von Zufallsvariablen und c eine Konstante. Dann sind auch $\{cX_n\}_{n\in\mathbb{N}}$, $\{c+X_n\}_{n\in\mathbb{N}}$, $\{X_n+Y_n\}_{n\in\mathbb{N}}$ und $\{\max\{|X_n|,|Y_n|\}\}_{n\in\mathbb{N}}$ gleichgradig integrierbar.

Übungsaufgabe 3.5.3 Zeige die Aussage der obigen Bemerkung.

Lemma 3.5.4 Eine Folge von Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}}$ ist gleichgradig integrierbar genau dann, wenn folgende zwei Bedingungen erfüllt sind:

1. Gleichmäßige Beschränktheit: $\sup\limits_{n\in\mathbb{N}} \mathbb{E}|X_n| < \infty$.
2. Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$, sodass $\mathbb{E}(|X_n| I(A)) < \varepsilon$ für alle $n \in \mathbb{N}$ und alle $A \in \mathcal{F}$ mit $P(A) < \delta$ gilt.

Beweis Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen. Wir müssen folgende Äquivalenz zeigen:

$$\sup_{n\in\mathbb{N}} \mathbb{E}(|X_n| I(|X_n| > x)) \underset{x\to\infty}{\longrightarrow} 0 \iff \begin{cases} 1) \sup\limits_{n\in\mathbb{N}} \mathbb{E}|X_n| < \infty \\ 2) \forall \varepsilon > 0 \ \exists \delta > 0 : \mathbb{E}(|X_n| I(A)) < \varepsilon \\ \qquad \forall n \in \mathbb{N}, \ \forall A \in \mathcal{F} : \ P(A) < \delta \end{cases}$$

"$\Leftarrow$" Definiere $A_n = \{|X_n| > x\}$ für alle $n \in \mathbb{N}$ und $x > 0$. Mit der Markow-Ungleichung folgt $P(A_n) \le \frac{1}{x} \mathbb{E}|X_n|$ und deshalb

$$\sup_{n\in\mathbb{N}} P(A_n) \le \frac{1}{x} \sup_{n\in\mathbb{N}} \mathbb{E}|X_n| \le \frac{c}{x} \underset{x\to\infty}{\longrightarrow} 0.$$

Insbesondere existiert also ein $N > 0$ so dass für alle $x > N$: $P(A_n) < \delta$, und aus 2) folgt dann $\sup_{n\in\mathbb{N}} \mathbb{E}(|X_n| I(A_n)) \le \varepsilon$. Da $\varepsilon > 0$ beliebig klein ist, folgt

$$\sup_{n\in\mathbb{N}} \mathbb{E}(|X_n| I(|X_n| > x)) \underset{x\to\infty}{\longrightarrow} 0.$$

"$\Rightarrow$"

1. Es gilt

$$\begin{aligned} \sup_n \mathbb{E}|X_n| &\le \sup_{n\in\mathbb{N}} (\mathbb{E}(|X_n| I(|X_n| > x)) + \mathbb{E}(|X_n| I(|X_n| \le x))) \\ &\le \sup_{n\in\mathbb{N}} (\mathbb{E}(|X_n| I(|X_n| > x)) + x \underbrace{P(|X_n| \le x)}_{\le 1}) \\ &\le \varepsilon + x < \infty. \end{aligned}$$

2. Für alle $\varepsilon > 0$ existiert ein $x > 0$ mit $\mathbb{E}(|X_n| I(|X_n| > x)) < \frac{\varepsilon}{2}$ nach Definition der gleichgradigen Integrierbarkeit. Wähle $\delta > 0$

so, dass $x\delta < \frac{\varepsilon}{2}$ gilt, dann folgt

$$\begin{aligned}\mathbb{E}(|X_n|I(A)) =& \mathbb{E}(\underbrace{|X_n|}_{\leq x}\underbrace{I(|X_n| \leq x)}_{\leq 1}I(A)) + \mathbb{E}(|X_n|I(|X_n| > x)\underbrace{I(A)}_{\leq 1}) \\ &\leq \underbrace{xP(A)}_{\leq \varepsilon/2} + \underbrace{\mathbb{E}(|X_n|I(|X_n| > x))}_{\leq \varepsilon/2} \leq \varepsilon.\end{aligned}$$

□

Satz 3.5.5 Eine Folge von Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}_0}$ ist gleichgradig integrierbar genau dann, wenn eine Funktion $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ existiert für die gilt: $\frac{\psi(x)}{x} \uparrow \infty$ für $x \to \infty$ und $\sup_{n\in\mathbb{N}} \mathbb{E}\psi(|X_n|) < \infty$. Für die Rückrichtung kann ψ konvex gewählt werden.

Beweis Wir können o.B.d.A. annehmen, dass $X_n \geq 0$ f.s. für alle $n \in \mathbb{N}$.

"$\Leftarrow$" Sei $v(x) = \psi(x)/x$. Dann gilt

$$\begin{aligned}\mathbb{E}(X_n I(X_n > x)) &\leq \frac{1}{v(x)}\mathbb{E}(X_n v(X_n) I(X_n > x)) \\ &\leq \frac{1}{v(x)}\mathbb{E}\psi(X_n) \leq \frac{1}{v(x)} \sup_{n\in\mathbb{N}_0} \mathbb{E}\psi(X_n) \underset{x\to+\infty}{\longrightarrow} 0,\end{aligned}$$

da nach Voraussetzung $v(x)$ von unten gegen ∞ geht. Für $X_n > x$ gilt also $v(X_n) > v(x) \Rightarrow \frac{v(X_n)}{v(x)} \geq 1$, und $xv(x) = \psi(x)$.

"$\Rightarrow$" Da $\{X_n\}_{n\in\mathbb{N}_0}$ gleichgradig integrierbar ist, gilt

$$u(x) := \sup_{n\in\mathbb{N}} \mathbb{E}(X_n I(X_n > x)) \underset{x\to\infty}{\longrightarrow} 0.$$

Wähle eine Folge $y_0 = 0$, $y_k \underset{k\to\infty}{\longrightarrow} \infty$ so, dass

$$\sum_{k=0}^{\infty} \sqrt{u(y_k)} < c < \infty.$$

Definiere $g(x) := \frac{x}{\sqrt{u(y_k)}}$ für $x \in [y_k, y_{k+1})$, dann gilt $\frac{g(x)}{x} \uparrow \infty$ für $x \to \infty$, weil

$$\frac{g(y_k - 0)}{y_k} \leq \frac{1}{\sqrt{u(y_{k-1})}} \leq \frac{1}{\sqrt{u(y_k)}} = \frac{g(y_k)}{y_k}, \quad k \in \mathbb{N}.$$

Weiter gilt

$$\begin{aligned}\mathbb{E}g(X_n) &= \sum_{k=0}^{\infty} \mathbb{E}g(X_n) I(X_n \in [y_k, y_{k+1})) \\ &= \sum_{k=0}^{\infty} \mathbb{E}\left(\frac{X_n}{\sqrt{u(y_k)}} I(X_n \in [y_k, y_{k+1}))\right) \\ &\le \sum_{k=0}^{\infty} \frac{1}{\sqrt{u(y_k)}} u(y_k) = \sum_{k=0}^{\infty} \sqrt{u(y_k)} < c < \infty, \quad n \in \mathbb{N}\end{aligned}$$

nach der Definition von $u(y_k)$. Um die Aussage zu beweisen, ist es ausreichend eine konvexe Funktion $\psi \le g$ so zu konstruieren, dass $\psi(x)/x \uparrow \infty$ für $x \to \infty$. Wähle $\psi(x)$ für $x \ge 0$ als die lineare Interpolation zwischen den Punkten $(y_k, g(y_k - 0)), k \in \mathbb{N}_0$.
Da $\frac{g(y_k-0)}{y_k} = (u(y_{k-1}))^{-1/2}$ monoton wachsend für $k \to \infty$ ist, so ist

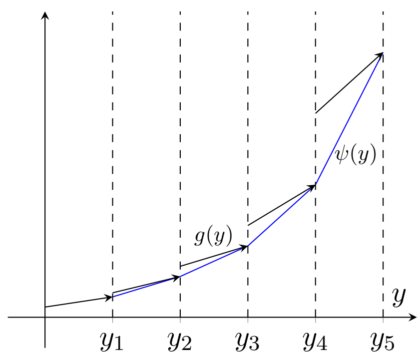


Abbildung 3.4: Funktionen ψ und g

der Epigraph von ψ eine konvexe Hülle des Epigraphen der unstetigen Funktion $g(x)$. Deshalb gilt $g(x) \ge \psi(x)$ $\forall x \ge 0$ und $\psi(\cdot)$ ist nach Konstruktion konvex. Nun folgt

$$\frac{\psi(y_k)}{y_k} = \frac{g(y_k - 0)}{y_k} = \frac{1}{\sqrt{u(y_k)}} \xrightarrow[k\to\infty]{f.s.} \infty,$$

also insbesondere auch $\frac{\psi(x)}{x} \xrightarrow[x\to\infty]{} \infty$. Nach der Konstruktion von g und ψ gilt, dass $\frac{\psi(x)}{x} = a_k - \frac{b_k}{x}$ für $x \in [y_k, y_{k+1})$ mit $b_k > 0$. Somit ist die Monotonie bewiesen und es folgt

$$\sup_{n\in\mathbb{N}} \mathbb{E}\psi(X_n) \le \sup_{n\in\mathbb{N}} \mathbb{E}g(X_n) \le c < \infty.$$

□

Lemma 3.5.6 Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen mit $\mathbb{E}|X_n| < \infty$, $n \in \mathbb{N}$ und $X_n \underset{n\to\infty}{\xrightarrow{P}} X$. Dann gilt $X_n \underset{n\to\infty}{\xrightarrow{L^1}} X$ genau dann, wenn $\{X_n\}_{n\in\mathbb{N}}$ gleichgradig integrierbar ist.
Insbesondere impliziert $X_n \underset{n\to\infty}{\xrightarrow{P}} X$, dass $\mathbb{E}X_n \underset{n\to\infty}{\longrightarrow} \mathbb{E}X$ gilt.

Beweis

"$\Leftarrow$" Sei $\{X_n\}_{n\in\mathbb{N}}$ gleichgradig integrierbar. Wir müssen also zeigen, dass $\mathbb{E}|X_n - X| \underset{n\to\infty}{\longrightarrow} 0$. Da $X_n \underset{n\to\infty}{\xrightarrow{P}} X$ gilt, hat $\{X_n\}_{n\in\mathbb{N}}$ eine Teilfolge $\{X_{n_k}\}_{k\in\mathbb{N}}$, die f.s. gegen X konvergiert. Das Lemma von Fatou besagt, dass

$$\mathbb{E}|X| \leq \liminf_{k\to\infty} \mathbb{E}|X_{n_k}| \leq \sup_{n\in\mathbb{N}} \mathbb{E}|X_n|,$$

und deshalb gilt $\mathbb{E}|X| < \infty$ wegen Lemma 3.5.4, 1). Nach Bemerkung 3.5.2 folgt durch die gleichgradige Integrierbarkeit von $\{X_n\}_{n\in\mathbb{N}}$, dass auch $\{X_n - X\}_{n\in\mathbb{N}}$ gleichgradig integrierbar ist. Weiter folgern wir aus $\lim\limits_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$ und Lemma 3.5.4, 2), dass

$$\lim_{n\to\infty} \mathbb{E}(|X_n - X| I\left(|X_n - X| > \varepsilon\right)) = 0, \quad \varepsilon > 0.$$

Daher gilt

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}|X_n - X| &= \lim_{n\to\infty} [\, \mathbb{E}(|X_n - X| I\left(|X_n - X| > \varepsilon\right)) \\ &\qquad + \mathbb{E}(|X_n - X| I\left(|X_n - X| \leq \varepsilon\right))] \\ &\leq 2\varepsilon\end{aligned}$$

für alle $\varepsilon > 0$, d.h.

$$X_n \underset{n\to\infty}{\xrightarrow{L^1}} X.$$

"$\Rightarrow$" Sei nun $\mathbb{E}|X_n - X| \underset{n\to\infty}{\longrightarrow} 0$. Wir zeigen nun die Eigenschaften 1) und 2) aus Lemma 3.5.4:

1. $\sup\limits_{n\in\mathbb{N}} \mathbb{E}|X_n| \leq \sup\limits_{n\in\mathbb{N}} \mathbb{E}|X_n - X| + \mathbb{E}|X| < \infty$, da $X_n \underset{n\to\infty}{\xrightarrow{L^1}} X$.
2. Für alle $A \in \mathcal{F}$ mit $P(A) \leq \delta$ gilt:

$$\begin{aligned}\mathbb{E}(|X_n| I(A)) &\leq \mathbb{E}(|X_n - X| \underbrace{I(A)}_{\leq 1}) + \mathbb{E}(|X| I(A)) \\ &\leq \underbrace{\mathbb{E}|X_n - X|}_{< \frac{\varepsilon}{2}} + \frac{\varepsilon}{2} = \varepsilon\end{aligned}$$

bei passender Wahl von δ, da $\mathbb{E}|X| < \infty$, und weil $\forall \varepsilon > 0\ \exists N$, sodass $\mathbb{E}|X_n - X| < \frac{\varepsilon}{2}$ für alle $n > N$ gilt.

$\square$

Folgerung 3.5.7 Es gelte $X_n \xrightarrow[n\to\infty]{P} X$.

1. Falls $f : \mathbb{R} \to \mathbb{R}$ stetig und entweder
   (a) beschränkt ist oder
   (b) $\{f(X_n)\}_{n\in\mathbb{N}}$ gleichgradig integrierbar ist,

   dann gilt $f(X_n) \xrightarrow[n\to\infty]{L^1} f(X)$.

2. Falls $\{|X_n|^r\}_{n\in\mathbb{N}}$ für $r \geq 1$ gleichgradig integrierbar ist, dann gilt $X_n \xrightarrow[n\to\infty]{L^r} X$.
   Falls $\mathbb{E}|X_n|^r < \infty$ und $X_n \xrightarrow[n\to\infty]{L^r} X$ für $r \geq 1$ gilt, dann ist $\{|X_n|^r\}_{n\in\mathbb{N}}$ gleichgradig integrierbar.

3. Falls $\mathbb{E}|X_n|^{r+\alpha} < c < \infty$ für alle $n \in \mathbb{N}$ und bestimmte $r \geq 1, \alpha > 0$, dann gilt $X_n \xrightarrow[n\to\infty]{L^r} X$.

Beweis

1. b) Nach dem Stetigkeitssatz 3.4.4 gilt $f(X_n) \xrightarrow[n\to\infty]{P} f(X)$. Die Aussage folgt aus Lemma 3.5.6.

1. a) Wir zeigen, dass $\{f(X_n)\}_{n\in\mathbb{N}}$ gleichgradig integrierbar ist. Da f beschränkt ist, ist $\{f(X_n)\}_{n\in\mathbb{N}}$ nach Theorem 3.5.5 gleichgradig integrierbar mit $\psi(x) = x^{1+\delta}$ und $\delta > 0$. Mit Hilfe von 1 b) folgt die Aussage.

2. Definiere $Z_n := |X_n - X|^r$, $n \in \mathbb{N}$. Es gilt, dass $Z_n \xrightarrow[n\to\infty]{P} 0$ und $|X_n|^r \xrightarrow[n\to\infty]{P} |X|^r$ wegen des Stetigkeitssatzes. Wegen der Mikowski-Ungleichung ist $\{Z_n\}_{n\in\mathbb{N}}$ gleichgradig integrierbar, da $\{|X_n|^r\}_{n\in\mathbb{N}}$ und $|X|^r$ gleichgradig integrierbar sind. Nach Anwenden von Lemma 3.5.6 auf $\{Z_n\}_{n\in\mathbb{N}}$ ist die Aussage gezeigt.

3. Theorem 3.5.5 mit $\psi(x) = x^{1+\alpha/r}$ sagt, dass $\{|X_n|^r\}_{n\in\mathbb{N}}$ gleichgradig integrierbar ist. Aussage 2. schließt den Beweis ab.

$\square$

3.6 Drei–Reihen–Satz von Kolmogorow

Es ist bekannt, dass

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n^\alpha} < \infty \Longleftrightarrow \alpha > 1, \\ \sum_{n=1}^{\infty} \frac{(-1)^n}{n^\alpha} < \infty \Longleftrightarrow \alpha > 0, \end{aligned} \tag{3.12}$$

da die Differenz zweier aufeinander folgenden Summanden der zweiten Reihe die Ordnung $\frac{1}{n^{\alpha+1}}$, hat, d.h. es gilt

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^\alpha} = -1 + \sum_{k=1}^{\infty} \left(\frac{1}{(2k)^\alpha} - \frac{1}{(2k+1)^\alpha} \right).$$

Dabei

$$\begin{aligned} \frac{1}{(2k)^\alpha} - \frac{1}{(2k+1)^\alpha} &= \frac{(2k+1)^\alpha - (2k)^\alpha}{(2k)^\alpha (2k+1)^\alpha} = \frac{\left(1+\frac{1}{2k}\right)^\alpha - 1}{(2k+1)^\alpha} \\ &\overset{k\to\infty}{=} \frac{1 + \frac{\alpha}{2k} - 1 + o\left(\frac{1}{2k}\right)}{(2k+1)^\alpha} \\ &= \frac{\alpha + o(1)}{2k(2k+1)^\alpha} = O\left(\frac{1}{(2k)^{\alpha+1}}\right) \\ &= O\left(\frac{1}{n^{\alpha+1}}\right), \quad n = 2k. \end{aligned}$$

Es stellt sich nun folgende Frage: Für welche $\alpha > 0$ konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{\delta_n}{n^\alpha}$, wobei δ_n u.i.v. Zufallsvariablen mit $\mathbb{E}\delta_n = 0$ sind
(z.B. $P(\delta_n = \pm 1) = \frac{1}{2}$)? Oder noch allgemeiner: Unter welchen Voraussetzungen gilt für unabhängige Zufallsvariablen X_n, dass $\sum_{n=1}^{\infty} X_n < \infty$ f.s.?
Wir wissen bereits, dass für eine Folge von Zufallsvariablen $\{Y_n\}_{n\in\mathbb{N}}$ mit $Y_n \xrightarrow[n\to\infty]{f.s.} Y$ folgendes gilt:

$$Y_n \xrightarrow[n\to\infty]{P} Y.$$

Die Rückrichtung gilt im Allgemeinen nicht.

Satz 3.6.1 Für unabhängige Zufallsvariablen X_n, $n \in \mathbb{N}$, gilt

$$S_n = \sum_{j=1}^{n} X_j \xrightarrow[n\to\infty]{P} S \Rightarrow S_n \xrightarrow[n\to\infty]{f.s.} S.$$

Für den Beweis brauchen wir folgende Hilfsätze:

Lemma 3.6.2 Falls für unabhängige Zufallsvariablen X_n, $n \in \mathbb{N}$, gilt

$$P(|S_n - S_k| \geq y) \leq p < 1, \quad k \leq n,$$

für ein $y > 0$, dann

$$P(\max_{k\leq n} |S_k| \geq x) \leq \frac{1}{1-p} P(|S_n| > x - y), \quad x \in \mathbb{R}.$$

Beweis Sei $Y = \min\{k \geq 1 : |S_k| \geq x\}$, $A_j = \{Y = j\}$, $j \in \mathbb{N}$. Da A_j paarweise disjunkt sind, gilt

$$\begin{aligned} P(|S_n| > x - y) &\geq \sum_{k=1}^{n} P(\{|S_n| > x - y\} \cap A_k) \\ &\geq \sum_{k=1}^{n} P(\{|S_n - S_k| < y\} \cap A_k) \end{aligned}$$

wegen $\{|S_n - S_k| < y\} \cap A_k \subseteq \{|S_n| > x - y\} \cap A_k$. Jedoch sind $A_k \in \sigma(X_j, j \leq k)$ und $\{|S_n - S_k| < y\} \in \sigma(X_j, j \geq k+1)$ stochastisch unabhängig. Damit gilt per Voraussetzung

$$\begin{aligned} P(|S_n| > x - y) &\geq \sum_{k=1}^{n} P(|S_n - S_k| < y) P(A_k) \\ &\geq (1-p) \sum_{k=1}^{n} P(A_k) \\ &= (1-p) P(\max_{k\leq n} |S_k| \geq x). \end{aligned}$$

□

Folgerung 3.6.3 (Ungleichung von Kolmogorov) Falls $\{X_n\}$ eine Folge von unabhängigen quadratisch integrierbaren Zufallsvariablen mit $\mathbb{E}X_n = 0$, $n \in \mathbb{N}$ ist, dann gilt

$$P(\max_{k\leq n} |S_k| \geq x) \leq 2P(|S_n| > x - \sqrt{2\text{Var } S_n}).$$

Beweis Es folgt aus Lemma 3.6.1 mit $p = 1/2$ und $y = \sqrt{2\text{Var } S_n}$, weil nach Tschebyschew–Ungleichung

$$P(|S_n - S_k| \geq \sqrt{2\text{Var } S_n}) \leq \frac{\text{Var } (S_n - S_k)}{2\text{Var } S_n} \leq 1/2, \quad k \leq n.$$

□

Es folgt nun der Beweis von Satz 3.6.1.

Beweis Sei $\varepsilon > 0$, $A^{\varepsilon}_{n,m} = \{|S_n - S_m| > \varepsilon\}$ und $A^{\varepsilon}_m = \bigcup_{n\geq m} A^{\varepsilon}_{n,m}$. Zeige, dass $\{S_n\}$ eine f.s. Fundamentalfolge ist, d.h. nach Satz 3.1.1, $\forall \varepsilon > 0$

$$P(\sup_{n\geq m} |S_n - S_m| > 2\varepsilon) \underset{m\to\infty}{\longrightarrow} 0 \Leftrightarrow P(A^{2\varepsilon}_m) \underset{m\to\infty}{\longrightarrow} 0.$$

Da $S_n \xrightarrow[n\to\infty]{P} S$, gilt $P(B_n^\varepsilon) \to 0$, $n \to \infty$ mit $B_n^\varepsilon = \{|S_n - S| > \varepsilon\}$. Zusätzlich ist $\{S_n\}$ stochastisch eine Cauchy–Folge, also

$$p_{m,M} := \sup_{m\le n\le M} P(A_{n,m}^\varepsilon) \to 0, \quad m, M \to \infty.$$

Für große m, M gilt also $p_{m,M} < 1/2$. Lemma 3.6.2 mit $p = p_{m,M}$, $y = \varepsilon$, $x = 2\varepsilon$ ergibt

$$\begin{aligned} &P(\sup_{m\le n\le M} |S_n - S_m| > 2\varepsilon) = P(\bigcup_{n=m+1}^{M} A_{m,n}^{2\varepsilon}) \\ &\le \frac{1}{1 - p_{m,M}} P(A_{M,m}^\varepsilon) \le 2P(A_{M,m}^\varepsilon). \end{aligned}$$

Die Stetigkeit des Wahrscheinlichkeitsmaßes impliziert

$$P(A_m^{2\varepsilon}) = \lim_{M\to\infty} P(\bigcup_{n=m+1}^{M} A_{m,n}^{2\varepsilon}) \le 2 \limsup_{M\to\infty} P(A_{M,m}^\varepsilon).$$

Da $A_{M,m}^\varepsilon \subset B_M^{\varepsilon/2} \cup B_m^{\varepsilon/2}$, die letzte Ungleichung ergibt

$$P(A_m^{2\varepsilon}) \le 2P(B_m^{\varepsilon/2}) + 2\underbrace{\limsup_{M\to\infty} P(B_M^{\varepsilon/2})}_{=0} \longrightarrow 0, \quad m \to \infty.$$

$\square$

Folgerung 3.6.4 Für unabhängige Zufallsvariablen X_n, $n \in \mathbb{N}$ mit

- $\operatorname{Var} X_n < \infty$,
- $\mathbb{E}X_n = 0$, $n \in \mathbb{N}$ und
- $\sum_{n=1}^\infty \operatorname{Var} X_n < \infty$

konvergiert $\sum_{n=1}^\infty X_n$ f.s.

Beweis Seien $S_n = \sum_{i=1}^n X_i$ für $n \in \mathbb{N}$. Wir zeigen, dass $\{S_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L^2(\Omega, \mathcal{F}, P)$ ist. Für $n > m$ folgt

$$\mathbb{E}(S_n - S_m)^2 = \|S_n - S_m\|_{L^2}^2 = \sum_{j=m+1}^{n} \operatorname{Var} X_j \underset{n,m\to\infty}{\longrightarrow} 0,$$

da $\sum_{j=1}^\infty \operatorname{Var} X_j < \infty$. Deswegen ist $\{S_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L^2(\Omega, \mathcal{F}, P)$. Dann

$$\exists S = L^2 - \lim_{n\to\infty} S_n = \sum_{j=1}^{\infty} X_j$$

und also $S_n \xrightarrow[n\to\infty]{P} S$. Die Aussage folgt aus Theorem 3.6.1. $\square$

Folgerung 3.6.5 Wenn $\sum_{n=1}^{\infty} a_n^2 < \infty$ für eine deterministische Folge $\{a_n\}_{n\in\mathbb{N}}$ gilt und $\{\delta_n\}$ eine Folge von u.i.v. Zufallsvariablen mit $\mathbb{E}\delta_n = 0$, $\text{Var } \delta_n = \sigma^2 < \infty$, $n \in \mathbb{N}$ ist, dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n \delta_n$ fast sicher.

Übungsaufgabe 3.6.6 Leite Folgerung 3.6.5 aus Folgerung 3.6.5 her.

Für unser Eröffnungsbeispiel (3.12) mit δ_n u.i.v., $\mathbb{E}\delta_n = 0$, $\text{Var } \delta_n = \sigma^2 > 0$ (z.B. $P(\delta_n = \pm 1) = \frac{1}{2}$), $a_n = \frac{1}{n^\alpha}$, $n \in \mathbb{N}$ gilt:

$$\sum_{n=1}^{\infty} \frac{\delta_n}{n^\alpha} < \infty \quad \text{f.s.},$$

falls $\sum_{n=1}^{\infty} \frac{1}{n^{2\alpha}} < \infty$, d.h. für $\alpha > \frac{1}{2}$.

Folgerung 3.6.7 (Drei–Reihen–Satz von Kolmogorow) Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge unabhängiger Zufallsvariablen mit

$$\sum_{n=1}^{\infty} \mathbb{E}X_n < \infty, \qquad \sum_{n=1}^{\infty} \text{Var } X_n < \infty.$$

Dann gilt

$$\sum_{n=1}^{\infty} X_n < \infty \quad \text{f.s.}$$

Beweis Definiere $Y_n := X_n - \mathbb{E}X_n$. Also gilt $X_n = \mathbb{E}X_n + Y_n = a_n + Y_n$, $\mathbb{E}Y_n = 0$ für $n \in \mathbb{N}$ und $\sum_{n=1}^{\infty} a_n < \infty$ nach Voraussetzung. Dann folgt

$$\sum_{n=1}^{\infty} Y_n < \infty \quad \text{f.s.}$$

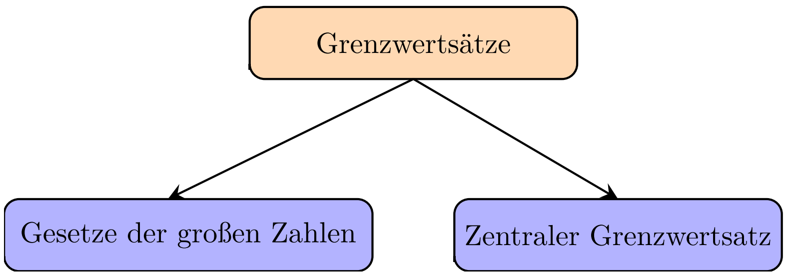
aus Folgerung 3.6.4, da $\text{Var } X_n = \text{Var } Y_n$ für $n \in \mathbb{N}$ und

$$\sum_{n=1}^{\infty} \text{Var } X_n < \infty \Rightarrow \sum_{n} X_n = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} Y_n < \infty \quad f.s.$$

□

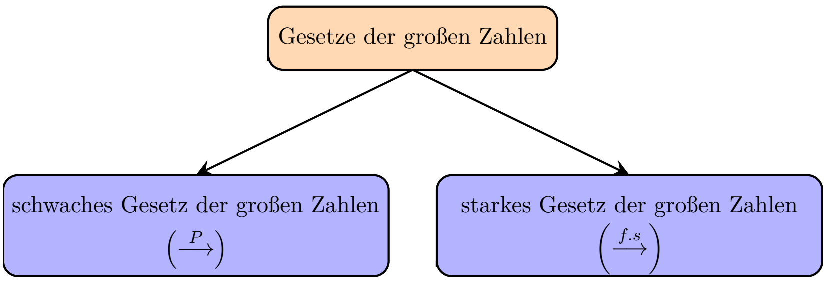
Kapitel 4

Grenzwertsätze



In diesem Kapitel betrachten wir Aussagen der Wahrscheinlichkeitsrechnung, die Näherungsformeln von großer anwendungsbezogener Bedeutung liefern. Dies wird an mehreren Beispielen erläutert.

4.1 Gesetze der großen Zahlen



Ein typisches Gesetz der großen Zahlen besitzt die Form

$$\frac{1}{n}\sum_{j=1}^{n} X_j \underset{n\to\infty}{\longrightarrow} \mathbb{E}X_0\,, \tag{4.1}$$

wobei $\{X_n\}_{n\in\mathbb{N}}$ unabhängige identisch verteilte Zufallsvariablen mit $X_n \stackrel{d}{=} X_0$, $\mathbb{E}|X_0| < \infty$ sind.
Die Konvergenz in (4.1) wird entweder in Wahrscheinlichkeit oder fast sicher verstanden. Wenn $\xrightarrow{P}$ gemeint ist, spricht man von dem *schwachen Gesetz der großen Zahlen*. Falls $\xrightarrow{\text{f.s.}}$ gemeint ist, heißt die Aussage (4.1) *starkes Gesetz der großen Zahlen*.

Im Folgenden verwenden wir die Bezeichnungen $S_n = \sum_{i=j}^n X_j$, $\overline{X}_n = \frac{S_n}{n}$, $\forall n \in \mathbb{N}$ für eine Folge $\{X_n\}_{n\in\mathbb{N}}$ von Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$.

4.1.1 Schwaches Gesetz der großen Zahlen

Satz 4.1.1 *(Markow)*
Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $\mathbb{E}X_n^2 < \infty$ für alle n. Falls

$$\operatorname{Var} \overline{X}_n \underset{n\to\infty}{\longrightarrow} 0\,, \tag{4.2}$$

dann gilt

$$\overline{X}_n - \frac{1}{n}\sum_{j=1}^n \mathbb{E}X_j \underset{n\to\infty}{\xrightarrow{P}} 0\,.$$

Beweis Aus der Ungleichung von Tschebyschew folgt für alle $\varepsilon > 0$

$$P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{j=1}^n \mathbb{E}X_j\right| > \varepsilon\right) \le \frac{\mathbb{E}(\overline{X}_n - \mathbb{E}\overline{X}_n)^2}{\varepsilon^2} = \frac{\operatorname{Var} \overline{X}_n}{\varepsilon^2} \underset{n\to\infty}{\longrightarrow} 0\,.$$

Somit gilt

$$\overline{X}_n - \mathbb{E}\overline{X}_n \underset{n\to\infty}{\xrightarrow{P}} 0\,.$$

□

Folgerung 4.1.2 Seien die Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}}$ im Satz 4.1.1 unabhängig. Dann gilt Folgendes:

1. Die Bedingung $\operatorname{Var} \overline{X}_n \underset{n\to\infty}{\longrightarrow} 0$ bekommt die Form

$$\frac{1}{n^2}\sum_{i=1}^n \operatorname{Var} X_i \underset{n\to\infty}{\longrightarrow} 0\,.$$

2. Falls $\operatorname{Var} X_n \le c = \mathit{const}$ für alle $n \in \mathbb{N}$, dann gilt die Bedingung (4.2) und somit die Aussage des Satzes 4.1.1 (Satz von Tschebyschew).

3. Insbesondere ist die Bedingung Var $X_n \leq c = const$ für alle $n \in \mathbb{N}$ erfüllt, falls $\{X_n\}_{n\in\mathbb{N}}$ unabhängige identisch verteilte Zufallsvariablen mit
$$\mathbb{E}X_n = \mu, \text{ Var } X_n = \sigma^2 < \infty$$
sind. Dann nimmt das schwache Gesetz der großen Zahlen die klassische Form
$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{P} \mu$$
an.

Die Existenz der zweiten Momente ist für das schwache Gesetz der großen Zahlen nicht entscheidend. So kann man mit Hilfe der charakteristischen Funktionen folgenden Satz beweisen:

Satz 4.1.3 *(Schwaches Gesetz der großen Zahlen von Kchintschin)*
Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsvariablen, $n \in \mathbb{N}$, $X_n \in L^1$, mit demselben Erwartungswert $\mathbb{E}X_n = \mu < \infty$. Dann gilt
$$\overline{X}_n \xrightarrow[n\to\infty]{P} \mu\,.$$

Beweis Wir haben
$$\begin{aligned}
\varphi_{\overline{X}_n}(t) &= \mathbb{E}e^{i\overline{X}_n t} = \mathbb{E}e^{i\frac{S_n}{n}t} = \mathbb{E}e^{iS_n\frac{t}{n}} \\
&\underbrace{=}_{X_k \text{ unabh.}} \prod_{k=1}^{n} \mathbb{E}e^{iX_k\frac{t}{n}} = \prod_{k=1}^{n} \varphi_{X_k}\left(\frac{t}{n}\right) \\
&\underbrace{=}_{\text{Satz 2.2.5, 2) für } n\to\infty} \prod_{k=1}^{n}\left(1 + i\mu\frac{t}{n} + o\left(\frac{t}{n}\right)\right) \\
&= \left(1 + i\mu\frac{t}{n} + o\left(\frac{1}{n}\right)\right)^n \xrightarrow[n\to\infty]{} e^{i\mu t} = \varphi_\mu(t),\ t \in \mathbb{R}\,.
\end{aligned}$$

Somit folgt aus dem Satz 3.3.6 die Konvergenz $\overline{X}_n \xrightarrow[n\to\infty]{d} \mu$, und daher $\overline{X}_n \xrightarrow[n\to\infty]{P} \mu$ (vgl. Satz 3.3.3). □

4.1.2 Starkes Gesetz der großen Zahlen

Satz 4.1.4 Seien $\{X_n\}_{n\in\mathbb{N}}$ unabhängige Zufallsvariablen, $\mathbb{E}X_n^2 < \infty$ für alle $n \in \mathbb{N}$ und $\sum_{n=1}^{\infty} \frac{\text{Var } X_n}{n^2} < \infty$. Dann gilt $\overline{X}_n - \mathbb{E}\overline{X}_n \xrightarrow[n\to\infty]{\text{f.s.}} 0$.

Um diesen Satz beweisen zu können, braucht man folgende Hilfsaussage:

Lemma 4.1.5 *(Ungleichung von Kolmogorow):*
Seien $\{X_n\}_{n\in\mathbb{N}}$ unabhängige Zufallsvariablen mit
$$\mathbb{E}X_n = 0 \text{ und } \mathbb{E}X_n^2 < \infty \quad n \in \mathbb{N}\,.$$

Dann gilt $P(\max_{1\le k\le n}|S_k|\ge\varepsilon)\le\frac{\mathbb{E}(S_n^2)}{\varepsilon^2}\quad\varepsilon>0,\,n\in\mathbb{N}\,.$

Beweis Führen wir Ereignisse $A=\{\max_{1\le k\le n}|S_k|\ge\varepsilon\}$ und

$$A_k=\{|S_j|<\varepsilon,\,j=1,\ldots,k-1,\,|S_k|\ge\varepsilon\},\quad k=1,\ldots,n$$

ein. Es gilt $A=\bigcup_{k=1}^n A_k$ und somit

$$\mathbb{E}(S_n^2)\ge\mathbb{E}(S_n^2\cdot I(A))=\sum_{k=1}^n\mathbb{E}(S_n^2\cdot I(A_k))\,.$$

Betrachten wir

$$\begin{aligned}\mathbb{E}(S_n^2\cdot I(A_k)) &= \mathbb{E}\left((S_k+(X_{k+1}+\ldots+X_n))^2\cdot I(A_k)\right)\\ &= \mathbb{E}(S_k^2\cdot I(A_k))+2\underbrace{\mathbb{E}\left(S_k\cdot(X_{k+1}+\ldots+X_n)I(A_k)\right)}_{=0}\\ &\quad+\underbrace{\mathbb{E}\left((X_{k+1}+\ldots+X_n)^2\cdot I(A_k)\right)}_{\ge0}\\ &\ge \mathbb{E}(S_k^2\cdot I(A_k))\,,\end{aligned}$$

weil $S_k\cdot I(A_k)$ und $X_{k+1}+\ldots+X_n$ unabhängig sind und somit

$$\begin{aligned}\mathbb{E}\left(S_k\cdot I(A_k)(X_{k+1}+\ldots+X_n)\right) &= \mathbb{E}(S_k\cdot I(A_k))\cdot\mathbb{E}(X_{k+1}+\ldots+X_n)\\ &= \mathbb{E}(S_k\cdot I_A(A_k))\cdot\sum_{j=k+1}^n\mathbb{E}X_j=0\end{aligned}$$

wegen $\mathbb{E}X_j=0$ für jedes i,j. Zusammengefasst ergibt sich

$$\begin{aligned}\mathbb{E}(S_n^2) &\ge \sum_{k=1}^n\mathbb{E}(S_k^2\cdot I(A_k))\ge\varepsilon^2\sum_{k=1}^n\mathbb{E}\,I(A_k)\\ &= \varepsilon^2\sum_{k=1}^n P(A_k)=\varepsilon^2P\left(\bigcup_{k=1}^n A_k\right)=\varepsilon^2P(A)\end{aligned}$$

und

$$P(A)\le\frac{\mathbb{E}(S_n^2)}{\varepsilon^2}\,.$$

□

Beweis des Satzes 4.1.4
Es reicht aus, wenn der Satz 4.1.4 für X_n mit $\mathbb{E}X_n=0$ bewiesen wird, sonst ersetzt man X_n durch $X_n-\mathbb{E}X_n$. Seien also X_n, $n\in\mathbb{N}$ unabhängig,

$$\mathbb{E}X_n=0,\,\mathbb{E}X_n^2<\infty,\,\sum_{n=1}^\infty\frac{\mathbb{E}X_n^2}{n^2}<\infty\,.$$

Zeigen wir, dass $\overline{X}_n \xrightarrow[n\to\infty]{\text{f.s.}} 0$. Laut Satz 3.1.1 genügt es zu zeigen, dass für

$$\tilde{A}_n = \{\sup_{k\geq n} |\overline{X}_k| > \varepsilon\} \text{ gilt } P(\tilde{A}_n) \underset{n\to\infty}{\longrightarrow} 0\,, \quad \varepsilon > 0\,.$$

Es gilt offensichtlich $\tilde{A}_m \subseteq \bigcup_{n\geq \log_2 m+1} \underbrace{\left\{\max_{2^{n-1}\leq k<2^n} |\overline{X}_k| > \varepsilon\right\}}_{=A_n}$,

weil $\sup_{k\geq m} |\overline{X}_k| > \varepsilon$ die Existenz von $n \geq \log_2 m + 1$ mit der Eigenschaft

$$\max_{2^{n-1}\leq k<2^n} |\overline{X}_k| > \varepsilon$$

impliziert. Daher genügt es $P\left(\bigcup_{n=m}^{\infty} A_n\right) \underset{m\to\infty}{\longrightarrow} 0$ zu zeigen, um $P(\tilde{A}_n) \underset{n\to\infty}{\longrightarrow} 0$ zu beweisen. Nach Lemma 4.1.5 gilt

$$\begin{aligned}
P(A_n) &= P\left(\max_{2^{n-1}\leq k<2^n} \frac{2^{n-1}}{k} \underbrace{\left|\sum_{j=1}^{k} X_j\right|}_{=S_k} > 2^{n-1}\cdot\varepsilon \right) \\
&\leq P(\max_{2^{n-1}\leq k<2^n} |S_k| > 2^{n-1}\cdot\varepsilon) \\
&\leq P(\max_{1\leq k\leq 2^n} |S_k| > 2^{n-1}\cdot\varepsilon) \\
&\leq \frac{\text{Var } S_{2^n}}{\varepsilon^2\cdot 2^{2n-2}} = \frac{\sum_{j=1}^{2^n}\sigma_j^2}{\varepsilon^2\cdot 2^{2n-2}}\,,
\end{aligned}$$

wegen Unabhängigkeit von X_j und $\frac{2^{n-1}}{k} \leq 1$, wobei $\mathbb{E}X_j^2 = \text{Var } X_j = \sigma_j^2 < \infty\,, \quad j\in\mathbb{N}$. Dann gilt

$$\sum_{k=1}^{\infty} P(A_k) \leq \frac{4}{\varepsilon^2}\sum_{k=1}^{\infty}\frac{1}{2^{2k}}\sum_{j=1}^{2^k}\sigma_j^2 \leq \frac{4}{\varepsilon^2}\sum_{j=1}^{\infty}\sigma_j^2 \underbrace{\sum_{k:\,2^k\geq j}\frac{1}{2^{2k}}}_{=\sum_{k\geq\log_2 j}\left(\frac14\right)^k=\frac{1}{j^2}\cdot\frac{1}{1-\frac14}=\frac{3}{4j^2}} = \frac{3}{\varepsilon^2}\sum_{j=1}^{\infty}\frac{\sigma_j^2}{j^2} < \infty$$

nach der Voraussetzung des Satzes. Aus $\sum_{k=1}^{\infty} P(A_k) < \infty$ folgt

$$P\left(\bigcup_{k=m}^{\infty} A_k\right) \leq \sum_{k=m}^{\infty} P(A_k) \underset{m\to\infty}{\longrightarrow} 0\,,$$

und der Satz 4.1.4 ist bewiesen. □

Bemerkung 4.1.6 Falls $\{X_n\}_{n\in\mathbb{N}}$ unabhängige, identisch verteilte Zufallsvariablen mit $\mathbb{E}X_n = \mu$, Var $X_n = \sigma^2 < \infty$ sind, dann gilt

$$\sum_{n=1}^{\infty}\frac{\text{Var } X_n}{n^2} = \sigma^2\sum_{n=1}^{\infty}\frac{1}{n^2} < \infty$$

und somit die Aussage des Satzes 4.1.4:

$$\overline{X}_n \xrightarrow[n\to\infty]{\text{f.s.}} \mu \,. \tag{4.3}$$

Man stellt jedoch fest, dass die Existenz von Var X_n nicht gebraucht wird, um die Aussage (4.3) zu beweisen.

Satz 4.1.7 *(Starkes Gesetz der großen Zahlen von Kolmogorow)*
Seien $\{X_n\}_{n\in\mathbb{N}}$ unabhängige identisch verteilte Zufallsvariablen.
Es gilt $\overline{X}_n \xrightarrow[n\to\infty]{\text{f.s.}} \mu$ genau dann, wenn $\mathbb{E}X_n = \mu < \infty$.

Für den Beweis des Satzes brauchen wir folgendes

Lemma 4.1.8 Sei $X \geq 0$ eine Zufallsvariable. Dann gilt

$$\sum_{n=1}^{\infty} P(X \geq n) \leq \mathbb{E}X \leq 1 + \sum_{n=1}^{\infty} P(X \geq n)\,.$$

Beweis Es gilt

$$\begin{aligned}
\sum_{n=1}^{\infty} P(X \geq n) &= \sum_{n=1}^{\infty}\sum_{k\geq n} P(k \leq X < k+1) = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P(k \leq X < k+1) \\
&= \sum_{k=1}^{\infty} kP(k \leq X < k+1) = \sum_{k=0}^{\infty} \mathbb{E}(k \cdot I(X \in [k, k+1))) \\
&\leq \sum_{k=0}^{\infty} \mathbb{E}(X \cdot I(X \in [k, k+1))) = \mathbb{E}X
\end{aligned}$$

und weiter

$$\begin{aligned}
\mathbb{E}X &\leq \sum_{k=0}^{\infty} \mathbb{E}((k+1) \cdot I(X \in [k, k+1)) \\
&= \sum_{k=0}^{\infty} (k+1) \cdot P(k \leq X < k+1) \\
&= \sum_{k=0}^{\infty} k \cdot P(k \leq X < k+1) + \underbrace{\sum_{k=0}^{\infty} P(k \leq X < k+1)}_{=1} \\
&= \sum_{n=1}^{\infty} P(X \geq n) + 1\,.
\end{aligned}$$

$\square$

Beweis des Satzes 4.1.7

"$\Rightarrow$" Falls $\overline{X}_n \xrightarrow[n\to\infty]{\text{f.s.}} \mu$ für ein $\mu \in \mathbb{R}$, dann

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n} \cdot \frac{S_{n-1}}{n-1} = \underbrace{\overline{X}_n}_{\to\mu} - \underbrace{\frac{n-1}{n}}_{\to 1} \underbrace{\overline{X}_{n-1}}_{\to\mu} \xrightarrow[n\to\infty]{\text{f.s.}} 0\,.$$

Deshalb tritt das Ereignis $A_n = \{|X_n|/n \geq 1\}$ nur endlich oft auf $\Longrightarrow$ $P\left(\limsup_{n\to\infty} A_n\right) = 0$ und nach dem Lemma von Borel–Cantelli (vgl. [20, Lemma 2.2.16])

$$\infty > \sum_{n=1}^{\infty} P\left(|X_n|/n \geq 1\right) = \sum_{n=1}^{\infty} P(|X_1| \geq n)$$

wegen der Tatsache, dass X_n identisch verteilt sind. Dann gilt nach Lemma 4.1.8

$$\mathbb{E}|X_n| = \mathbb{E}|X_1| \leq 1 + \sum_{n=1}^{\infty} P(|X_1| \geq n) < \infty\,.$$

"$\Leftarrow$" Sei $\mathbb{E}X_n = \mu < \infty$. Führen wir

$$X_k^* = X_k \cdot I(|X_k| \leq k) = \begin{cases} X_k, & |X_k| \leq k \\ 0, & |X_k| > k \end{cases}$$

ein. Es ist klar, dass die Zufallsvariablen $\{X_k^*\}$ auch unabhängig sind. Um zu zeigen, dass

$$\frac{\sum_{j=1}^n X_j - n\mu}{n} \xrightarrow[n\to\infty]{\text{f.s.}} 0\,,$$

genügt es zu zeigen, dass

$$\frac{\sum_{j=1}^n X_i^* - \mathbb{E}\sum_{j=1}^n X_i^*}{n} \xrightarrow[n\to\infty]{\text{f.s.}} 0\,.$$

Das Ereignis $A_k = \{X_k \neq X_k^*\}$ tritt mit Wahrscheinlichkeit 1 nach dem Lemma von Borel–Cantelli nur endlich oft auf. Um dies zu zeigen, benutzen wir die Unabhängigkeit von A_k. Mit

$$\begin{aligned} \sum_{k=1}^{\infty} P(A_k) &= \sum_{k=1}^{\infty} P(|X_k| > k) = \sum_{k=1}^{\infty} P(\underbrace{|X_1|}_{X_i \text{ ident.}} > k) \\ &\leq \sum_{k=1}^{\infty} P(|X_1| \geq k) \underbrace{\leq}_{\text{Lemma 4.1.8}} \mathbb{E}|X_1| < \infty \end{aligned}$$

folgt nach dem Lemma von Borel–Cantelli

$$P\Big(\bigcap_{n=1}^{\infty} \bigcup_{k\geq n} A_k\Big) = 0\,.$$

Daher folgt

$$\begin{aligned}
&\frac{\sum_{j=1}^n X_j - n\mu}{n} - \frac{\sum_{j=1}^n X_j^* - \sum_{j=1}^n \mathbb{E}X_j^*}{n} \\
&\qquad\qquad = \frac{\sum_{j=1}^n (X_j - X_j^*)}{n} - \frac{\sum_{j=1}^n (\mu - \mathbb{E}X_j^*)}{n} \\
&\qquad\qquad\qquad \underset{p=const}{=} \underbrace{\frac{\sum_{k=1}^p (X_{j_k} - X_{j_k}^*)}{n}}_{\to 0\,(n\to\infty)} - \underbrace{\frac{\sum_{j=1}^n (\mu - \mathbb{E}X_j^*)}{n}}_{\to 0\,(n\to\infty)} .
\end{aligned}$$

Die erste Summe hängt nicht von n ab. Die Konvergenz der zweiten Summe folgt sofort aus $\mathbb{E}X_j^* \underset{j\to\infty}{\longrightarrow} \mu$. Es ist so, weil

$$\mu - \mathbb{E}X_j^* = \mathbb{E}\left(X_1 \cdot I(|X_1| > j)\right) \underset{j\to\infty}{\longrightarrow} 0$$

wegen $\mathbb{E}|X_1| < \infty$. Da für alle $\varepsilon > 0$ ein n_0 existiert mit der Eigenschaft, dass für alle $k \geq n_0$ $\quad |\mathbb{E}X_k^* - \mu| < \varepsilon$ gilt, treffen wir die Abschätzung

$$\begin{aligned}
\left| \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}X_k^* - \mu \right| &\leq \left| \frac{1}{n} \sum_{k=1}^{n_0} (\mathbb{E}X_k^* - \mu) \right| + \frac{1}{n} \sum_{k=n_0}^{n} |\mathbb{E}X_k^* - \mu| \\
&\leq \frac{n_0}{n} \max_{k=1\ldots n_0} |\mathbb{E}X_k^* - \mu| + \frac{n-n_0}{n}\varepsilon < 2\varepsilon ,
\end{aligned}$$

falls n so gewählt wird, dass $\frac{n_0}{n} \max_{k=1\ldots n_0} |\mathbb{E}X_k^* - \mu| < \varepsilon$.

Nach dem Satz 4.1.4 genügt es zu zeigen, dass $\sum_{n=1}^{\infty} \frac{\text{Var } X_n^*}{n^2} < \infty$. Es gilt

$$\text{Var } X_n^* \leq \mathbb{E}X_n^{*2} \leq \sum_{k=1}^{n} k^2 P(k-1 \leq |X_1| < k) ,$$

weil $|X_n^*| \leq Z_n = \sum_{k=1}^n k \cdot I(k-1 \leq |X_1| < k)$ und

$$\mathbb{E}X_n^{*2} \leq \mathbb{E}Z_n^2 = \sum_{k=1}^{n} k^2 P(k-1 \leq |X_1| < k) .$$

Daraus folgt

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\text{Var } X_n^*}{n^2} &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{k^2}{n^2} P(k-1 \leq |X_1| < k) \\
&= \sum_{k=1}^{\infty} k^2 P(k-1 \leq |X_1| < k) \cdot \sum_{n=k}^{\infty} \frac{1}{n^2} \\
&= \underbrace{\sum_{k=1}^{\infty} P(k-1 \leq |X_1| < k)}_{=1} + \sum_{k=1}^{\infty} k^2 P(k-1 \leq |X_1| < k) \cdot \sum_{n \geq k+1} \frac{1}{n^2} .
\end{aligned}$$

Es gilt

$$\sum_{n=k+1}^{\infty} \frac{1}{n^2} \leq \sum_{n=k+1}^{\infty} \frac{1}{n(n-1)} = \sum_{n=k+1}^{\infty} \left(\frac{1}{n-1} - \frac{1}{n} \right) = \frac{1}{k}$$

$$\Longrightarrow \sum_{n=1}^{\infty} \frac{\text{Var } X_n^*}{n^2} \leq 1 + \sum_{k=1}^{\infty} kP(k-1 \leq |X_1| < k)$$

$$= 2 + \sum_{k=1}^{\infty} (k-1) P(k-1 \leq |X_1| < k) \leq 2 + \mathbb{E}|X_1| < \infty \,,$$

weil $\sum_{k=1}^{\infty} (k-1) \cdot I(k-1 \leq |X_1| < k) \leq |X_1|$.

□

4.1.3 Anwendung der Gesetze der großen Zahlen

Wahrscheinlichkeitstheoretischer Beweis des Approximationssatzes von Weierstrass

Satz 4.1.9 *(Weierstrass)*
Jede auf $[0,1]$ stetige Funktion g lässt sich gleichmäßig durch Polynome P_n des Grades n approximieren:

$\varepsilon > 0$ existiert ein $m \in \mathbb{N}$, sodass für alle $n \geq m \quad |g(x) - P_n(x)| < \varepsilon \,, x \in [0,1]$.

Beweis Für beliebiges $x \in [0,1]$ betrachten wir eine Folge von unabhängigen identisch verteilten Zufallsvariablen $\{X_n\}_{n \in \mathbb{N}}$ mit $X_n \sim Bernoulli(x)$. Es gilt $S_n = \sum_{i=1}^{n} X_i \sim Bin(n, x)$ und somit ist

$$\mathbb{E} g(\overline{X}_n) = \mathbb{E} g\left({}^{S_n}\!/\!{}_n \right) = \sum_{k=0}^{n} g\left({}^{k}\!/\!{}_n \right) \binom{n}{k} \cdot x^k (1-x)^{n-k}$$

ein Polynom in x des Grades n. Setzen wir $P_n(x) = \mathbb{E} g(\overline{X}_n)$ und zeigen, dass $P_n \underset{n\to\infty}{\longrightarrow} g$ gleichmäßig in $x \in [0,1]$. Aus dem Satz 4.1.7 folgt

$$\overline{X}_n \xrightarrow[n\to\infty]{\text{f.s.}} \mathbb{E} X_1 = x \,, \quad x \in [0,1] \,.$$

Aus der Konvergenzkette

$$\overline{X}_n \xrightarrow[n\to\infty]{\text{f.s.}} x \Longrightarrow \overline{X}_n \xrightarrow[n\to\infty]{P} x \Longrightarrow \overline{X}_n \xrightarrow[n\to\infty]{d} x \Leftrightarrow \overline{X}_n \xrightarrow[n\to\infty]{w} x$$

folgt

$$\mathbb{E} \varphi(\overline{X}_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E} \varphi(x)$$

für jede Funktion $h \in C_0(\mathbb{R})$. Sei $h(x) = g(x)$ für $x \in [0,1]$ (da g stetig auf $[0,1]$ ist, ist sie dort gleichmäßig stetig). Somit gilt die Konvergenz

$$P_n(x) = \mathbb{E}g(\overline{X}_n) \underset{n\to\infty}{\longrightarrow} g(x), \quad \forall x \in [0,1]\,.$$

Wir müssen jetzt zeigen, dass diese Konvergenz gleichmäßig in x ist. Da jede stetige Funktion auf $[0,1]$ beschränkt ist, gilt $b = \max_{x\in[0,1]} |g(x)| < \infty$. Wegen der gleichmäßigen Stetigkeit von g auf $[0,1]$ gilt zusätzlich: Für alle

$\varepsilon > 0$ existiert ein $\delta > 0$: für $x, y \in \mathbb{R}$ mit $|x-y| \le \delta$ gilt $|g(x)-g(y)| < \varepsilon$.

Dann

$$\begin{aligned}
|P_n(x) - g(x)| &= |\mathbb{E}g(\overline{X}_n) - \mathbb{E}g(x)| \le \mathbb{E}|g(\overline{X}_n) - g(x)| \\
&= \mathbb{E}\left|\Big(g(\overline{X}_n) - g(x)\Big) \cdot I(|\overline{X}_n - x| \le \delta)\right| + \\
&\quad +\mathbb{E}\left|\Big(g(\overline{X}_n) - g(x)\Big) \cdot I(|\overline{X}_n - x| > \delta)\right| \\
&\le \varepsilon + 2b \cdot P(|\overline{X}_n - x| > \delta) < \varepsilon + 2b\varepsilon\,,
\end{aligned}$$

weil aus $\overline{X}_n \xrightarrow{P} x$ folgt, dass $P(|\overline{X}_n - x| > \delta) \underset{n\to\infty}{\longrightarrow} 0$ für ausreichend große n.
Dass diese Konvergenz gleichmäßig in $x \in [0,1]$ ist, zeigt die Tschebyschew–Ungleichung (vgl. [20, Folgerung 4.6.2]):

$$\begin{aligned}
P\left(\left|\overline{X}_n - x\right| > \delta\right) &\le \frac{\mathrm{Var}\ \overline{X}_n}{\delta^2} = \frac{\frac{1}{n^2} \cdot nx(1-x)}{\delta^2} \\
&\le \frac{1}{n\delta^2} \underset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

gleichmäßig in $x \in [0,1]$. Somit haben wir gezeigt, dass $P_n(x) \underset{n\to\infty}{\longrightarrow} g(x)$ gleichmäßig in $x \in [0,1]$. $\square$

4.2 Zentraler Grenzwertsatz

Für die Gesetze der großen Zahlen wurde die Normierung $\frac{1}{n}$ der Summe $S_n = \sum_{j=1}^n X_j$ gewählt, um $\frac{S_n}{n} \underset{n\to\infty}{\longrightarrow} \mathbb{E}X_1$ zu beweisen. Falls jedoch eine andere Normierung gewählt wird, so sind andere Grenzwertaussagen möglich. Im Fall der Normierung $\frac{1}{\sqrt{n}}$ spricht man von zentralen Grenzwertsätzen: unter gewissen Voraussetzungen gilt also

$$\frac{S_n - n\mathbb{E}X_1}{\sqrt{n\mathrm{Var}\ X_1}} \underset{n\to\infty}{\xrightarrow{d}} Y \sim N(0,1)\,.$$

Die Voraussetzungen werden wir nach und nach abschwächen, um folgende Varianten des Zentralen Grenzwertsatzes zu beweisen:

1. Klassischer zentraler Grenzwertsatz,
2. Grenzwersatz von Lindeberg,
3. Grenzwertsatz von Ljapunow, Feller usw.

4.2.1 Klassischer zentraler Grenzwertsatz

Satz 4.2.1 Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen identisch verteilten Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $\mathbb{E}X_j = \mu$, Var $X_j = \sigma^2$, wobei $0 < \sigma^2 < \infty$. Dann gilt

$$P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq x\right) \underset{n\to\infty}{\longrightarrow} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-y^2}{2}}\, dy\,, \quad x \in \mathbb{R}\,,$$

wobei $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-y^2}{2}}\, dy$ die Verteilungsfunktion einer $N(0,1)$–verteilten Zufallsvariablen ist.

Beweis Ohne Beschränkung der Allgemeinheit betrachten wir statt X_i die standartisierten Zufallsvariablen $\bar{X}_j = \frac{X_j-\mu}{\sigma}$, für die gilt $\mathbb{E}\bar{X}_j = 0$, Var $\bar{X}_j = \mathbb{E}\bar{X}_j^2 = 1$. Es muss also gezeigt werden, dass

$$\sum_{j=1}^{n} \frac{\bar{X}_j}{\sqrt{n}} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)$$

gilt, was (nach dem Satz 3.3.6) äquivalent zu

$$\varphi_{\sum_{j=1}^{n} \frac{\bar{X}_j}{\sqrt{n}}}(t) \underset{n\to\infty}{\longrightarrow} \varphi_Y(t) = e^{-\frac{t^2}{2}}\,, t \in \mathbb{R}$$

ist. Wie im Beweis des Satzes 4.1.3 schreiben wir

$$\begin{aligned} \varphi_{\sum_{j=1}^{n} \frac{\bar{X}_j}{\sqrt{n}}}(t) &= \mathbb{E}\exp\left\{i \sum_{j=1}^{n} \frac{\bar{X}_j t}{\sqrt{n}}\right\} = \left(\varphi_{\bar{X}_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n \\ &= \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \underset{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}}\,, \quad t \in \mathbb{R}\,, \end{aligned}$$

wobei hier Satz 2.2.5, 2 bis zur Ordnung 2 benutzt wird. Somit ist unser Satz bewiesen. □

Satz 4.2.2 Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen identisch verteilten Zufallsvariablen mit $\mathbb{E}X_n = \mu$, $0 <$ Var $X_n = \sigma^2 < \infty$. Sei $\{N_n\}_{n\in\mathbb{N}}$ eine Folge von $\mathbb{N}$–wertigen Zufallsvariablen mit $N_n \xrightarrow[n\to\infty]{\text{f.s.}} \infty$. Falls es eine Folge $\{a_n\}_{n\in\mathbb{N}}$ gibt mit den Eigenschaften $a_n > 0\,, n \in \mathbb{N}$, $a_1 \leq a_2 \leq a_2 \leq \ldots$ und $a_n \underset{n\to\infty}{\longrightarrow} \infty$, sodass $\frac{N_n}{a_n} \xrightarrow[n\to\infty]{P} c > 0$, wobei c eine Konstante ist, dann gilt

$$\frac{S_{N_n} - N_n\mu}{\sigma\sqrt{N_n}} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)\,.$$

Beweis Ohne Beschränkung der Allgemeinheit setzen wir $\mu = 0$, $\sigma^2 = 1$ (vgl. den Beweis des Satzes 4.2.1). Außerdem sei $k_n = [ca_n]$ der ganze Teil von ca_n. Aus $\frac{N_n}{a_n} \xrightarrow{P} c$ folgt dann $\frac{N_n}{k_n} \xrightarrow{P} 1$ und $\frac{k_n}{N_n} \xrightarrow{P} 1$, $n \to \infty$. Es muss gezeigt werden, dass $\frac{S_{N_n}}{\sqrt{N_n}} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)$.

$$\frac{S_{N_n}}{\sqrt{N_n}} = \sqrt{\frac{k_n}{N_n}} \left(\frac{S_{k_n}}{\sqrt{k_n}} + \frac{S_{N_n} - S_{k_n}}{\sqrt{k_n}} \right) .$$

Aus dem Satz 4.2.1 folgt $\frac{S_{k_n}}{\sqrt{k_n}} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)$. Da $\sqrt{\frac{k_n}{N_n}} \xrightarrow[n\to\infty]{P} 1$, genügt es zu zeigen, dass

$$\frac{S_{N_n} - S_{k_n}}{\sqrt{k_n}} \xrightarrow[n\to\infty]{P} 0$$

(vgl. den Satz von Slutsky). Für beliebige $\varepsilon, \delta > 0$ gilt

$$\begin{aligned} P\left(\frac{|S_{N_n} - S_{k_n}|}{\sqrt{k_n}} > \varepsilon \right) &= \underbrace{P(|S_{N_n} - S_{k_n}| > \varepsilon\sqrt{k_n}\, , \, |N_n - k_n| \le \delta k_n)}_{I_n} \\ &\quad + P(|S_{N_n} - S_{k_n}| > \varepsilon\sqrt{k_n}\, , \, |N_n - k_n| > \delta k_n) \\ &\le I_n + P\left(\left| \frac{N_n}{k_n} - 1 \right| > \delta \right) \end{aligned}$$

und wegen $\frac{N_n}{k_n} \xrightarrow{P} 1$ somit $P\left(\left| \frac{N_n}{k_n} - 1 \right| > \delta \right) \xrightarrow[n\to\infty]{} 0$. Da

$$\begin{aligned} &\{|S_{N_n} - S_{k_n}| > \varepsilon\sqrt{k_n}\, , \, |N_n - k_n| \le \delta k_n\} \\ &\qquad \subseteq \left\{ \max_{k_n \le j \le k_n + \delta k_n} |S_j - S_{k_n}| > \varepsilon\sqrt{k_n} \right\} \\ &\qquad\qquad \cup \left\{ \max_{k_n - \delta k_n \le j \le k_n} |S_j - S_{k_n}| > \varepsilon\sqrt{k_n} \right\} , \end{aligned}$$

es folgt daraus

$$\begin{aligned} I_n &\le P\left(\max_{k_n \le j \le (1+\delta)k_n} |S_j - S_{k_n}| > \varepsilon\sqrt{k_n} \right) \\ &\quad + P\left(\max_{(1-\delta)k_n \le j \le k_n} |S_j - S_{k_n}| > \varepsilon\sqrt{k_n} \right) \\ &\underset{\text{Lemma 4.1.5}}{\le} \frac{\mathrm{Var}\ (S_{(1+\delta)k_n} - S_{k_n})}{k_n \varepsilon^2} + \frac{\mathrm{Var}\ (S_{(1-\delta)k_n} - S_{k_n})}{k_n \varepsilon^2} \\ &\le \frac{((1+\delta)k_n - k_n + 1) + (k_n - (1-\delta)k_n + 1)}{k_n \varepsilon^2} \\ &= \frac{2\delta k_n + 2}{k_n \varepsilon^2} = \frac{2}{\varepsilon^2}(\delta + \frac{1}{k_n}) < c_1 \delta \end{aligned}$$

für hinreichend große n, weil $\delta > 0$ beliebig ist und $k_n \to \infty$. Somit gilt

$$P\left(\frac{|S_{N_n} - S_{k_n}|}{\sqrt{N_n}} > \varepsilon\right) < c_2\delta$$

für große n, und der Satz ist bewiesen. □

4.2.2 Grenzwertsatz von Lindeberg

Im klassischen zentralen Grenzwertsatz wurden Folgen von unabhängigen und identisch verteilten Zufallsvariablen betrachtet. In diesem Abschnitt lassen wir die Voraussetzung der identischen Verteiltheit fallen und formulieren einen allgemeineren Grenzwertsatz in der Form von Lindeberg.

Satz 4.2.3 *(Lindeberg)*
Sei $\{X_{nk},\, 1 \le k \le n\,,\, n \in \mathbb{N}\}$ eine Folge von unabhängigen Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit den Eigenschaften

$$\mathbb{E}X_{nk} = 0,\; 0 < \sigma_{nk}^2 = \text{Var } X_{nk} < \infty\,, k \in \{1, \ldots, n\},\; \sum_{k=1}^{n} \sigma_{nk}^2 = 1, n \in \mathbb{N}\,.$$

Falls für alle $\varepsilon > 0$

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{E}\left(X_{nk}^2 I(|X_{nk}| > \varepsilon)\right) = 0\,, \tag{4.4}$$

dann gilt

$$\sum_{k=1}^{n} X_{nk} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)\,.$$

Bemerkung 4.2.4

1. Im Satz 4.2.3 wird lediglich die Unabhängigkeit von X_{nk} innerhalb der Serie $k \in \{1, \ldots, n\}$ gefordert. Für unterschiedliche n können die Zufallsvariablen X_{nk} jedoch abhängig sein.

2. Der klassische zentrale Grenzwertsatz stellt einen Spezialfall des Satzes 4.2.3 dar: Falls $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen identisch verteilten Zufallsvariablen mit $\mathbb{E}X_n = \mu\,,\; 0 < \sigma^2 = \text{Var } X_n < \infty$ $\forall n \in \mathbb{N}$ ist, setzen wir $X_{nk} = \frac{X_k - \mu}{\sqrt{n}\sigma}, \quad k = 1, \ldots, n \quad n \in \mathbb{N}\,.$ Es gilt

$\mathbb{E}X_{nk} = 0, \text{Var } X_{nk} = \frac{1}{n}, k = 1, \ldots, n, \quad \sum_{k=1}^n \frac{1}{n} = 1$ und

$$\begin{aligned}
&\sum_{k=1}^n \mathbb{E}\left(X_{nk}^2 I(|X_{nk}| > \varepsilon)\right) \\
&\qquad = \frac{1}{n\sigma^2} \sum_{k=1}^n \mathbb{E}\left((X_k - \mu)^2 I(|X_k - \mu| > \sqrt{n}\sigma)\right) \\
&\qquad \underset{X_k\text{–id. vert.}}{=} \frac{1}{\sigma^2} \mathbb{E}\left(\underbrace{(X_1 - \mu)}_{Z}{}^2 I(|X_1 - \mu| > \sigma\sqrt{n})\right) \\
&\qquad = \frac{1}{\sigma^2} \int_{|x| > \sigma\sqrt{n}} x^2 dF_Z(x) \underset{n\to\infty}{\longrightarrow} 0\,.
\end{aligned}$$

Somit sind alle Voraussetzungen des Satzes 4.2.3 erfüllt.

3. Die Lindeberg–Bedingung (4.4) impliziert auch *gleichmäßige asymptotische Kleinheit von* X_{nk}. Tatsächlich ergeben [20, Folgerung 4.6.2.] und Relation (4.4) für beliebige $\varepsilon, \delta > 0$, dass:

$$\begin{aligned}
&\max_{1\le k\le n} P(|X_{nk}| > \varepsilon) \underset{[20,\text{ Satz 4.7.1.}]}{\le} \max_{1\le k\le n} \frac{\mathbb{E}X_{nk}^2}{\varepsilon^2} \\
&\underset{\forall\delta>0}{\le} \frac{1}{\varepsilon^2}\left(\max_{1\le k\le n} \mathbb{E}\left(X_{nk}^2 I(|X_{nk}| \le \delta)\right) + \max_{1\le k\le n} \mathbb{E}\left(X_{nk}^2 I(|X_{nk}| > \delta)\right)\right) \\
&\qquad \le \frac{\delta^2}{\varepsilon^2} + \frac{1}{\varepsilon^2} \sum_{k=1}^n \mathbb{E}\left(X_{nk}^2 I(|X_{nk}| > \delta)\right) \underset{n\to\infty}{\longrightarrow} 0\,.
\end{aligned}$$

, weil $\delta > 0$ beliebig klein gewählt werden kann.

Für den Beweis des Satzes 4.2.3 brauchen wir eine Reihe von Hilfsergebnissen.

Lemma 4.2.5 Für $n \in \mathbb{N}$ seien $y_1, \ldots, y_n$ und $z_1, \ldots, z_n$ komplexe Zahlen mit $|y_k|, |z_k| \le 1\,, 1 \le k \le n\,.$ Dann gilt

$$\left|\prod_{k=1}^n z_k - \prod_{k=1}^n y_k\right| \le \sum_{k=1}^n |z_k - y_k|\,.$$

Beweis Verwenden wir eine Induktion bzgl. n. Der Fall $n = 1$ ist offensichtlich. Falls die Aussage des Lemmas gilt für n, zeigen wir, dass sie auch für

$n+1$ gilt.

$$\begin{aligned}
\left|\prod_{k=1}^{n+1} y_k - \prod_{k=1}^{n+1} z_k\right| &\leq \left|\prod_{k=1}^{n+1} y_k - y_{n+1}\prod_{k=1}^{n} z_k\right| + \left|y_{n+1}\prod_{k=1}^{n} z_k - \prod_{k=1}^{n+1} z_k\right| \\
&\leq \underbrace{|y_{n+1}|}_{\leq 1} \cdot \left|\prod_{k=1}^{n} y_k - \prod_{k=1}^{n} z_k\right| + \underbrace{\left|\prod_{k=1}^{n} z_k\right|}_{\leq \prod_{k=1}^{n}|z_k| \leq 1} \cdot |y_{n+1} - z_{n+1}| \\
&\underset{\text{I.S.}}{\leq} \sum_{k=1}^{n} |y_k - z_k| + |y_{n+1} - z_{n+1}| = \sum_{k=1}^{n+1} |y_k - z_k| \,.
\end{aligned}$$

□

Lemma 4.2.6 Sei X eine Zufallsvariable mit $\mathbb{E}|X|^n < \infty$ für ein $n \in \mathbb{N}$ und charakteristischer Funktion φ_X. Dann gilt

$$\varphi_X(t) = \sum_{k=0}^{n} \frac{(it)^k}{k!} \mathbb{E}X^k + R_n(t)\,, \quad t \in \mathbb{R}\,,$$

wobei

$$|R_n(t)| \leq \mathbb{E}\left(\min\left\{\frac{|tX|^{n+1}}{(n+1)!}, 2\frac{|tX|^n}{n!}\right\}\right) .$$

Beweis Zeigen wir zunächst induktiv bzgl. $n \in \mathbb{N}$, dass für alle $t \in \mathbb{R}$ gilt

$$\left|e^{it} - \left(1 + it + \ldots + \frac{(it)^n}{n!}\right)\right| \leq \min\left\{\frac{|t|^{n+1}}{(n+1)!}, 2\frac{|t|^n}{n!}\right\} . \tag{4.5}$$

Für $n = 0$ gilt $|e^{it} - 1| \leq |e^{it}| + 1 = 2$, andererseits

$$|e^{it} - 1| = \left|\int_0^t e^{is} ds\right| \leq \int_0^{|t|} \underbrace{|e^{is}|}_{=1} ds = |t|\,,$$

denn

$$\left|\int_\Gamma f(z)\,dz\right| \leq \int_\Gamma |f(z)|\,|dz|$$

aus der Komplexanalysis. Daher

$$\left|e^{it} - 1\right| \leq \min\{2, |t|\} .$$

Nun gelte (4.5) für ein $n \in \mathbb{N}$. Zeigen wir dessen Gültigkeit auch für $n+1$.

$$\begin{aligned}
\left| e^{it} - \sum_{k=0}^{n+1} \frac{(it)^k}{k!} \right| &= \left| \int_0^t \left(e^{is} - \sum_{k=0}^{n} \frac{(is)^k}{k!} \right) ds \right| \\
&\underset{\text{Ind. Annahme}}{\leq} \int_0^{|t|} \min\left\{ \frac{|s|^{n+1}}{(n+1)!}, 2\frac{|s|^n}{n!} \right\} ds \\
&\leq \min\left\{ \int_0^{|t|} \frac{|s|^{n+1}}{(n+1)!} ds, \, 2\int_0^{|t|} \frac{|s|^n}{n!} ds \right\} \\
&= \min\left\{ \frac{|t|^{n+2}}{(n+2)!}, 2\frac{|t|^{n+1}}{(n+1)!} \right\},
\end{aligned}$$

wobei die vorletzte Ungleichung aus

$$\int \min\{f,g\}\, dx \leq \min\left\{ \int f\, dx, \int g\, dx \right\} \quad \text{für } f, g \geq 0 \tag{4.6}$$

folgt (vgl. Abb. 4.1). Somit ist die Gültigkeit von (4.5) bewiesen. Nun erset-

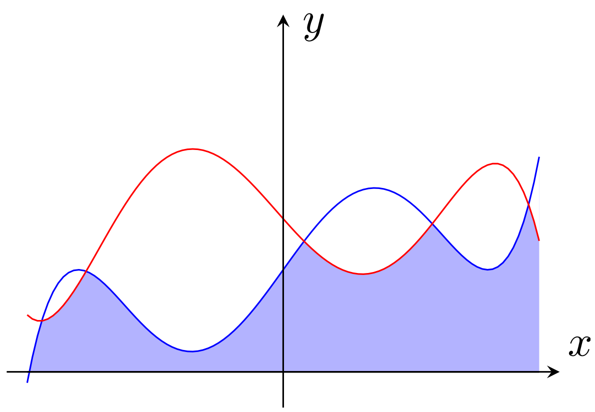


Abbildung 4.1: Illustration zur Ungleichung (4.6)

zen wir t in (4.5) durch tX und bilden Erwartungswerte an beiden Seiten. Dann bekommen wir

$$\begin{aligned}
\left| \varphi_X(t) - \sum_{k=0}^{n} \frac{(it)^k}{k!} \mathbb{E}X^k \right| &\leq \mathbb{E}\left| e^{itX} - \sum_{k=0}^{n} \frac{(it)^k}{k!} X^k \right| \\
&\underset{(4.5)}{\leq} \mathbb{E}\min\left\{ \frac{|tX|^{n+1}}{(n+1)!}, 2\frac{|tX|^n}{n!} \right\}.
\end{aligned}$$

$\square$

Beweis des Satzes 4.2.3
Sei $S_n = \sum_{k=1}^n X_{nk}\,, n \in \mathbb{N}$. Es genügt zu zeigen, dass $\varphi_{S_n}(t) \underset{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}}$ für alle $t \in \mathbb{R}$ nach dem Satz 3.3.6, wobei

$$\varphi_{S_n}(t) = \mathbb{E}e^{iS_n t} = \mathbb{E}\left(e^{i\sum_{k=1}^n X_{nk}t}\right) = \prod_{k=1}^n \mathbb{E}e^{iX_{nk}t} = \prod_{k=1}^n \varphi_{X_{nk}}(t)\,, \quad t \in \mathbb{R}$$

wegen der Unabhängigkeit von X_{nk}, $k = 1, \ldots, n$.
Aus Lemma 4.2.5 folgt

$$\begin{aligned}
\left|\varphi_{S_n}(t) - e^{-\frac{t^2}{2}}\right| &= \left|\prod_{k=1}^n \varphi_{X_{nk}}(t) - e^{-\frac{t^2}{2}}\right| \\
\underset{\sum_{k=1}^n \sigma_{nk}^2 = 1}{=} & \left|\prod_{k=1}^n \varphi_{X_{nk}}(t) - \prod_{k=1}^n e^{-\frac{\sigma_{nk}^2 t^2}{2}}\right| \le \sum_{k=1}^n \left|\varphi_{X_{nk}}(t) - e^{-\frac{\sigma_{nk}^2 t^2}{2}}\right| \\
&\le \sum_{k=1}^n \left|\varphi_{X_{nk}}(t) - \left(1 - \frac{\sigma_{nk}^2 t^2}{2}\right)\right| + \sum_{k=1}^n \left|1 - \frac{\sigma_{nk}^2 t^2}{2} - e^{-\frac{\sigma_{nk}^2 t^2}{2}}\right|\,.
\end{aligned}$$

Es muss gezeigt werden, dass beide Summen in der Abschätzung gegen Null konvergieren für $n \to \infty$. Wenden wir Lemma 4.2.6 an:

$$\begin{aligned}
& \left|\varphi_{X_{nk}}(t) - \left(1 - \frac{\sigma_{nk}^2 t^2}{2}\right)\right| \le \mathbb{E}\min\left\{\frac{|tX_{nk}|^3}{3!},\, 2\frac{|tX_{nk}|^2}{2!}\right\} \\
\le\ & \mathbb{E}\min\left\{|t|^3|X_{nk}|^3,\, |t|^2 X_{nk}^2\right\} \cdot \left(I(|X_{nk}| \le \varepsilon) + I(|X_{nk}| > \varepsilon)\right) \\
\le\ & |t|^3\mathbb{E}(\underbrace{|X_{nk}|^3}_{\le \varepsilon X_{nk}^2} \cdot I(|X_{nk}| \le \varepsilon)) + |t|^2\mathbb{E}\left(X_{nk}^2 \cdot I(|X_{nk}| > \varepsilon)\right) \\
\le\ & \varepsilon|t|^3 \underbrace{\mathbb{E}X_{nk}^2}_{=\sigma_{nk}^2} + t^2\mathbb{E}\left(X_{nk}^2 \cdot I(|X_{nk}| > \varepsilon)\right)\,, \quad \varepsilon > 0\,.
\end{aligned}$$

Aus $\sum_{k=1}^n \sigma_{nk}^2 = 1$ und Bedingung (4.4) folgt

$$\sum_{k=1}^n \left|\varphi_{nk}(t) - \left(1 - \frac{\sigma_{nk}^2 t^2}{2}\right)\right| \le \varepsilon|t|^3 \underbrace{\sum_{k=1}^n \sigma_{nk}^2}_{=1} + t^2 \sum_{k=1}^n \mathbb{E}\left(X_{nk}^2 \cdot I\{|X_{nk}| > \varepsilon\}\right)$$
$$\underset{n\to\infty}{\longrightarrow} 0\,,$$

weil ε beliebig klein gewählt werden kann. Aus der Ungleichung

$$|e^{-x} - 1 + x| \le x^2$$

für $|x| \le \frac{1}{2}$ und der Tatsache, dass $t^2 \max_{1\le k\le n} \sigma_{nk}^2 \le \frac{1}{2}$ für hinreichend große n (vgl. Bemerkung 4.2.4, 3): $\max_{1\le k\le n}\sigma_{nk}^2 \underset{n\to\infty}{\longrightarrow} 0$ folgt

$$\sum_{k=1}^{n}\left|1-\frac{\sigma_{nk}^2 t^2}{2}-e^{-\frac{\sigma_{nk}^2 t^2}{2}}\right| \le \frac{t^4}{4}\sum_{k=1}^{n}\sigma_{nk}^4 \le \frac{t^4}{4}\max_{1\le k\le n}\sigma_{nk}^2\cdot\underbrace{\sum_{k=1}^{n}\sigma_{nk}^2}_{=1}$$
$$= \frac{t^4}{4}\max_{1\le k\le n}\sigma_{nk}^2 \underset{n\to\infty}{\longrightarrow} 0\,.$$

Der Satz ist bewiesen. □

Bemerkung 4.2.7 Statt die Zufallsvariable X_{nk} in der Formulierung des Satzes 4.2.3 zu verwenden, können auch die Zufallsvariablen $\{X_n\}$ (wie im klassischen zentralen Grenzwertsatz) verwendet werden, wie folgende äquivalente Formulierung zeigt:

Satz 4.2.8 *(Lindeberg)*
Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsvariablen mit $\mathbb{E}X_n=\mu_n$, $0<\sigma_n^2=$Var $X_n<\infty\,, n\in\mathbb{N}$. Sei $S_n=\sum_{k=1}^n X_k$, $D_n^2=\sum_{k=1}^n\sigma_k^2$. Falls

$$\frac{1}{D_n^2}\sum_{k=1}^{n}\mathbb{E}\Big((X_k-\mu_k)^2\cdot I(|X_k-\mu_k|>\varepsilon D_n)\Big)\underset{n\to\infty}{\longrightarrow} 0\,,$$

dann gilt

$$\frac{S_n-\sum_{k=1}^n\mu_k}{D_n}\xrightarrow[n\to\infty]{d} Y\sim N(0,1)\,.$$

Beweis Setzen wir $X_{nk}=\frac{X_k-\mu_k}{D_n}$ und wenden wir den Satz 4.2.3 auf diese Folge von $\{X_{nk}\}_{k=1}^{n\in\mathbb{N}}$ an, die den Voraussetzungen des Satzes genügt. □

Folgerung 4.2.9 Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsvariablen mit $|X_n|\le c<\infty\,, n\in\mathbb{N}$ und $D_n\underset{n\to\infty}{\longrightarrow}\infty$. Dann gilt der zentrale Grenzwertsatz in der Form des Satzes 4.2.8.

Beweis Wir müssen die Gültigkeit der Lindeberg–Bedingung (4.4) prüfen: aus der Ungleichung von Tschebyschew folgt

$$\mathbb{E}\Big((X_k-\mu_k)^2\cdot I(|X_k-\mu_k|>\varepsilon D_n)\Big)\underset{|X_k|\le c,|\mathbb{E}X_k|\le c}{\le}(2c)^2 P(|X_k-\mu_k|>\varepsilon D_n)$$
$$\le 4c^2\frac{\sigma_k^2}{\varepsilon^2 D_n^2}$$

für $1\le k\le n$ und somit

$$\frac{1}{D_n^2}\sum_{k=1}^{n}\mathbb{E}\Big((X_k-\mu_k)^2\cdot I(|X_k-\mu_k|>\varepsilon D_n)\Big)\le 4c^2\frac{\overbrace{\sum_{k=1}^{n}\sigma_k^2}^{=D_n^2}}{\varepsilon^2 D_n^4}=\frac{4c^2}{\varepsilon^2 D_n^2}\underset{n\to\infty}{\longrightarrow} 0\,,$$

weil $D_n^2 \underset{n\to\infty}{\longrightarrow} \infty$. □

Bemerkung 4.2.10 Aus den Sätzen 4.2.3 und 4.2.8 ist ersichtlich, dass die Lindeberg–Bedingung (4.4) eine hinreichende Bedingung für die Gültigkeit des zentralen Grenzwertsatzes ist. Es gibt aber auch andere Bedingungen, die wir jetzt auflisten:

1. *Feller–Bedingung:*
$$\max_{1\le k\le n} \mathbb{E}X_{nk}^2 \underset{n\to\infty}{\longrightarrow} 0\,. \tag{4.7}$$

2. *Gleichmäßige asymptotische Kleinheit der Summanden:*
$$\max_{1\le k\le n} P(|X_{nk}|>\varepsilon) \underset{n\to\infty}{\longrightarrow} 0 \quad \forall\varepsilon>0\,. \tag{4.8}$$

3. *Ljapunow–Bedingung:*
$$\sum_{k=1}^{n} \mathbb{E}|X_{nk}|^{2+\delta} \underset{n\to\infty}{\longrightarrow} 0 \tag{4.9}$$
für ein $\delta>0$.

Zwischen Ihnen gilt folgende Relation:

$$(4.9) \Longrightarrow (4.4) \Longrightarrow (4.7) \Longrightarrow (4.8)$$

Die Relation (4.4) $\Longrightarrow$ (4.7) $\Longrightarrow$ (4.8) wurde bereits in der Bemerkung 4.2.4 gezeigt. Die Relation (4.9) $\Longrightarrow$ (4.4) zeigen wir im Satz von Ljapunow (vgl. Satz 4.2.13) Im folgenden Satz wird behauptet, dass unter gewissen Voraussetzungen die Lindeberg–Bedingung auch notwendig für die Gültigkeit des zentralen Grenzwertsatzes ist:

Satz 4.2.11 *(Lindeberg–Feller)*
Sei $\{X_{nk}\}_{k=1\ldots n,\, n\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen Zufallsvariablen mit

$$\begin{aligned} \mathbb{E}X_{nk} &= 0, & k&=1,\ldots,n\,, & n&\in\mathbb{N},\\ 0<\sigma_{nk}^2 &= \operatorname{Var} X_{nk}<\infty, & k&=1,\ldots,n\,, & n&\in\mathbb{N},\\ \sum_{k=1}^{n}\sigma_{nk}^2 &= 1\,, & n&\in\mathbb{N}\,. \end{aligned}$$

Falls eine der Bedingungen (4.7) oder (4.8) gilt, dann ist die Lindeberg–Bedingung (4.4) hinreichend und notwendig für die Gültigkeit des zentralen Grenzwertsatzes in der Form

$$\sum_{k=1}^{n} X_{nk} \xrightarrow{d} Y \sim N(0,1)\,.$$

Beweis siehe [19, 334–335]. □

Wie folgende Beispiele zeigen, ist die Lindeberg–Bedingung im Allgemeinen nicht notwendig:

Beispiel 4.2.12

1. *Lindeberg–Bedingung ist nicht notwendig:*
   Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen Zufallsvariablen mit
   $$X_n \sim N(0,\sigma_n^2),\ \sigma_1^2 = 1,\ \sigma_n^2 = 2^{n-2},\ n \geq 2\,.$$
   Für $S_n = \sum_{j=1}^n X_j$ gilt dann
   $$\begin{aligned}\mathbb{E}S_n &= 0,\\ D_n^2 = \text{Var } S_n &= \sum_{j=1}^n \sigma_j^2 = 1+1+2+2^2+\ldots+2^{n-2}\\ &= 1+\left(\frac{1-2^{n-1}}{1-2}\right) = 2^{n-1}\,,\end{aligned}$$
   somit
   $$X_{nk} = \frac{X_k}{D_n} \sim \begin{cases} N(0,2^{1-n}) & k=1\\ N(0,2^{k-n-1}) & k\geq 2\end{cases} \text{ und } \sum_{k=1}^n X_{nk} \sim N(0,1)\,, n\in\mathbb{N}\,,$$
   d.h. es gilt der zentrale Grenzwertsatz.
   Allerdings
   $$\max_{1\leq k\leq n} \mathbb{E}X_{nk}^2 = \frac{\sigma_n^2}{D_n^2} = \text{Var } X_{nn} = \frac{1}{2} \underset{n\to\infty}{\not\longrightarrow} 0\,,$$
   somit ist die Bedingung von Feller (und folglich die von Lindeberg) nicht erfüllt.

2. *Folge $\{X_n\}_{n\in\mathbb{N}}$, für die der zentrale Grenzwertsatz nicht gilt:*
   Sei $\{Y_n\}_{n\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen Zufallsvariablen
   $$Y_n = \begin{cases} 1, & \text{mit Wkt. } \frac{1}{2}\\ -1, & \text{sonst}\end{cases}, \quad n\in\mathbb{N}\,.$$
   Definieren wir $X_n = \frac{\sqrt{15}}{4^n} Y_n$, $n\in\mathbb{N}$. Die Folge $\{X_n\}_{n\in\mathbb{N}}$ besteht aus stochastisch unabhängigen Zufallsvariablen mit
   $$\mathbb{E}X_n = \frac{\sqrt{15}}{4^n}\mathbb{E}Y_n = 0 \text{ und Var } X_n = \frac{15}{4^{2n}}\text{Var } Y_n = \frac{15}{4^{2n}}\,.$$

Somit gilt für $S_n = \sum_{j=1}^n X_j$, $\mathbb{E}S_n = 0$,

$$D_n^2 = \text{Var } S_n = \sum_{j=1}^n \text{Var } X_j = 15\left(\frac{1}{16} + \frac{1}{(16)^2} + \ldots + \frac{1}{(16)^n}\right)$$
$$= \frac{1}{16}\frac{1-16^{-n}}{1-16^{-1}} \cdot 15 = 1 - \frac{1}{16^n},\ n \in \mathbb{N}\,.$$

Deshalb $D_n^2 \underset{n\to\infty}{\longrightarrow} 1$. Weiterhin hat man mit Wahrscheinlichkeit 1

$$|S_n| \le \sqrt{15}\left(\frac{1}{4} + \frac{1}{4^2} + \ldots + \frac{1}{4^n}\right) = \sqrt{15}\cdot\frac{1}{4}\cdot\frac{1-\frac{1}{4^n}}{1-\frac{1}{4}} = \frac{\sqrt{15}}{3}\cdot\left(1 - \frac{1}{4^n}\right)$$

und somit

$$\frac{|S_n|}{D_n} \le \frac{\sqrt{15}}{3}\frac{\left(1-\frac{1}{4^n}\right)}{\sqrt{1-\frac{1}{16^n}}} \le 2 \qquad \text{f.s.}$$

für große n, was darauf hindeutet, dass $\frac{S_n}{D_n} \underset{n\to\infty}{\overset{d}{\not\longrightarrow}} Y \sim N(0,1)$, da der Träger von $N(0,1)$ unbeschränkt ist. Mit Hilfe des 3-Reihen-Satzes (vgl. Folgerung 3.6.7) kann man hier zeigen, dass aus $\sum_{k=1}^\infty VarX_k < \infty$ gerade $\sum_{k=1}^\infty X_k < \infty$ f.s. folgt. Somit muss der Zentrale Grenzwertsatz nicht gelten!

Satz 4.2.13 *(Bedingung von Ljapunow)* Sei $\{X_{nk}\}_{k=1\ldots n\,,\,n\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen Zufallsvariablen mit

$$\mathbb{E}X_{nk} = 0\,,\ 0 < \sigma_{nk}^2 = \text{Var } X_{nk} < \infty \quad \forall k = 1,\ldots,n,\ \sum_{k=1}^n \sigma_{nk}^2 = 1 \quad \forall n \in \mathbb{N}\,.$$

Falls zusätzlich die Ljapunow–Bedingung 4.9) (Bemerkung 4.2.10) gilt für ein $\delta > 0$, dann gilt

$$\sum_{k=1}^n X_{nk} \underset{n\to\infty}{\overset{d}{\longrightarrow}} Y \sim N(0,1)\,.$$

Beweis Zeigen wir, dass aus der Ljapunow–Bedingung die Lindeberg–Bedingung folgt: Für alle $\delta > 0$, $\varepsilon > 0$ gilt

$$\mathbb{E}\left(X_{nk}^2 \cdot I(|X_{nk}| > \varepsilon)\right) = \varepsilon^2\mathbb{E}\left(\left(\frac{X_{nk}}{\varepsilon}\right)^2 \cdot I\left(\frac{|X_{nk}|}{\varepsilon} > 1\right)\right)$$
$$\le \varepsilon^2\mathbb{E}\left(\left(\frac{X_{nk}}{\varepsilon}\right)^{2+\delta} \cdot I\left(\left|\frac{X_{nk}}{\varepsilon}\right| > 1\right)\right) \le \varepsilon^{-\delta}\cdot\mathbb{E}|X_{nk}|^{2+\delta}$$

und somit

$$\sum_{k=1}^n \mathbb{E}\left(X_{nk}^2 \cdot I(|X_{nk}| > \varepsilon)\right) \le \varepsilon^{-\delta}\sum_{k=1}^n \mathbb{E}|X_{nk}|^{2+\delta} \underset{n\to\infty}{\longrightarrow} 0\,.$$

Somit ist die Lindeberg–Bedingung erfüllt, und die Behauptung des Satzes 4.2.13 folgt aus dem Satz 4.2.3. □

Beispiel 4.2.14 Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsvariablen mit

$$X_n = \begin{cases} n, & \text{mit Wahrscheinlichkeit } \frac{1}{2}, \\ -n, & \text{mit Wahrscheinlichkeit } \frac{1}{2}. \end{cases}$$

Es gilt $\mathbb{E}X_n = 0$, $\sigma_n^2 = \text{Var } X_n = n^2$, für $\delta = 1$

$$\mathbb{E}|X_n|^{2+\delta} = \mathbb{E}|X_n|^3 = n^3 \,, n \in \mathbb{N},$$

somit

$$\begin{aligned} D_n^2 &= \sum_{k=1}^n \sigma_k^2 = \sum_{k=1}^n k^2 = n^3 \left(\frac{1}{n} + \left(\frac{n-1}{n}\right)^2 \cdot \frac{1}{n} + \ldots + \left(\frac{1}{n}\right)^2 \cdot \frac{1}{n} \right) \\ &\underset{n\to\infty}{\sim} n^3 \int_0^1 x^2 \, dx = \frac{n^3}{3}, \\ \sum_{k=1}^n \mathbb{E}|X_k|^{2+\delta} &= \sum_{k=1}^n k^3 = n^4 \left(\frac{1}{n} + \left(\frac{n-1}{n}\right)^3 \cdot \frac{1}{n} + \ldots + \left(\frac{1}{n}\right)^3 \cdot \frac{1}{n} \right) \\ &\underset{n\to\infty}{\sim} n^4 \int_0^1 x^3 \, dx = \frac{n^4}{4} \,. \end{aligned}$$

Daraus folgt für $X_{nk} = \frac{X_k}{D_n}$ und $\delta = 1$

$$\sum_{k=1}^n \mathbb{E}|X_{nk}|^{2+\delta} = \frac{\sum_{k=1}^n \mathbb{E}|X_k|^3}{D_n^3} \underset{n\to\infty}{\sim} \frac{\frac{n^4}{4}}{\left(\sqrt{\frac{n^3}{3}}\right)^3} = const \cdot \frac{n^4}{n^{\frac{9}{2}}} \underset{n\to\infty}{\longrightarrow} 0 \,.$$

Somit ist die Bedingung von Ljapunow erfüllt und für $\sum_{k=1}^n X_{nk}$ gilt der zentrale Grenzwertsatz, d.h.

$$\sqrt{3}\frac{\sum_{k=1}^n X_k}{n^{\frac{3}{2}}} \xrightarrow[n\to\infty]{d} Y \sim N(0,1) \,.$$

4.2.3 Konvergenzgeschwindigkeit im zentralen Grenzwertsatz

In diesem Abschnitt möchten wir die *Schnelligkeit der Konvergenz im zentralen Grenzwertsatz* untersuchen. Damit aber diese Fragestellung überhaupt sinnvoll erscheint, muss die Konvergenz im zentralen Grenzwertsatz gleichmäßig sein:

$$\sup_{x\in\mathbb{R}} \left| P\left(\frac{\sum_{j=1}^n X_j - n\mu}{\sqrt{n}\sigma} \le x \right) - \Phi(x) \right| \underset{n\to\infty}{\longrightarrow} 0 \,.$$

Das ist tatsächlich der Fall, wie aus der Stetigkeit von $\Phi(x)$ und dem folgenden Lemma 4.2.15 hervorgeht.

Lemma 4.2.15 Sei $X_n \xrightarrow[n\to\infty]{d} X$, wobei $F_n(x) = P(X_n \le x)$ und $F(x) = P(X \le x)$. Falls $F(x)$ für alle $x \in \mathbb{R}$ stetig ist, dann konvergiert F_n *gleichmäßig* gegen F:

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \underset{n\to\infty}{\longrightarrow} 0\,.$$

Das Hauptergebnis des Abschnittes 4.2.3 ist folgende Abschätzung der Konvergenzgeschwindigkeit im klassischen zentralen Grenzwertsatz.

Satz 4.2.16 *(Berry–Esséen)*
Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen identisch verteilten Zufallsvariablen mit $\mathbb{E}X_n = \mu$, Var $X_n = \sigma^2 > 0$, $\mathbb{E}|X_n|^3 < \infty$. Sei

$$F_n(x) = P\left(\frac{\sum_{j=1}^n X_j - n\mu}{\sqrt{n}\sigma} \le x\right), \quad x \in \mathbb{R},\, n \in \mathbb{N}\,.$$

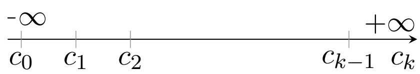
Dann gilt

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \le \frac{c \cdot \mathbb{E}|X_1 - \mu|^3}{\sigma^3\sqrt{n}}\,,$$

wobei c eine Konstante ist, $\frac{1}{\sqrt{2\pi}} \le c < 0{,}4785$, $\frac{1}{\sqrt{2\pi}} \approx 0{,}39894$.

Beweis des Lemmas 4.2.15
Sei $\varepsilon > 0$ beliebig, ein $k > \frac{1}{\varepsilon}$ ($\Longleftrightarrow \frac{1}{k} < \varepsilon$). Wählen wir

$$c_0 = -\infty, \quad c_j = F^{-1}\left(\frac{j}{k}\right), \quad j = 1, \dots, k-1, \quad c_k = \infty\,.$$



Es gilt

$$F(c_j) - F(c_{j-1}) = \frac{j}{k} - \frac{j-1}{k} = \frac{1}{k} < \varepsilon \quad \forall 1 \le j \le k\,. \tag{4.10}$$

Zu zeigen ist es, dass $\forall \varepsilon > 0 \quad \exists\, n_0 : \forall n \ge n_0\, \forall a < b$

$$|F_n(b) - F(b) - F_n(a) + F(a)| < const \cdot \varepsilon\,.$$

Es existiert ein n_0 :

$$\forall n \ge n_0 \quad |F_n(c_j) - F_n(c_i) - F(c_j) + F(c_i)| < \varepsilon \quad \forall 1 \le i < j \le k\,, \tag{4.11}$$

wegen $F_n(x) \underset{n\to\infty}{\longrightarrow} F(x) \quad \forall x \in \mathbb{R}$.
Für alle $a < b$ existieren i, j : $\quad a \in [c_i, c_{i+1})$, $b \in [c_j, c_{j+1})$. Sei X_n eine Zufallsvariable mit der Verteilungsfunktion F_n. Dann gilt:

1. *Untere Schranke*

   (a) Falls $i < j$, dann
   $$\begin{aligned} F_n(b) - F_n(a) &= P(a < X_n \le b) \underset{[a,b]\supset[c_{i+1},c_j]}{\ge} P(c_{i+1} \le X_n \le c_j) \\ &\underset{(4.11)}{>} F(c_j) - F(c_{i+1}) - \varepsilon \underset{(4.10)}{>} F(b) - F(a) - 3\varepsilon. \end{aligned}$$

   (b) Falls $i = j$, gilt analog dazu
   $$F_n(b) - F_n(a) \ge 0$$
   und da $[a, b] \subset [c_i, c_{i+1}]$ und $F(b) - F(a) < \varepsilon$ auch
   $$F_n(b) - F_n(a) \ge 0 > F(b) - F(a) - \varepsilon\,.$$

2. *Obere Schranke*

   (a) Falls $i > 0$, $j + 1 < k$, dann
   $$F_n(b) - F_n(a) \le P(c_i \le X_n \le c_{j+1}) < F(b) - F(a) + 3\varepsilon\,,$$
   wie in 1a).

   (b) Falls $i = 0$, $j + 1 < k$, dann
   $$\begin{aligned} F_n(b) - F_n(a) &\le P(X_n < c_{j+1}) = 1 - P(X_n \ge c_{j+1}) \\ &\le 1 - P(c_{j+1} \le X_n \le c_{k-1}) \\ &< \underbrace{1 - F(c_{k-1})}_{<\varepsilon} + F(c_{j+1}) + \varepsilon \\ &< \underbrace{F(c_{j+1}) - F(b)}_{<\varepsilon} + F(b) \quad F(b) + 3\varepsilon < F(b) - F(a) + \underbrace{4\varepsilon}_{\text{weil } F(a)<\varepsilon}. \end{aligned}$$

   (c) Falls $i > 0$, $j + 1 = k$, analog zu 2a), 2b):
   $$\begin{aligned} F_n(b) - F_n(a) &\le P(c_i < X_n) = 1 - P(X_n \le c_i) \le 1 - F(c_i) + \varepsilon \\ &< F(b) - F(c_i) + 2\varepsilon \underset{|c_i - a| < \frac{1}{k}}{<} F(b) - F(a) + 3\varepsilon\,, \end{aligned}$$
   weil $1 - F(b) < \varepsilon$.

   (d) Falls $i = 0$, $j + 1 = k$, dann
   $$F_n(b) - F_n(a) \underset{\text{weil } F(a)<\varepsilon,\, F(b)>1-\varepsilon}{\le} 1 < F(b) - F(a) + 2\varepsilon\,.$$

   Also
   $$\forall \varepsilon > 0 \quad \exists n_0 : \forall n \ge n_0 \quad |F_n(b) - F(b) - F_n(a) + F(a)| < 4\varepsilon\,.$$

□

Für den Beweis des Satzes 4.2.22 wird eine Reihe von Hilfssätzen benötigt:

Lemma 4.2.17 *(Fourier–Umkehrformel)*
Sei X eine Zufallsvariable mit der charakteristischen Funktion $\varphi_X(t) = \mathbb{E}e^{itX}$. Falls $|\varphi_X| \in L^1(\mathbb{R})$, dann besitzt die Verteilung von X eine Dichte $f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ixy}\varphi_X(y)\,dy$, die beschränkt ist:

$$|f(x)| \le \frac{1}{2\pi}\int_{\mathbb{R}} |\varphi_X(y)|\,dy\,, \quad x \in \mathbb{R}\,.$$

Beweis folgt aus Satz 2.2.7, 2).

Lemma 4.2.18 *(Riemann–Lebesgue)*
Ist $f : \mathbb{R} \to \mathbb{C}$ eine Funktion mit $|f| \in L^1(\mathbb{R})$, so gilt

$$\varphi(\lambda) = \int_{\mathbb{R}} e^{i\lambda x} f(x)\,dx \underset{\lambda\to\pm\infty}{\longrightarrow} 0\,.$$

(Ohne Beweis, vgl. [8, Satz 1.17.14]).

Lemma 4.2.19 Seien F und G zwei Verteilungsfunktionen, wobei G differenzierbar auf $\mathbb{R}$ ist mit $G'(x) \le m < \infty$, $x \in \mathbb{R}$. Für ein $T > 0$ sei $v_T(x)$ die Dichte

$$v_T(x) = \frac{1}{\pi}\frac{1-\cos(Tx)}{Tx^2}\,, \quad x \in \mathbb{R}$$

der Verteilung v_T.
Es gilt

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)| \le 2\left(\sup_{t\in\mathbb{R}}\left|\int_{\mathbb{R}} \Delta(t-x)v_T(x)\,dx\right| + \frac{12m}{\pi T}\right)\,,$$

wobei $\Delta(x) = F(x) - G(x)$ für jedes $x \in \mathbb{R}$.

Beweis Führen wir folgende Bezeichnungen ein:

$$\begin{aligned}
\varrho &= \sup_{x\in\mathbb{R}} |F(x) - G(x)|\,,\\
\Delta^T(t) &= \int_{\mathbb{R}} \Delta(t-x)v_T(x)\,dx = (\Delta * v_T)(t)\,,\\
\varrho^T &= \sup_{t\in\mathbb{R}} |\Delta^T(t)|\,.
\end{aligned}$$

Für ein beliebiges $\varepsilon > 0$ existiert ein $x_0 \in \mathbb{R}:\quad \varrho \le |F(x_0) - G(x_0)| + \varepsilon$. Setzen wir ohne Beschränkung der Allgemeinheit voraus, dass $F(x_0) \ge G(x_0)$. Es gilt

$$\begin{aligned}
\Delta(x_0+s) =& F(x_0+s) - G(x_0+s)\\
&\ge F(x_0) - G(x_0) - ms\\
&\ge \varrho - \varepsilon - ms \quad \forall s \ge 0\,,
\end{aligned}$$

weil

$$F(x_0+s) \geq F(x_0),\ s \geq 0, \qquad \frac{G(x_0)-G(x_0+s)}{s} \geq -m-\delta \text{ für } s \to 0\,.$$

Für $h = \frac{\varrho}{2m}$, $t = x_0 + h$ und $x = h - s$ (somit $x_0 + s = t - x$) erhält man

$$\Delta(\underbrace{t-x}_{=x_0+s}) \geq \varrho - \varepsilon - m\left(\frac{\varrho}{2m} - x\right) = \varrho - \frac{\varrho}{2} - \varepsilon + mx = \frac{\varrho}{2} - \varepsilon + mx$$

für alle $|x| \leq h$, da $s = h - x \geq 0$ sein muss. Für $|x| > h$ können wir einfach $\Delta(t-x) \geq -\varrho$ schreiben. Somit gilt

$$\begin{aligned}
|\Delta^T(t)| &\geq \underbrace{\int_{\mathbb{R}} \Delta(t-x) v_T(x)\, dx}_{=\Delta^T(t)} \\
&\geq \int_{|x|\leq h} \left(\frac{\varrho}{2} - \varepsilon + mx\right) v_T(x)\, dx - \varrho \int_{|x|>h} v_T(x)\, dx.
\end{aligned}$$

Da aus Symmetriegründen $\int_{|x|\leq h} x v_T(x)\, dx = 0$ ist, folgt

$$\begin{aligned}
|\Delta^T(t)| &\geq \left(\frac{\varrho}{2} - \varepsilon\right) \underbrace{\left(\int_{|x|\leq h} + \int_{|x|>h}\right) v_T(x)\, dx}_{=1} \\
&\quad - \left(\frac{\varrho}{2} - \varepsilon\right) \int_{|x|>h} v_T(x)\, dx - \varrho \int_{|x|>h} v_T(x)\, dx \\
&\geq \frac{\varrho}{2} - \varepsilon - \frac{3\varrho}{2} \int_{|x|>h} v_T(x)\, dx \geq \frac{\varrho}{2} - \varepsilon - \frac{3}{2} \varrho \frac{4}{\pi T h} \\
&= \frac{\varrho}{2} - \varepsilon - \frac{3}{2} \cdot 4 \cdot \frac{\overbrace{2m}^{\frac{\varrho}{h}=2m}}{\pi T} = \frac{\varrho}{2} - \varepsilon - \frac{12m}{\pi T}\,,
\end{aligned}$$

weil

$$\begin{aligned}
\frac{1}{\pi T} \int_{|x|>h} \frac{1-\cos(Tx)}{x^2}\, dx &= \frac{2}{\pi} \int_{T|x|>Th} \frac{\sin^2\left(\frac{Tx}{2}\right)}{(Tx)^2} d(Tx) \\
&\overset{y=Tx}{=} \frac{2}{\pi} \int_{|x|>Th} \frac{\sin^2(y/2)}{y^2}\, dy \leq \frac{2}{\pi} \int_{|y|>Th} \frac{1}{y^2}\, dy \\
&= \frac{4}{\pi} \int_{y>Th} \frac{1}{y^2}\, dy = \frac{4}{\pi} \left(-\frac{1}{y}\right)\Big|_{Th}^{\infty} = \frac{4}{\pi T h}\,, \\
&\Longrightarrow\ -\int_{|x|>h} v_T(x)\, dx \geq -\frac{4}{\pi T h}\,.
\end{aligned}$$

Also zusammengefasst gilt $|\Delta^T(t)| \geq \frac{\varrho}{2} - \varepsilon - \frac{12m}{\pi T}$. Somit erhält man

$$\frac{\varrho}{2} \leq \max_t |\Delta^T(t)| + \frac{12m}{\pi T} + \varepsilon\,.$$

Da $\varepsilon > 0$ beliebig ist, folgt für $\varepsilon \to 0$ die Behauptung. □

Übungsaufgabe 4.2.20 Beweisen Sie, dass die charakteristische Funktion der Verteilung mit der Dichte

$$v_T(x) = \frac{1-\cos(Tx)}{\pi T x^2}\,, \quad x \in \mathbb{R}$$

folgendermaßen aussieht:

$$\varphi_T(t) = \int_{\mathbb{R}} e^{itx} v_T(x)\,dx = \begin{cases} 1 - \frac{|t|}{T}, & |t| \le T \\ 0, & \text{sonst.} \end{cases}$$

Bezeichnen wir $\nu_T(x) = \int_{-\infty}^{x} v_T(y)\,dy$ als die Verteilungsfunktion mit Dichte v_T.

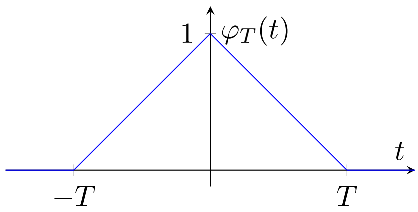


Abbildung 4.2: Charakteristische Funktion $\varphi_T(t)$

Lemma 4.2.21 *(Ungleichung von Esséen)*
Seien $F(x), G(x)$ zwei Verteilungsfunktionen mit charakteristischen Funktionen $f(t)$, $g(t)$. Falls $G(x)$ eine Ableitung $G'(x)$, $x \in \mathbb{R}$ mit $\sup_{x\in\mathbb{R}} G'(x) < \infty$ besitzt, dann gilt für alle $T > 0$

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)| \le \frac{2}{\pi}\int_0^T \left|\frac{f(t)-g(t)}{t}\right| dt + \frac{24}{\pi T}\sup_{x\in\mathbb{R}} |G'(x)|.$$

Beweis Wir wollen annehmen, dass die rechte Seite der Ungleichung endlich ist, also

$$\frac{f(t)-g(t)}{t}$$

integrierbar, sonst ist die Ungleichung trivial. Wir werden die Verteilungsfunktion "glätten", indem wir $\mu = F * v_T$ und $\nu = G * v_T$ betrachten. Seien φ_μ bzw. φ_ν die dazugehörigen charakteristischen Funktionen. Da für $Z \sim \mu$ $Z \stackrel{d}{=} X + Y$ gilt, wobei X, Y unabhängige Zufallsvariablen mit $X \sim F$, $Y \sim \nu_T$, bekommt man somit $\varphi_\mu = f \cdot \varphi_T$, $\varphi_\nu = g \cdot \varphi_T$, wobei φ_T die oben genannte charakteristische Funktion von ν_T ist. Da φ_T und somit φ_μ, φ_ν außerhalb von $[-T, T]$ verschwinden, sind sie insbesondere integrierbar, somit hat man mit Hilfe des Lemmas 4.2.17, dass μ und ν absolut stetig sind mit Dichten

$$\begin{aligned} f_\mu(t) &= \frac{1}{2\pi}\int_{-T}^{T} e^{-itx}\varphi_T(x)f(x)\,dx\,, \\ f_\nu(t) &= \frac{1}{2\pi}\int_{-T}^{T} e^{-itx}\varphi_T(x)g(x)\,dx\,. \end{aligned}$$

$$\Longrightarrow f_\mu(t)-f_\nu(t) = \frac{1}{2\pi}\int_{-T}^{T} e^{-itx}(f(x)-g(x))\varphi_T(x)\,dx\,. \tag{4.12}$$

Um die Ungleichung von Esséen zu beweisen, benutzen wir Lemma 4.2.19, wobei

$$\Delta^T(t) = \int_{\mathbb{R}} \Delta(t-x)v_T(x)\,dx$$

$$= (\underbrace{F*v_T}_{\mu}-\underbrace{G*v_T}_{\nu})(t) = \frac{1}{2\pi}\int_{-T}^{T} e^{-itx}\frac{f(x)-g(x)}{-ix}\varphi_T(x)\,dx. \tag{4.13}$$

Somit folgt daraus

$$\begin{aligned} \varrho^T &= \sup_{t\in\mathbb{R}}|\Delta^T(t)| \le \frac{1}{2\pi}\int_{-T}^{T}\left|e^{-itx}\frac{f(x)-g(x)}{-ix}\varphi_T(x)\right|dx \\ &\le \frac{1}{2\pi}\int_{-T}^{T}\underbrace{|e^{-itx}|}_{=1}\cdot\left|\frac{f(x)-g(x)}{x}\right|\cdot\underbrace{|\varphi_T(x)|}_{\le 1}\,dx \le \frac{1}{2\pi}\int_{-T}^{T}\left|\frac{f(x)-g(x)}{x}\right|dx \\ &= \frac{1}{\pi}\int_0^T\frac{|f(x)-g(x)|}{x}\,dx \end{aligned}$$

und aus Lemma 4.2.19 die Behauptung des Satzes. Der Integrand auf der rechten Seite lässt sich nach oben durch

$$\left|\frac{f(x)-g(x)}{x}\right|$$

abschätzen, wie wir soeben gesehen haben. Diese Schranke ist integrierbar. Somit kann man die beiden Seiten der Formel (4.13) bzgl. t differenzieren (nach dem Satz von Lebesgue). Dabei bekommt man genau die Gleichung (4.12). Daraus folgt, dass

$$\Delta^T(t) = \frac{1}{2\pi}\int_{-T}^{T} e^{-itx}\frac{f(x)-g(x)}{-ix}\varphi_T(x)\,dx + c$$

für eine Konstante $c\in\mathbb{C}$. Da $\Delta^T(t) = F*\nu_T(t) - G*\nu_T(t) \underset{t\to\infty}{\longrightarrow} 1-1=0$ und nach Lemma 4.2.18 dasselbe für

$$\int_{-T}^{T} e^{-itx}\frac{f(x)-g(x)}{-ix}\varphi_T(x)\,dx$$

gilt, folgt $c=0$, und der Satz ist bewiesen. $\square$

Den Satz von Berry–Esséen formulieren und beweisen wir in einer etwas allgemeineren Form, woraus Satz 4.2.16 folgen wird.

Satz 4.2.22 *(Berry–Esséen)*
Sei $\{X_n\}$ eine Folge von unabhängigen Zufallsvariablen mit $\mathbb{E}X_n = 0$ und $\sigma_n^2 = \text{Var } X_n \in (0,\infty)$. Dann gilt für alle n:

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{c}{\left(\sum_{j=1}^n \sigma_j^2\right)^{\frac{3}{2}}} \sum_{j=1}^n \mathbb{E}|X_j|^3 ,$$

wobei $c = 8.74$ und

$$F_n(x) = P\left(\frac{\sum_{j=1}^n X_j}{\sqrt{\sum_{j=1}^n \sigma_j^2}} \leq x\right) .$$

Bemerkung 4.2.23
Falls X_j unabhängig identisch verteilt sind mit $0 < \sigma^2 = \text{Var } X_j < \infty$, $\mu = \mathbb{E}X_i$, dann gilt

$$\sum_{j=1}^n \sigma_j^2 = n\sigma^2, \quad \sum_{j=1}^n \mathbb{E}|Y_j|^3 = n\mathbb{E}|Y_1|^3$$

für $Y_j = X_j - \mu$ und

$$\frac{\sum_{j=1}^n \mathbb{E}|Y_j|^3}{\left(\sum_{j=1}^n \sigma_j^2\right)^{\frac{3}{2}}} = \frac{n\mathbb{E}|Y_1|^3}{n\sqrt{n}\sigma^3} = \frac{\mathbb{E}|X_1-\mu|^3}{\sqrt{n}\sigma^3} .$$

Somit ist der Satz 4.2.16 für $c = 8.74$ bewiesen. Diese Überlegungen können verfeinert werden. Dann folgt $\frac{1}{\sqrt{2\pi}} \leq c < 0,4785$, vgl. [23].

Beweis des Satzes 4.2.22
Nach Lemma 4.2.21 gilt

$$\varrho_n \;=\; \sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{2}{\pi}\int_0^T \frac{\left|\varphi_{\frac{\sum_{j=1}^n X_j}{D_n}}(x) - e^{-\frac{x^2}{2}}\right|}{|x|}\,dx + \frac{24}{T\pi\cdot\sqrt{2\pi}} ,$$

wobei

$$D_n^2 = \sum_{j=1}^n \sigma_j^2 \text{ und } \max_{x\in\mathbb{R}} |\Phi'(x)| = \frac{1}{\sqrt{2\pi}}, \qquad T > 0 .$$

Desweiteren gilt

$$\varphi_{\frac{\sum_{j=1}^n X_j}{D_n}}(x) = \varphi_{\sum_{j=1}^n X_j}\left(\frac{x}{D_n}\right) = \prod_{j=1}^n \varphi_{X_j}\left(\frac{x}{D_n}\right) ,$$

weil X_i unabhängig sind und genauso

$$e^{-\frac{x^2}{2}} = e^{-\sum_{j=1}^{n} \frac{\sigma_j^2}{D_n^2} \frac{x^2}{2}}.$$

Bezeichnen wir $\alpha_j = \mathbb{E}|X_j|^3$, $\alpha = \sum_{j=1}^{n} \alpha_j^3$, $\beta_j = \varphi_{X_j}\left(\frac{x}{D_n}\right)$, $\gamma_j = e^{-\frac{1}{2}\frac{\sigma_j^2}{D_n^2}x^2}$. Sei n beliebig aber fest. Falls ein j existiert, sodass $\alpha_j = \infty$, dann ist die Ungleichung im Satz 4.2.22 trivialerweise erfüllt. Deswegen setzen wir voraus, dass $\alpha_j < \infty \, \forall j = 1 \dots n$. Setze $T = \frac{8D_n^3}{9\alpha}$. Somit gilt

$$\varrho_n \leq \frac{2}{\pi} \int_0^T \frac{\left|\prod_{j=1}^{n} \beta_j - \prod_{j=1}^{n} \gamma_j\right|}{x}\, dx + \frac{24}{T\pi\sqrt{2\pi}}.$$

Zeigen wir, dass

$$\left|\prod_{j=1}^{n} \beta_j - \prod_{i=j}^{n} \gamma_j\right| \leq \frac{8}{9T}\left(\frac{|x|^3}{6} + \frac{5|x|^4}{72}\right) e^{-\frac{x^2}{8}}, \quad |x| \leq T \tag{4.14}$$

gilt. Daraus folgt die Behauptung des Satzes, nämlich

$$\begin{aligned}
\frac{2}{\pi}\int_0^T \frac{\left|\prod_{j=1}^{n} \beta_j - \prod_{j=1}^{n} \gamma_j\right|}{x}\, dx &\leq \frac{8}{9\cdot T}\frac{2}{\pi}\int_0^\infty e^{-\frac{x^2}{8}}\left(\frac{x^2}{6} + \frac{5}{72}x^3\right) dx \\
&= \frac{16}{9\cdot T\pi} \underbrace{\cdot}_{*} \left(\frac{1}{6}\cdot\frac{1}{2}\left(\frac{1}{8}\right)^{-\frac{3}{2}}\Gamma\left(\frac{3}{2}\right) + \frac{5}{72}\cdot\frac{1}{2}\left(\frac{1}{8}\right)^{-2}\underbrace{\Gamma(2)}_{=1}\right) \\
&= \frac{16}{9\cdot T\pi}\left(\frac{8^{\frac{3}{2}}}{12}\cdot\frac{1}{2}\sqrt{\pi} + \frac{32\cdot 5}{72}\right) = \frac{16}{9\cdot T\pi}\left(\frac{8\cdot 2\sqrt{2}}{4\cdot 3\cdot 2}\sqrt{\pi} + \frac{5\cdot 16\cdot 2}{2\cdot 4\cdot 9}\right) \\
&= \frac{16}{9\cdot T\pi}\left(\frac{2\sqrt{2\pi}}{3} + \frac{5\cdot 4}{9}\right) = \frac{32}{27\cdot T\pi}\left(\sqrt{2\pi} + \frac{5\cdot 2}{3}\right)
\end{aligned}$$

($*$ mit Hilfe der Formel $\int_0^\infty x^k e^{-\lambda x^2}\, dx = \frac{1}{2}\lambda^{-\frac{k+1}{2}}\Gamma\left(\frac{k+1}{2}\right)$, $k \in \mathbb{Z}_+$, $\lambda > 0$). Daraus folgt

$$\begin{aligned}
\varrho_n &\leq \frac{1}{T}\frac{32}{27}\underbrace{\left(\frac{\sqrt{2}}{\sqrt{\pi}} + \frac{5\cdot 2}{\pi}\right)}_{=\frac{2\pi + 5\cdot 2\sqrt{2\pi}}{\sqrt{2\pi}\cdot\pi}} + \frac{24}{T\pi\sqrt{2\pi}} = \frac{8}{T\pi\sqrt{2\pi}}\left(\frac{8}{27}(\pi + 5\sqrt{2\pi}) + 3\right) \\
&= \frac{8\alpha}{9\cdot\frac{8}{9}D_n^3}\cdot\frac{1}{\pi\sqrt{2\pi}}\left(\frac{8}{3}(\pi + 5\sqrt{2\pi}) + 27\right) = \frac{\alpha}{D_n^3}\cdot c,
\end{aligned}$$

wobei $c = \frac{1}{\pi\sqrt{2\pi}}\left(\frac{8}{3}(\pi + 5\sqrt{2\pi}) + 27\right) \approx 8,74$.
Jetzt beweisen wir die Abschätzung (4.14). Nach dem Lemma 4.2.6 erhalten

wir

$$|\beta_j| = \left|\varphi_{X_j}\left(\frac{x}{D_n}\right)\right| \leq \left|1 - \frac{\sigma_j^2 x^2}{2 \cdot D_n^2}\right| + \frac{\alpha_j |x|^3}{6 \cdot D_n^3}$$

bis zur 3. Ordnung, weil $\mathbb{E}X_j = 0$. Betrachten wir ausschließlich $|x| \leq T$.

Falls $\sigma_j T \leq D_n\sqrt{2}$, dann gilt $\sigma_j|x| \leq \sigma_j T \leq D_n\sqrt{2}$ und somit

$$1 - \frac{\sigma_j^2 x^2}{2D_n^2} \geq 0\,, \quad |x| \leq T\,.$$

Unter der Anwendung der Ungleichung $1 + x \leq e^x$, $x \in \mathbb{R}$ folgt daraus

$$|\beta_j| \leq e^{\left(-\frac{\sigma_j^2}{2 \cdot D_n^2} + \frac{\alpha_j T}{6 \cdot D_n^3}\right)x^2}\,,$$

weil $|x| \leq T$.

Falls $\sigma_j T > D_n\sqrt{2}$, dann gilt

$$|\beta_j| \leq 1 \leq e^{\left(-\frac{\sigma_j^2}{2D_n^2} + \frac{3}{8}\frac{\alpha_j T}{D_n^3}\right)x^2} = \delta_j\,, \quad j = 1 \ldots n\,,$$

was aus der Ungleichung $-\frac{\sigma_j^2}{2D_n^2} + \frac{3}{8}\frac{\alpha_j T}{D_n^3} \geq 0$ folgt. Tatsächlich gilt aufgrund der Ljapunow–Ungleichung

$$\begin{aligned}
\alpha_j \quad &\geq \quad \sigma_j^3 \underbrace{>}_{\sigma_j T > D_n\sqrt{2}} \frac{\sigma_j^2 D_n \sqrt{2}}{T} \underbrace{>}_{\frac{2\sqrt{2}}{3} < 1} \frac{\sigma_j^2 D_n\sqrt{2}}{T} \cdot \frac{2\sqrt{2}}{3} = \frac{4}{3} \cdot \frac{\sigma_j^2 D_n}{T} \\
&\Longrightarrow \quad \sigma_j^2 \leq \frac{3 \cdot \alpha_j T}{4D_n} \Longrightarrow -\frac{\sigma_j^2}{2 \cdot D_n^2} \geq -\frac{3 \cdot \alpha_j T}{8 \cdot D_n^3} \\
&\Longrightarrow \quad -\frac{\sigma_j^2}{2 \cdot D_n^2} + \frac{3}{8} \cdot \frac{\alpha_j T}{D_n^3} \geq 0\,.
\end{aligned}$$

Daher gilt in jedem Fall $|\beta_j| \leq \sigma_j$, $|x| \leq T$.
Trivialerweise folgt

$$|\gamma_j| = e^{-\frac{1}{2}\cdot\frac{\sigma_j^2}{D_n^2}x^2} \leq e^{\left(-\frac{\sigma_j^2}{2 \cdot D_n^2} + \frac{3}{8} \cdot \frac{\alpha_j T}{D_n^3}\right)x^2} = \delta_j\,, \quad \forall |x| \leq T, \quad j = 1 \ldots n\,.$$

Da $\varrho_n = \sup_{\mathbb{R}} |F_1(x) - \Phi(x)| \leq 1$, so ist die Berry–Esséen–Ungleichung trivialerweise erfüllt, falls $\frac{\alpha}{D_n^3} \geq \frac{1}{6}$ gilt. Daher werden wir im Folgenden stets

$\frac{\alpha}{D_n^3} < \frac{1}{6} \Longrightarrow T = \frac{8D_n^3}{9\alpha} = \frac{8}{9}\left(\frac{\alpha}{D_n^3}\right)^{-1} > \frac{8}{9}\cdot 6 = \dfrac{16}{3}$ annehmen. In diesem Fall gilt für alle $j = 1\dots n$

$$
\begin{aligned}
\prod_{\substack{l=1\\ l\neq j}}^{n} \delta_l \quad & \leq \quad \exp\left\{\left(\sum_{l\neq j}\sigma_l^2\frac{1}{D_n^2} + \frac{3\sum_{l\neq j}\alpha_l T}{8D_n^3}\right)x^2\right\} \\
& \leq \quad \exp\left\{\left(-\frac{1}{2}\left(1-\frac{\sigma_j^2}{D_n^2}\right)+\frac{3\alpha T}{8D_n^3}\right)x^2\right\} \\
& \underset{T=\frac{8D_n^3}{9\alpha}}{\leq} \quad \exp\left\{\left(-\frac{1}{2}+\frac{1}{2}\frac{\sigma_j^2}{D_n^2}+\frac{3}{8}\frac{\alpha\cdot 8}{9}D_n^{-3}\right)x^2\right\} \\
& \underset{\frac{\sigma_j}{D_n}\leq\frac{1}{4}}{\leq} \quad \exp\left\{\left(-\frac{1}{2}+\frac{1}{32}+\underbrace{\frac{1}{3}}_{\frac{\alpha}{D_n^3}<\frac{1}{6}<1}\right)x^2\right\} \leq e^{-\frac{x^2}{8}},
\end{aligned}
$$

weil

$$
-\frac{1}{2}+\frac{1}{3}+\frac{1}{32} = -\frac{1}{6}+\frac{1}{32} = \frac{-32+1}{6\cdot 32} = -\frac{31}{6\cdot 32} \leq -\frac{5}{32} \leq -\frac{4}{32} = -\frac{1}{8},
$$

falls $\dfrac{\sigma_l}{D_n} \leq \frac{4}{3}T^{-1} \leq \dfrac{1}{4}$. Falls $\frac{\sigma_l}{D_n} > \dfrac{4}{3T}$, dann gilt

$$
\delta_l = e^{-\frac{\sigma_l^2}{2D_n^2}+\frac{3\alpha_l T}{8D_n^3}} \underset{\alpha_l\geq\sigma_l^3}{\geq} e^{-\frac{\sigma_l^2}{2D_n^2}+\frac{3}{8}\frac{\sigma_l^3 T}{D_n^3}} > \exp\left\{\frac{1}{2}\frac{\sigma_l^2}{D_n^2}\underbrace{\left(-1+\frac{3}{4}\frac{\sigma_l}{D_n}T\right)}_{>0,\text{ da }\frac{\sigma_l}{D_n}>\frac{4}{3}T}\right\} \geq e^0 = 1,
$$

und somit

$$
\begin{aligned}
\prod_{\substack{l=1\\ l\neq j}}^{n}\delta_l \leq \prod_{l=1}^{n}\delta_l &= \exp\left\{\left(-\frac{1}{2}\underbrace{\frac{\sum_{l=1}^n\sigma_l^2}{D_n^2}}_{=1}+\frac{3\sum_{l=1}^n\alpha_l}{8D_n^2}\right)x^2\right\} \\
&\leq \exp\left\{\left(-\frac{1}{2}+\frac{3}{8}\frac{\alpha T}{D_n^3}+\frac{1}{32}\right)x^2\right\} \leq e^{-\frac{x^2}{8}}.
\end{aligned}
$$

Es gilt zusätzlich

$$
\begin{aligned}
|\beta_l-\gamma_l| \quad & \leq \quad \left|\beta_l - 1 + \frac{x^2\sigma_l^2}{2D_n^2}\right| + \left|\gamma_l - 1 + \frac{x^2\sigma_l^2}{2D_n^2}\right| \\
& \underset{\text{Lemma 4.2.6}}{\leq} \quad \frac{|x|^3\alpha_l}{3!\cdot D_n^3} + \frac{x^4\sigma_l^4}{\underbrace{4!}_{=8\cdot 3}D_n^4}\underbrace{\int_{\mathbb{R}}X^4\Phi(x)}_{=3} \\
& = \quad \frac{|x|^3\alpha_l}{6D_n^3} + \frac{x^4\sigma_l^4}{8\cdot D_n^4}.
\end{aligned}
$$

Daher bekommt man

$$\begin{aligned}
\sum_{l=1}^{n}|\beta_l-\gamma_l| &\le \frac{|x|^3\sum_{l=1}^n\alpha_l}{6D_n^3}+\frac{x^4\sum_{l=1}^n\sigma_l^4}{8D_n^4}=\frac{|x|^3\alpha}{6D_n^3}+\frac{x^4\overbrace{\sum_{l=1}^{n}\sigma_l^4}^{\le\alpha^{\frac13}\sum\alpha_l=\alpha^{\frac13}\cdot\alpha=\alpha^{\frac43}}}{8D_n^4}\\
&\underbrace{\le}_{(*)} \frac{|x|^3\alpha}{6D_n^3}+\frac{x^4}{8}\left(\frac{\alpha}{D_n^3}\right)^{\frac43} \underset{\text{weil } \frac{\alpha}{D_n^3}<\frac16}{\le} \frac{|x|^3\alpha}{6D_n^3}+\frac{x^4}{8\sqrt[3]{6}}\cdot\frac{\alpha}{D_n^3}\\
&= \left(\frac{|x|^3}{6}+\frac{x^4}{8\sqrt[3]{6}}\right)\frac{\alpha}{D_n^3}=\left(\frac{|x|^3}{6}+\frac{x^4}{8\sqrt[3]{6}}\right)\frac{8}{9}\frac{1}{T}\,,
\end{aligned}$$

weil $\frac{\alpha}{D_n^3}=\frac{8}{9}\frac{1}{T}$.

($(*)$ $\sigma_l^4\le\alpha_l^{\frac43}\le\alpha^{\frac13}\cdot\alpha_l$ nach der Ljapunow–Ungleichung, weil $\alpha=\sum_{l=1}^n\alpha_l$ ist.)

Zusammenfassend erhalten wir

$$\begin{aligned}
\left|\prod_{l=1}^n\beta_l-\prod_{l=1}^n\gamma_l\right| &\le \sum_{l=1}^n|\beta_1\cdot\ldots\cdot\beta_{j-1}\cdot(\beta_j-\gamma_j)\cdot\gamma_{j+1}\cdot\ldots\cdot\gamma_n|\\
&\le \sum_{l=1}^n\delta_1\cdot\ldots\cdot\delta_{j-1}\cdot\delta_{j+1}\cdot\ldots\cdot\delta_n\cdot|\beta_j-\gamma_j|\\
&\le e^{-\frac{x^2}{8}}\sum_{l=1}^n|\beta_l-\gamma_l|\le\frac{8}{9T}\left(\frac{|x|^3}{6}+\frac{5}{72}x^4\right)e^{-\frac{x^2}{8}}\,,
\end{aligned}$$

weil $\frac{1}{8\sqrt[3]{6}}=0,068\ldots<0,0694\ldots=\frac{5}{72}$, und der Satz ist bewiesen. □

Bemerkung 4.2.24 Die Konstante c im Satz 4.2.16 kann in der Tat nicht kleiner als $\frac{1}{\sqrt{2\pi}}\approx 0,4$ sein, wie folgendes Beispiel zeigt. Somit ist die untere Berry–Esséen–Schranke scharf ohne weitere Voraussetzungen an die Zufallsvariablen X_n.

Sei $\{X_n\}$ eine Folge von unabhängig identisch verteilten Zufallsvariablen, wobei

$$X_n=\begin{cases}1, & \text{mit Wkt. } \frac12\\ -1, & \text{mit Wkt. } \frac12\end{cases}.$$

Somit gilt für gerade n

$$2P\left(\sum_{j=1}^nX_j<0\right)+P\left(\sum_{j=1}^nX_j=0\right)=1\,,$$

weil aus Symmetriegründen $P\left(\sum_{j=1}^n X_j < 0\right) = P\left(\sum_{j=1}^n X_j > 0\right)$ gilt. Somit gilt nach der Stirlingschen Formel $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, n \to \infty$, dass

$$\begin{aligned}
\left| P\left(\sum_{j=1}^n X_j < 0\right) - \frac{1}{2} \right| &= \frac{1}{2} P\left(\sum_{j=1}^n X_j = 0\right) = \binom{n}{\frac{n}{2}} \frac{1}{2^n} \cdot \frac{1}{2} \\
&= \frac{n!}{\left(\frac{n}{2}\right)! \cdot \left(\frac{n}{2}\right)!} \frac{1}{2^{n+1}} \\
&\approx \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}}{\left(\left(\frac{n}{2e}\right)^{\frac{n}{2}} \sqrt{\pi n}\right)^2} \frac{1}{2^{n+1}} \\
&= \frac{n^n e^{-n} \sqrt{\pi n} \sqrt{2}}{\left(n^{\frac{n}{2}}\right)^2 (2e)^{-n} \pi n} \frac{1}{2^{n+1}} \\
&= \frac{\sqrt{2}}{2\sqrt{\pi n}} = \frac{1}{\sqrt{2\pi} \cdot \sqrt{n}}\,.
\end{aligned}$$

Da $\mathbb{E}X_n = 0$, Var $X_n = 1 = \mathbb{E}|X_n|^3$, gilt somit

$$\begin{aligned}
\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| &\geq \max\{|F_n(0) - \Phi(0)|\,, |\underbrace{F_n(0-0)}_{P\left(\frac{\sum_{j=1}^n X_j}{\sqrt{n}} < 0\right)} - \underbrace{\Phi(0)}_{=\frac{1}{2}}|\} \\
&\geq \left| P\left(\sum_{j=1}^n X_j < 0\right) - \frac{1}{2} \right| \underset{n\to\infty}{\sim} \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{n}}\,,
\end{aligned}$$

somit ist $c \geq \frac{1}{\sqrt{2\pi}} \approx 0,4$.

In den Sätzen 4.2.16 und 4.2.22 haben wir immer die Existenz des *dritten Momentes* von X_n vorausgesetzt. Der zentrale Grenzwertsatz in Form von Lindeberg geht lediglich von der Existenz der *zweiten Momente* aus. Deshalb geben wir im nächsten Satz die Konvergenzgeschwindigkeit im zentralen Grenzwertsatz für den Fall, wenn lediglich die Lindeberg–Bedingung erfüllt ist.

Satz 4.2.25 *(Berry)*
Sei $\{X_n\}_{n=1}^\infty$ eine Folge von unabhängigen Zufallsvariablen mit $\mathbb{E}X_j = \mu_j$ und $0 <$ Var $X_j = \sigma_i^2 < \infty$. Sei

$$D_n^2 = \sum_{j=1}^n \sigma_j^2 \text{ und } L_n(\delta) = \frac{1}{D_n^2} \sum_{j=1}^n \mathbb{E}\left((X_j - \mu_j)^2 \cdot I(|X_j - \mu_j| > \delta \cdot D_n)\right)$$

für alle $\delta > 0$. Dann existiert eine Konstante $D > 0$ mit der Eigenschaft: Aus $L_n(\delta) \leq \delta^3$ folgt

$$\varrho_n = \sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| < D\delta$$

für alle $\delta > 0$, wobei

$$F_n(x) = P\left(\frac{\sum_{j=1}^n X_j - \sum_{j=1}^n \mu_j}{D_n} \leq x\right), \quad x \in \mathbb{R}$$

ist.

Ohne Beweis. Siehe den Beweis in [8, 171].

Bemerkung 4.2.26 Da die Lindeberg–Bedingung (in einer ihrer Formen)

$$L_n(\delta) \underset{n\to\infty}{\longrightarrow} 0\,, \quad \delta > 0$$

ist, kann somit "quasi" behauptet werden, dass

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \leq const \cdot \sqrt[3]{L_n(\delta)}\,.$$

Kapitel 5

Theorie zufälliger Funktionen

In diesem Kapitel werden zufällige Funktionen in ihrer allgemeinen Form vorgestellt. Dieses Konzept ist vor allem in den weiterführenden Kapiteln wichtig, da es die Grundlage für das Verständnis von Stochastischen Prozessen liefert.

5.1 Zufällige Funktionen

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeits-Raum und $(\mathcal{S}, \mathcal{B})$ ein Meßraum, $\Omega, \mathcal{S} \neq \emptyset$.

Definition 5.1.1 Ein *zufälliges Element* $X : \Omega \to \mathcal{S}$ ist eine $\mathcal{F}|\mathcal{B}$-messbare Funktion (Schreibweise: $X \in \mathcal{F}|\mathcal{B}$), d.h.,

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}, \quad B \in \mathcal{B}.$$

Falls X ein zufälliges Element ist, dann ist $X(\omega)$ eine *Realisierung von* X für beliebiges $\omega \in \Omega$.

Die σ-Algebra $\mathcal{B}$ bestehend aus Teilmengen von $\mathcal{S}$ heißt von $\mathcal{M}$ (Elemente von $\mathcal{M}$ sind Teilmengen von $\mathcal{S}$) *induziert*, falls

$$\mathcal{B} = \bigcap_{\substack{\mathcal{F} \supset \mathcal{M} \\ \mathcal{F}\text{-}\sigma\text{-Algebra von } \mathcal{S}}} \mathcal{F}$$

(Schreibweise: $\mathcal{B} = \sigma(\mathcal{M})$).
Falls $\mathcal{S}$ ein topologischer oder metrischer Raum ist, so wird $\mathcal{M}$ meist als Familie aller offenen Mengen von $\mathcal{S}$ gewählt und $\sigma(\mathcal{M})$ wird *Borel σ-Algebra* genannt (Notation: $\mathcal{B} = \mathcal{B}(\mathcal{S})$).

Beispiel 5.1.2

1. Falls $\mathcal{S} = \mathbb{R}$, $\mathcal{B} = \mathcal{B}(\mathbb{R})$ dann nennt man das zufällige Element X *Zufallsvariable.*

2. Falls $\mathcal{S} = \mathbb{R}^m$, $\mathcal{B} = \mathcal{B}(\mathbb{R}^m)$, $m > 1$, dann nennt man X *Zufallsvektor*. Zufallsvariablen und Zufallsvektoren wurden bereits in „Elementare WR und Statistik“[20] betrachtet .

3. Sei $\mathcal{S}$ die Familie aller abgeschlossenen Mengen von $\mathbb{R}^m$. Sei

$$\mathcal{M} = \{\{A \in \mathcal{S} : A \cap B \neq \emptyset\}, B \text{ belieb. kompakte Menge in } \mathbb{R}^m\},$$

$\mathcal{B} = \sigma(\mathcal{M})$. Dann nennt man $X : \Omega \to \mathcal{S}$ eine *zufällige abgeschlossene Menge*.

Ein Beispiel für eine zufällige Menge kann wie folgt konstruiert werden: Seien $Y_1, \ldots, Y_n \in [0,1]^m$ n unabhängige gleichverteilte Punkte und $R_1, \ldots, R_n$ stochastisch unabhängige Zufallsvariablen auf $(\Omega, \mathcal{F}, P)$ mit $R_1, \ldots, R_n > 0$ fast sicher. Sei $X = \cup_{i=1}^n B_{R_i}(Y_i)$ mit $B_r(x) = \{y \in \mathbb{R}^m : \|y - x\| \leq r\}$. Offensichtlich ist dies eine abgeschlossene zufällige Menge.

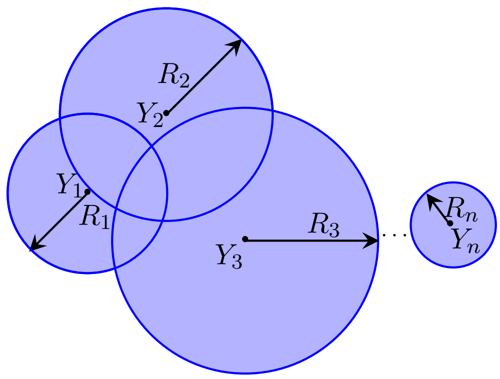


Abbildung 5.1: Beispiel einer zufälligen abgeschlossenen Menge $X = \cup_{j=1}^n B_{R_j}(Y_j)$

Übungsaufgabe 5.1.3 Seien $(\Omega, \mathcal{F})$ und $(\mathcal{S}, \mathcal{B})$ messbare Räume mit $\mathcal{B} = \sigma(\mathcal{M})$, wobei $\mathcal{M}$ eine Familie von Teilmengen von $\mathcal{S}$ ist. Zeige, dass $X : \Omega \to \mathcal{S}$ $\mathcal{F}|\mathcal{B}$-messbar ist genau dann, wenn $X^{-1}(C) \in \mathcal{F}$ mit $C \in \mathcal{M}$ gilt.

Definition 5.1.4 Sei T eine beliebige Indexmenge und $(\mathcal{S}_t, \mathcal{B}_t)_{t \in T}$ eine Familie von messbaren Räumen. Eine Familie $X = \{X(t), t \in T\}$ von zufälligen Elementen $X(t) : \Omega \to \mathcal{S}_t$ auf $(\Omega, \mathcal{F}, P)$ und $\mathcal{F}|\mathcal{B}_t$-messbar für alle $t \in T$ heißt *zufällige Funktion* (assoziiert mit $(\mathcal{S}_t, \mathcal{B}_t)_{t \in T}$).

Deshalb gilt, dass $X : \Omega \times T \to (\mathcal{S}_t,\ t \in T)$, d.h. $X(\omega, t) \in \mathcal{S}_t$ für alle $\omega \in \Omega$ mit $t \in T$ und $X(\cdot, t) \in \mathcal{F}|\mathcal{B}_t$, $t \in T$. Oft schreiben wir $X(t)$ anstatt $X(\omega, t)$. Manchmal hängt $(\mathcal{S}_t, \mathcal{B}_t)$ nicht von $t \in T$ ab: $(\mathcal{S}_t, \mathcal{B}_t) = (\mathcal{S}, \mathcal{B})$ für alle $t \in T$.

Beispiel 5.1.5

1. $T \subseteq \mathbb{Z}$: X heißt *zufällige Folge* oder *(zeit-)diskreter stochastischer Prozess* (Zeitreihe).
   Zum Beispiel $T = \mathbb{Z}, \mathbb{N}$.

2. $T \subseteq \mathbb{R}$: X heißt *(zeit-)stetiger stochastischer Prozess.*
   Zum Beispiel $T = \mathbb{R}_+$, $[a, b]$, $-\infty < a < b < \infty$, $\mathbb{R}$.

3. $T \subseteq \mathbb{R}^d, d \geq 2$: X heißt *Zufallsfeld.*
   Zum Beispiel $T = \mathbb{Z}^d$, $\mathbb{R}^d_+$, $\mathbb{R}^d$, $[a, b]^d$.

4. $T \subseteq \mathcal{B}(\mathbb{R}^d)$: X heißt *Mengen-indizierter Prozess.*
   Falls $X(\cdot) \geq 0$ f.s. und σ-additiv auf der σ-Algebra T ist, dann heißt X ein *zufälliges Maß.*

In den Fällen 1. und 2. kann $t \in T$ als Zeitparameter interpretiert werden. Für jedes $\omega \in \Omega$ nennt man $\{X(\omega, t),\ t \in T\}$ *Trajektorie* oder *Pfad* der zufälligen Funktion X. Wir möchten zeigen, dass $X = \{X(t),\ t \in T\}$ ein zufälliges Element im entsprechenden Funktionenraum mit folgender σ-Algebra ist:
Sei $\mathcal{S}_T = \prod_{t \in T} \mathcal{S}_t$ das kartesische Produkt von $\mathcal{S}_t$, $t \in T$, d.h., $X \in \mathcal{S}_T$ falls $X(t) \in \mathcal{S}_t$, $t \in T$. Die *elementare Zylindermenge* in $\mathcal{S}_T$ ist definiert als

$$C_T(B_t) = \{X \in \mathcal{S}_T : X(t) \in B_t\},$$

wobei $t \in T$ ein ausgewählter Punkt in T ist und $B_t \in \mathcal{B}_t$ eine Teilmenge von $\mathcal{S}_t$. Also enthält $C_T(B_t)$ alle Pfade von X, die durch das „Tor“ B_t gehen (vgl. Abbildung 5.2).

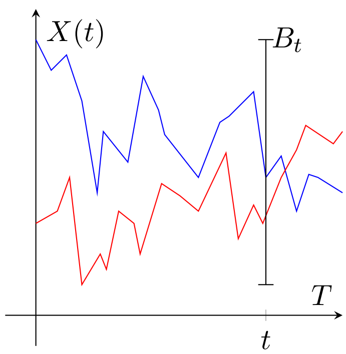


Abbildung 5.2: Zwei Pfade durch das „Tor“ B_t.

Definition 5.1.6 Die *zylindrische σ-Algebra $\mathcal{B}_T$* ist die σ-Algebra induziert in $\mathcal{S}_T$ durch die Familie aller elementaren Zylinder.
(Schreibweise: $\mathcal{B}_T = \otimes_{t \in T} \mathcal{B}_t$). Falls $\mathcal{B}_t = \mathcal{B}$ für alle $t \in T$, dann schreiben wir $\mathcal{B}^T$ anstatt $\mathcal{B}_T$.

Lemma 5.1.7 Die Familie $X = \{X(t),\ t \in T\}$ ist eine zufällige Funktion auf $(\Omega, \mathcal{F}, P)$ mit Zustandsraum $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$ genau dann, wenn für $\omega \in \Omega$ die Funktion $\omega \mapsto X(\omega, \cdot)$ $\mathcal{F}|\mathcal{B}_T$-messbar ist.

Übungsaufgabe 5.1.8 Beweise Lemma 5.1.7.

Definition 5.1.9 Sei X ein zufälliges Element $X : \Omega \to \mathcal{S}$, d.h. X ist $\mathcal{F}|\mathcal{B}$-messbar. Die *Verteilung von* X ist das Wahrscheinlichkeitsmaß P_X auf $(\mathcal{S}, \mathcal{B})$, sodass

$$P_X(B) = P(X^{-1}(B)), \quad B \in \mathcal{B}.$$

Lemma 5.1.10 Für jedes Wahrscheinlichkeitsmaß μ auf $(\mathcal{S}, \mathcal{B})$ existiert ein zufälliges Element X mit μ als Verteilung.

Beweis Wähle $\Omega = \mathcal{S}$, $\mathcal{F} = \mathcal{B}$, $P = \mu$ und $X(\omega) = \omega$, $\omega \in \Omega$. □

Unter welchen Bedingungen existiert eine zufällige Funktion mit vorgegebenen Eigenschaften? Eine zufällige Funktion bestehend aus unabhängigen zufälligen Elementen existiert immer. Diese Tatsache ist auch bekannt als

Satz 5.1.11 *(Lomnicki-Ulam)*
Sei $(\mathcal{S}_t, \mathcal{B}_t, \mu_t)_{t\in T}$ eine Folge von W-Räumen. Es existiert eine zufällige Funktion $X = \{X(t),\ t \in T\}$ auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ (mit Zustandsraum $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$), sodass

1. $X(t),\ t \in T$ sind stochastisch unabhängige zufällige Elemente,
2. $P_{X(t)} = \mu_t$ auf $(\mathcal{S}_t, \mathcal{B}_t)$, $t \in T$.

Viele Klassen von zufälligen Prozessen basieren auf unabhängigen zufälligen Elementen, wie wir in Abschnitt 5.2 sehen werden.

Definition 5.1.12 Auf $(\Omega, \mathcal{F}, P)$ sei $X = \{X(t), t \in T\}$ eine zufällige Funktion mit Zustandsraum $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$. Die *endlich-dimensionale Verteilung von* X ist definiert als die Verteilung $P_{t_1,\dots,t_n}$ von $(X(t_1), \dots, X(t_n))^T$ für beliebige $n \in \mathbb{N}$, $t_1, \dots, t_n \in T$ auf $(\mathcal{S}_{t_1,\dots,t_n}, \mathcal{B}_{t_1,\dots,t_n})$.
Dabei ist $\mathcal{S}_{t_1,\dots,t_n} = \mathcal{S}_{t_1} \times \dots \times \mathcal{S}_{t_n}$ und $\mathcal{B}_{t_1,\dots,t_n} = \mathcal{B}_{t_1} \otimes \dots \otimes \mathcal{B}_{t_n}$ die σ-Algebra auf $\mathcal{S}_{t_1,\dots,t_n}$, die induziert wird durch alle Mengen $B_{t_1} \times \dots \times B_{t_n}$, $B_{t_i} \in \mathcal{B}_{t_i}$, $i = 1, \dots, n$, d.h., $P_{t_1,\dots,t_n}(C) = P((X(t_1), \dots, X(t_n))^T \in C)$, $C \in \mathcal{B}_{t_1,\dots,t_n}$. Insbesondere erhält man für $C = B_1 \times \dots \times B_n$, $B_k \in \mathcal{B}_{t_k}$, dass

$$P_{t_1,\dots,t_n}(B_1 \times \dots \times B_n) = P(X(t_1) \in B_1, \dots, X(t_n) \in B_n).$$

Übungsaufgabe 5.1.13 Beweisen Sie, dass $X_{t_1,\dots,t_n} = (X(t_1), \dots, X(t_n))^T$ ein $\mathcal{F}|\mathcal{B}_{t_1,\dots,t_n}$– messbares zufälliges Element ist.

Definition 5.1.14 Sei $\mathcal{S}_t = \mathbb{R}$ für alle $t \in T$. Die zufällige Funktion $X = \{X(t), t \in T\}$ nennt man *symmetrisch*, falls alle ihren endlich-dimensionalen Verteilungen symmetrische Wahrscheinlichkeitsmaße sind, d.h.,

$P_{t_1,\dots,t_n}(A) = P_{t_1,\dots,t_n}(-A)$ für $A \in \mathcal{B}_{t_1,\dots,t_n}$ und alle $n \in \mathbb{N}$, $t_1, \dots, t_n \in T$, wobei

$$P_{t_1,\dots,t_n}(-A) = P((-X(t_1), \dots, -X(t_n))^T \in A).$$

Übungsaufgabe 5.1.15 Beweisen Sie, dass die endlich-dimensionalen Verteilungen einer zufälligen Funktion X die folgenden Eigenschaften haben: Für beliebiges $n \in \mathbb{N}$, $n \geq 2$, $\{t_1, \dots, t_n\} \subset T$, $B_k \in \mathcal{S}_{t_k}$, $k = 1, \dots, n$ und eine beliebige Permutation $(i_1, \dots, i_n)$ von $(1, \dots, n)$ gilt:

1. *Symmetrie:*

$$P_{t_1,\dots,t_n}(B_1 \times \dots \times B_n) = P_{t_{i_1},\dots,t_{i_n}}(B_{i_1} \times \dots \times B_{i_n})$$

2. *Konsistenz:*

$$P_{t_1,\dots,t_n}(B_1 \times \dots \times B_{n-1} \times \mathcal{S}_{t_n}) = P_{t_1,\dots,t_{n-1}}(B_1 \times \dots \times B_{n-1})$$

Das folgende Theorem besagt, dass die erwähnten Eigenschaften hinreichend sind für die Existenz einer zufälligen Funktion mit gegebenen endlich-dimensionalen Verteilungen.

Satz 5.1.16 *(Kolmogorow)*
Sei $\{P_{t_1,\dots,t_n},\ n \in \mathbb{N},\ \{t_1, \dots, t_n\} \subset T\}$ eine Familie von Wahrscheinlichkeitsmaßen auf

$$(\mathbb{R}^m \times \dots \times \mathbb{R}^m, \mathcal{B}(\mathbb{R}^m) \otimes \dots \otimes \mathcal{B}(\mathbb{R}^m)),$$

die die Voraussetzungen 1 und 2 von Übungsaufgabe 5.1.15 erfüllen. Dann existiert eine zufällige Funktion $X = \{X(t), t \in T\}$ auf einem W-Raum $(\Omega, \mathcal{F}, P)$ mit endlich-dimensionalen Verteilungen $P_{t_1,\dots,t_n}$.

Zunächst brauchen wir einige Hilfssätze, die wir nicht beweisen werden.

Lemma 5.1.17 Sei $\mathcal{F}(T)$ die Menge aller endlichen Teilmengen von T. Seien $\mathcal{S}_t$ mit $t \in T$ separable metrische Räume mit Metriken $d_t(\cdot, \cdot)$ und Borel–σ–Algebren $\mathcal{B}_t$.
Dann ist der Raum $\mathcal{S}_J = \{\text{Funktion } y \text{ auf } J : y(t) \in \mathcal{S}_t\}$, $J \in \mathcal{F}(T)$ mit der Metrik $d_J(x, y) = \max\limits_{t \in J} d_t(x(t), y(t))$, $x, y \in \mathcal{S}_J$ ein separabler metrischer Raum mit zylindrischer σ-Algebra $\mathcal{B}_J$, die gleich der Borel-σ-Algebra $\mathcal{B}(\mathcal{S}_J)$ ist.
Falls $\mathcal{S}_t$ vollständig ist, d.h. ein polnischer Raum ist, für alle $t \in T$, dann ist auch $\mathcal{S}_J$ ebenfalls vollständig, also auch ein polnischer Raum.

Lemma 5.1.18 Sei $\mathcal{S}$ ein metrischer Raum und μ ein Maß auf der Borel-σ-Algebra $\mathcal{B}(\mathcal{S})$. Dann gilt

$$\mu(B) = \sup_{\substack{A \subset B \\ A \text{ abgeschlossen}}} \mu(A) = \inf_{\substack{A \supset B \\ A \text{ offen}}} \mu(A), \quad B \in \mathcal{B}(\mathcal{S}).$$

Lemma 5.1.19 Seien $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$ und $(\mathcal{H}_t, \mathcal{F}_t)_{t\in T}$ zwei Familien isomorpher messbarer Räume:

$$(\mathcal{S}_t, \mathcal{B}_t) \sim (\mathcal{H}_t, \mathcal{F}_t), \quad t \in T,$$
$$\text{d.h. } \exists h : \mathcal{S}_t \to \mathcal{H}_t \text{ bijektiv} : h \in \mathcal{B}_t|\mathcal{F}_t \text{ und } h^{-1} \in \mathcal{F}_t|\mathcal{B}_t.$$

Dann gilt für alle $J \subset T$, dass $(\mathcal{S}_J, \mathcal{B}_J) \sim (\mathcal{H}_J, \mathcal{F}_J)$, wobei $\mathcal{B}_J$ und $\mathcal{F}_J$ zylindrische σ-Algebren auf $\mathcal{S}_J = \prod_{t\in J} \mathcal{S}_t$ und $\mathcal{H}_J = \prod_{t\in J} \mathcal{H}_t$ sind.

Beweis des Satzes 5.1.16
Wir geben hier eine Beweisskizze an.

1. Sei $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$ eine Familie von Messräumen. Zur Vereinfachung, betrachte $\mathcal{S}_t = [0,1]$, $\mathcal{B}_t = \mathcal{B}([0,1])$ für alle $t \in T$. Per Definition 5.1.12

$$\mathcal{S}_{t_1,\ldots,t_n} = \mathcal{S}_{t_1} \times \ldots \times \mathcal{S}_{t_n},$$
$$\mathcal{B}_{t_1,\ldots,t_n} = \mathcal{B}_{t_1} \otimes \ldots \otimes \mathcal{B}_{t_n},$$

für alle $n \in \mathbb{N}$, $t_1, \ldots, t_n \in T$, $t_i \neq t_j$ für $i \neq j$. Sei $\{P_{t_1,\ldots,t_n}\}$ eine Familie von Maßen auf $(\mathcal{S}_{t_1,\ldots,t_n}, \mathcal{B}_{t_1,\ldots,t_n})$, die alle symmetrisch und konsistent sind wie in Übungsaufgabe 5.1.15. Sei $\pi_{T,J} : \mathcal{S} \to \mathcal{S}_J$ die Koordinatenprojektion einer Funktion $\{y(t), t \in T\} \in \mathcal{S}$ auf die Koordinaten mit Indizes in J, d.h.

$$\pi_{T,J}y = (y(t_1), \ldots, y(t_n)) \quad \text{oder}$$
$$\pi_{T,J}\{y(t), t \in T\} = \{y(t), t \in J\}, \text{ falls } J = \{t_1, \ldots, t_n\}.$$

Mit „Projektion von μ“ meinen wir, dass

$$\mu \circ \pi_{T,J}^{-1}(B_{t_1} \times \ldots \times B_{t_n}) = P_J(B_{t_1} \times \ldots \times B_{t_n})$$

für jedes $n \in \mathbb{N}$, $J = \{t_1, \ldots, t_n\} \in \mathcal{F}(T)$, $B_{t_i} \in \mathcal{B}_{t_i}$, $i = 1, \ldots, n$, wobei $\mathcal{F}(T)$ die Menge aller endlichen Teilmengen von T ist. Falls ein solches Wahrscheinlichkeitsmaß μ auf $(\mathcal{S}, B_T)$ existiert, dann ist die zufällige Funktion X mit Verteilung μ gegeben durch die Koordinatenprojektion auf dem kanonischem Raum:

2. Wähle $\Omega = \mathcal{S}$, $\mathcal{F} = \mathcal{B}_T$ und $P = \mu$ und setze $X(t,\omega) = \omega(t)$ für $\omega \in \mathcal{S}$ und $t \in T$. Wir transformieren die Maße P_J von $(\mathcal{S}_J, \mathcal{B}_J)$ in den Raum $(\mathcal{S}, \mathcal{B}_{T,J})$ mit σ-Algebra $\mathcal{B}_{T,J} = \pi_{T,J}^{-1}\mathcal{B}_J \subset \mathcal{B}_T$ wie folgt: Definiere ein neues Maß

$$\mu_J(\tilde{B}) = P_J(B)$$

auf $(\mathcal{S}, \mathcal{B}_{T,J})$, wobei $B = B_{t_1} \times \ldots \times B_{t_n} \in \mathcal{B}_J$, $J = \{t_1, \ldots, t_n\}$ und $\widetilde{B} = B \times \prod_{t\in T\setminus J} \mathcal{S}_t \in \mathcal{B}_{T,J}$. Es ist offensichtlich, dass

$$B \in \mathcal{B}_J \Leftrightarrow \widetilde{B} \in \mathcal{B}_{T,J}.$$

Unser Ziel ist es, ein Maß μ auf $(\mathcal{S}, \mathcal{B}_T)$ zu finden, sodass $\mu = \mu_J$ auf jedem $\mathcal{B}_{T,J}$, $J \in \mathcal{F}(T)$. Entsprechend definieren wir μ auf der zylindrischen Algebra $\mathcal{C}_T = \bigcup\limits_{J \in \mathcal{F}(T)} \mathcal{B}_{T,J}$ indem wir $\mu(B) = \mu_J(B)$ setzen, falls ein $J \in \mathcal{F}(T)$ existiert mit $B \in \mathcal{B}_{T,J}$. Dies führt zu einer endlich-additiver Funktion auf $\mathcal{C}_T$. Nach dem Satz von Caratheodory (vgl. [20, Satz 2.2.9]) kann μ eindeutig fortgesetzt werden auf $\mathcal{B}_T = \sigma(\mathcal{C}_T)$, falls μ σ-additiv auf $\mathcal{C}_T$ ist.

3. Wir weisen also die abzählbare Additivität nach, was äquivalent dazu ist, die endliche Additivität und Stetigkeit auf $\emptyset$ zu zeigen, d.h.

$$\forall \{C_n\}_{n \in \mathbb{N}} \subset \mathcal{C}_T, \qquad \underbrace{C_n \searrow \emptyset}_{C_{n+1} \subset C_n, \bigcap_{n=1}^\infty C_n = \emptyset} \qquad \Rightarrow \mu(C_n) \underset{n \to \infty}{\longrightarrow} 0.$$

Wir zeigen die Negation: Falls für ein $\epsilon_0 > 0$ $\mu(C_n) \geq \epsilon_0$ für n groß genug gilt, dann gilt

$$C_n \underset{n \to \infty}{\not\to} \emptyset. \tag{5.1}$$

Da $C_n \in \mathcal{C}_T$, so $\exists\, J_n \in \mathcal{F}(T)$ mit $C_n \in \mathcal{B}_{T,J_n}$.
O.B.d.A. $J_n \subset J_{n+1}$. Wir beschränken uns auf Mengen $C_n = \pi^{-1}_{T,J_n} B_n$ mit kompakten Mengen B_n in $\mathcal{S}_{J_n}$, $J_n \in \mathcal{F}(T)$. Zeigen wir, wieso:

a) Nach Lemma 5.1.17 gilt $\mathcal{B}_{J_n} = \mathcal{B}(\mathcal{S}_{J_n})$. Nach Lemma 5.1.18 gilt: $\forall B_n \in \mathcal{B}_{J_n}$ $\exists$ abgeschlossenes $K_n \subset B_n$ in $(\mathcal{S}_{J_n}, d_{J_n})$ mit

$$\mu_{J_n}(B_n \setminus K_n) < 2^{-(n+1)} \epsilon_0, \quad n \in \mathbb{N}.$$

Da $\mathcal{S}_{J_n} = [0,1]^{J_n}$ kompakt ist, ist K_n es auch.
Für $C_n = \pi^{-1}_{T,J_n} B_n$ und $\widetilde{K_n} = \pi^{-1}_{T,J_n} K_n$ gilt dann

$$\mu\left(C_n \backslash \widetilde{K_n}\right) = \mu_{J_n}\left(B_n \backslash K_n\right) < z_0 \frac{1}{2^{n+1}}.$$

Sei $L_n = \bigcap_{i=1}^n \widetilde{K_i}$. Die Basis $D_n = \cap_{i=1}^n \pi^{-1}_{J_n,J_i} K_i$ der Zylinder $L_n = \pi^{-1}_{T,J_n} D_n$ ist kompakt,
und es gilt $L_n \subset K_n \subset C_n$, $L_{n+1} \subset L_n$ $\forall\, n$.

$$\mu\left(C_n \setminus L_n\right) = \mu\left(C_n \cap \bigcup_{i=1}^n \widetilde{K_i}^c\right) = \mu\left(\bigcup_{i=1}^n (C_n \setminus \widetilde{K_i})\right)$$
$$\leq \sum_{i=1}^n \mu\left(C_i \setminus \widetilde{K_i}\right) < \sum_{i=1}^n 2^{-(i+1)} \epsilon_0 < \epsilon_0/2, \text{ da } C_i \supset C_n,$$

und daher

$$\epsilon_0 \leq \mu(C_n) = \mu(L_n) + \mu(C_n \setminus L_n) \leq \mu(L_n) + \epsilon_0/2$$
$$\Rightarrow \mu(L_n) \geq \epsilon_0/2, \quad n \in \mathbb{N}.$$

Sonst reicht es, $C_n = \pi^{-1}_{T,J_n} B_n$ für kompakte Mengen B_n in $\mathcal{S}_{J_n}$ zu betrachten.

b) Definiere $J = \cup_{n=1}^{\infty} J_n$ und die Metrik

$$d_J(x,y) = \sum_{n=1}^{\infty} 2^{-n} \frac{d_{J_n}(x|_{J_n}, y|_{J_n})}{1 + d_{J_n}(x|_{J_n}, y|_{J_n})}$$

auf $\mathcal{S}_J$. Ähnlich wie in Lemma 5.1.17 kann man zeigen, dass (S_J, d_J) ein polnischer Raum mit zylindrischer σ-Algebra $\mathcal{B}_J = \mathcal{B}(\mathcal{S}_J)$ (Borel σ-Algebra) ist.
Die Mengen $\widehat{C_n} = \pi^{-1}_{J,J_n} B_n$, $n \in \mathbb{N}$ sind kompakt in $\mathcal{S}_J$, da sie abgeschlossene Teilmengen einer kompakten Menge sind (da π_{J,J_n} stetig). Es gilt

$$C_n = \pi^{-1}_{T,J_n} B_n = \pi^{-1}_{T,J} \pi^{-1}_{J,J_n} B_n = \pi^{-1}_{T,J} \widehat{C_n}.$$

Da $C_n \subset C_m$, gilt auch $\widehat{C_n} \subset \widehat{C_m}$ für $m \leq n$ und

$$\bigcap_{n=1}^{\infty} C_n = \pi^{-1}_{T,J} \left(\bigcap_{n=1}^{\infty} \widehat{C_n} \right).$$

Für eine absteigende Folge von kompakten Mengen gilt

$$\bigcap_{n=1}^{\infty} \widehat{C_n} \neq \emptyset \quad \text{und daher} \quad \bigcap_{n=1}^{\infty} C_n \neq \emptyset.$$

Dies ist ein Widerspruch zu der Annahme $\cap_{n=1}^{\infty} C_n = \emptyset$.

□

Bemerkung 5.1.20 Mit Hilfe von Lemma 5.1.19 kann Satz 5.1.16 auf allgemeinere als $[0,1]$ oder $\mathbb{R}^m$, sogenannte *Borel-Räume*, verallgemeinert werden. Diese sind auf gewisse Art isomorph zu $([0,1], \mathcal{B}([0,1]))$ oder zu einem von dessen Unterräumen.

Definition 5.1.21 Sei $X = \{X_t, t \in T\}$ eine zufällige Funktion mit Werten in $(\mathcal{S}, \mathcal{B})$, d.h. $X_t \in \mathcal{S}$ fast sicher für $t \in T$. Sei (T, C) auch ein messbarer Raum. Dann heißt X *messbar*, falls die Funktion $X : (\omega, t) \mapsto X(\omega, t) \in \mathcal{S}$, $(\omega, t) \in \Omega \times T$, $\mathcal{F} \otimes C|\mathcal{B}$-messbar ist.

Diese Definition enthält nicht nur die Messbarkeit von X bzgl. $\omega \in \Omega$: $X(\cdot, t) \in \mathcal{F}|\mathcal{B}$ für alle $t \in T$, sondern beiden Argumenten
$X(\cdot, \cdot) \in \mathcal{F} \otimes C|\mathcal{B}$. Die Messbarkeit von X ist von entscheidender Rolle, wenn $X(\omega, t)$ an zufälligen Zeitpunkten $\tau : \Omega \to T$ betrachtet wird, d.h. $X(\omega, \tau(\omega))$. In der Theorie der Martingale wird τ *Stoppzeit* für X genannt. Die Verteilungen von $X(\omega, t)$ und $X(\omega, \tau(\omega))$ können stark voneinander abweichen.

5.2 Elementare Beispiele

Der Satz von Kolmogorow (vgl. Theorem 5.1.16) kann nur in wenigen Fällen direkt verwendet werden, um einen stochastischen Prozess zu konstruieren, da für viele zufällige Funktionen endlich dimensionalen Verteilungen nicht explizit bekannt sind. In diesen Fällen definieren wir eine neue zufällige Funktion $X = \{X(t), t \in T\}$ als

$$X(t) = g(t, Y_1, Y_2, \ldots), \quad t \in T,$$

wobei g eine messbare Funktion ist und $\{Y_n\}$ eine Folge von zufälligen Elementen (also auch zufällige Funktionen), dessen Existenz bereits bewiesen wurde. Im Folgenden möchten wir Beispiele solcher Fälle besprechen. Sei $X = \{X(t), t \in T\}$ eine reellwertige zufällige Funktion auf dem W-Raum $(\Omega, \mathcal{F}, P)$.

1. *Weißes Rauschen*:

   Definition 5.2.1 Wir nennen $X = \{X(t), t \in T\}$ ein *weißes Rauschen*, falls alle $X(t)$, $t \in T$, u.i.v. Zufallsvariablen sind.

   Satz 5.1.11 besagt, dass ein solches X existiert. Es wird oft verwendet um das Rauschen in (elektromagnetischen oder akustischen) Signalen zu modellieren.

   - Falls $X(t) \sim \text{Ber}(p)$, $p \in (0, 1)$, $t \in T$, so nennt man dieses Rauschen auch *Salz-und-Pfeffer-Rauschen*. Dies entsteht beim Übersenden von binären Daten in Computernetzwerken.
   - Falls $X(t) \sim \mathcal{N}(0, \sigma^2)$, $\sigma^2 > 0$, $t \in T$, so nennt man X *weißes Gaußsches Rauschen*. Man kann solches Rauschen in akustischen Signalen beobachten.

2. *Gaußsche zufällige Funktion*:

   Definition 5.2.2 Eine *Gaußsche* zufällige Funktion $X = \{X(t), t \in T\}$ ist eine zufällige Funktion, deren endlich-dimensionalen Verteilungen alle multivariat Normal sind, d.h. für alle $n \in \mathbb{N}$, $t_1, \ldots, t_n \subset T$ gilt

   $$X_{t_1,\ldots,t_n} = ((X(t_1), \ldots, X(t_n))^\top \sim \mathcal{N}(\mu_{t_1,\ldots,t_n}, \Sigma_{t_1,\ldots,t_n}),$$

wobei

$$\mu_{t_1,\dots,t_n} = (\mathbb{E}X(t_1),\dots,\mathbb{E}X(t_n))^\top \quad \text{und}$$
$$\Sigma_{t_1,\dots,t_n} = ((\mathrm{cov}(X(t_i),X(t_j))_{i,j=1}^n.$$

Übungsaufgabe 5.2.3 Beweisen Sie, dass die Verteilung einer Gaußschen zufälligen Funktion X eindeutig bestimmt ist durch ihre Erwartungswertfunktion $\mu(t) = \mathbb{E}X(t), t \in T$, und Kovarianzfunktion $C(s,t)=\mathbb{E}[X(s)X(t)]$, $s,t \in T$.

Beispiele für Gaußsche Prozesse:

- *Wiener-Prozess* (oder die *Brownsche Bewegung*):
  $X = \{X(t), t \geq 0\}$ mit Erwartungswertfunktion $\mu(t) \equiv 0, t \geq 0$ und Kovarianzfunktion $C(s,t) = \min\{s,t\}$, $s,t \geq 0$. Meistens wird zusätzlich gefordert, dass die Pfade von X stetige Funktionen sind.
- *Gaußsches weißes Rauschen*:
  Gaußsches weißes Rauschen ist eine Gaußsche zufällige Funktion.

3. *Lognormal- und χ^2-Funktionen:*
   Eine zufällige Funktion $X = \{X(t),\ t \in T\}$ nennt man *lognormal*, falls $X(t) = e^{Y(t)}$ und $Y = \{Y(t), t \in T\}$ eine Gaußsche zufällige Funktion ist.
   Falls $X(t) = \|Y(t)\|^2$ und $Y = \{Y(t), t \in T\}$ eine Gaußsche zufällige Funktion mit Werten in $\mathbb{R}^n$ ist, für die gilt $Y(t) \sim \mathcal{N}(0, \mathrm{I})$, $t \in T$, so nennen wir X eine *χ^2-Funktion*. Wir meinen mit I die $(n \times n)$-Einheitsmatrix. Dann gilt

$$X(t) \sim \chi_n^2, \quad t \in T.$$

4. *Kosinus-Welle:*
   Wir definieren $X = \{X(t), t \in \mathbb{R}\}$ durch $X(t) = \sqrt{2}\cos(2\pi Y + tZ)$, mit $Y \sim \mathcal{U}([0,1])$ und einer von Y unabhängigen ZV Z. Dann nennen wir X eine *Kosinus-Welle*.

   Übungsaufgabe 5.2.4 Seien $X_1, X_2, \dots$ u.i.v. Kosinus-Wellen. Bestimmen Sie den schwachen Grenzwert der endlich-dimensionalen Verteilungen der zufälligen Funktion $\left\{\frac{1}{\sqrt{n}}\sum_{k=1}^n X_k(t),\ t \in \mathbb{R}\right\}$ für $n \to \infty$.

5. *Poisson-Prozess:*
   Sei $\{Y_n\}_{n\in\mathbb{N}}$ eine Folge von u.i.v. ZVn $Y_n \sim \mathrm{Exp}(\lambda), \lambda > 0$. Der stochastische Prozess X definiert durch $X(t) = \max\{n \in \mathbb{N} : \sum_{k=1}^n Y_k \leq t\}$ heißt *Poisson-Prozess* mit Intensität $\lambda > 0$. $X(t)$ zählt die Anzahl gewisser Ereignisse bis zum Zeitpunkt $t > 0$, wobei das typische Intervall zwischen zwei benachbarten Ereignissen $\mathrm{Exp}(\lambda)$-verteilt ist.

Beispiele: Schadensmeldung eines Versicherers, das Entdecken eines Elementarteilchens im Geigerzähler, etc.

$$X(t) = \#\text{ Schäden bzw. Teilchen im Intervall } [0,t]\,.$$

5.3 Regularitätseigenschaften

Der Satz von Kolmogorow (Satz 5.1.16) besagt die Existenz der Verteilung einer zufälligen Funktion mit vorgegebenen endlich-dimensionalen Verteilungen. Jedoch macht der Satz keine Aussage über die Eigenschaften der Pfade aus. Dies ist auch verständlich, da alle zufälligen Objekte in der Wahrscheinlichkeitstheorie nur im f.s. Sinn definiert sind, also bis auf eine Menge

$$A \subset \Omega \text{ mit } P(A) = 0.$$

Beispiel 5.3.1 Sei $(\Omega, \mathcal{F}, P) = ([0,1]\,, \mathcal{B}([0,1]), \nu_1)$ mit Lebesgue Maß ν_1 auf $[0,1]$. Definiere $X = \{X(t),\ t \in [0,1]\}$ durch $X(t) \equiv 0,\ t \in [0,1]$ und $Y = \{Y(t),\ t \in [0,1]\}$ durch

$$Y(t) = \begin{cases} 1, & t = U, \\ 0, & \text{sonst,} \end{cases}$$

wobei $U(\omega) = \omega$, $\omega \in [0,1]$, eine $\mathcal{U}([0,1])$-verteilte ZV auf $(\Omega, \mathcal{F}, P)$.
Es gilt: $P(U = t) = 0,\ t \in T \Rightarrow P(Y(t) = 0) = 1,\ t \in T \Rightarrow X \stackrel{d}{=} Y$. X hat allerdings stetige und Y unstetige Pfade und

$$P(X(t) = 0,\ \forall t \in T) = 1 \neq 0 = P(Y(t) = 0,\ \forall t \in T).$$

Die Ausnahmemenge A kann für $X(t)$ für jedes $t \in T$ ganz unterschiedlich aussehen.
Deswegen fordern wir, dass alle $X(t)$, $t \in T$ auf der selben Teilmenge $\Omega_0 \subseteq \Omega$ mit $P(\Omega_0) = 1$ definiert sind. Die neue zufällige Funktion $\widetilde{X} : \Omega_0 \times T \to \mathbb{R}$ wird später *Modifikation* von $X : \Omega \times T \to \mathbb{R}$ genannt. X und $\widetilde{X}$ unterscheiden sich auf der Nullmenge Ω/Ω_0. Wenn wir sagen „Pfade der zufälligen Funktion X haben Eigenschaft C“, implizieren wir die Existenz einer Modifikation von X mit Eigenschaft C.
Das soeben Gesagte motiviert folgende Definitionen:

Definition 5.3.2 Die zufälligen Funktionen $X = \{X(t),\ t \in T\}$ und $Y = \{Y(t),\ t \in T\}$ auf $(\Omega, \mathcal{F}, P)$ nennt man *(stochastisch) äquivalent*, falls

$$B_t = \{\omega \in \Omega : X(\omega,t) \neq Y(\omega,t)\} \in \mathcal{F},\ t \in T,$$

und $P(B_t) = 0,\ t \in T$. Solche Zufallsfunktionen X und Y nennt man *Versionen oder Modifikationen* der selben zufälligen Funktion.

Definition 5.3.3 Die zufälligen Funktionen $X = \{X(t),\ t \in T\}$ und $Y = \{Y(t),\ t \in T\}$ auf dem selben W-Raum $(\Omega, \mathcal{F}, P)$ assoziiert mit $(\mathcal{S}_t, \mathcal{B}_t)_{t \in T}$ haben *äquivalente Pfade* (oder heißen *stochastisch nicht unterscheidbar*), falls

$$A = \{\omega \in \Omega : X(\omega, t) \neq Y(\omega, t) \ \text{ für ein } \ t \in T\} \in \mathcal{F}$$

und $P(A) = 0$.

Falls $(\Omega, \mathcal{F}, P)$ *vollständig* ist (d.h. aus $A \in \mathcal{F} : P(A) = 0$ folgt für alle $B \subset A$: $B \in \mathcal{F}$ und $P(B) = 0$), dann sind nicht unterscheidbare Prozesse stochastisch äquivalent. Die Rückrichtung gilt i.A. nicht (sie gilt für s.g. *separable Prozesse*. Dies ist der Fall, wenn T abzählbar ist).

Übungsaufgabe 5.3.4 Beweisen Sie, dass die zufälligen Funktionen X und Y in Beispiel 5.3.1 stochastisch äquivalent sind.

Definition 5.3.5 Wir nennen zwei zufällige Funktionen $X = \{X(t),\ t \in T\}$ und $Y = \{Y(t),\ t \in T\}$ (nicht unbedingt auf dem selben W-Raum) *äquivalent in Verteilung*, falls $P_X = P_Y$ auf $(\mathcal{S}_t, \mathcal{B}_t)_{t \in T}$.
Schreibweise: $X \stackrel{d}{=} Y$.

Nach dem Satz von Kolmorogow (Satz 5.1.16) ist es ausreichend für die Äquivalenz in Verteilung von X und Y, wenn sie dieselben endlich-dimensionalen Verteilungen haben. Offensichtlich impliziert die stochastische Äquivalenz die Äquivalenz in Verteilung. Umgekehrt jedoch nicht.

Definition 5.3.6 Eine zufällige Funktion $X = \{X(t),\ t \in T\}$ nennt man

a) *stochastisch stetig auf* T, falls $X(s) \xrightarrow[s \to t]{P} X(t)$, für beliebige $t \in T$, d.h.

$$P(|X(s) - X(t)|_{\mathcal{S}} > \varepsilon) \xrightarrow{P} 0, \text{für alle } \varepsilon > 0.$$

b) L^p*-stetig auf* T, $p \geq 1$, falls $X(s) \xrightarrow[s \to t]{L^p} X(t)$, $t \in T$, d.h.

$$\mathbb{E}|X(s) - X(t)|^p_{\mathcal{S}} \xrightarrow[s \to t]{} 0.$$

Für $p = 2$ sprechen wir von „Stetigkeit im quadratischen Mittel".

c) *f.s. stetig auf* T, falls $X(s) \xrightarrow[s \to t]{f.s.} X(t)$, $t \in T$, d.h.,

$$P(X(s) \xrightarrow[s \to t]{} X(t)) = 1, \quad t \in T.$$

d) *stetig*, falls alle Pfade von X stetige Funktionen sind.

Meistens sind die Fälle c) und d) am interessantesten, obwohl die stochastische Stetigkeit die schwächste ist. Die folgende Grafik veranschaulicht die Relationen zwischen den verschiedenen Formen der Stetigkeit:

$$\boxed{L^p\text{-Stetigk.}} \Longrightarrow \boxed{\text{stoch. Stetigk.}} \Longleftarrow \boxed{\text{f.s. Stetigk.}} \Longleftarrow \boxed{\text{Stetigk. aller Pfade}}$$

Anhand des folgenden Beispiels sehen wir warum die Fälle c) und d) so wichtig sind.

Beispiel 5.3.7 Sei $T = [0,1]$, $(\Omega, \mathcal{F}, P)$ ein kanonischer W-Raum mit $\Omega = \mathbb{R}^{[0,1]} = \prod_{t\in[0,1]} \mathbb{R}$ und $X = \{X(t),\ t \in [0,1]\}$ ein stochastischer Prozess auf $(\Omega, \mathcal{F}, P)$.
Nicht alle Ereignisse sind Elemente von $\mathcal{F}$, z.B. ist

$$A = \{\omega \in \Omega : X(\omega, t) = 0 \,\forall\, t \in [0,1]\} = \cap_{t\in[0,1]} \{X(\omega, t) = 0\}$$

nicht in $\mathcal{F}$, da es ein Durchschnitt von überabzählbar vielen messbaren Ereignissen aus $\mathcal{F}$ ist.
Falls jedoch X stetig ist, dann sind dessen Pfade stetige Funktionen und wir schreiben $A = \cap_{t\in D} \{X(\omega, t) = 0\}$, wobei D eine dichte abzählbare Teilmenge von $[0,1]$ ist, z.B. $D = \mathbb{Q} \cap [0,1]$. Dann gilt $A \in \mathcal{F}$.

Allerdings ist es in vielen Anwendungen (z.B. Finanzmathematik) nicht realistisch, Phänomene durch stochastische Prozesse mit stetigen Pfaden zu modellieren. Deshalb betrachten wir eine größere Klasse von Pfaden von X: die sogenannte *càdlàg-Klasse* (càdlàg = continue à droite, limitée à gauche (fr.)).

Definition 5.3.8 Ein stochastischer Prozess $X = \{X(t),\ t \in \mathbb{R}\}$ heißt *càdlàg*, falls alle seine Pfade rechtstetige Funktionen sind und linksseitige Grenzwerte besitzen.

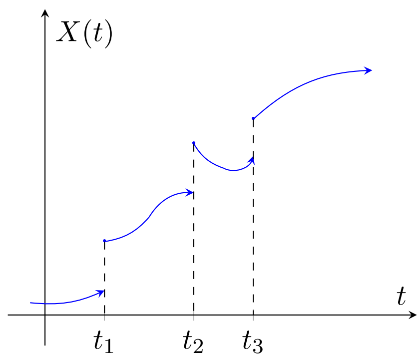


Abbildung 5.3: Beispiel eines càdlàg–Prozess

Ein beispielhafter Pfad für einen Prozess wie in Definition 5.3.8 kann in Abbildung 5.3 gefunden werden. Wir untersuchen die bisher eingeführten Eigenschaften nun etwas genauer. Zum Beispiel ist die stochastische Stetigkeit eine Eigenschaft der zwei-dimensionalen Verteilung $P_{s,t}$ von X. Dies halten wir im folgendem Lemma fest.

Lemma 5.3.9 Sei $X = \{X(t),\ t \in T\}$ eine zufällige Funktion mit Werten in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, wobei T ein Banach-Raum ist. Folgende Aussagen sind dann äquivalent:

a) $X(s) \xrightarrow[s\to t_0]{P} Y$,

b) $P_{s,t} \xrightarrow[s,t\to t_0]{d} P_{(Y,Y)}$,

wobei $t_0 \in T$ und Y eine Zufallsvariable ist. Für die stochastische Stetig. von X, wählen wir $t_0 \in T$ beliebig und $Y = X(t_0)$.

Beweis *a) ⇒ b)*
$X(s) \xrightarrow[s\to t_0]{P} Y$ bedeutet $(X(s), X(t))^\top \xrightarrow[s,t\to t_0]{P} (Y,Y)^\top$:

$$P(\underbrace{|(X(s),X(t)) - (Y,Y)|_2}_{(|X(s)-Y|^2+|X(t)-Y|^2)^{1/2}} > \varepsilon) \leqslant P(|X(s)-Y| > \varepsilon/\sqrt{2}) \\ + P(|X(t)-Y| > \varepsilon/\sqrt{2}) \xrightarrow[s,t\to t_0]{} 0.$$

Dann folgt $P_{s,t} \xrightarrow{d} P_{(Y,Y)}$, da $\xrightarrow{P}$-Konvergenz stärker ist als $\xrightarrow{d}$-Konvergenz.
b) ⇒ a)
Für $\varepsilon > 0$ betrachten wir $g_\varepsilon : \mathbb{R} \to [0,1]$ mit $g_\varepsilon(0) = 0$, $g_\varepsilon(x) = 1$, $x \notin B_\varepsilon(0)$. Es gilt für alle $s, t \in T$, dass

$$\mathbb{E} g_\varepsilon(|X(s) - X(t)|) = P(|X(s) - X(t)| > \varepsilon) \\ + \mathbb{E}(g_\varepsilon(|X(s) - X(t)|) I(|X(s) - X(t)| \le \varepsilon)),$$

und daher

$$\begin{aligned} P(|X(s) - X(t)| > \varepsilon) &\le \mathbb{E} g_\varepsilon(|X(s) - X(t)|) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} g_\varepsilon(|x-y|) P_{s,t}(d(x,y)) \\ &\xrightarrow[\substack{s\to t_0 \\ t\to t_0}]{} \int_{\mathbb{R}} \int_{\mathbb{R}} g_\varepsilon(|x-y|) P_{(Y,Y)}(d(x,y)) = 0, \end{aligned}$$

da $P_{(Y,Y)}$ konzentriert ist auf $\{(x,y) \in \mathbb{R}^2 : x = y\}$ und $g_\varepsilon(0) = 0$. Also ist $\{X(s)\}_{s\to t_0}$ eine Fundamentalfolge (in Wahrscheinlichkeit) und daher gilt $X(s) \xrightarrow[s\to t_0]{P} Y$. □

Es kann sein, dass X stochastisch stetig ist, obwohl alle Pfade von X Sprünge haben, d.h. es kann keine f.s. stetige Modifikation von X existieren.
Erklärung: X kann einen Sprung für ein konkretes $t \in T$ mit Wahrscheinlichkeit 0 haben. Daher treten die Sprünge auf den Pfaden von X immer an verschiedenen Stellen auf.

Übungsaufgabe 5.3.10 Beweisen Sie, dass ein Poisson-Prozess stochastisch stetig ist, obwohl es keine f.s stetige Modifikation von ihm gibt.

Übungsaufgabe 5.3.11 Sei T kompakt. Zeigen Sie:
Falls X stochastisch stetig ist auf T, dann ist X auch gleichmäßig stochastisch stetig, d.h, für alle $\varepsilon, \eta > 0$ $\exists \delta > 0$, s.d. für alle $s, t \in T$ mit $|s-t|_T < \delta$ gilt

$$P(|X(s) - X(t)|_{\mathcal{S}} > \varepsilon) < \eta.$$

Nun sei $\mathcal{S} = \mathbb{R}$, $\mathbb{E}X^2(t) < \infty$, $t \in T$, $\mathbb{E}X(t) = 0$, $t \in T$ und
$C(s,t) = \mathbb{E}\left[X(s)X(t)\right]$ die Kovarianzfunktion von X.

Lemma 5.3.12 Sei $t_0 \in T$. Es gelten folgende Aussagen:

a) Falls $X(s) \xrightarrow[s\to t_0]{L^2} Y$ für eine ZV Y mit $\mathbb{E}Y^2 < \infty$, dann folgt $C(s,t) \xrightarrow[s,t\to t_0]{} \mathbb{E}Y^2$.

b) Falls $C(s,t) \xrightarrow[s,t\to t_0]{} a > 0$, dann existiert eine ZV Y mit $\mathbb{E}Y^2 = a$ und $X(s) \xrightarrow[s\to t_0]{L^2} Y$.

Beweis

a) Es gilt

$$\begin{aligned}
&|C(s,t) - \mathbb{E}Y^2| = |\mathbb{E}(X(s)X(t)) - \mathbb{E}Y^2| \\
&= |\mathbb{E}\left[(X(s) - Y + Y)(X(t) - Y + Y)\right] - \mathbb{E}Y^2| \\
&\leq \mathbb{E}|(X(s) - Y)(X(t) - Y)| + \mathbb{E}|(X(s) - Y)Y| + \mathbb{E}|(X(t) - Y)Y| \\
&\leq \sqrt{\underbrace{\mathbb{E}(X(s) - Y)^2\ \mathbb{E}(X(t) - Y)^2}_{||X(s)-Y||^2_{L^2}\cdot||X(t)-Y||^2_{L^2}}} + \sqrt{\mathbb{E}Y^2\underbrace{\mathbb{E}(X(s) - Y)^2}_{||X(s)-Y||^2_{L^2}}} \\
&+ \sqrt{\mathbb{E}Y^2\underbrace{\mathbb{E}(X(t) - Y)^2}_{||X(t)-Y||^2_{L^2}}} \xrightarrow[s,t\to t_0]{} 0
\end{aligned}$$

nach Voraussetzung a).

b) Es gilt

$$\begin{aligned}
\mathbb{E}(X(s) - X(t))^2 &= \mathbb{E}(X(s))^2 - 2\mathbb{E}[X(s)X(t)] + \mathbb{E}(X(t))^2 \\
&= C(s,s) + C(t,t) - 2C(s,t) \\
&\xrightarrow[s,t\to t_0]{} 2a - 2a = 0.
\end{aligned}$$

Also ist $\{X(s),\ s \to t_0\}$ eine Fundamentalfolge im L^2-Sinn, und wir erhalten $X(s) \xrightarrow[s\to t_0]{L^2} Y$ für eine ZV Y. Ferner kann man zeigen, dass $\mathbb{E}Y^2 = a$.

□

Folgerung 5.3.13 Die zentrierte zufällige Funktion X, die die Bedingungen von Lemma 5.3.12 erfüllt, ist stetig auf T im L^2-Sinn genau dann, wenn ihre Kovarianzfunktion $C : T^2 \to \mathbb{R}$ stetig auf der Diagonalen diag $T^2 = \{(s,t) \in T^2 : s = t\}$ ist, d.h., $\lim_{s,t\to t_0} C(s,t) = C(t_0,t_0) = \text{Var } X(t_0)$ für alle $t_0 \in T$.

Beweis Wähle $Y = X(t_0)$ in Lemma 5.3.12. □

Bemerkung 5.3.14 Falls X nicht zentriert ist, dann ist die Stetigkeit von $\mu(t) = \mathbb{E}X(t), t \in T$ und C auf diag T^2 notwendig für die L^2-Stetigkeit von X auf T.

Eine zufällige Funktion X, die stetig im L^2-Sinn ist, kann trotzdem unstetige Pfade haben. In den meisten relevanten Anwendungsfällen hat X eine f.s. stetige Modifikation. Wir werden dies später in einem entsprechendem Satz genauer formulieren.

Übungsaufgabe 5.3.15 Geben Sie ein Beispiel eines stochastischen Prozesses mit f.s. unstetigen Pfaden, der aber L^2-stetig ist.

Wir untersuchen nun die Eigenschaft der f.s. Stetigkeit genauer. Wie bereits erwähnt, können wir lediglich von einer stetigen Modifikation eines Prozesses sprechen. Die Existenz einer solchen Modifikation hängt auch von den Eigenschaften der zwei-dimensionalen Verteilung des Prozesses ab. Dies ist die Aussage des folgenden Satzes (ursprünglich bewiesen von A. Kolmogorow).

Satz 5.3.16 *(Chentsov, Loève, Kolmogorow)*
Sei $X = \left\{X(t),\ t \in \mathbb{R}^d\right\}$ ein reellwertiges Zufallsfeld. X hat eine stetige Modifikation, falls $\exists \alpha, c, \delta > 0$, sodass

$$\mathbb{E}|X(t+h) - X(t)|^\alpha < c|h|^{d+\delta},\ t \in \mathbb{R}^d, \tag{5.2}$$

für hinreichend kleine $|h|$. Diese Modifikation ist f.s. lokal Hölder-stetig mit Hölder-Exponenten $\gamma \in (0, \delta/\alpha)$.

Beweis Siehe [13, Theorem 3.23]. □

Als Nächstes möchten wir Prozesse mit càdlàg-Pfaden diskutieren. Sei $(\Omega, \mathcal{F}, P)$ ein vollständiger W-Raum.

Lemma 5.3.17 Seien $X = \{X(t),\ t \geq 0\}$ und $Y = \{Y(t),\ t \geq 0\}$ zwei Modifikationen einer zufälligen Funktion auf $(\Omega, \mathcal{F}, P)$. Falls X and Y f.s. rechtsseitig stetig sind, dann sind X and Y nicht unterscheidbar.

Beweis Seien Ω_X, Ω_Y zwei „Ausnahmemengen“, auf denen die Pfade von X bzw. Y nicht rechtsseitig stetig sind. Es gilt

$$P(\Omega_X) = P(\Omega_Y) = 0.$$

Betrachte $A_t = \{\omega \in \Omega : X(\omega, t) \neq Y(\omega, t)\}$, $t \in [0, +\infty)$ und $A = \cup_{t \in \mathbb{Q}_+} A_t$, wobei $\mathbb{Q}_+ = \mathbb{Q} \cap [0, +\infty)$. Da X und Y stochastisch äquivalent sind, gilt $P(A) = 0$ und deshalb auch

$$P(\widetilde{A}) \leq P(A) + P(\Omega_X) + P(\Omega_Y) = 0,$$

wobei $\widetilde{A} = A \cup \Omega_X \cup \Omega_Y$. Also gilt $X(\omega, t) = Y(\omega, t)$ für $t \in \mathbb{Q}_+$ und $\omega \in \Omega \backslash \tilde{A}$. Als nächstes zeigen wir dies für alle $t \geq 0$.
Für ein beliebiges $t \geq 0$ existiert eine Folge $\{t_n\} \subset \mathbb{Q}_+$, s.d. $t_n \downarrow t$. Da $X(\omega, t_n) = Y(\omega, t_n)$ für alle $n \in \mathbb{N}$ und alle $\omega \in \Omega \setminus \widetilde{A}$, folgt

$$X(\omega, t) = \lim_{n \to \infty} X(\omega, t_n) = \lim_{n \to \infty} Y(\omega, t_n) = Y(\omega, t)$$

für $t \geq 0$ und $\omega \in \Omega \setminus \widetilde{A}$.
Also sind X und Y nicht unterscheidbar. □

Folgerung 5.3.18 Falls zwei càdlàg-Prozesse $X = \{X(t),\ t \geq 0\}$ und $Y = \{Y(t),\ t \geq 0\}$ Modifikationen der selben zufälligen Funktion sind, dann sind sie nicht unterscheidbar.

5.4 Differenzierbarkeit von Pfaden

Definition 5.4.1 Eine reellwertige zufällige Funktion $X = \{X(t),\ t \in T\}$ ist *differenzierbar auf T in Richtung $h \in T$ stochastisch, im L^p-Sinn*, $p \geq 1$, oder *f.s.*, falls

$$\lim_{l \to 0} \frac{X(t + hl) - X(t)}{l} = X_h'(t),\ t \in T$$

im jeweiligen Sinn (stochastisch, L^p oder f.s.) existiert.

Lemma 5.3.9 besagt, dass die stochastische Differenzierbarkeit durch die drei-dimensionalen Verteilungen von X bestimmt ist (da die gemeinsame Verteilung von $\frac{X(t+hl)-X(t)}{l}$ und $\frac{X(t+hl')-X(t)}{l'}$ schwach konvergieren soll). Die Differenzierbarkeit in L^2-Sinn ist bestimmt durch die Glattheit der Kovarianzfunktion $C(s, t)$ (vgl. Korollar 5.3.13).

Übungsaufgabe 5.4.2 Zeige, dass

1. ein Wiener Prozess nicht stochastisch differenzierbar auf $[0, \infty)$ ist.
2. ein Poisson Prozess stochastisch differenzierbar auf $[0, \infty)$ ist, aber nicht im L^p-Sinn, $p > 1$.

Lemma 5.4.3 Eine zentrierte zufällige Funktion $X = \{X(t),\ t \in T\}$ (d.h., $\mathbb{E}X(t) \equiv 0$, $t \in T$) mit $\mathbb{E}[X^2(t)] < \infty, t \in T$ ist L^2-differenzierbar in $t \in T$

in Richtung $h \in T$, falls ihre Kovarianzfunktion C zwei mal differenzierbar in (t,t) in Richtung h ist, d.h., falls

$$\exists C''_{hh}(t,t) = \left.\frac{\partial^2 C(s,t)}{\partial s_h \partial t_h}\right|_{s=t}.$$

$X'_h(t)$ ist L^2-stetig in $t \in T$, falls $C''_{hh}(s,t) = \frac{\partial^2 C(s,t)}{\partial s_h \partial t_h}$ stetig ist in $s = t$. Außerdem ist $C''_{hh}(s,t)$ die Kovarianzfunktion von $X'_h = \{X'_h(t),\ t \in T\}$.

Beweis Nach Lemma 5.3.12 genügt es zu zeigen, dass

$$J = \lim_{l,l' \to 0} \mathbb{E}\left(\frac{X(t+lh) - X(t)}{l} \cdot \frac{X(s+l'h) - X(s)}{l'}\right) \qquad (*)$$

für $s = t$ existiert. Der Erwartungswert auf der rechten Seite von $(*)$ ist gleich

$$\begin{aligned}
&\frac{1}{ll'}\Big(C(t+lh, s+l'h) - C(t+lh, s) - C(t, s+l'h) + C(t,s)\Big)\\
&= \frac{1}{l}\left(\frac{C(t+lh, s+l'h) - C(t+lh, s)}{l'} - \frac{C(t, s+l'h) - C(t,s)}{l'}\right)\\
&\xrightarrow[l,l' \to 0]{} C''_{hh}(s,t)\,.
\end{aligned}$$

Alle anderen Aussagen folgen ebenfalls aus dieser Relation. □

Übungsaufgabe 5.4.4 Zeige, dass aus der L^p-Differenzierbarkeit, $p \geq 1$, die L^p-Stetigkeit des stochastischen Prozesses folgt.

Bemerkung 5.4.5 Die Eigenschaft der L^p-Differenzierbarkeit, $p \geq 1$ und f.s. Differenzierbarkeit von zufälligen Funktionen sind unabhängig im folgenden Sinn: $\exists$ stoch. Prozess mit L^2-differenzierbaren Pfaden, obwohl sie f.s. unstetig sind. Umgekehrt sind Prozesse mit f.s. differenzierbaren Pfaden nicht immer L^2-differenzierbar, da z.B. die erste Ableitung der Kovarianzfunktion unstetig sein kann.

5.5 Momente und Kovarianz

Sei $X = \{X(t),\ t \in T\}$ eine reellwertige zufällige Funktion und T eine beliebige Indexmenge.

Definition 5.5.1 Das *gemischte Moment* $\mu^{(j_1,\ldots,j_n)}(t_1,\ldots,t_n)$ von X der Ordnung $(j_1,\ldots,j_n) \in \mathbb{N}^n$, $t_1,\ldots,t_n \in T$ ist gegeben durch

$$\mu^{(j_1,\ldots,j_n)}(t_1,\ldots,t_n) = \mathbb{E}\left[X^{j_1}(t_1) \cdot \ldots \cdot X^{j_n}(t_n)\right],$$

wobei die Existenz und Endlichkeit des Erwartungswertes vorausgesetzt wird. Dann ist es ausreichend vorauszusetzen, dass $\mathbb{E}|X(t)|^j < \infty$ für alle $t \in T$ und $j = j_1 + \ldots + j_n$.

Beispiel 5.5.2

$\mu(t) = \mu^{(1)}(t) = \mathbb{E}X(t)$, $t \in T$ ist die *Erwartungswertfunktion von* X.

$\mu^{(1,1)}(s,t) = \mathbb{E}[X(s)X(t)] = C(s,t)$ ist die *(nicht-zentrierte) Kovarianzfunktion von* X. Wohingegen die *zentrierte Kovarianzfunktion* gegeben ist als

$$K(s,t) = \text{Cov}\ ((X(s), X(t)) = \mu^{(1,1)}(s,t) - \mu(s)\mu(t),\ s,t \in T.$$

Übungsaufgabe 5.5.3 Zeige, dass die zentrierte Kovarianzfunktion einer reellwertigen zufälligen Funktion X

1. *symmetrisch* ist, d.h., $K(s,t) = K(t,s)$, $s,t \in T$.
2. *positiv semidefinit* ist, d.h., für $n \in \mathbb{N}$, $t_1, \ldots, t_n \in T$, $z_1, \ldots, z_n \in \mathbb{C}$ gilt
$$\sum_{k,j=1}^{n} K(t_k, t_j) z_k \bar{z}_j \geq 0.$$
3. $K(t,t) = \text{Var}\ X(t)$, $t \in T$ genügt.

Die Eigenschaften 1. und 2. gelten auch für nicht-zentrierte Kovarianzfunktionen $C(s,t)$.

Bemerkung 5.5.4

1. Die Erwartungswertfunktion $\mu(t)$ zeigt einen (nicht-zufälligen) Trend. Falls $\mu(t)$ bekannt ist, kann die zufällige Funktion X zentriert werden indem man die zentrierte zufällige Funktion $Y = \{Y(t),\ t \in T\}$ mit $Y(t) = X(t) - \mu(t)$, $t \in T$ betrachtet.
2. Die Kovarianzfunktion $K(s,t)$ (bzw. $C(s,t)$) enthählt Informationen über die Struktur der Abhängigkeit von X.
3. Manchmal betrachtet man die Korrelationsfunktion $R(s,t) = \frac{K(s,t)}{\sqrt{K(s,s)K(t,t)}}$ für alle $s,t \in T$: $K(s,s) = \text{Var}\ X(s) > 0$, $K(t,t) = \text{Var}\ X(t) > 0$ anstelle von K und C. Aus der Cauchy-Schwarz-Ungleichung folgt $|R(s,t)| \leq 1$, $s,t \in T$.
4. Im Allgemeinen bestimmt die Menge aller gemischten Momente nicht (eindeutig) die Verteilung einer zufälligen Funktion.

Übungsaufgabe 5.5.5 Gib ein Beispiel für zwei verschiedene zufällige Funktionen $X = \{X(t),\ t \in T\}$ und $Y = \{Y(t),\ t \in T\}$ an, für die gilt $\mathbb{E}X(t) = \mathbb{E}Y(t)$, $t \in T$ und $\mathbb{E}(X(s)X(t)) = \mathbb{E}(Y(s)Y(t))$, $s,t \in T$.

Übungsaufgabe 5.5.6 Sei $\mu : T \to \mathbb{R}$ eine messbare Funktion und $K : T \times T \to \mathbb{R}$ eine positiv semidefinite symmetrische Funktion. Weise die Existenz einer zufälligen Funktion $X = \{X(t),\ t \in T\}$ mit $\mathbb{E}X(t) = \mu(t)$, $\text{Cov}\ (X(s), X(t)) = K(s,t)$, $s,t \in T$ nach.

Sei $X = \{X(t),\ t \in T\}$ eine reellwertige zufällige Funktion mit $\mathbb{E}\,|X(t)|^k < \infty$, $t \in T$ für ein $k \in \mathbb{N}$.

Definition 5.5.7 Der *mittlere Zuwachs der Ordnung* k von X ist gegeben durch $\gamma_k(s,t) = \mathbb{E}(X(s) - X(t))^k$, $s,t \in T$

Besonders wichtig ist die Funktion

$$\gamma(s,t) = \frac{1}{2}\gamma_2(s,t) = \frac{1}{2}\mathbb{E}(X(s) - X(t))^2 \quad s,t \in T,$$

die wir *Variogramm von* X nennen. In der Geostatistik wird das Variogramm anstelle der Kovarianzfunktion verwendet. Oft ersetzt man die Bedingung $\mathbb{E}X^2(t) < \infty$, $t \in T$ durch $\gamma(s,t) < \infty$ für alle $s,t \in T$.

Übungsaufgabe 5.5.8 Beweise die Existenz einer zufälligen Funktion ohne endliche zweite Momente mit $\gamma(s,t) < \infty$, $s,t \in T$.

Übungsaufgabe 5.5.9 Zeige, dass für eine zufällige Funktion $X = \{X(t),\ t \in T\}$ mit Erwartungswertfunktion μ und Kovarianzfunktion K gilt:

$$\gamma(s,t) = \frac{K(s,s) + K(t,t)}{2} - K(s,t) + \frac{1}{2}(\mu(s) - \mu(t))^2, \quad s,t \in T.$$

Falls die zufällige Funktion X *komplexwertig* ist, d.h., $X : \Omega \times T \to \mathbb{C}$ mit $\mathbb{E}\,|X(t)|^2 < \infty$, $t \in T$, dann ist die Kovarianzfunktion von X gegeben durch

$$K(s,t) = \mathbb{E}(X(s) - \mathbb{E}X(s))(\overline{X(t) - \mathbb{E}X(t)}) \quad s,t \in T,$$

wobei $\overline{z}$ das komplex-konjugierte von $z \in \mathbb{C}$ ist. Dann gilt $K(s,t) = \overline{K(t,s)}$, $s,t \in T$, und K ist positiv semidefinit.

5.6 Stationarität und Unabhängigkeit

Sei T eine Teilmenge eines linearen Vektorraums mit Operationen $+, -$ über $\mathbb{R}$.

Definition 5.6.1 Eine zufällige Funktion $X = \{X(t),\ t \in T\}$ heißt *stationär* (*stationär im strengen Sinn*), falls für alle $n \in \mathbb{N}$, $h, t_1, \ldots, t_n \in T$ mit $t_1 + h, \ldots, t_n + h \in T$ gilt:

$$P_{(X(t_1),\ldots,X(t_n))} = P_{(X(t_1+h),\ldots,X(t_n+h))},$$

d.h., alle endlich-dimensionalen Verteilungen von X sind translationsinvariant in T.

Definition 5.6.2 Eine (komplexwertige) zufällige Funktion $X = \{X(t),\ t \in T\}$ heißt *stationär zweiter Ordnung* (oder *stationär im weiten Sinn*), falls $\mathbb{E}|X(t)|^2 < \infty$, $t \in T$, $\mathbb{E}X(t) \equiv \mu$, $t \in T$ und

$$K(s,t) = \text{Cov}\ (X(s), X(t)) = K(s+h, t+h)$$

für alle $h, s, t \in T : s+h, t+h \in T$.

Bemerkung 5.6.3 1. Falls X stationär zweiter Ordnung ist, ist es praktisch, eine Funktion $K(t) := K(0,t)$, $t \in T$ einzuführen (wobei $0 \in T$).

2. Stationarität im strengen Sinn und zweiter Ordnung implizieren sich nicht gegenseitig.

3. Falls eine komplex-wertige zufällige Funktion stationär ist und endliche zweite Momente hat, dann ist die Funktion auch stationär zweiter Ordnung.

Definition 5.6.4 Eine reellwertige Funktion $X = \{X(t),\ t \in T\}$ heißt *intrinsisch stationär zweiter Ordnung*, falls $\gamma_k(s,t)$, $s,t \in T$ existiert für $k \le 2$, und

$$\gamma_1(s,t) = 0, \quad \gamma_2(s,t) = \gamma_2(s+h, t+h)$$

für alle $s,t,h \in T$, $s+h, t+h \in T$ gilt.

Für reellwertige zufällige Funktionen ist intrinsische Stationarität zweiter Ordnung allgemeiner als die Stationarität zweiter Ordnung, da die Existenz von $\mathbb{E}|X(t)|^2$, $t \in T$ nicht gefordert ist. Ein analoges Konzept der Stationarität der Zuwächse von X existiert im strengen Sinn.

Definition 5.6.5 Sei $X = \{X(t),\ t \in T\}$ ein reellwertiger stochastischer Prozess, $T \subset \mathbb{R}$. Wir sagen, dass X

1. *stationäre Zuwächse* hat, falls für alle $n \in \mathbb{N}$, $h, t_0, t_1, t_2, \ldots, t_n \in T$ mit
$$t_0 < t_1 < t_2 < \ldots < t_n, \quad t_i + h \in T, \quad i = 0, \ldots, n$$
die Verteilung von
$$(X(t_1+h) - X(t_0+h), \ldots, X(t_n+h) - X(t_{n-1}+h))^\top$$
nicht von h abhängt.

2. *unabhängige Zuwächse* hat, falls für alle $n \in \mathbb{N}$, $t_0, t_1, \ldots, t_n \in T$ mit $t_0 < t_1 < \ldots < t_n$ die Zufallsvariablen $X(t_0), X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1})$ unabhängig sind.

Bemerkung 5.6.6 Seien $(\mathcal{S}_1, \mathcal{B}_1)$ und $(\mathcal{S}_2, \mathcal{B}_2)$ messbare Räume. Im allgemeinen gilt für zwei zufällige Elemente $X : \Omega \to \mathcal{S}_1$ und $Y : \Omega \to \mathcal{S}_2$, dass diese sind *unabhängig* auf dem selbem W-Raum (Ω, F, P) sind, falls
$P(X \in A_1, Y \in A_2) = P(X \in A_1)P(Y \in A_2)$ für alle $A_1 \in \mathcal{B}_1$, $A_2 \in \mathcal{B}_2$.Diese Definition kann angewendet auf die Unabhängigkeit von zufälligen Funktionen X und Y mit Zustandsraum $(\mathcal{S}_T, \mathcal{B}_T)$, da sie als zufällige Elemente mit $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{S}_T$, $\mathcal{B}_1 = \mathcal{B}_2 = \mathcal{B}_T$ interpretiert werden können. Dasselbe gilt für die Unabhängigkeit eines zufälligen Elementes (oder einer zufälligen Funktion) X und einer Teil-σ-Algebra $\mathcal{G} \subseteq \mathcal{F}$: dies ist der Fall, falls $P(\{X \in A\} \cap G) = P(X \in A)P(G)$ für alle $A \in \mathcal{B}_1$, $G \in \mathcal{G}$ (oder $A \in \mathcal{B}_T$, $G \in \mathcal{G}$) gilt.

5.7 Prozesse mit unabhängigen Zuwächsen

Sei $\{\varphi_{s,t},\ s,t \geq 0\}$ eine Familie charakteristische Funktionen von Wahrscheinlichkeitsmaßen $Q_{s,t},\ s,t \geq 0$ auf $\mathcal{B}(\mathbb{R})$, d.h., für $z \in \mathbb{R},\ s,t \geq 0$ gilt $\varphi_{s,t}(z) = \int_{\mathbb{R}} e^{izx} Q_{s,t}(dx)$.

Satz 5.7.1 Es existiert ein stochastischer Prozess $X = \{X(t),\ t \geq 0\}$ mit unabhängigen Zuwächsen und der Eigenschaft, dass für alle $s,t \geq 0$ die charakteristische Funktion von $X(t) - X(s)$ gegeben ist durch $\varphi_{s,t}$ genau dann, wenn

$$\varphi_{s,t} = \varphi_{s,u}\varphi_{u,t} \tag{5.3}$$

für alle $0 \leq s < u < t < \infty$. Dabei kann die Verteilung von $X(0)$ beliebig gewählt werden.

Beweis $\Rightarrow$ Klar, da für alle $s,u,t \in (0,\infty) : s < u < t$ gilt

$$X(t) - X(s) = \underbrace{X(t) - X(u)}_{Y_1} + \underbrace{X(u) - X(s)}_{Y_2}$$

und $X(t) - X(u)$ und $X(u) - X(s)$ unabhängig sind. Dann folgt, dass $\varphi_{s,t} = \varphi_{Y_1+Y_2} = \varphi_{Y_1}\varphi_{Y_2} = \varphi_{s,u}\varphi_{u,t}$.

$\Leftarrow$ Falls die Existenz eines Prozesses X mit unabhängigen Zuwächsen und $\varphi_{X(t)-X(s)} = \varphi_{s,t}$ auf $(\Omega, \mathcal{F}, P)$ bereits gezeigt wurde, kann man die charakteristische Funktion der endlich-dimensionalen Verteilung mithilfe von $\{\varphi_{s,t}\}$ wie folgt definieren:
Sei $n \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_n < \infty$ und $Y = (X(t_0), X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1}))^\top$.
Die Unabhängigkeit impliziert

$$\begin{aligned}\varphi_Y(\underbrace{z_0, z_1, \ldots, z_n}_{z}) &= \mathbb{E}e^{i\langle z,Y\rangle} \\ &= \varphi_{X(t_0)}(z_0)\varphi_{t_0,t_1}(z_1)\ldots\varphi_{t_{n-1},t_n}(z_n),\ z \in \mathbb{R}^{n+1},\end{aligned}$$

wobei die Verteilung von $X(t_0)$ ein beliebiges Wahrscheinlichkeitsmaß Q_0 auf $\mathcal{B}(\mathbb{R})$ ist. Für $X_{t_0,\ldots,t_n} = (X(t_0), X(t_1), \ldots, X(t_n))^\top$ jedoch gilt $X_{t_0,\ldots,t_n} = AY$, wobei

$$A = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix}.$$

Dann folgt $\varphi_{X_{t_0,\dots,t_n}}(z) = \varphi_{AY}(z) = \mathbb{E}e^{i\langle z,AY\rangle} = \mathbb{E}e^{i\langle A^\top z,Y\rangle} = \varphi_Y(A^\top z)$. Daher hat die Verteilung von $X_{t_0,\dots,t_n}$ die charakteristische Funktion

$$\varphi_{X_{t_0,\dots,t_n}}(z) = \varphi_{Q_0}(l_0)\varphi_{t_0,t_1}(l_1)\dots\varphi_{t_{n-1},t_n}(l_n),$$

wobei $l = (l_0, l_1, \dots, l_n)^\top = A^\top z$ und daher gilt

$$\begin{cases} l_0 &= z_0 + \dots + z_n \\ l_1 &= z_1 + \dots + z_n \\ & \vdots \\ l_n &= z_n \end{cases}$$

Dabei ist $\varphi_{X(t_0)} = \varphi_{Q_0}$ und $\varphi_{X_{t_1,\dots,t_n}}(z_1,\dots,z_n) = \varphi_{X_{t_0,\dots,t_n}}(0, z_1, \dots, z_n)$ gilt für alle $z_i \in \mathbb{R}$.
Wir zeigen nun die Existenz eines solchen Prozesses X.
Dafür konstruieren wir folgende Familie von charakteristische Funktionen

$$\{\varphi_{t_0}, \varphi_{t_0,t_1,\dots,t_n}, \varphi_{t_1,\dots,t_n}, \quad 0 = t_0 < t_1 < \dots < t_n < \infty,\ n \in \mathbb{N}\}$$

aus φ_{Q_0} und $\{\varphi_{s,t},\ 0 \le s < t\}$ wie zuvor:

$$\varphi_{t_0} = \varphi_{Q_0},\ \varphi_{t_1,\dots,t_n}(z_1,\dots,z_n) = \varphi_{t_0,t_1,\dots,t_n}(0, z_1, \dots, z_n),\ z_i \in \mathbb{R},$$

$$\varphi_{t_0,\dots,t_n}(z) = \varphi_{t_0}(z_0 + \dots + z_n)\varphi_{t_0,t_1}(z_1 + \dots + z_n)\dots\varphi_{t_{n-1},t_n}(z_n).$$

Wir müssen prüfen, ob das dazugehörige Wahrscheinlichkeitsmaß der charakteristische Funktionen die Bedingungen des von Kolmogorow (vgl. Satz 5.1.16) erfüllt. Man kann zeigen, dass die Symmetrie und Konsistenz im Satz von Kolmogorow äquivalent sind zu

a) $\varphi_{t_{i_0},\dots,t_{i_n}}(z_{i_0},\dots,z_{i_n}) = \varphi_{t_0,\dots,t_n}(z_0,\dots,z_n)$ für eine beliebige Permutation $(0, 1, \dots, n) \mapsto (i_0, i_1, \dots, i_n)$,

b)
$$\varphi_{t_0,\dots,t_{m-1},t_{m+1},\dots,t_n}(z_0,\dots,z_{m-1},z_{m+1},\dots,z_n) = \\ = \varphi_{t_0,\dots,t_n}(z_0,\dots,0,\dots,z_n)$$

für alle $z_0,\dots,z_n \in \mathbb{R}$, $m \in \{1,\dots,n\}$.

a) ist offensichtlich.

b) folgt aus (5.3), denn

$$\varphi_{t_{m-1},t_m}(0 + z_{m+1} + \dots + z_n)\varphi_{t_m,t_{m+1}}(z_{m+1} + \dots + z_n) \\ = \varphi_{t_{m-1},t_{m+1}}(z_{m+1} + \dots + z_n)$$

für alle $m \in \{1,\dots,n\}$.

Also ist die Existenz von X bewiesen.

□

Beispiel 5.7.2

1. Falls $T = \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, dann hat $X = \{X(t),\ t \in \mathbb{N}_0\}$ unabhängige Zuwächse genau dann, wenn $X(n) \stackrel{d}{=} \sum_{i=0}^{n} Y_i$, wobei $\{Y_i\}$ unabhängige Zufallsvariablen sind und $Y_n \stackrel{d}{=} X(n) - X(n-1)$, $n \in \mathbb{N}$. Ein solcher Prozess X heißt *zufällige Irrfahrt* oder *random Walk*. Ein solcher Prozess kann auch für Y_i mit Werten in $\mathbb{R}^m$ definiert werden.

2. Ein Poisson–Prozess mit Intensität λ hat unabhängige Zuwächse.

3. Ein Wiener–Prozess hat unabhängige Zuwächse.

Übungsaufgabe 5.7.3 Beweise die Aussagen im Beispiel 1.

Übungsaufgabe 5.7.4 Sei $X = \{X(t),\ t \geq 0\}$ ein Prozess mit unabhängigen Zuwächsen und $g : [0, \infty) \to \mathbb{R}$ eine beliebige (deterministische) Funktion. Zeige, dass der Prozess $Y = \{Y(t),\ t \geq 0\}$ mit $Y(t) = X(t) + g(t)$, $t \geq 0$ auch unabhängige Zuwächse hat.

Kapitel 6

Zählprozesse

In diesem Kapitel werden verschiedene Beispiele von stochastischen Prozessen, welche das Zählen von Ereignissen modellieren diskutiert. Die sogenannten Zählprozesse, welche diese Ereignisse modellieren, besitzen stückweise konstante Pfade, deren Werte auf $\mathbb{N}_0$ liegen.

6.1 Erneuerungsprozesse

Sei $(\Omega, \mathcal{F}, P)$ W-Raum und $\{S_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge von f.s. nicht-negativen Zufallsvariablen, d.h. $0 \le S_1 \le S_2 \le \ldots \le S_n \le \ldots$.

Definition 6.1.1 Den stochastischen Prozess $N = \{N(t),\ t \ge 0\}$ nennt man *Zählprozess*, falls

$$N(t) = \sum_{n=1}^{\infty} I(S_n \le t),$$

wobei $I(A)$ die Indikatorfunktion vom Ereignis $A \in \mathcal{F}$ ist.

Beispiel 6.1.2 $N(t)$ zählt die Ereignisse, welche an den Zeitpunkten S_n bis zum Zeitpunkt t eintreten. Zum Beispiel können folgende Ereignisse zu den Zeitpunkten S_n eintreten:

1. das n-te Elementarteilchen im Geigerzähler,
2. ein Schadensfall in der Sachschadensversicherung,
3. ein Datenpaket in einem Server innerhalb eines Computernetzwerkes, usw.

Einen wichtigen Fall der Zählprozesse bilden die Erneuerungsprozesse.

Definition 6.1.3 Sei $\{T_n\}_{n\in\mathbb{N}}$ eine Folge von u.i.v. nicht-negativen Zufallsvariablen mit $P(T_1 > 0) > 0$. Ein Zählprozess $N = \{N(t),\ t \ge 0\}$ mit $N(0) = 0$ f.s., $S_n = \sum_{k=1}^{n} T_k$, $n \in \mathbb{N}$, heißt *Erneuerungsprozess*. Wir nennen S_n den *n-ten Erneuerungszeitpunkt*, $n \in \mathbb{N}$.

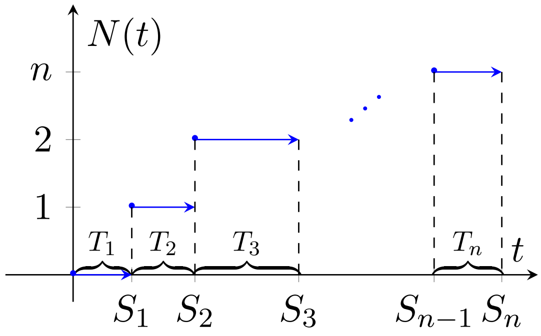


Abbildung 6.1: Konstruktion und Pfade eines Erneuerungsprozesses

Abbildung 6.1 kann wie folgt interpretiert werde:

- „Zwischenankunftszeiten" T_n entspricht der Lebensdauer eines technischen (Ersatz-)Teils oder Mechanismus innerhalb eines Systems. Also $S_n =$ Zeitpunkte, an denen das System versagt.

- Defektes Teil wird umgehend durch ein baugleiches Teil ersetzt (z.B. Glühbirne auswechseln).

- $N(t) = \#$ Reparaturen (sog. „Erneuerungen") des Systems bis zum Zeitpunkt t.

Bemerkung 6.1.4

1. Wir setzen $N(t) = \infty$, falls $S_n \le t$ für alle $n \in \mathbb{N}$.

2. Oft sind nur $T_2, T_3, \ldots$ i.v. mit $\mathbb{E}T_n < \infty$. Die Verteilung von T_1 ist frei wählbar. Ein solcher Prozess $N = \{N(t),\ t \ge 0\}$ heißt *verzögerter Erneuerungsprozess* (mit Verzögerung T_1).

3. Manchmal wird $T_n \ge 0$ nicht gefordert.

4. Klar, dass $\{S_n\}_{n \in \mathbb{N}_0}$ mit $S_0 = 0$ f.s., $S_n = \sum_{k=1}^n T_k$, $n \in \mathbb{N}$ eine *zufällige Irrfahrt* ist.

5. Falls der n-te Austausch eines defekten Teils eine Zeit T'_n braucht, dann ist durch $\tilde{T}_n = T_n + T'_n$, $n \in \mathbb{N}$ ein neuer Erneuerungsprozess gegeben mit den selben stochastischen Eigenschaften wie in Definition 6.1.1.

Im Folgenden nehmen wir an, dass $\mu = \mathbb{E}T_n \in (0, \infty)$, $n \in \mathbb{N}$.

Satz 6.1.5 *(Individueller Ergodensatz)*
Für einen Erneuerungsprozess $N = \{N(t),\ t \ge 0\}$ gilt

$$\lim_{t \to \infty} \frac{N(t)}{t} = \frac{1}{\mu} \quad \text{f.s.}$$

Beweis Für alle $t \geq 0$ und $n \in \mathbb{N}$ gilt $\{N(t) = n\} = \{S_n \leq t < S_{n+1}\}$ und daher $S_{N(t)} \leq t < S_{N(t)+1}$ und

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{S_{N(t)+1}}{N(t)+1} \cdot \frac{N(t)+1}{N(t)}.$$

Wenn wir zeigen, dass $\frac{S_{N(t)}}{N(t)} \xrightarrow[t\to\infty]{f.s} \mu$ und $N(t) \xrightarrow[t\to\infty]{f.s.} \infty$, dann gilt $\frac{t}{N(t)} \xrightarrow[t\to\infty]{f.s} \mu$ und damit auch die Aussage des Satzes.
Nach dem starken Gesetz der großen Zahlen von Kolmogorow (vgl. Satz 4.1.7) gilt $\frac{S_n}{n} \xrightarrow[n\to\infty]{f.s.} \mu$ und daher auch $S_n \xrightarrow[n\to\infty]{a.s.} \infty$ und $P(N(t) < \infty) = 1$, da

$$\begin{aligned} P(N(t) = \infty) &= P(\ S_n \leq t, \forall n) \\ &= 1 - \underbrace{P(\exists n : \forall m \in \mathbb{N}_0\ S_{n+m} > t)}_{=1,\text{ falls } S_n \xrightarrow[n\to\infty]{f.s} \infty} = 1 - 1 = 0. \end{aligned}$$

Dann ist $N(t)$, $t \geq 0$ eine reellwertige Zufallsvariable.
Wir zeigen, dass $N(t) \xrightarrow[t\to\infty]{a.s.} \infty$.
Da alle Pfade von $N(t)$ monoton nicht-fallend in $t \geq 0$ sind, $\exists \lim_{t\to\infty} N(\omega, t)$ für alle $\omega \in \Omega$. Weiter gilt

$$\begin{aligned} P(\lim_{t\to\infty} N(t) < \infty) &= \lim_{n\to\infty} P(\lim_{t\to\infty} N(t) < n) \\ &\overset{(*)}{=} \lim_{n\to\infty} \lim_{t\to\infty} P(N(t) < n) \\ &= \lim_{n\to\infty} \lim_{t\to\infty} P(S_n > t) \\ &= \lim_{n\to\infty} \lim_{t\to\infty} P(\sum_{k=1}^{n} T_k > t) \\ &\leq \lim_{n\to\infty} \lim_{t\to\infty} \sum_{k=1}^{n} \underbrace{P(T_k > \frac{t}{n})}_{\xrightarrow[t\to\infty]{} 0} = 0. \end{aligned}$$

$(*)$ gilt, weil

$$\begin{aligned} \left\{\lim_{t\to\infty} N(t) < n\right\} &= \{\exists t_0 \in \mathbb{Q}_+ : \forall t \geq t_0\ N(t) < n\} \\ &= \cup_{t_0 \in \mathbb{Q}_+} \cap_{\substack{t \in \mathbb{Q}_+ \\ t \geq t_0}} \{N(t) < n\} = \liminf_{\substack{t \in \mathbb{Q}_+ \\ t\to\infty}} \{N(t) < n\}. \end{aligned}$$

Dann nutzt man die Stetigkeit des W-Maßes, wobei
$\mathbb{Q}_+ = \mathbb{Q} \cap \mathbb{R}_+ = \{q \in \mathbb{Q} : q \geq 0\}$. Da für alle $\omega \in \Omega$, gilt, dass $\lim_{n\to\infty} \frac{S_n}{n} = \lim_{t\to\infty} \frac{S_{N(t)}}{N(t)}$ (eine Realisierung von $N(\cdot)$ ist eine Teilfolge von $\mathbb{N}$), folgt $\lim_{t\to\infty} \frac{S_{N(t)}}{N(t)} \overset{a.s}{=} \mu$. $\square$

Bemerkung 6.1.6 Der individuelle Ergodensatz kann auf den Fall von nicht identisch verteilten T_n verallgemeinert werden. Dabei wird gefordert, dass $\mu_n = \mathbb{E}T_n$, $\{T_n - \mu_n\}_{n\in\mathbb{N}}$ gleichgradig integrierbar ist und
$\frac{1}{n}\sum_{k=1}^{n}\mu_k \xrightarrow[n\to\infty]{} \mu > 0$.
Dann kann gezeigt werden, dass $\frac{N(t)}{t} \xrightarrow[t\to\infty]{P} \frac{1}{\mu}$ (vgl. [3], Seite 276).

Satz 6.1.7 *(Zentraler Grenzwertsatz)*
Für $\mu \in (0,\infty)$ und $\sigma^2 = \operatorname{Var} T_1 \in (0,\infty)$ gilt

$$\mu^{\frac{3}{2}} \cdot \frac{N(t) - \frac{t}{\mu}}{\sigma\sqrt{t}} \xrightarrow[t\to\infty]{d} Y,$$

mit $Y \sim \mathcal{N}(0,1)$.

Beweis Nach dem Zentralen Grenzwersatz für Summen von u.i.v. Zufallsvariablen (vgl. Satz 4.2.1) gilt

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow[n\to\infty]{d} Y. \tag{6.1}$$

Sei $[x]$ der ganzzahlige Anteil von $x \in \mathbb{R}$. Es gilt für $a = \frac{\sigma^2}{\mu^3}$, dass

$$P\left(\frac{N(t) - \frac{t}{\mu}}{\sqrt{at}} \le x\right) = P\left(N(t) \le x\sqrt{at} + \frac{t}{\mu}\right) = P\left(S_{m(t)} > t\right),$$

mit $m(t) = \left[x\sqrt{at} + \frac{t}{\mu}\right] + 1$, $t \ge 0$ und $\lim_{t\to\infty} m(t) = \infty$. Daher erhalten wir

$$\left|P\left(\frac{N(t) - \frac{t}{\mu}}{\sqrt{at}} \le x\right) - \Phi(x)\right| = \left|P\left(S_{m(t)} > t\right) - \Phi(x)\right|$$
$$= \left|P\left(\frac{S_{m(t)} - \mu m(t)}{\sigma\sqrt{m(t)}} > \frac{t - \mu m(t)}{\sigma\sqrt{m(t)}}\right) - \Phi(x)\right| =: I_t(x)$$

für beliebiges $t \ge 0$ und $x \in \mathbb{R}$, wobei Φ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist. Für ein festes $x \in \mathbb{R}$ definieren wir $Z_t = -\frac{t-\mu m(t)}{\sigma\sqrt{m(t)}} - x$, $t \ge 0$. Dann gilt

$$I_t(x) = \left|P\left(\frac{S_{m(t)} - \mu m(t)}{\sigma\sqrt{m(t)}} + Z_t > -x\right) - \Phi(x)\right|.$$

Wenn wir $Z_t \xrightarrow[t\to\infty]{} 0$ beweisen und danach (6.1) und den Satz von Slutsky (vlg. [20, Theorem 6.4.1]) anwenden, erhalten wir

$$\frac{S_{m(t)} - \mu m(t)}{\sigma\sqrt{m(t)}} + Z_t \xrightarrow[t\to\infty]{d} Y \sim \mathcal{N}(0,1),$$

da aus $Z_t \xrightarrow[t\to\infty]{f.s.} 0$ folgt, dass $Z_t \xrightarrow[t\to\infty]{d} 0$. Daher schreiben wir

$$I_t(x) \xrightarrow[t\to\infty]{} \left|\bar{\Phi}(-x) - \Phi(x)\right| = |\Phi(x) - \Phi(x)| = 0,$$

wobei $\overline{\Phi}(x) = 1 - \Phi(x)$ die Tail-Funktion der $\mathcal{N}(0,1)$-Verteilung ist und wir die Symmetrieeigenschaft von $\mathcal{N}(0,1) : \overline{\Phi}(-x) = \Phi(x),\ x \in \mathbb{R}$ verwendet haben. Nun zeigen wir, dass $Z_t \xrightarrow[t\to\infty]{} 0$, also $\frac{t-\mu m(t)}{\sigma\sqrt{m(t)}} \xrightarrow[t\to\infty]{} -x$. Es gilt $m(t) = x\sqrt{at} + \frac{t}{\mu} + \varepsilon(t)$, wobei $\varepsilon(t) \in [0,1)$. Dann folgt

$$\begin{aligned}
\frac{t-\mu m(t)}{\sigma\sqrt{m(t)}} &= \frac{t - \mu x\sqrt{at} - t - \mu\varepsilon(t)}{\sigma\sqrt{m(t)}} \\
&= -x\frac{\sqrt{at}\mu}{\sigma\sqrt{x\sqrt{at} + \frac{t}{\mu} + \varepsilon(t)}} - \frac{\mu\varepsilon(t)}{\sigma\sqrt{m(t)}} \\
&= -\frac{x\mu}{\sigma\sqrt{\frac{x}{\sqrt{at}} + \frac{1}{\mu a} + \frac{\varepsilon(t)}{at}}} - \frac{\mu\varepsilon(t)}{\sigma\sqrt{m(t)}} \\
&= \underbrace{-\frac{x\frac{\mu}{\sigma}}{\sqrt{\frac{\mu^2}{\sigma^2} + \frac{x}{\sqrt{at}} + \frac{\varepsilon(t)}{at}}}}_{\xrightarrow[t\to\infty]{} -x} - \underbrace{\frac{\mu\varepsilon(t)}{\sigma\sqrt{m(t)}}}_{\xrightarrow[t\to\infty]{} 0} \xrightarrow[t\to\infty]{} -x.
\end{aligned}$$

□

Bemerkung 6.1.8 Der Zentrale Grenzwersatz lässt sich auch in Lindeberg-Form formulieren und für nicht identisch verteilte T_n beweisen (vgl. [3] S. 276 f.).

Definition 6.1.9 Die Funktion $H(t) = \mathbb{E}N(t),\ t \geq 0$ heißt *Erneuerungsfunktion* des Prozesses N (oder der Folge $\{S_n\}_{n\in\mathbb{N}}$). Sei $F_T(x) = P(T_1 \leq x)$, $x \in \mathbb{R}$ die Verteilungsfunktion von T_1. Für beliebige Verteilungsfunktionen $F, G : \mathbb{R} \to [0,1]$ ist die *Faltung* $F * G$ definiert als
$F * G(x) = \int_{-\infty}^{+\infty} F(x-y)dG(y)$. Die *k-fache Faltung* F^{*k} der Verteilung F mit sich selbst, $k \in \mathbb{N}_0$, ist induktiv definiert:

$$\begin{aligned}
F^{*0}(x) &= I(x \in [0,\infty)),\ x \in \mathbb{R}, \\
F^{*1}(x) &= F(x),\ x \in \mathbb{R}, \\
F^{*(k+1)}(x) &= F^{*k} * F(x),\ x \in \mathbb{R}.
\end{aligned}$$

Lemma 6.1.10 Die Erneuerungsfunktion H eines Erneuerungsprozesses N ist monoton wachsend und rechtsseitig stetig auf $\mathbb{R}_+$. Weiter gilt

$$H(t) = \sum_{n=1}^{\infty} P(S_n \leq t) = \sum_{n=1}^{\infty} F_T^{*n}(t),\ t \geq 0. \tag{6.2}$$

Beweis Die Monotonie und rechtsseitige Stetigkeit von H folgen aus den f.s. Pfadeigenschaften von N. Wir zeigen (6.2):

$$\begin{aligned} H(t) = \mathbb{E}N(t) = \mathbb{E}\sum_{n=1}^{\infty} I(S_n \le t) &\overset{(*)}{=} \sum_{n=1}^{\infty} \mathbb{E}I(S_n \le t) \\ &= \sum_{n=1}^{\infty} P(S_n \le t) = \sum_{n=1}^{\infty} F_T^{*n}(t), \end{aligned}$$

da $P(S_n \le t) = P(T_1 + \ldots + T_n \le t) = F_T^{*n}(t)$, $t \ge 0$. Die Relation $(*)$ gilt für alle Partialsummen auf beiden Seiten, also auch im Grenzwert. $\square$

Bis auf wenige Ausnahmen ist es unmöglich, die Erneuerungsfunktion H mit Hilfe der Formel (6.2) analytisch zu berechnen. Deshalb wird die *Laplace-Transformation* von H oft verwendet.

Definition 6.1.11 Für $G : [0, \infty) \to \mathbb{R}$ monoton und rechtsseitig stetig ist die *Laplace-Transformation* gegeben durch

$$\hat{l}_G(s) = \int_0^{\infty} e^{-sx} dG(x), \quad s \ge 0,$$

wobei das obige Integral ein *Lebesgue-Stieltjes Integral* ist, also ein Lebesgue Integral bezüglich des Maßes μ_G auf $\mathcal{B}_{\mathbb{R}_+}$ definiert durch
$\mu_G((x, y]) = G(y) - G(x)$, $0 \le x < y < \infty$, falls G monoton wachsend ist. Die Laplace-Transformierte $\hat{l}_X$ einer Zufallsvariablen $X \ge 0$ ist definiert als

$$\hat{l}_X(s) = \int_0^{\infty} e^{-sx} dF_X(x), \quad s \ge 0.$$

Lemma 6.1.12 Für $s > 0$ gilt:

$$\hat{l}_H(s) = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_1}(s)}.$$

Beweis Es gilt

$$\begin{aligned} \hat{l}_H(s) &= \int_0^{\infty} e^{-sx} dH(x) \overset{(6.2)}{=} \int_0^{\infty} e^{-sx} d\left(\sum_{n=1}^{\infty} F_T^{*n}(x)\right) \\ &= \sum_{n=1}^{\infty} \int_0^{\infty} e^{-sx} dF_T^{*n}(x) = \sum_{n=1}^{\infty} \hat{l}_{T_1+\ldots+T_n}(s) \\ &= \sum_{n=1}^{\infty} \left(\hat{l}_{T_1}(s)\right)^n = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_1}(s)}, \end{aligned}$$

wobei für $s > 0$ gilt, dass $\hat{l}_{T_1}(s) < 1$ und daher die geometrische Reihe $\sum_{n=1}^{\infty} \left(\hat{l}_{T_1}(s)\right)^n$ konvergiert. $\square$

Bemerkung 6.1.13 Falls $N = \{N(t),\ t \geq 0\}$ ein verzögerter Erneuerungsprozess (mit Verzögerung T_1) ist, dann gelten die Aussagen der Lemmata 6.1.10 und 6.1.12 in folgender Form:

1.
$$H(t) = \sum_{n=0}^{\infty} (F_{T_1} * F_{T_2}^{*n})(t),\ t \geq 0,$$
wobei F_{T_1} bzw. F_{T_2} die Verteilungsfunktionen von T_1 bzw. T_n sind, $n \geq 2$.

2.
$$\hat{l}_H(s) = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_2}(s)},\ s \geq 0, \tag{6.3}$$
wobei $\hat{l}_{T_1}$ bzw. $\hat{l}_{T_2}$ die Laplace-Transformierten der Verteilungen von T_1 bzw. T_n, $n \geq 2$ sind.

Für weitere Untersuchungen brauchen wir einen Satz (von Wald) über den Erwartungswert einer Summe (einer zufälligen Anzahl) von unabhängigen Zufallsvariablen.

Definition 6.1.14 Sei ν eine $\mathbb{N}$-wertige Zufallsvariable und $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf dem selben W-Raum. ν heißt *unabhängig von der Zukunft*, falls für alle $n \in \mathbb{N}$ das Ereignis $\{\nu \leq n\}$ nicht von der σ-Algebra $\sigma(\{X_k,\ k > n\})$ abhängt.

Satz 6.1.15 *(Waldsche Identität)*
Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariable mit $\sup \mathbb{E}|X_n| < \infty$, $\mathbb{E}X_n = a$, $n \in \mathbb{N}$, und sei ν eine $\mathbb{N}$-wertige Zufallsvariable, die von der Zukunft unabhängig ist mit $\mathbb{E}\nu < \infty$. Dann gilt

$$\mathbb{E}(\sum_{n=1}^{\nu} X_n) = a \cdot \mathbb{E}\nu.$$

Beweis Definiere $S_n = \sum_{k=1}^{n} X_k$, $n \in \mathbb{N}$. Da $\mathbb{E}\nu = \sum_{n=1}^{\infty} P(\nu \geq n)$, folgt die Aussage aus dem Lemma 6.1.16. $\square$

Lemma 6.1.16 *(Kolmogorow-Prokhorov)*
Sei ν eine $\mathbb{N}$-wertige Zufallsvariable, die unabhängig von der Zukunft ist und es gelte
$$\sum_{n=1}^{\infty} P(\nu \geq n)\mathbb{E}|X_n| < \infty. \tag{6.4}$$
Dann gilt
$$\mathbb{E}S_\nu = \sum_{n=1}^{\infty} P(\nu \geq n)\mathbb{E}X_n.$$
Falls $X_n \geq 0$ f.s., dann ist die Voraussetzung (6.4) nicht notwendig.

Beweis Es gilt
$$S_\nu = \sum_{n=1}^{\nu} X_n = \sum_{n=1}^{\infty} X_n I(\nu \geq n).$$
Definiere $S_{\nu,n} = \sum_{k=1}^n X_k I(\nu \geq k)$, $n \in \mathbb{N}$. Zuerst, beweisen wir die Aussage für $X_n \geq 0$ f.s., $n \in \mathbb{N}$. Für jedes $\omega \in \Omega$ gilt $S_{\nu,n} \uparrow S_\nu$, $n \to \infty$ und nach dem Satz über die monotone Konvergenz:
$$\mathbb{E}S_\nu = \lim_{n\to\infty} \mathbb{E}S_{\nu,n} = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{E}(X_k I(\nu \geq k)).$$
Da $\{\nu \geq k\} = \{\nu \leq k-1\}^c$ nicht von $\sigma(X_k) \subset \sigma(\{X_n,\ n \geq k\})$ abhängt, gilt $\mathbb{E}(X_k I(\nu \geq k)) = \mathbb{E}X_k P(\nu \geq k)$, $k \in \mathbb{N}$ und $\mathbb{E}S_\nu = \sum_{n=1}^\infty P(\nu \geq n)\mathbb{E}X_n$.
Nun sei X_n beliebig.
Wähle

- $Y_n = |X_n|$,
- $Z_n = \sum_{k=1}^n Y_k$,
- $Z_{\nu,n} = \sum_{k=1}^n Y_k I(\nu \geq k)$, $n \in \mathbb{N}$.

Da $Y_n \geq 0$, $n \in \mathbb{N}$, folgt
$$\mathbb{E}Z_\nu = \sum_{n=1}^{\infty} \mathbb{E}(|X_n|)P(\nu \geq n) < \infty$$
aus (6.4). Da $|S_{\nu,n}| \leq Z_{\nu,n} \leq Z_\nu$, $n \in \mathbb{N}$, folgt aus dem Satz von Lebesgue über die dominierte Konvergenz, dass
$\mathbb{E}S_\nu = \lim_{n\to\infty} \mathbb{E}S_{\nu,n} = \sum_{n=1}^\infty \mathbb{E}X_n P(\nu \geq n)$, wobei die Reihe absolut konvergiert. $\square$

Folgerung 6.1.17

1. Für eine beliebige Borel-messbare Funktion $g : \mathbb{R}_+ \to \mathbb{R}_+$ und den Erneuerungsprozess $N = \{N(t),\ t \geq 0\}$ mit Zwischenankunftszeiten $\{T_n\}$, T_n u.i.v., $\mu = \mathbb{E}T_n \in (0, \infty)$ gilt
$$\mathbb{E}\left(\sum_{n=1}^{N(t)+1} g(T_n)\right) = (1 + H(t))\mathbb{E}g(T_1),\ t \geq 0.$$
2. $H(t) < \infty$, $t \geq 0$.

Beweis

1. Für jedes $t \geq 0$ ist offensichtlich, dass $\nu = 1 + N(t)$ nicht von der Zukunft von $\{T_n\}_{n\in\mathbb{N}}$ abhängt. Der Rest folgt aus dem Satz 6.1.3 mit $X_n = g(T_n)$, $n \in \mathbb{N}$.

2. Für $s > 0$ betrachte $T_n^{(s)} = \min\{T_n, s\}$, $n \in \mathbb{N}$. Wähle $s > 0$ so, dass für beliebiges, aber festes $\varepsilon > 0$:

$$\mu^{(s)} = \mathbb{E}T_1^{(s)} \geq \mu - \varepsilon > 0.$$

Sei $N^{(s)}$ der Erneuerungsprozess basierend auf der Folge $\{T_n^{(s)}\}_{n\in\mathbb{N}}$ der Zwischenankunftszeiten:

$$N^{(s)}(t) = \sum_{n=1}^{\infty} I(S_n^{(s)} \leq t), \quad t \geq 0,$$

wobei $S_n^{(s)} = T_1^{(s)} + \ldots + T_n^{(s)}$, $n \in \mathbb{N}$. Es gilt dann $N(t) \leq N^{(s)}(t)$, $t \geq 0$, f.s., und aus Teil 1. folgt:

$$\begin{aligned}(\mu - \varepsilon)(\mathbb{E}N^{(s)}(t) + 1) \leq \mu^{(s)}(\mathbb{E}N^{(s)}(t) + 1) &= \mathbb{E}S^{(s)}_{N^{(s)}(t)+1} \\ &= \mathbb{E}(\underbrace{S^{(s)}_{N^{(s)}(t)}}_{\leq t} + \underbrace{T^{(s)}_{N^{(s)}(t)+1}}_{\leq s}) \leq t + s, \quad t \geq 0.\end{aligned}$$

Daher $H(t) = \mathbb{E}N(t) \leq \mathbb{E}N^{(s)}(t) \leq \frac{t+s}{\mu-\varepsilon}$, $t \geq 0$.
Es folgt $H(t) < \infty$, $t \geq 0$. Da $\varepsilon > 0$ beliebig ist, folgt auch

$$\limsup_{t\to\infty} \frac{H(t)}{t} \leq \frac{1}{\mu}.$$

□

Folgerung 6.1.18 *(Elementarer Erneuerungssatz)*
Für einen Erneuerungsprozess N wie in Folgerung 6.1.17.1., definiert, gilt:

$$\lim_{t\to\infty} \frac{H(t)}{t} = \frac{1}{\mu}.$$

Beweis In Folgerung 6.1.17.2. haben wir bewiesen, dass

$$\limsup_{t\to\infty} \frac{H(t)}{t} \leq \frac{1}{\mu}.$$

Wenn wir zeigen, dass

$$\liminf_{t\to\infty} \frac{H(t)}{t} \geq \frac{1}{\mu},$$

ist die Aussage bewiesen. Nach dem Satz 6.1.5 gilt $\frac{N(t)}{t} \xrightarrow[t\to\infty]{} \frac{1}{\mu}$ f.s., und daher folgt mit Fatous Lemma

$$\frac{1}{\mu} = \mathbb{E}\liminf_{t\to\infty} \frac{N(t)}{t} \leq \liminf_{t\to\infty} \frac{\mathbb{E}N(t)}{t} = \liminf_{t\to\infty} \frac{H(t)}{t}.$$

□

Bemerkung 6.1.19

1. Falls $\mu_2 = \mathbb{E}T_1^2 < \infty$, kann eine genauere Asymptotik für $H(t)$, $t \to \infty$ hergeleitet werden:
$$\frac{H(t)}{t} = \frac{1}{\mu} + \frac{\mu_2}{2\mu^2 t} + o(1/t),\ t \to \infty.$$

2. Der elementare Erneuerungssatz gilt auch für verzögerte Erneuerungsprozesse mit $\mu = \mathbb{E}T_2$.

3. Definieren wir das *Erneuerungsmaß* H auf $\mathcal{B}(\mathbb{R}_+)$ durch
$$H(B) = \sum_{n=1}^{\infty} \int_B dF_T^{*n}(x), B \in \mathcal{B}(\mathbb{R}_+),$$
wobei $F_T^{*n}(x) = F_{T_1} * F_{T_2}^{*(n-1)}(x)$. Es gilt dann
$$\begin{aligned} &H([0,t]) = H(t), \\ &H((s,t]) = H(t) - H(s), \quad s,t \geq 0, \end{aligned}$$
falls H sowohl die Erneuerungsfunktion als auch das Erneuerungsmaß ist.

Satz 6.1.20 *(Hauptsatz der Erneuerungstheorie)*
Sei $N = \{N(t),\ t \geq 0\}$ ein (verzögerter) Erneuerungsprozess assoziiert mit der Folge $\{T_n\}_{n\in\mathbb{N}}$ mit unabhängigen Zufallsvariablen T_n, $n \in \mathbb{N}$ und i.v. $\{T_n,\ n \geq 2\}$. Weiter sei die Verteilung von T_2 nicht arithmetisch d.h., nicht konzentriert auf einem regulären Gitter mit Wahrscheinlichkeit 1. Die Verteilung von T_1 sei beliebig und $\mathbb{E}T_2 = \mu \in (0,\infty)$. Dann gilt
$$\int_0^t g(t-x)dH(x) \xrightarrow[t\to\infty]{} \frac{1}{\mu}\int_0^\infty g(x)dx,$$
wobei $g : \mathbb{R}_+ \to \mathbb{R}$ Riemann-integrierbar auf $[0,n]$ für alle $n \in \mathbb{N}$ ist, und $\sum_{n=0}^\infty \max\limits_{n\leq x\leq n+1} |g(x)| < \infty$.

Insbesondere gilt $H((t-u,t]) \xrightarrow[t\to\infty]{} \frac{u}{\mu}$ für beliebiges $u \in \mathbb{R}_+$ und daher verhält sich H asymptotisch (für $t \to \infty$) wie das Lebesgue-Maß.

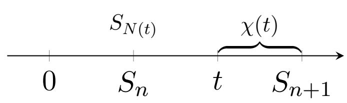


Abbildung 6.2: Exzess von N

Definition 6.1.21 Eine Zufallsvariable $\chi(t) = S_{N(t)+1} - t$ heißt *Exzess* von N zum Zeitpunkt $t \geq 0$ (Offensichtlich gilt $\chi(0) = T_1$).

Im folgenden wollen wir einen Erneuerungsprozess mit stationären Zuwächsen betrachten. Sei hierfür $N = \{N(t),\ t \geq 0\}$ ein verzögerter Erneuerungsprozess assoziiert mit der Folge von Zwischenankunftszeiten $\{T_n\}_{n\in\mathbb{N}}$. Seien F_{T_1} und F_{T_2} Verteilungsfunktionen der Verzögerungen T_1 und T_n, $n \geq 2$. Wir nehmen an, dass $\mu = \mathbb{E}T_2 \in (0,\infty)$, $F_{T_2}(0) = 0$ und daher $T_2 > 0$ f.s. und

$$F_{T_1}(x) = \frac{1}{\mu}\int_0^x \overline{F}_{T_2}(y)dy,\ x \geq 0. \tag{6.5}$$

In diesem Fall heißt F_{T_1} *integrierte Tailverteilungsfunktion* von T_2.

Satz 6.1.22 Unter den obigen Voraussetzungen ist N ein Prozess mit stationären Zuwächsen.

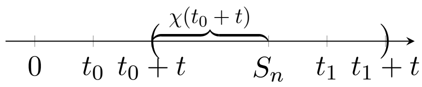


Abbildung 6.3: Skizze zum Beweis des Satzes 6.1.22

Beweis Sei $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \ldots < t_n < \infty$. Weil $\{T_n,\ n \in \mathbb{N}\}$ unabhängig sind, ist die gemeinsame Verteilung von $(N(t_1+t) - N(t_0+t), \ldots, N(t_n+t) - N(t_{n-1}+t))^\top$ unabhängig von t, falls die Verteilung von $\chi(t)$ nicht von t abhängt. Daher $\chi(t_0+t) \stackrel{d}{=} \chi(t_i+t) \stackrel{d}{=} \chi(0) = T_1$, $t \geq 0$, (vgl. Abbildung 6.3)
Wir zeigen, dass $F_{T_1} = F_{\chi(t)}$, $t \geq 0$ gilt:

$$\begin{aligned}
F_{\chi(t)}(x) &= P(\chi(t) \leq x) = \sum_{n=0}^{\infty} P(S_n \leq t,\ t < S_{n+1} \leq t+x) \\
&= P(S_0 = 0 \leq t,\ t < S_1 = T_1 \leq t+x) \\
&+ \sum_{n=1}^{\infty} \mathbb{E}(\mathbb{E}(I(S_n \leq t,\ t < S_n + T_{n+1} \leq t+x) \mid S_n)) \\
&= F_{T_1}(t+x) - F_{T_1}(t) \\
&+ \sum_{n=1}^{\infty} \int_0^t P(t-y < T_{n+1} \leq t+x-y)\ dF_{S_n}(y) \\
&= F_{T_1}(t+x) - F_{T_1}(t) \\
&+ \int_0^t P(t-y < T_2 \leq t+x-y)\ d(\underbrace{\sum_{n=1}^{\infty} F_{S_n}(y)}_{H(y)}).
\end{aligned}$$

Wenn wir zeigen, dass $H(y) = \frac{y}{\mu}$, $y \geq 0$, dann folgt auch

$$\begin{aligned}
F_{\chi(t)}(x) &\overset{z=t-y}{=} F_{T_1}(t+x) - F_{T_1}(t) \\
&+ \frac{1}{\mu}\int_t^0 (F_{T_2}(z+x) - 1 + 1 - F_{T_2}(z))d(-z) \\
&= F_{T_1}(t+x) - F_{T_1}(t) + \frac{1}{\mu}\int_0^t (\bar{F}_{T_2}(z) - \bar{F}_{T_2}(z+x))dz \\
&= F_{T_1}(t+x) - F_{T_1}(t) + F_{T_1}(t) - \frac{1}{\mu}\int_x^{t+x} \bar{F}_{T_2}(y)dy \\
&= F_{T_1}(t+x) - F_{T_1}(t+x) + F_{T_1}(x) = F_{T_1}(x),\ x \geq 0,
\end{aligned}$$

nach der Form (6.5) der Verteilung von T_1. Nun zeigen wir $H(t) = \frac{t}{\mu}$, $t \geq 0$ und benutzen dazu (6.5):

$$\begin{aligned}
\hat{l}_{T_1}(s) &= \frac{1}{\mu}\int_0^\infty e^{-st}(1 - F_{T_2}(t))dt \\
&= \frac{1}{\mu}\underbrace{\int_0^\infty e^{-st}dt}_{\frac{1}{s}} - \frac{1}{\mu}\int_0^\infty e^{-st}F_{T_2}(t)dt \\
&= \frac{1}{\mu s}\left(1 + \int_0^\infty F_{T_2}(t)de^{-st}\right) \\
&= \frac{1}{\mu s}(1 + \underbrace{e^{-st}F_{T_2}(t)\,|_0^\infty}_{-F_{T_2}(0)=0} - \underbrace{\int_0^\infty e^{-st}dF_{T_2}(t))}_{\hat{l}_{T_2}(s)} \\
&= \frac{1}{\mu s}(1 - \hat{l}_{T_2}(s)),\ s \geq 0.
\end{aligned}$$

Mit dem Lemma 6.1.12 erhalten wir

$$\hat{l}_H(s) = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_2}(s)} = \frac{1}{\mu s} = \frac{1}{\mu}\int_0^\infty e^{-st}dt = \hat{l}_{\frac{t}{\mu}}(s),\ s \geq 0.$$

Da die Laplace-Transformierte einer Funktion diese eindeutig bestimmt, gilt $H(t) = \frac{t}{\mu}$, $t \geq 0$. □

Bemerkung 6.1.23 Im Beweis des Satzes 6.1.22 haben wir gezeigt, dass für verzögerte Erneuerungsprozesse mit Verteilung (6.5), $H(t) \sim \frac{t}{\mu}$ nicht nur asymptotisch für $t \to \infty$ (wie im elementaren Erneuerungssatz) gilt, sondern $H(t) = \frac{t}{\mu}$ für alle $t \geq 0$. Das heißt, im Durchschnitt erhalten wir $\frac{1}{\mu}$ Erneuerungen per Einheitszeitintervall. Deshalb nennt man einen solchen Prozess N einen *homogenen Erneuerungsprozess*.

Man kann auch folgenden Satz beweisen:

Satz 6.1.24 Falls $N = \{N(t),\ t \geq 0\}$ ein verzögerter Erneuerungsprozess mit beliebiger Verzögerung T_1 ist und nicht-arithmetischer Verteilung von T_n, $n \geq 2$, $\mu = \mathbb{E}T_2 \in (0, \infty)$, dann gilt, dass

$$\lim_{t\to\infty} F_{\chi(t)}(x) = \frac{1}{\mu}\int_0^x \overline{F}_{T_2}(y)dy,\ x \geq 0.$$

D.h., dass bei der Definition eines homogenen Erneuerungsprozesses die Grenzverteilung des Exzesses $\chi(t)$, $t \to \infty$ gleich der Verteilung von T_1 ist.

6.2 Inhomogene Poisson-Prozesse

In diesem Abschnitt wird die Definition eines homogenen Poisson-Prozesses verallgemeinert.

Definition 6.2.1 Der Zählprozess $N = \{N(t),\ t \geq 0\}$ heißt *Poisson-Prozess* mit Intensitätsmaß Λ, falls

1. $N(0) = 0$ f.s.
2. Λ ist ein lokal endliches Maß auf $\mathbb{R}_+$, d.h., für das Maß $\Lambda : \mathcal{B}(\mathbb{R}_+) \to \mathbb{R}_+$ gilt $\Lambda(B) < \infty$ für jede beschränkte Menge $B \in \mathcal{B}(\mathbb{R}_+)$.
3. N hat unabhängige Zuwächse.
4. $N(t) - N(s) \sim \text{Pois}(\Lambda((s,t]))$ für alle $0 \leq s < t < \infty$.

Manchmal wird der Poisson-Prozess $N = \{N(t),\ t \geq 0\}$ definiert durch das dazugehörige zufällige Poisson-Zählmaß $N = \{N(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$ via $N = N([0,t])$, $t \geq 0$, wobei ein Zählmaß ein lokal endliches Maß mit Werten in $\mathbb{N}_0$ ist.

Definition 6.2.2 Ein zufälliges Zählmaß $N = \{N(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$ heißt *Poissonsch* mit lokal endlichem Intensitätsmaß Λ, falls

1. für beliebiges $n \in \mathbb{N}$ und für beliebige paarweise disjunkte beschränkte Mengen $B_1, B_2, \ldots, B_n \in \mathcal{B}(\mathbb{R}_+)$ die Zufallsvariablen $N(B_1), N(B_2), \ldots, N(B_n)$ unabhängig sind.
2. $N(B) \sim \text{Pois}(\Lambda(B))$, $B \in \mathcal{B}(\mathbb{R}_+)$, B-beschränkt.

Offensichtlich folgen 3. und 4. in der Definition 6.2.1 aus 1. und 2. in der Definition 6.2.2. Punkt 1. aus Definition 6.2.1 ist jedoch eigenständig.
$N(B)$, $B \in \mathcal{B}(\mathbb{R}_+)$ wird als die Anzahl der Punkte von N in der Menge B interpretiert.

Bemerkung 6.2.3 Genau wie in Definition 6.2.2, kann ein Poisson-Zählmaß auch auf beliebigen metrischen Räumen E mit einer Borel-σ-Algebra definiert werden. Sehr oft wird $E = \mathbb{R}^d$, $d \geq 1$ gewählt.

Lemma 6.2.4 Für jedes lokal endliches Maß Λ auf $\mathbb{R}_+$ existiert ein Poisson-Prozess mit Intensitätsmaß Λ.

Beweis Würde ein solcher Poisson-Prozess existieren, wäre die charakteristische Funktion $\varphi_{N(t)-N(s)}(\cdot)$ der Zuwächse $N(t) - N(s)$, $0 \leq s < t < \infty$ gleich

$$\varphi_{s,t}(z) = \varphi_{\text{Pois}(\Lambda((s,t]))}(z) = e^{\Lambda((s,t])(e^{iz}-1)}, \quad z \in \mathbb{R}$$

nach 4. der Definition 6.2.1.
Wir zeigen, dass die Familie der charakteristischen Funktionen $\{\varphi_{s,t},\ 0 \leq s < t < \infty\}$ die Voraussetzung des Satzes 5.7.1 erfüllt: Für alle u so, dass $0 \leq s < u < t$ gilt

$$\begin{aligned}
\varphi_{s,u}(z)\varphi_{u,t}(z) &= e^{\Lambda((s,u])(e^{iz}-1)} e^{\Lambda((u,t])(e^{iz}-1)} \\
&= e^{(\Lambda((s,u])+\Lambda((u,t]))(e^{iz}-1)} \\
&= e^{\Lambda((s,t])(e^{iz}-1)} = \varphi_{s,t}(z), \quad z \in \mathbb{R},
\end{aligned}$$

da das Maß Λ additiv ist. Daher folgt die Existenz des Poisson-Prozesses N aus dem Satz 5.7.1. $\square$

Bemerkung 6.2.5

1. Die Existenz eines Poisson-Zählmaßes kann mit Hilfe des Satzes von Kolmogorow (vgl. Satz 5.1.16) bewiesen werden (in seiner allgemeineren Version).

2. Aus den Eigenschaften der Poisson-Verteilung folgt

   $$\mathbb{E}N(B) = \text{Var } N(B) = \Lambda(B), \quad B \in \mathcal{B}(\mathbb{R}_+).$$

   Daher kann $\Lambda(B)$ als die durchschnittliche Anzahl der Punkten von N in der Menge B, $B \in \mathcal{B}(\mathbb{R}_+)$ interpretiert werden. Wir zeigen bald, dass in diesem Fall N ein homogener Poisson-Prozess mit Intensität λ ist. Zur Erinnerung: In Abschnitt 5.2 war der homogene Poisson-Prozess definiert als ein Erneuerungsprozess mit u.i.v. Zwischenankunftszeiten $T_N \sim \text{Exp}(\lambda)$: $N(t) = \sup\{n \in \mathbb{N} \ : S_n \leq t\}$, $S_n = T_1 + \ldots + T_n$, $n \in \mathbb{N}$, $t \geq 0$.

Übungsaufgabe 6.2.6 Zeigen Sie, dass ein homogener Poisson-Prozess ein homogener Erneuerungsprozess mit $T_1 \stackrel{d}{=} T_2 \sim \text{Exp}(\lambda)$ ist.[1]

[1]Hinweis: Man muss zeigen, dass für eine beliebige Exponential-verteilte ZV X die integrierte Tailverteilungsfunktion von X gleich F_X ist.

Satz 6.2.7 Sei $N = \{N(t),\ t \geq 0\}$ ein Zählprozess.
Die folgenden Aussagen sind äquivalent:

1. N ist eine homogener Poisson-Prozess mit Intensität $\lambda > 0$.
2. (a) $N(t) \sim \text{Pois}(\lambda t)$, $t \geq 0$
   (b) für ein beliebiges $n \in \mathbb{N}$, $t \geq 0$, gilt, dass der Zufallsvektor $(S_1, \ldots, S_n)$ unter der Bedingung $\{N(t) = n\}$ die selbe Verteilung wie ein Vektor der Ordnungsstatistiken von u.i.v. ZVn $U_i \in \mathcal{U}([0,t])$, $i = 1, \ldots, n$ hat.
3. (a) N hat unabhängige Zuwächse,
   (b) $\mathbb{E}N(1) = \lambda$, und
   (c) Eigenschaft 2b) gilt.
4. (a) N hat stationäre und unabhängige Zuwächse,
   (b) es gilt $P(N(t) = 0) = 1 - \lambda t + o(t)$, $P(N(t) = 1) = \lambda t + o(t)$, $t \downarrow 0$.
5. (a) N hat stationäre und unabhängige Zuwächse,
   (b) Eigenschaft 2a) gilt.

Bemerkung 6.2.8

1. Nach Satz 6.2.7.5, ist Definition 6.2.1 mit $\Lambda(dx) = \lambda dx$, $\lambda \in (0, \infty)$ offensichtlich eine äquivalente Definition homogener Poisson-Prozesse.
2. Der homogene Poisson-Prozess wurde Anfang des 20. Jh. von den Physikern A. Einstein and M. Smoluchowski eingeführt, um den Zählprozess von Elementarteilchen in Geigerzählern zu modellieren.
3. Aus 6.2.7.4b) folgt $P(N(t) > 1) = o(t)$, $t \downarrow 0$.
4. Die Intensität von N kann wie folgt interpretiert werden: $\lambda = \mathbb{E}N(1) = \frac{1}{\mathbb{E}T_n}$, also die durchschnittliche Anzahl von Erneuerungen von N innerhalb eines Zeitintervalls der Länge 1.
5. Die Erneuerungsfunktion eines homogenen Poisson-Prozesses ist $H(t) = \lambda t$, $t \geq 0$. Dabei gilt $H(t) = \Lambda([0,t])$, $t > 0$ im Falle des nicht-homogenen Poisson-Prozesses.

Beweis Schema: 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 5) $\Rightarrow$ 1).
1) $\Rightarrow$ 2):

Aus 1) folgt $S_n = \sum_{k=1}^n T_k \sim \mathrm{Erl}(n,\lambda)$, da $T_k \sim \mathrm{Exp}(\lambda)$, $n \in \mathbb{N}$ und daher $P(N(t)=0) = P(T_1 > t) = e^{-\lambda t}$, $t \geq 0$, und für $n \in \mathbb{N}$

$$\begin{aligned}
P(N(t)=n) &= P(\{N(t) \geq n\} \setminus \{N(t) \geq n+1\}) \\
&= P(N(t) \geq n) - P(N(t) \geq n+1) \\
&= P(S_n \leq t) - P(S_{n+1} \leq t) \\
&= \int_0^t \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} dx - \int_0^t \frac{\lambda^{n+1} x^n}{n!} e^{-\lambda x} dx \\
&= \int_0^t \frac{d}{dx}\left(\frac{(\lambda x)^n}{n!} e^{-\lambda x}\right) dx = \frac{(\lambda t)^n}{n!} e^{-\lambda t},\ t \geq 0.
\end{aligned}$$

Damit ist 2a) gezeigt.
Wir folgern 2b): Nach dem Transformationssatz für Zufallsvariablen (vgl. [20, Theorem 3.6.2]) folgt aus

$$\begin{cases}
S_1 &= T_1 \\
S_2 &= T_1 + T_2 \\
&\vdots \\
S_{n+1} &= T_1 + \ldots + T_{n+1}
\end{cases}$$

dass die Dichte $f_{(S_1,\ldots,S_{n+1})}$ von $(S_1,\ldots,S_{n+1})^\top$ ausgedrückt werden kann durch die Dichte von $(T_1,\ldots,T_{n+1})^\top$, $T_i \sim \mathrm{Exp}(\lambda)$,u.i.v.:

$$\begin{aligned}
f_{(S_1,\ldots,S_{n+1})}(t_1,\ldots,t_{n+1}) &= \prod_{k=1}^{n+1} f_{T_k}(t_k - t_{k-1}) \\
&= \prod_{k=1}^{n+1} \lambda e^{-\lambda(t_k - t_{k-1})} \\
&= \lambda^{n+1} e^{-\lambda t_{n+1}}
\end{aligned}$$

für beliebige $0 \leq t_1 \leq \ldots \leq t_{n+1}$, $t_0 = 0$. Für alle anderen $t_1,\ldots,t_{n+1}$ gilt

$f_{(S_1,\ldots,S_{n+1})}(t_1,\ldots,t_{n+1}) = 0$. Also gilt weiter

$$\begin{aligned}
& f_{(S_1,\ldots,S_n)}(t_1,\ldots,t_n|N(t)=n) \\
= & f_{(S_1,\ldots,S_n)}(t_1,\ldots,t_n|S_k \le t,\ k \le n,\ S_{n+1} > t) \\
= & \frac{\int_t^\infty f_{(S_1,\ldots,S_{n+1})}(t_1,\ldots,t_{n+1})dt_{n+1} I(0 \le t_1 \le t_2 \le \ldots \le t_n \le t)}{\int_0^t \int_0^t \ldots \int_0^t \int_t^\infty f_{(S_1,\ldots,S_{n+1})}(t_1,\ldots,t_{n+1}) ds_{n+1} ds_n \ldots ds_1} \\
= & \frac{\int_t^\infty \lambda^{n+1} e^{-\lambda t_{n+1}} dt_{n+1}}{\int_0^t \ldots \int_0^t \int_t^\infty \lambda^{n+1} e^{-\lambda s_{n+1}} I(0 \le s_1 \le \ldots \le s_n \le t) ds_{n+1} ds_n \ldots ds_1} \\
\times & I(0 \le t_1 \le t_2 \le \ldots \le t_n \le t) \\
= & \frac{n!}{t^n} I(0 \le t_1 \le t_2 \le \ldots \le t_n \le t),
\end{aligned}$$

da man induktiv zeigen kann, dass

$$\int_0^t \int_0^t \ldots \int_0^t \mathbb{I}(0 \le s_1 \le \ldots \le s_n \le t) ds_1 \ldots ds_n = \frac{t^n}{n!}.$$

Dies ist genau die Dichte der Ordnungsstatistik von n u.i.v. $\mathcal{U}([0,t])$-Zufallsvariablen.

Übungsaufgabe 6.2.9 Zeige, dass das die Dichte der Ordnungsstatistik von n u.i.v. $\mathcal{U}([0,t])$-Zufallsvariablen ist.

2) $\Rightarrow$ 3)
Aus 2a) folgt offensichtlich 3b). Wir müssen lediglich die Unabhängigkeit der Zuwächse von N beweisen. Für ein beliebiges $n \in \mathbb{N}$, $x_1,\ldots,x_n \in \mathbb{N}$, $t_0 = 0 < t_1 < \ldots < t_n$ für $x = x_1 + \ldots + x_n$ gilt

$$\begin{aligned}
& P(\cap_{k=1}^n \{N(t_k) - N(t_{k-1}) = x_k\}) \\
= & \underbrace{P(\cap_{k=1}^n \{N(t_k) - N(t_{k-1}) = x_k\}|N(t_n)=x)}_{\frac{x!}{x_1!\ldots x_n!}\prod_{k=1}^n \left(\frac{t_k - t_{k-1}}{t_n}\right)^{x_k} \text{ nach 2b)}} \times \underbrace{P(N(t_n)=x)}_{e^{-\lambda t_n}\frac{(\lambda t_n)^x}{x!} \text{ nach 2a)}} \\
= & \prod_{k=1}^n \frac{(\lambda(t_k - t_{k-1}))^{x_k}}{x_k!} e^{-\lambda(t_k - t_{k-1})}, \qquad (6.6)
\end{aligned}$$

wobei die zweite Wahrscheinlichkeit in der zweiten Zeile die Polynomialverteilung mit Parametern n, $\left\{\frac{t_k - t_{k-1}}{t_n}\right\}_{k=1}^n$ ist. Dieses Ereignis bedeutet, dass x unabhängige gleich-verteilte Punkte in $[0,t]$, auf n Körbe so verteilt sind, dass genau x_k Punkte in den Korb der Länge $t_k - t_{k-1}$, $k = 1,\ldots,n$ fallen, vgl. Zeichnung. Also ist 3a) gezeigt, denn

$$P(\cap_{k=1}^n \{N(t_k) - N(t_{k-1}) = x_k\}) = \prod_{k=1}^n P(\{N(t_k) - N(t_{k-1}) = x_k\}).$$

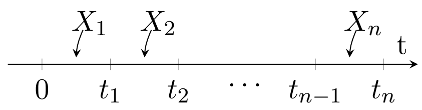


Abbildung 6.4: Skizze zu dem Fall 2) $\Rightarrow$ 3).

3) $\Rightarrow$ 4)
Wir zeigen, dass N stationäre Zuwächse hat. Für ein beliebiges $n \in \mathbb{N}_0$, $x_1, \ldots, x_n \in \mathbb{N}$, $t_0 = 0 < t_1 < \ldots < t_n$ und $h > 0$ betrachten wir

$$I(h) = P(\cap_{k=1}^n \{N(t_k + h) - N(t_{k-1} + h) = x_k\})$$

und zeigen, dass $I(h)$ nicht von $h \in \mathbb{R}$ abhängt. Nach der Formel (6.6) gilt

$$\begin{aligned} I(h) &= \sum_{m=0}^{\infty} \frac{(m+x)!}{m! x_1! \ldots x_n!} \prod_{k=1}^{n} \left(\frac{t_k + h - t_{k-1} - h}{t_n + h} \right)^{x_k} \left(\frac{h}{t_n + h} \right)^m \\ &\times P(N(t_n + h) = m + x) \\ &= \sum_{m=0}^{\infty} P\left(\cap_{k=1}^n \{N(t_k) - N(t_{k-1}) = x_k\} | N(t_n + h) = m + x\right) \\ &\times P\left(N(t_n + h) = m + x\right) \\ &= I(0) \end{aligned}$$

für alle $h > 0$. Nun gilt für die Eigenschaft 4b) für $h \in (0,1)$:

$$\begin{aligned} P(N(h) = 0) &= \sum_{k=0}^{\infty} P(N(h) = 0, N(1) = k) \\ &= \sum_{k=0}^{\infty} P(N(h) = 0, N(1) - N(h) = k) \\ &= \sum_{k=0}^{\infty} P(N(1) - N(h) = k, N(1) = k) \\ &= \sum_{k=0}^{\infty} P(N(1) = k) P(N(1) - N(h) = k \mid N(1) = k) \\ &= \sum_{k=0}^{\infty} P(N(1) = k)(1-h)^k. \end{aligned}$$

Insbesondere gilt, dass $P(N(h) = 0) = 1 - \lambda h + o(h)$, d.h.,

$$\lim_{h \to 0} \frac{1}{h}(1 - P(N(h) = 0)) = \lambda,$$

denn

$$\begin{aligned}
\frac{1}{h}\left(1-P(N(h)=0)\right) &= \frac{1}{h}\left(1-\sum_{k=0}^{\infty}P(N(1)=k)(1-h)^k\right)\\
&= \sum_{k=1}^{\infty}P(N(1)=k)\cdot\frac{1-(1-h)^k}{h}\\
&\xrightarrow[h\to 0]{} \sum_{k=1}^{\infty}P(N(1)=k)\underbrace{\lim_{h\to 0}\frac{1-(1-h)^k}{h}}_{k}\\
&= \sum_{k=0}^{\infty}P(N(1)=k)k=\mathbb{E}N(1)=\lambda,
\end{aligned}$$

da die Reihe gleichmäßig in h konvergiert, weil sie dominiert wird durch $\sum_{k=0}^{\infty}P(N(1)=k)k=\lambda<\infty$, was aus der Ungleichung $(1-h)^k\geq 1-kh$, $h\in(0,1)$, $k\in\mathbb{N}$ folgt. Ebenso kann man zeigen, dass

$$\lim_{h\to 0}\frac{P(N(h)=1)}{h}=\lim_{h\to 0}\sum_{k=1}^{\infty}P(N(1)=k)k(1-h)^{k-1}=\lambda.$$

4) $\Rightarrow$ 5)
Wir müssen zeigen, dass für ein beliebiges $n\in\mathbb{N}$ und $t\geq 0$

$$p_n(t)=P(N(t)=n)=e^{-\lambda t}\frac{(\lambda t)^n}{n!} \tag{6.7}$$

gilt. Wir zeigen dies per Induktion bzgl. n.
Zuerst zeigen wir, dass $p_0(t)=e^{-\lambda t}$, $n=0$. Dazu betrachte

$$\begin{aligned}
p_0(t+h) &= P(N(t+h)=0)\\
&= P(N(t)=0, N(t+h)-N(t)=0)\\
&= p_0(t)p_0(h)\\
&= p_0(t)(1-\lambda h+o(h)),\quad h\to +0.
\end{aligned}$$

Ähnlich zeigt man, dass

$$p_0(t)=p_0(t-h)(1-\lambda h+o(h)),\quad h\to +0.$$

Daher gilt

$$p_0'(t)=\lim_{h\to 0}\frac{p_0(t+h)-p_0(t)}{h}=-\lambda p_0(t),\quad t>0.$$

Aus $p_0(0)=P(N(0)=0)=1$ folgt

$$\begin{cases} p_0'(t) &= -\lambda p_0(t),\\ p_0(0) &= 1,\end{cases}$$

so dass eine eindeutige Lösung $p_0(t) = e^{-\lambda t}$, $t \geq 0$ existiert. Formel (6.7) gelte nun für n. Wir zeigen dies nun für $n+1$:

$$\begin{aligned}
p_{n+1}(t+h) &= P(N(t+h) = n+1) \\
&= P(N(t) = n, N(t+h) - N(t) = 1) \\
&+ P(N(t) = n+1, N(t+h) - N(t) = 0) \\
&+ \sum_{k=0}^{n-1} P(N(t) = k, N(t+h) - N(t) = n+1-k) \\
&\overset{\text{Bemerkung 6.2.8.3}}{=} p_n(t) \cdot p_1(h) + p_{n+1}(t) \cdot p_0(h) + o(h) \\
&= p_n(t)(\lambda h + o(h)) \\
&+ p_{n+1}(t)(1 - \lambda h + o(h)) + o(h), \;\; h \to +0.
\end{aligned}$$

Daraus folgt

$$\begin{cases} p'_{n+1}(t) = -\lambda p_{n+1}(t) + \lambda p_n(t), \;\; t > 0, \\ p_{n+1}(0) = 0. \end{cases} \tag{6.8}$$

Da $p_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$ gilt, sehen wir, dass $p_{n+1}(t) = e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!}$ eine Lösung von (6.8) ist. Tatsächlich gilt

- $p_{n+1}(t) = C(t) e^{-\lambda t} \Rightarrow C'(t) e^{-\lambda t} = \lambda C(t) e^{-\lambda t} - \lambda C(t) e^{-\lambda t} + \lambda p_n(t)$,
- $C'(t) = \frac{\lambda^{n+1} t^n}{n!} \Rightarrow C(t) = \frac{\lambda^{n+1} t^{n+1}}{(n+1)!}, \quad C(0) = 0.$

5) $\Rightarrow$ 1)
Sei N ein Zählprozess $N(t) = \max\{n : S_n \leq t\}$, $t \geq 0$, der die Bedingungen 5a) and 5b) erfüllt. Wir zeigen, dass $S_n = \sum_{k=1}^n T_k$ mit u.i.v. T_k und $T_k \sim \text{Exp}(\lambda)$, $k \in \mathbb{N}$. Da $T_k = S_k - S_{k-1}$, $k \in \mathbb{N}$, $S_0 = 0$, betrachten wir für $b_0 = 0 \leq a_1 < b_1 \leq \ldots \leq a_n < b_n$

$$\begin{aligned}
&P\Big(\bigcap_{k=1}^n \{a_k < S_k \leq b_k\} \Big) \\
&= P\Big(\bigcap_{k=1}^{n-1} \{N(a_k) - N(b_{k-1}) = 0, N(b_k) - N(a_k) = 1\} \\
&\qquad \bigcap \{N(a_n) - N(b_{n-1}) = 0, N(b_n) - N(a_n) \geq 1\} \Big)
\end{aligned}$$

$$
\begin{aligned}
&= \prod_{k=1}^{n-1} (\underbrace{P(N(a_k - b_{k-1}) = 0)}_{e^{-\lambda(a_k - b_{k-1})}} \underbrace{P(N(b_k - a_k) = 1)}_{\lambda(b_k - a_k)e^{-\lambda(b_k - a_k)}}) \\
&\qquad \times \underbrace{P(N(a_n - b_{n-1}) = 0)}_{e^{-\lambda(a_n - b_{n-1})}} \underbrace{P(N(b_n - a_n) \geq 1)}_{(1 - e^{-\lambda(b_n - a_n)})} \\
&= e^{-\lambda(a_n - b_{n-1})}(1 - e^{-\lambda(b_n - a_n)}) \prod_{k=1}^{n-1} \lambda(b_k - a_k) e^{-\lambda(b_k - b_{k-1})} \\
&= \lambda^{n-1}(e^{-\lambda a_n} - e^{-\lambda b_n}) \prod_{k=1}^{n-1} (b_k - a_k) \\
&= \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} \lambda^n e^{-\lambda y_n} dy_n \ldots y_1.
\end{aligned}
$$

Die gemeinsame Dichte von $(S_1, \ldots, S_n)^\top$ ist deshalb gegeben durch $\lambda^n e^{-\lambda y_n} I(0 \leq y_1 \leq y_2 \leq \ldots \leq y_n)$.
Wie in 1) $\Rightarrow$ 2), Teil 2,b) kann man mit Hilfe der Dichtetransformationsformel zeigen, dass $(T_1, \ldots, T_n)^\top$ die Dichte $\prod_{k=1}^n \lambda e^{-\lambda t_k}$ hat. $\square$

Definition 6.2.10 Sei $N = \{N(t),\ t \geq 0\}$ eine homogener Poisson-Prozess mit Intensität $\lambda > 0$ konstruiert durch die Folge von Zwischenankufts-zeiten $\{T_n\}_{n \in \mathbb{N}}$. Sei $\{U_n\}_{n \in \mathbb{N}}$ eine Folge von u.i.v. ZVn, undabhängig von $\{T_n\}_{n \in \mathbb{N}}$ und F_U die Verteilungsfunktion von U_1. Für ein beliebiges $t \geq 0$ sei $X(t) = I(N(t) > 0) \sum_{k=1}^{N(t)} U_k$. $X = \{X(t),\ t \geq 0\}$ heißt *zusammengesetzter Poisson-Prozess mit Parametern λ, F_U*. Die Verteilung von $X(t)$ heißt *zusammengesetzte Poisson-Verteilung mit Parametern λt, F_U*.

Bemerkung 6.2.11

1. Der zusammengesetzte Poisson-Prozess kann wie folgt interpretiert werden:
   $X(t) = \sum$ „Marken“ U_n des homogenen Poisson-Prozesses (N, U) bis Zeitpunkt t.

2. Klassische Anwendungen der zusammengesetzten Poisson-Prozesse sind
   - Warteschlangentheorie: $X(t) =$ Gesamtarbeitsbelastung eines Servers bis zum Zeitpunkt t, falls Aufträge zum Zeitpunkt $S_n = \sum_{k=1}^n T_k$, $n \in \mathbb{N}$ ankommen und U_n, $n \in \mathbb{N}$ Zeit verbrauchen.
   - Versicherungsmathematik: $X(t) =$ Gesamtschaden eines Portfolios bis zum Zeitpunkt $t \geq 0$ mit
     $N(t)=$ # Schäden und $U_n =$ jeweilige Schadenshöhe.

Satz 6.2.12 Sei $X = \{X(t),\ t \geq 0\}$ ein zusammengesetzter Poisson-Prozess mit Parametern λ, F_U. Es gilt:

1. X hat unabhängige und stationäre Zuwächse.
2. Falls $\hat{m}_U(s) = \mathbb{E}e^{sU_1}$, $s \in \mathbb{R}$ die momentenerzeugende Funktion von U_1 ist, s.d. $\hat{m}_U(s) < \infty$, $s \in \mathbb{R}$, dann gilt

$$\hat{m}_{X(t)}(s) = e^{\lambda t(\hat{m}_U(s)-1)},\ s \in \mathbb{R},\ t \geq 0,$$

$$\mathbb{E}X(t) = \lambda t \mathbb{E}U_1,\ t \geq 0,$$

$$\text{Var } X(t) = \lambda t \mathbb{E}U_1^2,\ t \geq 0.$$

Beweis 1. Z.z. Für beliebiges $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \ldots < t_n$ und $h \geq 0$ gilt

$$P\left(\sum_{i_1=N(t_0+h)+1}^{N(t_1+h)} U_{i_1} \leq x_1, \ldots, \sum_{i_n=N(t_{n-1}+h)+1}^{N(t_n+h)} U_{i_n} \leq x_n\right)$$
$$= \prod_{k=1}^{n} P\left(\sum_{i_k=N(t_{k-1})+1}^{N(t_k)} U_{i_k} \leq x_k\right)$$

für beliebige $x_1, \ldots, x_n \in \mathbb{R}$. Es gilt

$$P\left(\sum_{i_1=N(t_0+h)+1}^{N(t_1+h)} U_{i_1} \leq x_1, \ldots, \sum_{i_n=N(t_{n-1}+h)+1}^{N(t_n+h)} U_{i_n} \leq x_n\right)$$
$$= \sum_{k_1,\ldots,k_n=0}^{\infty} \left(\prod_{j=1}^{n} F_U^{*k_j}(x_j)\right) P\left(\bigcap_{m=1}^{n} \{N(t_m+h) - N(t_{m-1}+h) = k_m\}\right)$$
$$= \sum_{k_1,\ldots,k_n=0}^{\infty} \left(\prod_{j=1}^{n} F_U^{*k_j}(x_j)\right) \left(\prod_{m=1}^{n} P(N(t_m) - N(t_{m-1}) = k_m)\right)$$
$$= \prod_{m=1}^{n} \sum_{k_m=0}^{\infty} F_U^{*k_m}(x_m) P(N(t_m) - N(t_{m-1}) = k_m)$$
$$= \prod_{m=1}^{n} P\left(\sum_{k_m=N(t_{m-1})+1}^{N(t_m)} U_{k_m} \leq x_m\right).$$

2. Übungsaufgabe.

□

Ein Cox-Prozess ist ein (i.A. inhomogener) Poisson-Prozess mit zufälligem Intensitäts-Maß Λ.

Definition 6.2.13 Sei $\Lambda = \{\Lambda(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$ ein f.s. lokal endliches Zufallsmaß. Das zufällige Zählmaß $N = \{N(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$ heißt *Cox*

Zählmaß (oder doppelt-stochastisches Poisson-Maß) mit zufälligem Intensitätsmaß Λ, falls für beliebige $n \in \mathbb{N}$, $k_1, \ldots, k_n \in \mathbb{N}_0$ und $0 \leq a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_n < b_n$ gilt, dass

$$P(\cap_{i=1}^n \{N((a_i, b_i]) = k_i\}) = \mathbb{E}\left(\prod_{i=1}^n e^{-\Lambda((a_i,b_i])} \frac{\Lambda^{k_i}((a_i, b_i])}{k_i!}\right).$$

Der Prozess $\{N(t),\ t \geq 0\}$ mit $N(t) = N((0,t])$ heißt *Cox-Prozess* (oder *doppelt-stochastischer Poisson-Prozess*) mit zufälligem Intensitätsmaß Λ.

Beispiel 6.2.14

1. Falls das zufällige Maß Λ f.s. absolutstetig ist bzgl. des Lebesgue-Maßes, d.h., $\Lambda(B) = \int_B \lambda(t)dt$, B - beschränkt, $B \in \mathcal{B}(\mathbb{R}_+)$, wobei $\{\lambda(t), t \geq 0\}$ ein stochastischer Prozess mit f.s. Borel-messbaren, Lebesgue-integrierbaren Pfaden ist, $\lambda(t) \geq 0$ f.s. für alle $t \geq 0$, dann heißt $\{\lambda(t), t \geq 0\}$ *Intensitäts-Prozess* von N.

2. Es kann sein, dass $\lambda(t) \equiv Y$ mit einer nicht-negativen ZV Y. Dann gilt $\Lambda(B) = Y\nu_1(B)$, also hat N eine zufällige Intensität Y. Solche Cox-Prozesse heißen *gemischte Poisson-Prozesse.*

Ein Cox-Prozess $N = \{N(t),\ t \geq 0\}$ mit Intensitäts-Prozess $\{\lambda(t),\ t \geq 0\}$ kann wie folgt konstruiert und simuliert werden:

1. Sei $\tilde{N} = \{\tilde{N}(t),\ t \geq 0\}$ ein homogener Poisson-Prozess mit Intensität 1 unabhängig von $\{\lambda(t),\ t \geq 0\}$.

2. Es gilt $N \stackrel{d}{=} N_1$, wobei der Prozess $N_1 = \{N_1(t),\ t \geq 0\}$ durch $N_1(t) = \tilde{N}(\int_0^t \lambda(y)dy)$, $t \geq 0$, gegeben ist (die Behauptung $N \stackrel{d}{=} N_1$ muss bewiesen werden, was wir hier aber nicht tun werden).

Kapitel 7

Wiener-Prozess

Die korrekte Erklärung der Natur chaotischer Bewegungen von winzigen Partikeln in Flüssigkeiten oder Gasen durch die atomare Struktur der Masse, wurde bereits von dem römischen Philosophen Lucretius 60 v.Chr. in seinem Buch „Über die Natur der Dinge” gegeben: „Beobachte was passiert, wenn Sonnenstrahlen in ein Gebäude eintreten und Licht auf schattige Oberflächen strahlt. Zu sehen, ist eine Vielzahl von Partikeln, welche sich in der Luft bewegen... Ihre Bewegungen sind ein Indikator der unterliegenden Bewegung der Materie, welche uns verbogen bleibt... Der Ursprung findet sich in den Atomen, welche sich selbst (d.h. zufällig) Bewegen. Diese kleinen zusammengesetzten Körper, welche am wenigsten von dem Impetus der Atome entfernt sind, werden durch den Aufprall ihrer unsichtbaren Stöße in Bewegung gesetzt und prallen gegen die minimal größeren Körper. Diese Bewegungen addieren sich durch die Atome schrittweise bis zu dem Punkt auf, an welchem auch wir die Bewegungen im Sonnenlicht wahrnehmen können.”
Ähnliche Bewegungen von Körnern auf der Wasseroberfläche, wurden auch von dem Robert Brown beobachtet. Dieser erklärte die Bewegung allerdings durch eine bestimmte „Lebenskraft” in den Körnern. Aufgrund seines Beitrages erhielt der stochastische Prozess auch den Name „Brownsche Bewegung” (engl. ”Brownian Motion”). Ihren mathematischen Ursprung findet die Brownsche Bewegung im 20 Jh. durch Louis Bachelier (1900), Marian Smoluchowski und Albert Einstein (1905). Einstein zeigte, dass falls $(u(x,t))$ die Dichte (Anzahl der Partikel per Volumeneinheit) an einem Punkt x zum Zeitpunkt t ist, diese Funktion der Wärmeleitungsgleichung $u_t'(x,t) = \frac{1}{2}u_x''(x,t)$ mit Lösung $u(x,t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2}\right)$ $x \in \mathbb{R}$, $t > 0$ genügt. u ist insbesondere die Dichte einer $\mathcal{N}(0,t)$ verteilten Zufallsvariable. Einsteins Erklärung wurde durch Experimente 1908 bestätigt, welche einen Beweis für die atomare Struktur der Natur bewiesen.
Der französische Physiker Perrin beschrieb die Pfade der Brownschen Bewegung als natürliche Funktionen, welche zwar stetig aber nirgends differenzierbar sind. Der amerikanische Mathematiker Norbert Wiener (1928)

fasste diesen Gedankengang auf und betrachtete daraufhin die Pfade eines einzigen Korns, im Vergleich zum vorherigen Modell, welches die gesamte Ansammlung der Körner betrachtete. Wiener formulierte daraufhin die mathematische Grundlage der Brownschen Bewegung. Insbesondere definierte er diese als eine zufällige Funktion und bewies deren Existenz, was dazu führte, dass diese heute auch den Namen „Wiener-Prozess" besitzt.

7.1 Elementare Eigenschaften

In Abschnitt 5.2 haben wir die *Brownsche Bewegung* (oder *Wiener-Prozess*) $W = \{W(t),\ t \geq 0\}$ definiert als einen Gaußschen Prozess mit $\mathbb{E}W(t) = 0$ und $\operatorname{Cov}(W(s), W(t)) = \min\{s,t\}$, $s,t \geq 0$.
Frage: Wieso existiert die Brownsche Bewegung?
Nach dem Satz 5.1.16 existiert ein reell-wertiger Gaußscher Prozess $X = \{X(t),\ t \geq 0\}$ mit Erwartungswert $\mathbb{E}X(t) = \mu(t)$, $t \geq 0$ und Kovarianzfunktion $\operatorname{Cov}(X(s), X(t)) = C(s,t)$, $s,t \geq 0$ für jede Funktion $\mu : \mathbb{R}_+ \to \mathbb{R}$ und jede positiv semidefinite Funktion $C : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$.

Übungsaufgabe 7.1.1 Zeige, dass $C(s,t) = \min\{s,t\}$, $s,t \geq 0$ positiv semidefinit ist.

Wir geben nun eine neue (äquivalente) Definition:

Definition 7.1.2 Ein stochastischer Prozess $W = \{W(t),\ t \geq 0\}$ heißt *Wiener-Prozess* (oder *Brownsche Bewegung*), falls

1. $W(0) = 0$ f.s.,
2. W hat unabhängige Zuwächse,
3. $W(t) - W(s) \sim \mathcal{N}(0, t-s)$, $0 \leq s < t$.

Die Existenz von W (nach der Definition 7.1.2) folgt aus Satz 5.7.1, da

$$\varphi_{s,t}(z) = \mathbb{E}e^{iz(W(t)-W(s))} = e^{-\frac{(t-s)z^2}{2}}, \quad z \in \mathbb{R}$$

und

$$e^{-\frac{(t-u)z^2}{2}} e^{-\frac{(u-s)z^2}{2}} = e^{-\frac{(t-s)z^2}{2}}$$

für $0 \leq s < u < t$ und daher

$$\varphi_{s,u}(z)\varphi_{u,t}(z) = \varphi_{s,t}(z), \quad z \in \mathbb{R}.$$

Aus dem Satz 5.3.16 folgt die Existenz einer Modifikation mit stetigen Pfaden.

Übungsaufgabe 7.1.3 Zeige, dass Satz 5.3.16 für $\alpha = 3$, $\delta = \frac{1}{2}$ gilt.

Satz 7.1.4 Die beiden Definitionen des Wiener-Prozesses sind äquivalent.

Beweis

1. Definition 7.1.2 $\Leftarrow$ Definition in Abschnitt 5.2:
   $W(0) = 0$ f.s. folgt aus Var $(W(0)) = \min\{0, 0\} = 0$.
   Wir zeigen, dass die Zuwächse von W unabhängig sind:
   Falls $Y \sim \mathcal{N}(\mu, K)$ ein n-dimensionaler Gaußscher Zufallsvektor ist und A eine $(n \times n)$-Matrix, dann gilt

$$AY \sim \mathcal{N}(A\mu, AKA^\top).$$

   Dies folgt aus der expliziten Form der charakteristischen Funktion von Y. Nun sei $n \in \mathbb{N}$, $0 = t_0 \leq t_1 < \ldots < t_n$, $Y = (W(t_0), W(t_1), \ldots, W(t_n))^\top$.
   Für $Z = (W(t_0), W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1}))^\top$ gilt $Z = AY$, wobei

$$A = \begin{pmatrix} 1 & 0 & 0 & \ldots & \ldots & 0 \\ -1 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & -1 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{pmatrix}.$$

   Daher ist auch Z Gaußsch mit einer diagonalen Kovarianzmatrix. Es gilt

$$\begin{aligned} & \text{Cov}\ (W(t_{i+1}) - W(t_i), W(t_{j+1}) - W(t_j)) \\ = \ & \min\{t_{i+1}, t_{j+1}\} - \min\{t_{i+1}, t_j\} - \min\{t_i, t_{j+1}\} + \min\{t_i, t_j\} \\ = \ & 0 \end{aligned}$$

   für $i \neq j$. Daher sind die Koordinaten von Z unkorreliert, was äquivalent zur Unabhängigkeit im Falle von multivariater Normalerverteilung ist.
   Daher sind die Zuwächse von W unabhängig und es gilt für beliebige $0 \leq s < t$

$$W(t) - W(s) \sim \mathcal{N}(0, t - s).$$

   Die Normalverteilung folgt, da $Z = AY$ Gaußsch ist.
   Offensichtlich gilt $\mathbb{E}W(t) - \mathbb{E}W(s) = 0$ und

$$\begin{aligned} & \text{Var}\ (W(t) - W(s)) \\ = \ & \text{Var}\ (W(t)) - 2\text{Cov}\ (W(s), W(t)) + \text{Var}\ (W(s)) \\ = \ & t - 2\min\{s, t\} + s \\ = \ & t - s. \end{aligned}$$

2. Definition 7.1.2 $\Rightarrow$ Definition in Abschnitt 5.2:
Da $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ für $0 \le s < t$, gilt

$$\begin{aligned} & \text{Cov}\ (W(s), W(t)) = \mathbb{E}[W(s)(W(t) - W(s) + W(s))] \\ = \ & \mathbb{E}W(s)\mathbb{E}(W(t) - W(s)) + \text{Var}\ W(s) \\ = \ & s, \end{aligned}$$

und daher gilt auch Cov $(W(s), W(t)) = \min\{s, t\}$.
Aus $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ und $W(0) = 0$ folgt $\mathbb{E}W(t) = 0$, $t \ge 0$.
Dass W ein Gaußscher Prozess ist, folgt aus Teil 1) des Beweises: $Y = A^{-1}Z$.

$\square$

Definition 7.1.5 Der Prozess $\{W(t),\ t \ge 0\}$, $W(t) = (W_1(t), \ldots, W_d(t))^\top$, $t \ge 0$ heißt *d-dimensionale Brownsche Bewegung*, falls $W_i = \{W_i(t),\ t \ge 0\}$ unabhängige Wiener-Prozesse sind für $i = 1, \ldots, d$.

Definition 7.1.5 und Übungsaufgabe 7.1.3 garantieren die Existenz eines Wiener-Prozesses mit stetigen Pfaden.
Frage: Wie konstruieren wir diese Pfade?
Um die Antwort dieser Frage geht es im nächsten Abschnitt.

7.2 Explizite Konstruktion des Wiener-Prozess

Wir konstruieren den Wiener-Prozess zunächst auf dem Intervall $[0, 1]$.
Idee: Führe eine stochastischen Prozess $X = \{X(t),\ t \in [0,1]\}$ definiert auf einem W-Unterraum $(\Omega, \mathcal{A}, P)$ von $(\Omega, \mathcal{F}, P)$ mit $X \stackrel{d}{=} W$ ein, wobei $X(t) = \sum_{n=1}^{\infty} c_n(t) Y_n$, $t \in [0,1]$ und $\{Y_n\}_{n \in \mathbb{N}}$ eine Folge von u.i.v. $\mathcal{N}(0,1)$-ZVn ist und $c_n(t) = \int_0^t H_n(s)ds$, $t \in [0,1]$, $n \in \mathbb{N}$. Wir bezeichnen mit $\{H_n\}_{n \in \mathbb{N}}$ die orthonormale Haar-Basis in $L^2([0,1])$, die wir nun definieren:

Definition 7.2.1 Die Funktionen $H_n : [0,1] \to \mathbb{R}$, $n \in \mathbb{N}$ heißen *Haar-Funktionen*, falls

- $H_1(t) = 1, \qquad t \in [0,1]$,
- $H_2(t) = I_{[0,\frac{1}{2}]}(t) - I_{(\frac{1}{2},1]}(t), \qquad t \in [0,1]$,
- $H_k(t) = 2^{\frac{n}{2}}(I_{I_{n,k}}(t) - I_{J_{n,k}}(t)), \quad t \in [0,1]$, $2^n < k \le 2^{n+1}$,

wobei

- $I_{n,k} = [a_{n,k}, a_{n,k} + 2^{-n-1}]$,
- $J_{n,k} = (a_{n,k} + 2^{-n-1}, a_{n,k} + 2^{-n}]$,
- $a_{n,k} = 2^{-n}(k - 2^n - 1)$, $n \in \mathbb{N}$.

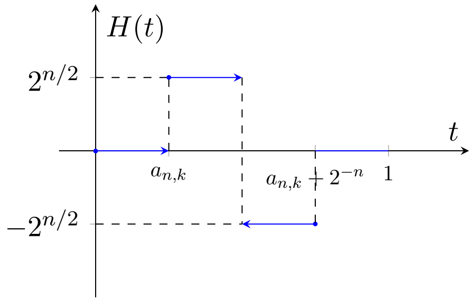


Abbildung 7.1: Haar-Funktionen

Lemma 7.2.2 Das Funktionensystem $\{H_n\}_{n\in\mathbb{N}}$ ist eine Orthonormalbasis des $L^2([0,1])$ mit Sklarprodukt

$$\langle f,g\rangle = \int_0^1 f(t)g(t)dt, \quad f,g\in L^2([0,1]).$$

Beweis Die Orthonormalität des Systems $\langle H_k,H_n\rangle = \delta_{kn},\ k,n\in\mathbb{N}$ folgt direkt aus der Defintion 7.2.1.
Wir beweisen nun die Vollständigkeit von $\{H_n\}_{n\in\mathbb{N}}$:
Es genügt zu zeigen, dass für eine beliebige Funktion $g\in L^2([0,1])$ mit $\langle g,H_n\rangle = 0$, $n\in\mathbb{N}$, fast überall auf $[0,1]$

$$g = 0$$

gilt. Wir können die Indikatorfunktion eines Intervalls $I_{[a_{n,k},a_{n,k}+2^{-n-1}]}$ schreiben als Linearkombination von H_n, $n\in\mathbb{N}$:

$$\begin{aligned}
I_{[0,\frac{1}{2}]} &= \frac{(H_1+H_2)}{2},\\
I_{(\frac{1}{2},1]} &= \frac{(H_1-H_2)}{2},\\
I_{[0,\frac{1}{4}]} &= \frac{(I_{[0,\frac{1}{2}]}+\frac{1}{\sqrt{2}}H_3)}{2}, \quad n=1, \quad k=3\\
I_{(\frac{1}{4},\frac{1}{2}]} &= \frac{(I_{[0,\frac{1}{2}]}-\frac{1}{\sqrt{2}}H_3)}{2}, \quad n=1, \quad k=3\\
&\vdots\\
I_{[a_{n,k},a_{n,k}+2^{-n-1}]} &= \frac{(I_{[a_{n,k},a_{n,k}+2^{-n}]}+2^{-\frac{n}{2}}H_k)}{2}, \quad 2^n<k\le 2^{n+1}.
\end{aligned}$$

Daher gilt

$$\int_{\frac{k}{2^n}}^{\frac{(k+1)}{2^n}} g(t)dt = 0, \quad n\in\mathbb{N}_0,\ k=0,\dots,2^n-1$$

und deswegen auch

$$G(t) = \int_0^t g(s)ds = 0, \quad t = \frac{k}{2^n},\, n \in \mathbb{N}_0,\, k = 0, \dots, 2^n - 1.$$

Da G stetig auf $[0,1]$ ist, folgt $G(t) = 0$, $t \in [0,1]$, und daher $g(s) = G'(s) = 0$ für fast alle $s \in [0,1]$. $\square$

Aus Lemma 7.2.2 folgt, dass zwei beliebige Funktionen $f, g \in L^2([0,1])$ dargestellt werden können als

$$f = \sum_{n=1}^{\infty} \langle f, H_n \rangle H_n \quad \text{und} \quad g = \sum_{n=1}^{\infty} \langle g, H_n \rangle H_n$$

(beide Reihen konvergieren in $L^2([0,1])$) und es gilt

$$\langle f, g \rangle = \sum_{n=1}^{\infty} \langle f, H_n \rangle \langle g, H_n \rangle$$

(Parseval–Identität).

Definition 7.2.3 Die Funktionen

$$S_n(t) = \int_0^t H_n(s)ds = \langle I_{[0,t]}, H_n \rangle, \quad t \in [0,1],\, n \in \mathbb{N}$$

heißen *Schauder-Funktionen*.

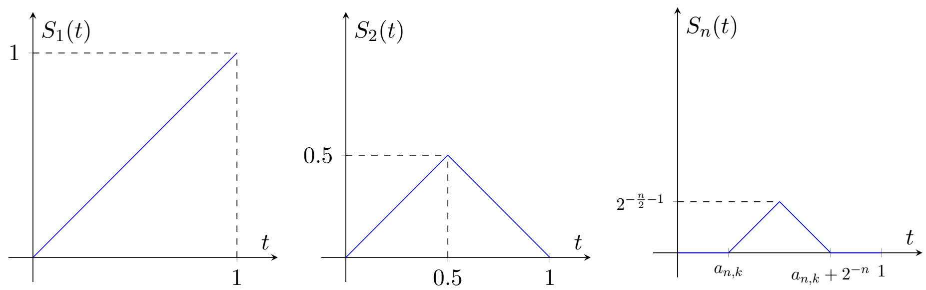


Abbildung 7.2: Schauder-Funktionen

Lemma 7.2.4 Es gilt:

1. $S_n(t) \geq 0$, $t \in [0,1]$, $n \in \mathbb{N}$,
2. $\sum_{k=1}^{2^n} S_{2^n+k}(t) \leq \frac{1}{2} 2^{-\frac{n}{2}}$, $t \in [0,1]$, $n \in \mathbb{N}$,
3. Sei $\{a_n\}_{n \in \mathbb{N}}$ eine Folge reeller Zahlen mit $a_n = O(n^{\varepsilon})$, $\varepsilon < \frac{1}{2}$, $n \to \infty$. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n S_n(t)$ absolut und gleichmäßig in $t \in [0,1]$ und ist daher eine stetige Funktion auf $[0,1]$.

Beweis

1. Folgt direkt aus der Definition 7.2.1.

2. Gilt, da die Funktionen S_{2^n+k} für $k = 1, \ldots, 2^n$ disjunkte Träger haben und
$$S_{2^n+k}(t) \leq S_{2^n+k}(a_{n,k} + 2^{-n-1}) = 2^{-\frac{n}{2}-1}, \quad t \in [0,1].$$

3. Es genügt zu zeigen, dass
$$R_n = \sup_{t\in[0,1]} \sum_{k>2^n} |a_k| S_k(t) \xrightarrow[n\to\infty]{} 0.$$
Für jedes $k \in \mathbb{N}$ und $c > 0$ gilt $|a_k| \leq ck^\varepsilon$. Daher gilt auch für alle $t \in [0,1]$ und $n \in \mathbb{N}$, dass
$$\begin{aligned} \sum_{2^n<k\leq 2^{n+1}} |a_k| S_k(t) &\leq c \cdot 2^{(n+1)\varepsilon} \cdot \sum_{2^n<k\leq 2^{n+1}} S_k(t) \\ \leq \quad c \cdot 2^{(n+1)\varepsilon} \cdot 2^{-\frac{n}{2}-1} &\leq c \cdot 2^{\varepsilon - n(\frac{1}{2}-\varepsilon)}. \end{aligned}$$
Da $\varepsilon < \frac{1}{2}$, folgt
$$R_m \leq c \cdot 2^\varepsilon \sum_{n\geq m} 2^{-n(\frac{1}{2}-\varepsilon)} \xrightarrow[m\to\infty]{} 0.$$

□

Lemma 7.2.5 Sei $\{Y_n\}_{n\in\mathbb{N}}$ eine Folge von (nicht notwendigerweise unabhängigen) Zufallsvariablen definiert auf $(\Omega, \mathcal{A}, P)$ mit $Y_n \sim \mathcal{N}(0,1)$, $n \in \mathbb{N}$. Dann gilt
$$|Y_n| = O((\log n)^{\frac{1}{2}}), \quad n \to \infty, \text{ f.s.}$$

Beweis Wir müssen zeigen, dass für $c > \sqrt{2}$ und fast alle $\omega \in \Omega$ ein $n_0 = n_0(\omega, c) \in \mathbb{N}$ existiert, s.d. $|Y_n| \leq c(\log n)^{\frac{1}{2}}$ für $n \geq n_0$. Falls $Y \sim \mathcal{N}(0,1)$, $x > 0$, dann gilt
$$\begin{aligned} \overline{\Phi}(x) := P(Y > x) &= \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_x^\infty \left(-\frac{1}{y}\right) d\left(e^{-\frac{y^2}{2}}\right) \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{x} e^{-\frac{x^2}{2}} - \int_x^\infty e^{-\frac{y^2}{2}} \frac{1}{y^2} dy\right) \\ &\leq \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}} \end{aligned}$$

(Wir können sogar zeigen, dass $\overline{\Phi}(x) \sim \frac{1}{\sqrt{2\pi}}\frac{1}{x}e^{-\frac{x^2}{2}}$, $x \to \infty$). Daher gilt für $c > \sqrt{2}$, dass

$$\begin{aligned}\sum_{n\geq 2} P(|Y_n| > c(\log n)^{\frac{1}{2}}) &\leq c^{-1}\frac{2}{\sqrt{2\pi}}\sum_{n\geq 2}(\log n)^{-\frac{1}{2}}e^{-\frac{c^2}{2}\log n}\\ &= \frac{c^{-1}\sqrt{2}}{\sqrt{\pi}}\sum_{n\geq 2}(\log n)^{-\frac{1}{2}}n^{-\frac{c^2}{2}} < \infty.\end{aligned}$$

Nach dem Lemma von Borel-Cantelli (vlg. [20, Lemma 2.2.1]) gilt

$$P\left(\bigcap_n \bigcup_{k\geq n} A_k\right) = 0,$$

falls $\sum_k P(A_k) < \infty$ mit

$$A_k = \{|Y_k| > c\cdot(\log k)^{\frac{1}{2}}\}, \quad k \in \mathbb{N}.$$

Daher tritt A_k in unendlicher Anzahl nur mit Wkt. 0 ein, d.h.

$$|Y_n| \leq c(\log n)^{\frac{1}{2}}$$

für $n \geq n_0$. □

Lemma 7.2.6 Sei $\{Y_n\}_{n\in\mathbb{N}}$ eine Folge von unabhängigen $\mathcal{N}(0,1)$-verteilten Zufallsvariablen. Seien $\{a_n\}_{n\in\mathbb{N}}$ und $\{b_n\}_{n\in\mathbb{N}}$ Folgen von Zahlen mit

$$\sum_{k=1}^{2^m}|a_{2^m+k}| \leq 2^{-\frac{m}{2}} \quad \text{und} \quad \sum_{k=1}^{2^m}|b_{2^m+k}| \leq 2^{-\frac{m}{2}}, \quad m \in \mathbb{N}.$$

Dann existieren $U = \sum_{n=1}^{\infty} a_n Y_n$ und $V = \sum_{n=1}^{\infty} b_n Y_n$ f.s., wobei

$$U \sim \mathcal{N}(0, \sum_{n=1}^{\infty} a_n^2) \text{ und } V \sim \mathcal{N}(0, \sum_{n=1}^{\infty} b_n^2),$$

mit Cov $(U,V) = \sum_{n=1}^{\infty} a_n b_n$. U und V sind außerdem unabhängig genau dann, wenn Cov $(U,V) = 0$.

Beweis Nach Lemma 7.2.4 und 7.2.5 existieren die Grenzwerte U und V f.s. (ersetze dazu a_n durch Y_n und S_n durch z.B. b_n in Lemma 7.2.4).
Aus der Faltungsstabilität der Normalverteilung folgt für $U^{(m)} = \sum_{n=1}^{m} a_n Y_n$ und $V^{(m)} = \sum_{n=1}^{m} b_n Y_n$, dass

$$U^{(m)} \sim \mathcal{N}(0, \sum_{n=1}^{m} a_n^2) \quad \text{und} \quad V^{(m)} \sim \mathcal{N}(0, \sum_{n=1}^{m} b_n^2).$$

Da $U^{(m)} \xrightarrow{d} U$ und $V^{(m)} \xrightarrow{d} V$, folgt $U \sim \mathcal{N}(0, \sum_{n=1}^{\infty} a_n^2)$ und $V \sim \mathcal{N}(0, \sum_{n=1}^{\infty} b_n^2)$. Es gilt sogar

$$\begin{aligned}
\text{Cov }(U,V) &= \mathbb{E}\left(\sum_{n,m=1}^{\infty} a_n b_m Y_n Y_m\right) = \sum_{n,m=1}^{\infty} a_n b_m \underbrace{\mathbb{E}(Y_n Y_m)}_{\delta_{nm}} \\
&= \sum_{n=1}^{\infty} a_n b_n,
\end{aligned}$$

nach dem Satz über die dominierte Konvergenz von Lebesgue, da nach Lemma 7.2.5 für $n \geq \mathbb{N}_0$

$$|Y_n| \leq c(\log n)^{\frac{1}{2}} \leq c \cdot cn^{\varepsilon}, \quad \varepsilon < \frac{1}{2}$$

gilt und die dominierende Reihe konvergiert nach Lemma 7.2.4:

$$\begin{aligned}
\sum_{n,k=2^m}^{2^{m+1}} a_n b_k Y_n Y_k \overset{f.s.}{\leq} \sum_{n,k=2^m}^{2^{m+1}} a_n b_k c^2 n^{\varepsilon} k^{\varepsilon} &\leq c^2 2^{2\varepsilon(m+1)} \cdot 2^{-\frac{m}{2}} \cdot 2^{-\frac{m}{2}} \\
&\leq 2c^2 2^{-(1-2\varepsilon)m}, \quad 1 - 2\varepsilon > 0.
\end{aligned}$$

Für hinreichend großes m gilt

$$\sum_{n,k=2^m}^{\infty} a_n b_k Y_n Y_k \leq 2c^2 \sum_{j=m}^{\infty} 2^{-(1-2\varepsilon)j} < \infty,$$

und diese Reihe konvergiert f.s. Nun zeigen wir

$$\text{Cov }(U,V) = 0 \iff U \text{ und } V \text{ sind unabhängig.}$$

Aus Unabhängigkeit folgt stets die Unkorreliertheit von Zufallsvariablen. Wir zeigen nun die Umkehrung:

Aus $(U^{(m)}, V^{(m)}) \xrightarrow[m\to\infty]{d} (U,V)$ folgt $\varphi_{(U^{(m)},V^{(m)})} \xrightarrow[m\to\infty]{} \varphi_{(U,V)}$ und daher

$$\begin{aligned}
\varphi_{(U,V)}(s,t) &= \lim_{m\to\infty} \mathbb{E}\exp\left\{ i\left(t\sum_{k=1}^m a_k Y_k + s\sum_{n=1}^m b_n Y_n \right)\right\} \\
&= \lim_{m\to\infty} \mathbb{E}\exp\left\{ i\sum_{k=1}^m (ta_k + sb_k) Y_k \right\} \\
&= \lim_{m\to\infty} \prod_{k=1}^m \mathbb{E}\exp\{i(ta_k+sb_k)Y_k\} \\
&= \lim_{m\to\infty} \prod_{k=1}^m \exp\left\{ -\frac{(ta_k+sb_k)^2}{2} \right\} \\
&= \exp\left\{ -\sum_{k=1}^\infty \frac{(ta_k+sb_k)^2}{2} \right\} \\
&= \exp\left\{ -\frac{t^2}{2}\sum_{k=1}^\infty a_k^2 \right\} \exp\left\{ ts \underbrace{\sum_{k=1}^\infty a_k b_k}_{\text{Cov }(U,V)=0} \right\} \exp\left\{ -\frac{s^2}{2}\sum_{k=1}^\infty b_k^2 \right\} \\
&= \varphi_U(t)\varphi_V(s),
\end{aligned}$$

$s,t \in \mathbb{R}$.
Daher sind U und V unabhängig, wenn Cov $(U,V) = 0$. □

Bemerkung 7.2.7

1. Lemma 7.2.6 ist ein Spezialfall der Theorie der Gaußschen linearen zufälligen Funktionen indiziert durch Elemente eines Hilbertraums L.

2. Wir setzten $L = l_2$, den Raum der Folgen reeller Zahlen $b = \{b_n\}_{n=1}^\infty$ mit der Eigenschaft $\|b\|_2 := \sqrt{\sum_{n=1}^\infty b_n^2} < \infty$ ausgestattet mit dem Skalarprodukt $\langle a,b\rangle_2 := \sum_{n=1}^\infty a_n b_n$ für $a,b \in l_2$.

3. Sei $X = \{X(a),\ a \in l_2\}$ definiert als $X(a) = \langle a, Y\rangle_2$, wobei $Y = \{Y_n\}_{n=1}^\infty$ eingeführt wurde in Lemma 7.2.6.

4. Dann folgt einfach, dass $X(a) \sim N(0, \|a\|_2^2)$ und Cov $(X(a), X(b)) = \langle a,b\rangle_2$.

5. Man kann leicht zeigen, dass X eine Gaußsche zufällige Funktion ist, die man *linear* nennt.

6. Gaußsche lineare zufällige Funktionen spielen eine große Rolle in der Theorie der L-wertigen Gaußschen zufälligen Elemente, vgl. [?] Kapitel. I, Abschnitt 4.

Satz 7.2.8 Sei $\{Y_n,\ n \in \mathbb{N}\}$ eine Folge von unabhängigen $\mathcal{N}(0,1)$-verteilten ZVn und auf dem W-Raum $(\Omega, \mathcal{F}, P)$.
Dann existiert ein W-Unterraum $(\Omega_0, \mathcal{F}_0, P)$ von $(\Omega, \mathcal{F}, P)$ und ein stochastischer Prozess $X = \{X(t),\ t \in [0,1]\}$ darauf, s.d.

$$X(t,\omega) = \sum_{n=1}^{\infty} Y_n(\omega) S_n(t), \quad t \in [0,1],\, \omega \in \Omega_0$$

und $X \stackrel{d}{=} W$ gilt.
Dabei ist $\{S_n\}_{n\in\mathbb{N}}$ die Familie der Schauder-Funktionen.

Beweis Nach Lemma 7.2.4.2 erfüllen die Koeffizienten $S_n(t)$ die Bedingungen von Lemma 7.2.6 für jedes $t \in [0,1]$.
Nach 7.2.5 existiert eine Teilmenge $\Omega_0 \subset \Omega$, $\Omega_0 \in \mathcal{F}$ mit $P(\Omega_0) = 1$, s.d. für jedes $\omega \in \Omega_0$ die Relation

$$|Y_n(\omega)| = O(\sqrt{\log n}), \quad n \to \infty$$

gilt.
Sei $\mathcal{F}_0 = \mathcal{F} \cap \Omega_0$. Wir beschränken uns auf den W-Raum $(\Omega_0, \mathcal{F}_0, P)$. Die Bedingung

$$a_n = Y_n(\omega) = O(n^{\varepsilon})$$

ist für $\varepsilon < \frac{1}{2}$ erfüllt, da

$$\sqrt{\log n} < n^{\varepsilon}$$

für hinreichend großes n gilt und nach Lemma 7.2.2, 3) die Reihe

$$\sum_{n=1}^{\infty} Y_n(\omega) S_n(t)$$

absolut und gleichmäßig auf $t \in [0,1]$ gegen die Funktion $X(\omega,t)$, $\omega \in \Omega_0$ konvergiert, die stetig ist in t für jedes $\omega \in \Omega_0$. $X(\cdot,t)$ ist eine Zufallsvariable, da in Lemma 7.2.6 die Konvergenz der Reihe f.s. gilt.
Es gilt $X(t) \sim \mathcal{N}(0, \sum_{n=1}^{\infty} S_n^2(t))$ für $t \in [0,1]$. Wir zeigen, dass der stochastische Prozess definiert auf $(\Omega_0, \mathcal{F}_0, P)$ ein Wiener-Prozess ist. Dazu überprüfen wir die Bedingungen von 7.1.2.

Betrachte beliebige $0 \le t_1 < t_2$ und $t_3 < t_4 \le 1$ und berechne:

$$\begin{aligned}
& \text{Cov}\ (X(t_2) - X(t_1), X(t_4) - X(t_3)) \\
= \ & \text{Cov}\ \left(\sum_{n=1}^{\infty} Y_n(S_n(t_2) - S_n(t_1)), \sum_{n=1}^{\infty} Y_n(S_n(t_4) - S_n(t_3)) \right) \\
= \ & \sum_{n=1}^{\infty} (S_n(t_2) - S_n(t_1)) \cdot (S_n(t_4) - S_n(t_3)) \\
= \ & \sum_{n=1}^{\infty} (\langle H_n, I_{[0,t_2]} \rangle - \langle H_n, I_{[0,t_1]} \rangle) \cdot (\langle H_n, I_{[0,t_4]} \rangle - \langle H_n, I_{[0,t_3]} \rangle) \\
= \ & \sum_{n=1}^{\infty} \langle H_n, I_{[0,t_2]} - I_{[0,t_1]} \rangle \langle H_n, I_{[0,t_4]} - I_{[0,t_3]} \rangle \\
= \ & \langle I_{[0,t_2]} - I_{[0,t_1]}, I_{[0,t_4]} - I_{[0,t_3]} \rangle \\
= \ & \langle I_{[0,t_2]}, I_{[0,t_4]} \rangle - \langle I_{[0,t_1]}, I_{[0,t_4]} \rangle - \langle I_{[0,t_2]}, I_{[0,t_3]} \rangle + \langle I_{[0,t_1]}, I_{[0,t_3]} \rangle \\
= \ & \min\{t_2, t_4\} - \min\{t_1, t_4\} - \min\{t_2, t_3\} + \min\{t_1, t_3\},
\end{aligned}$$

wegen der Parseval- Identität und da

$$\langle I_{[0,s]}, I_{[0,t]} \rangle = \int_0^{\min\{s,t\}} du = \min\{s,t\}, \quad s,t \in [0,1].$$

Falls $0 \le t_1 < t_2 \le t_3 < t_4 < 1$ gilt, folgt

$$\text{Cov}\ (X(t_2) - X(t_1), X(t_4) - X(t_3)) = t_2 - t_1 - t_2 + t_1 = 0.$$

Also sind Zuwächse von X (nach Lemma 7.2.6) unabhängig. Außerdem gilt

$$X(t) - X(s) = \sum_{n=1}^{\infty} Y_n(S_n(t) - S_n(s)) \sim N(0, \text{Var}\ (X(t) - X(s)))$$

gilt nach Lemma 7.2.6: $X(t) - X(s) \sim \mathcal{N}(0, t-s)$ und $X \stackrel{d}{=} W$ nach 7.1.2. □

Bemerkung 7.2.9

1. Satz 7.2.8 ist die Grundlage für die approximative Simulation der Pfade einer Brownschen Bewegung durch die Partialsummen $X^{(n)}(t) = \sum_{k=1}^{n} Y_k S_k(t)$, $t \in [0,1]$ für hinreichend großes $n \in \mathbb{N}$.

2. Die Konstruktion in Satz 7.2.8 kann genutzt werden für die Konstruktion des Wiener-Prozesses mit stetigen Pfaden auf dem Intervall $[0, t_0]$ für beliebiges $t_0 > 0$. Falls $W = \{W(t),\ t \in [0,1]\}$ ein Wiener-Prozess auf $[0,1]$ ist, dann ist auch $Y = \{Y(t),\ t \in [0,t_0]\}$ mit $Y(t) = \sqrt{t_0} W(\frac{t}{t_0})$, $t \in [0,t_0]$ ein Wiener-Prozess auf $[0,t_0]$.

3. Ein Wiener-Prozess W mit stetigen Pfaden auf $\mathbb{R}_+$ kann wie folgt konstruiert werden: Seien $W^{(n)} = \{W^{(n)}(t),\ t \in [0,1]\}$ unabhängige Kopien des Wiener-Prozesses wir in Satz 7.2.8. Definiere

$$W(t) = \sum_{n=1}^{\infty} I(t \in [n-1, n]) \left[\sum_{k=1}^{n-1} W^{(k)}(1) + W^{(n)}(t-(n-1)) \right]$$

für $t \geq 0$. Dann gilt

$$W(t) = \begin{cases} W^{(1)}(t), & t \in [0,1], \\ W^{(1)}(1) + W^{(2)}(t-1), & t \in [1,2], \\ W^{(1)}(1) + W^{(2)}(1) + W^{(3)}(t-2), & t \in [2,3], \\ \text{usw.} \end{cases}$$

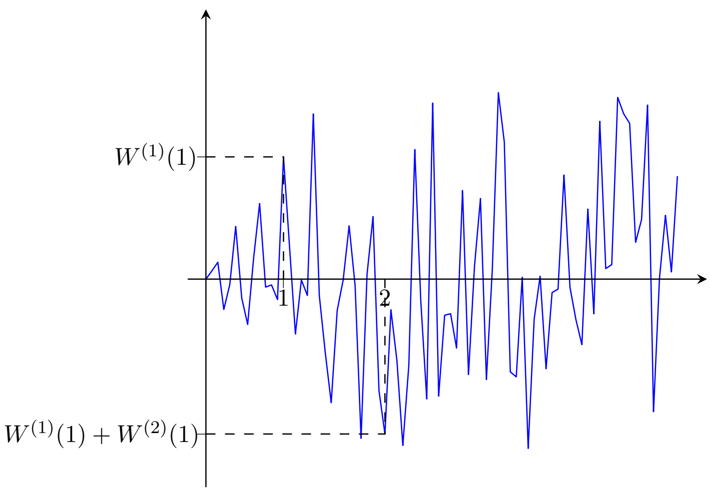


Abbildung 7.3: Konstruktion des Wiener Prozess auf $[0, \infty)$

Übungsaufgabe 7.2.10 Beweise 7.2.9.2.

Übungsaufgabe 7.2.11 Zeige, dass der eingeführte stochastische Prozess $W = \{W(t),\ t \geq 0\}$ ein Wiener-Prozess auf $\mathbb{R}_+$ ist.

7.3 Verteilungs- und Pfadeigenschaften vom Wiener - Prozess

7.3.1 Donskers Invarianzprinzip

Sei $\{W(t),\ t \in [0,1]\}$ ein Wiener-Prozess und $Z_1, Z_2, \ldots$ eine Folge von unabhängige Zufallsvariablen mit $\mathbb{E}Z_i = 0$, $\text{Var } Z_i = 1$, $i \geq 1$ z.B.

$P(Z_i = 1) = P(Z_i = -1) = \frac{1}{2}$.
Für jedes $n \in \mathbb{N}$ definieren wir $\{\widetilde{W}^{(n)}(t),\ t \in [0,1]\}$ durch

$$\widetilde{W}^{(n)}(t) = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} + (nt - \lfloor nt \rfloor)\frac{Z_{\lfloor nt \rfloor + 1}}{\sqrt{n}}, \tag{7.1}$$

wobei $S_i = Z_1 + \ldots + Z_i$, $i \geq 1$, $S_0 = 0$.
Konstruiere eine Approximation von W durch eine zufällige Irrfahrt $\widetilde{W}^{(n)}$ mit Schrittgröße Z_i für $n \to \infty$.

Satz 7.3.1 Seien $P_{\widetilde{W}^{(n)}}$ bzw. P_W Verteilungen von $\widetilde{W}^{(n)}$ bzw. W in $C[0,1]$. Dann gilt $P_{\widetilde{W}^{(n)}} \overset{W}{\Rightarrow} P_W$, für $n \to \infty$, wobei mit $\overset{W}{\Rightarrow}$ die schwache Konvergenz in $C[0,1]$, gemeint ist, d.h. für jede beschränkte, stetige Funktion $f : C[0,1] \to \mathbb{R}$ gilt

$$\int f dP_{\widetilde{W}^{(n)}} \underset{n\to\infty}{\longrightarrow} \int f dP_W.$$

Bemerkung 7.3.2

- Bei der Konstruktion der Approximation $\widetilde{W}^{(n)}(t)$, kann jede Folge von u.i.v. $\{Z_n\}_{n\in\mathbb{N}}$ mit $\mathbb{E}Z_n = 0$ und Var $Z_n = 1$ verwendet werden.
- Dies wurde von Monroe Donsker (1951) bewiesen und wird *Invarianzprinzip* genannt, da die Approximation $\widetilde{W}^{(n)}(\cdot)$ von $W(\cdot)$ nicht von der Verteilung von Z_1 abhängt.

Satz 7.3.1 besagt, dass die Verteilung der ganzen Pfade von $\widetilde{W}^{(n)}(\cdot)$ gegen die von $W(\cdot)$ konvergiert. Der Beweis ist sehr theoretisch (vgl.[15] 21.6-21.8). Wir zeigen nur die Konvergenz der endlich-dimensionalen Verteilungen.

Lemma 7.3.3 Für jedes $k \geq 1$ und beliebige $t_1, \ldots, t_k \in [0,1]$ gilt:

$$\left(\widetilde{W}^{(n)}(t_1), \ldots, \widetilde{W}^{(n)}(t_k)\right)^\top \underset{n\to\infty}{\overset{d}{\longrightarrow}} (W(t_1), \ldots, W(t_k))^\top.$$

Beweis Für $k = 1$, ist die Aussage ein einfaches Korollar des Zentralen Grenzwertsatzes. Betrachte den Spezialfall $k = 2$ (für $k > 2$ ist der Beweis analog): Sei $t_1 < t_2$. Für alle $s_1, s_2 \in \mathbb{R}$ gilt:

$$\begin{aligned}
s_1\widetilde{W}^{(n)}(t_1) + s_2\widetilde{W}^{(n)}(t_2) = \ & (s_1 + s_2)\frac{S_{\lfloor nt_1 \rfloor}}{\sqrt{n}} \\
+ \ & s_2\frac{(S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor + 1})}{\sqrt{n}} \\
+ \ & Z_{\lfloor nt_1 \rfloor + 1}\left((nt_1 - \lfloor nt_1 \rfloor)\frac{s_1}{\sqrt{n}} + \frac{s_2}{\sqrt{n}}\right) \\
+ \ & Z_{\lfloor nt_2 \rfloor + 1}(nt_2 - \lfloor nt_2 \rfloor)\frac{s_2}{\sqrt{n}},
\end{aligned}$$

da $S_{\lfloor nt_2 \rfloor} = S_{\lfloor nt_1 \rfloor} + S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor + 1} + Z_{\lfloor nt_1 \rfloor + 1}$. Beachte, dass die vier Summanden auf der rechten Seite der Gleichung unabhängig sind und, dass die letzten beiden gegen 0 konvergieren (f.s. und deshalb auch in Verteilung). Es gilt

$$Y_n = Z_{\lfloor nt \rfloor + 1} \frac{1}{\sqrt{n}} \left(nt - \lfloor nt \rfloor\right) \xrightarrow[n\to\infty]{d} 0,$$

da

$$\varphi_{Y_n}(s) = \varphi_Z\left(\frac{\left(nt - \lfloor nt \rfloor\right) s}{\sqrt{n}}\right) \xrightarrow[n\to\infty]{} \varphi_Z(0) = 1 = \varphi_0(s),$$

also auch

$$Y_n \xrightarrow[n\to\infty]{d} 0.$$

Daraus folgt auch

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E} e^{i(s_1 \widetilde{W}^{(n)}(t_1) + s_2 \widetilde{W}^{(n)}(t_2))} \\
&= \lim_{n\to\infty} \mathbb{E} e^{i \frac{s_1+s_2}{\sqrt{n}} S_{\lfloor nt_1 \rfloor}} \mathbb{E} e^{i \frac{s_2}{\sqrt{n}} (S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor + 1})} \\
&= \lim_{n\to\infty} \mathbb{E} e^{i(s_1+s_2)\sqrt{\frac{\lfloor nt_1 \rfloor}{n}} \frac{S_{\lfloor nt_1 \rfloor}}{\sqrt{\lfloor nt_1 \rfloor}}} \mathbb{E} e^{i \frac{s_2}{\sqrt{n}} S_{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor - 1}} \\
&\overset{\text{ZGWS}}{=} e^{-\frac{t_1}{2}(s_1+s_2)^2} e^{-\frac{t_2-t_1}{2} s_2^2} \\
&= e^{-\frac{1}{2}\left(s_1^2 t_1 + 2 s_1 s_2 t_1 + s_2^2 t_2\right)} \\
&= e^{-\frac{1}{2}\left(s_1^2 t_1 + 2 s_1 s_2 \min\{t_1, t_2\} + s_2^2 t_2\right)} = \varphi_{(W(t_1), W(t_2))}(s_1, s_2),
\end{aligned}$$

wobei $\varphi_{(W(t_1),W(t_2))}$ charakteristische Funktion von $(W(t_1), W(t_2))^\top$ ist. Genauer gilt

$$\sqrt{\frac{\lfloor nt_1 \rfloor}{n}} \xrightarrow[n\to\infty]{} \sqrt{t_1} \quad \text{und} \quad \frac{S_{\lfloor nt_1 \rfloor}}{\sqrt{\lfloor nt_1 \rfloor}} \xrightarrow[n\to\infty]{d} Y_1 \sim N(0,1)$$

nach dem Zentralen Grenzwertsatz und analog

$$\frac{1}{\sqrt{n}} S_{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor - 1} = \sqrt{\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor - 1}{n}} \frac{S_{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor - 1}}{\sqrt{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor - 1}},$$

$$\frac{S_{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor - 1}}{\sqrt{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor - 1}} \xrightarrow[n\to\infty]{d} \sqrt{t_2 - t_1} Y_2,$$

wobei $Y_2 \sim N(0,1)$ unabhängig von Y_1 ist. Die Äquivalenz der Konvergenz in Verteilung und der schwachen Konvergenz und die Tatsache, dass

$$\varphi_Y(s) = \mathbb{E} e^{isY} = e^{-s^2\sigma^2/2},$$

falls $Y \sim N(0, \sigma^2)$, schließen den Beweis. □

7.3.2 Gesetz der großen Zahlen

Definiere

$$M_t = \max_{s \in [0,t]} W(s), t > 0, \tag{7.2}$$

wobei $W = \{W(t), t \geq 0\}$ ein Wiener-Prozess ist. Die Abbildung $M_t : \Omega \to [0, \infty)$ gegeben in (7.2) ist eine wohldefinierte Zufallsvariable, da

$$\max_{s \in [0,t]} W(s, \omega) = \lim_{n \to \infty} \max_{i=1,\ldots,n} W\left(\frac{it}{n}, \omega\right)$$

für alle $\omega \in \Omega$ gilt, weil die Pfade von $\{W(t),\ t \geq 0\}$ stetig sind.
Den folgenden Satz nehmen wir ohne Beweis hin:

Satz 7.3.4 Sei $W = \{W(t),\ t \geq 0\}$ ein Wiener-Prozess definiert auf dem W-Raum $(\Omega, \mathcal{F}, P)$. Dann gilt:

$$P(M_t > x) = 2P(W(t) > x) = \sqrt{\frac{2}{\pi t}} \int_x^\infty e^{-\frac{y^2}{2t}} dy, \quad x \geq 0,\, t > 0. \tag{7.3}$$

Aus (7.3) folgt, dass $\max_{t \in [0,1]} W(t)$ einen exponentiell beschränkten Tail hat. Deshalb hat $\max_{t \in [0,1]} W(t)$ endliche k-te Momente für $k \in \mathbb{N}$.

Folgerung 7.3.5 Sei $\{W(t),\ t \geq 0\}$ ein Wiener-Prozess. Dann gilt

$$\frac{W(t)}{t} \xrightarrow[t \to \infty]{f.s.} 0.$$

Beweis Für jedes $t \geq 0$ $\exists!$ $n \in \mathbb{N} : t \in [n, n+1)$. Wir zeigen, dass

$$\frac{W(t)}{t} \underset{n,t \to \infty}{\overset{f.s.}{\sim}} \frac{W(n)}{n} \quad \text{und} \quad \frac{W(n)}{n} \xrightarrow[n \to \infty]{f.s.} 0.$$

Mit dem Starken Gesetz der großen Zahlen (vgl. Satz 4.1.7) folgt

$$\frac{1}{n} W(n) = \frac{1}{n} \sum_{i=1}^n (W(i) - W(i-1)) \xrightarrow[n \to \infty]{f.s.} \mathbb{E}W(1) = 0$$

aufgrund der Unabhängigkeit und Stationarität der Zuwächse

$$W(i) - W(i-1) \stackrel{d}{=} W(1) - W(0) = W(1), \quad i \in \mathbb{N}$$

von W und weil $W(1) \sim N(0,1)$. Nun gilt

$$\begin{aligned}
& \left| \frac{W(t)}{t} - \frac{W(n)}{n} \right| \\
\leq \; & \left| \frac{W(t)}{t} - \frac{W(n)}{t} \right| + \left| \frac{W(n)}{t} - \frac{W(n)}{n} \right| \\
\leq \; & \left| W(n) \left(\frac{1}{t} - \frac{1}{n} \right) \right| + \frac{1}{n} \sup_{s \in [0,1]} |W(n+s) - W(n)| \\
\leq \; & \left| \frac{2}{n} W(n) \right| + \frac{Z(n)}{n},
\end{aligned}$$

wobei

$$Z(n) = \sup_{s\in[0,1]} |W(n+s) - W(n)|, \quad n \in \mathbb{N}$$

eine Folge von u.i.v. Zufallsvariablen mit

$$Z(n) \stackrel{d}{=} Z(0) = \sup_{s\in[0,1]} |W(s)|$$

ist. Wir zeigen, dass $\mathbb{E}Z(0) < \infty$. Falls dies gilt, dann folgt auch

$$\frac{Z(n)}{n} = \frac{1}{n}\sum_{i=1}^{n} Z(i) - \frac{1}{n}\sum_{i=1}^{n-1} Z(i) \xrightarrow[n\to\infty]{f.s.} \mathbb{E}Z(0) - \mathbb{E}Z(0) = 0$$

nach dem starken Gesetz der großen Zahlen (vgl. Satz 4.1.7).

$$\begin{aligned} P(Z(0) > x) &\leq P\left(\max_{s\in[0,1]} W(s) > x\right) + P\left(\max_{s\in[0,1]} (-W(s)) > x\right) \\ &= 2P\left(\max_{s\in[0,1]} W(s) > x\right) = 2\sqrt{\frac{2}{\pi}} \int_x^\infty e^{-y^2/2} dy \end{aligned}$$

nach Satz 7.3.4, da $\{-W(s),\ s \geq 0\}$ auch ein Wiener-Prozess ist aufgrund seiner Symmetrie (vgl. Satz 7.3.6). Dann gilt

$$\begin{aligned} \mathbb{E}Z(0) &= \int_0^\infty P(Z(0) > x) dx \leq 2\sqrt{\frac{2}{\pi}} \int_0^\infty \int_x^\infty e^{-y^2/2} dy dx \\ &= 4\mathbb{E}(XI(X \geq 0)) < \infty \end{aligned}$$

für $X \sim N(0,1)$. □

7.3.3 Invarianzeigenschaften

Spezielle Transformationen des Wiener-Prozesses sind wiederum Wiener-Prozesse.

Satz 7.3.6 Sei $\{W(t),\ t \geq 0\}$ ein Wiener-Prozess. Dann sind die stochastischen Prozesse $\{Y^{(i)}(t),\ t \geq 0\}$, $i = 1, \ldots, 4$ mit

$$\begin{aligned} Y^{(1)}(t) &= -W(t), && \text{(Symmetrie)} \\ Y^{(2)}(t) &= W(t+t_0) - W(t_0), && \text{(Verschiebung des Ursprungs)} \\ Y^{(3)}(t) &= \sqrt{c}W\left(\frac{t}{c}\right), && \text{(Skalierung)} \\ Y^{(4)}(t) &= \begin{cases} tW(\frac{1}{t}) & t > 0, \\ 0 & t = 0 \end{cases} && \text{(Spiegelung im Ursprung)} \end{aligned}$$

für ein $t_0 > 0$ und ein $c > 0$ auch Wiener-Prozesse.

Beweis

1. $Y^{(i)}(0) = 0$ f.s. für $i = 1, \dots, 4$.

2. $Y^{(i)}$, $i = 1, \dots, 4$, haben unabhängige Zuwächse mit $Y^{(i)}(t_2) - Y^{(i)}(t_1) \sim \mathcal{N}(0, t_2 - t_1)$.

3. $Y^{(i)}$, $i = 1, \dots, 3$, haben stetige Pfade. $\{Y^{(4)}(t),\ t \geq 0\}$ hat stetige Pfade für $t > 0$.

4. Wir müssen zeigen, dass $Y^{(4)}(t)$ f.s. stetig in $t = 0$ ist, d.h. dass $\lim_{t\to 0} tW(\frac{1}{t}) \overset{f.s.}{=} 0$. Nach Folgerung 7.3.5 gilt
$$\lim_{t\to 0} tW\left(\frac{1}{t}\right) = \lim_{t\to\infty} \frac{W(t)}{t} \overset{f.s.}{=} 0.$$

Also sind $Y^{(i)}$, $i = 1, \dots, 4$, auch Wiener-Prozesse nach Definition 7.1.2 □

Folgerung 7.3.7 Sei $\{W(t),\ t \geq 0\}$ ein Wiener-Prozess. Dann gilt

$$P\left(\sup_{t\geq 0} W(t) = \infty\right) = P\left(\inf_{t\geq 0} W(t) = -\infty\right) = 1,$$

und deswegen auch

$$P\left(\sup_{t\geq 0} W(t) = \infty,\ \inf_{t\geq 0} W(t) = -\infty\right) = 1.$$

Beweis Für $x, c > 0$ gilt, dass

$$\begin{aligned} P\left(\sup_{t\geq 0} W(t) > x\right) &= P\left(\sup_{t\geq 0} W\left(\frac{t}{c}\right) > \frac{x}{\sqrt{c}}\right) \\ &= P\left(\sup_{t\geq 0} W(t) > \frac{x}{\sqrt{c}}\right) \\ &\underset{c\to+\infty}{\longrightarrow} P(\sup_{t\geq 0} W(t) > 0) \\ &\underset{c\to+0}{\longrightarrow} P(\sup_{t\geq 0} W(t) = +\infty) \end{aligned} \tag{7.4}$$

nicht von x abhängt. Also gilt

$$\begin{aligned} 1 &= P\left(\left\{\sup_{t\geq 0} W(t) = 0\right\} \bigcup \left\{\sup_{t\geq 0} W(t) > 0\right\}\right) \\ &= P\left(\sup_{t\geq 0} W(t) = 0\right) + \underbrace{P\left(\sup_{t\geq 0} W(t) > 0\right)}_{P\left(\sup_{t\geq 0} W(t) = +\infty\right)}, \end{aligned}$$

und weiter

$$\begin{aligned}
P\left(\sup_{t\geq 0} W(t)=0\right) &= P\left(\sup_{t\geq 0} W(t)\leq 0\right)\\
&\leq P\left(W(1)\leq 0,\ \sup_{t\geq 1} W(t)\leq 0\right)\\
&= P\left(W(1)\leq 0,\ \sup_{t\geq 1}(W(t)-W(1))\leq -W(1)\right)\\
&= \int_{-\infty}^{0} P\left(\sup_{t\geq 1}(\underbrace{W(t)-W(1)}_{Y^{(2)}(t)\stackrel{d}{=}W(t)})\leq -W(1)\mid W(1)=x\right) P_{W(1)}(dx)\\
&\stackrel{Y^{(2)}\stackrel{d}{=}W}{=} \int_{-\infty}^{0} P\left(\sup_{t\geq 0} W(t)\leq -x\right) P_{W(1)}(dx)\\
&= \int_{-\infty}^{0} P\left(\sup_{t\geq 0} W(t)=0\right) P_{W(1)}(dx)\\
&= P\left(\sup_{t\geq 0} W(t)=0\right)\frac{1}{2},
\end{aligned}$$

da $\sup_{t\geq 0} W(t)\geq 0$ f.s. und $P\left(\sup_{t\geq 0} W(t)\leq -x\right)$ nicht von x abhängt nach (7.4), also sei $x=0$. Dann gilt

$$P\left(\sup_{t\geq 0} W(t)=0\right)=0 \quad \text{und} \quad P\left(\sup_{t\geq 0} W(t)=\infty\right)=1.$$

Analog können wir zeigen, dass

$$P\left(\inf_{t\geq 0} W(t)=-\infty\right)=1.$$

Der Rest folgt aus $P(A\cap B)=1$ für alle $A,B\in\mathcal{F}$ mit $P(A)=P(B)=1$, da

$$P(A\cap B)=P(A)+P(B)-P(A\cup B)=1+1-1=1.$$

□

Bemerkung 7.3.8

1. $P\left(\sup_{t\geq 0} X(t)=\infty, \inf_{t\geq 0} X(t)=-\infty\right)=1$ impliziert, dass die Pfade von W unendlich oft zwischen positiven und negativen Werten auf $[0,\infty)$ oszillieren.

2. Zusätzlich zum starken Gesetz der großen Zahlen (vgl. Folgerung 7.3.5) erfüllt der Wiener-Prozess das Gesetz des iterierten Logarithmus:
$$\limsup_{t\to\infty} \frac{W(t)}{\sqrt{2t\log\log(t)}} \stackrel{f.s.}{=} 1, \quad \liminf_{t\to\infty} \frac{W(t)}{\sqrt{2t\log\log(t)}} \stackrel{f.s.}{=} -1.$$
Der Beweis ist zu finden z.B. in [3] Theorem 3.2.

Folgerung 7.3.9 Sei $\{W(t),\ t \geq 0\}$ ein Wiener-Prozess. Dann gilt
$$P\left(\omega \in \Omega : W(t,\omega) \text{ ist nirgends diff'bar auf } [0,\infty)\right) = 1.$$

Beweis Es gilt
$$\begin{aligned}
&\{\omega \in \Omega : W(t,\omega) \text{ ist nirgends diff'bar auf } [0,\infty)\}\\
=& \bigcap_{n=0}^{\infty} \{\omega \in \Omega : W(t,\omega) \text{ ist nirgends diff'bar auf } [n,n+1)\}.
\end{aligned}$$
Es genügt zu zeigen, dass
$$P(\omega \in \Omega : W(t,\omega) \text{ ist diff'bar für ein } t_0 = t_0(\omega) \in [0,1]) = 0.$$
Definiere die Menge
$$\begin{aligned}
A_{nm} =& \Big\{\omega \in \Omega :\ \exists\ t_0 = t_0(\omega) \in [0,1] \text{ mit}\\
& |W(t_0(\omega)+h,\omega) - W(t_0(\omega),\omega))| \leq mh,\ \forall h \in \left[0,\frac{4}{n}\right]\Big\}.
\end{aligned}$$
Dann gilt
$$\{\omega \in \Omega : W(t,\omega) \text{ diff'bar für ein } t_0 = t_0(\omega)\} \subseteq \bigcup_{m\geq 1}\bigcup_{n\geq 1} A_{nm}.$$
Es ist zu zeigen, dass $P(\cup_{m\geq 1}\cup_{n\geq 1} A_{nm}) = 0$. Da
$$P\left(\bigcup_{m\geq 1}\bigcup_{n\geq 1} A_{nm}\right) \leq \sum_{m\geq 1}\sum_{n\geq 1} P(A_{nm}),$$
genügt es zu zeigen, dass
$$P(A_{nm}) = 0 \quad \forall n,m \in \mathbb{N}.$$
Sei $k_0(\omega) = \arg\min_{k=1,\ldots,n}\{\frac{k}{n} \geq t_0(\omega)\}$. Dann gilt für $\omega \in A_{nm}$ und $j = 0,1,2$:
$$\begin{aligned}
&\left|W\left(\frac{k_0(\omega)+j+1}{n},\omega\right) - W\left(\frac{k_0(\omega)+j}{n},\omega\right)\right|\\
\leq& \left|W\left(\frac{k_0(\omega)+j+1}{n},\omega\right) - W\left(t_0(\omega),\omega\right)\right|\\
+& \left|W\left(\frac{k_0(\omega)+j}{n},\omega\right) - W\left(t_0(\omega),\omega\right)\right|\\
\leq& \frac{8m}{n}.
\end{aligned}$$

Sei

$$\Delta_n(k) = W\left(\frac{k+1}{n}\right) - W\left(\frac{k}{n}\right) \sim N(0, 1/n).$$

Dann gilt

$$P(|\Delta_n(k)| \leq x) = \sqrt{\frac{n}{2\pi}} \int\limits_{-x}^{x} e^{-\frac{nt^2}{2}} dt \leq \frac{2x\sqrt{n}}{\sqrt{2\pi}}, \quad \forall\, x \geq 0$$

und weiter auch

$$\begin{aligned} P(A_{nm}) &\leq P\left(\bigcup_{k=0}^{n} \bigcap_{j=0}^{2} \left\{|\Delta_n(k+j)| \leq \frac{8m}{n}\right\}\right) \\ &\leq \sum_{k=0}^{n} P\left(\bigcap_{j=0}^{2} \left\{|\Delta_n(k+j)| \leq \frac{8m}{n}\right\}\right) \\ &= \sum_{k=0}^{n} \left(P\left(|\Delta_n(0)| \leq \frac{8m}{n}\right)\right)^3 \\ &\leq (n+1)\left(\frac{16m}{\sqrt{2\pi n}}\right)^3 \underset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

aufgrund der Unabhängigkeit und Stationarität der Zuwächse des Wiener-Prozesses. Da $A_{nm} \subset A_{n+1,m}$, folgt $P\left(A_{nm}\right) \nearrow$ und deshalb

$$P(A_{nm}) = 0 \quad \forall\, n, m \in \mathbb{N}.$$

□

Folgerung 7.3.10 Mit Wahrscheinlichkeit 1 gilt

$$\sup_{n\geq 1} \sup_{0\leq t_0 < \ldots < t_n \leq 1} \sum_{i=1}^{n} |W(t_i) - W(t_{i-1})| = \infty,$$

d.h. die Pfade von $\{W(t),\ t \in [0,1]\}$ haben f.s. unbeschränkte Variation.

Beweis Da jede Funktion $g : [0,1] \to \mathbb{R}$ mit beschränkter Variation fast überall differenzierbar ist[1], folgt die Behauptung aus der Folgerung 7.3.9.
Ein alternativer Beweis kann wie folgt formuliert werden:
Es genügt zu zeigen, dass

$$\lim_{n\to\infty} \sum_{i=1}^{2^n} \left| W\left(\frac{it}{2^n}\right) - W\left(\frac{(i-1)t}{2^n}\right) \right| = \infty \text{ für } t = 1.$$

[1]Jede Funktion mit beschränkter Variation kann dargestellt werden als Differenz von zwei nicht-fallenden montonen Funktionen, wobei beide fast überall diff'bar auf $[0,1]$ sind nach Satz von Lebesgue (vgl. [?] S. 335).

Da W unabhängige und stationäre Zuwächse hat, gilt

$$\sum_{i=1}^{2^n} \underbrace{\left| W\left(\frac{it}{2^n}\right) - W\left(\frac{(i-1)t}{2^n}\right) \right|}_{Y_i \stackrel{d}{=} \sqrt{\frac{t}{2^n}} X_i,\ X_i \sim N(0,1) \text{ - u.i.v.}} = \frac{\sqrt{t}}{2^{n/2}} \sum_{i=1}^{2^n} |X_i|$$

$$= \ 2^{n/2}\sqrt{t} \ \underbrace{\frac{1}{2^n} \sum_{i=1}^{2^n} |X_i|}_{\xrightarrow{f.s.} \mathbb{E}|X_1|,\ n \to \infty} \ \underset{n\to\infty}{\overset{f.s}{\sim}} \ 2^{n/2}\sqrt{t}\mathbb{E}|X_1|$$

nach dem starken Gesetz der großen Zahlen 4.1.7, wobei $X_1 \sim N(0,1)$. Deshalb gilt

$$\sum_{i=1}^{2^n} \left| W\left(\frac{it}{2^n}\right) - W\left(\frac{(i-1)t}{2^n}\right) \right| \xrightarrow[n\to\infty]{f.s} \infty.$$

□

Bemerkung 7.3.11 Die quadratische Variation von W über $[s,t]$ ist gleich $t-s$:

$$\lim_{n\to\infty} \sum_{i=1}^{n} \left| W(t_i^{(n)}) - W(t_{i-1}^{(n)}) \right|^2 = t - s$$

f.s. oder in L^2, wobei $\Gamma_n = \{t_i^{(n)}\}_{i=1}^n$ eine Folge von Zerlegungen von $[s,t]$ ist, s.d.

$$s \le t_1 \le \cdots \le t_n \le t \quad \text{und} \quad \Gamma_n \subseteq \Gamma_{n+1}, \quad \forall\, n \in \mathbb{N}.$$

Literaturverzeichnis

[1] Dehling, H. B. Haupt. *Einfürung in die Wahrscheinlichkeitstheorie und Statistik.* Springer, Berlin, 2003.

[2] H. Bauer. *Wahrscheinlichkeitstheorie.* de Gruyter, Berlin, 1991.

[3] A. A. Borovkov. *Probability theory.* Universitext. Springer, London, 2013. Translated from the 2009 Russian fifth edition by O. B. Borovkova and P. S. Ruzankin, Edited by K. A. Borovkov.

[4] A.A. Borovkov. *Wahrscheinlichkeitstheorie: eine Einführung.* Birkhäuser, Basel, 1976.

[5] S. Cambanis, S. Huang, and G. Simons. On the theory of elliptically contoured distributions. *J. Multivariate Anal.*, 11(3):368–385, 1981.

[6] K. T. Fang, S. Kotz, and K. W. Ng. *Symmetric multivariate and related distributions*, volume 36 of *Monographs on Statistics and Applied Probability.* Chapman and Hall, Ltd., London, 1990.

[7] W. Feller. *An introduction to probability theory and its applications. Vol I/II.* J. Wiley & Sons, New York, 1970/71.

[8] P. Gänßler and W. Stute. *Wahrscheinlichkeitstheorie.* Springer, Berlin, 1977.

[9] H. O. Georgii. *Stochastik.* de Gruyter, Berlin, 2002.

[10] B. W. Gnedenko. *Einfürung in die Wahrscheinlichkeitstheorie.* Akademie, Berlin, 1991.

[11] C. Hesse. *Angewandte Wahrscheinlichkeitstheorie.* Vieweg, Braunschweig, 2003.

[12] J. Jacod and P. Protter. *Probability essentials.* Springer, Berlin, 2003.

[13] Olav Kallenberg. *Foundations of modern probability.* Springer Science & Business Media, 2006.

[14] A.F. Karr. *Probability.* Springer, New York, 1993.

[15] A. Klenke. *Wahrscheinlichkeitstheorie.* Universitext. Springer, Berlin, 2. edition, 2008.

[16] U. Krengel. *Einführung in die Wahrscheinlichkeitstheorie.* Vieweg, Braunschweig, 2002.

[17] L. Sachs. *Angewandte Statistik.* Springer, 2004.

[18] R. Schmidt. Tail dependence for elliptically contoured distributions. *Mathematical Methods of Operations Research*, 55(2):301–327, May 2002.

[19] A. N. Shiryaev. *Probability.* Springer, New York, 1996.

[20] E. Spodarev. *Elementare Wahrscheinlichkeitsrechnung und Statistik.* Ulm University, 2023. Vorlesungsskript.

[21] J. M. Stoyanov. *Counterexamples in probability.* Wiley & Sons, 1987.

[22] H. Tijms. *Understanding probability. Chance rules in everyday life.* Cambridge University Press, 2004.

[23] I. S. Tyurin. Refinement of the upper bounds of the constants in Lyapunov's theorem. *Russian Mathematical Surveys*, 65:586–588, 2010.

Index

Äquivalenz in Verteilung, 119

Approximation
durch einfache Funktionen, 4
Approximationssatz
von Weierstrass, *siehe* Weierstrass

Bayes–Formel
bedingte Dichten, 24
bedingte Verteilung
regulär, 23
bedingte Wahrscheinlichkeit, 22
regulär, 23
bedingter Erwartungswert, 16
Beispiele stochastischer Prozesse, 116–118
Lognormal- und χ^2-Funktionen, 117
Rauschen
Salz-und-Pfeffer-Rauschen, 116
weißes Gaußsches Rauschen, 116
weißes Rauschen, 116
Wiener-Prozess, 117, 156
zufällige Irrfahrt, 131
Bernoulli–Verteilung
charakteristische Funktion, 32
Berry
Satz von, 106
Satz von Berry–Esséen, 95
Binomialverteilung
erzeugende Funktion, *siehe* erzeugende Funktion
Bochner–Kchintschin, Satz von, *siehe* charakteristische Funktion
Braunsche Bewegung, *siehe* Wiener-Prozess

charakteristische Funktion, 29–41
Bernoulli–Verteilung, 32
Charakterisierungsaussagen, 40
Bochner–Kchintschin, Satz von, 40
Marcinkiewicz, Satz von, 40
Pólya, Satz von, 40
Definition, 30
Eigenschaften, 30, 33
Eindeutigkeitssatz, 39
Normalverteilung, 32
Poisson–Verteilung, 32
Umkehrformel, 37
Cox-Prozess, 153
Cox-Zählmaß, 153

Dichte
aus charakteristischer Funktion, 37
bedingt, 24
Donsker
Satz von, 168
doppelt-stochastisches Poisson-Maß, *siehe* Cox-Zählmaß

Eindeutigkeitssatz, *siehe* charakteristische Funktion
einfache Funktion, *siehe* Lebesgue
Ergodensatz
individueller, 133
Erneuerungsprozesse, 132–144
Elementarer Erneuerungssatz, 140
Erneuerungsfunktion, 136
Erneuerungsprozess, 132
Exzess, 142
Hauptsatz der Ernerungstheorie, 141
homogener Erneuerungsprozess, 143
Zählprozesse, 132
Erwartungswert, 2, 6
bedingter, 16
Eigenschaften, 17

unter Zufallsvariable, 21
Eigenschaften, 7
Funktion, 126
erzeugende Funktion, 41–44
Binomialverteilung, 43
Definition, 42
Eigenschaften, 42
Esséen
Lemma von, 99
Satz von Berry–Esséen, 95

Faltungsstabilität
Normalverteilung, 41
Poissonverteilung, 41
Fatou
Lemma von, 11
Feller
Feller–Bedingung, 91
Satz von Lindeberg–Feller, 91
Fourier–Umkehrformel, 97

Gesetz der großen Zahlen
schwaches, 75
starkes, 78
gleichmäßige asymptotische Kleinheit, 91
Grenzwertsätze, 73
Gesetz der großen Zahlen, 73–82
Anwendungen, 81–82
schwaches Gesetz der großen Zahlen, 74–75
starkes Gesetz der großen Zahlen, 75–81
zentraler Grenzwertsatz, 82–107
Grenzwertsatz von Lindeberg, 85–94
klassischer, 83
Konvergenzgeschwindigkeit, 94–107

Haar-Funktionen, *siehe* Wiener-Prozess
Helly
Lemma von, 55

integrierte Tailfunktion, 142

Kchintschin, schwaches Gesetz der großen Zahlen von, *siehe* Grenzwertsätze
Kolmogorow
Drei–Reihen–Satz, 72
Existenzsatz, 112
Gesetz der großen Zahlen, 78
Lemma von, *siehe* Grenzwertsätze
Lemma von Kolmogorov-Prokhorov, 138
Satz von, 123
Ungleichung, 70
komplexwertige Zufallsvariable, *siehe* Zufallsvariable
Konvergenz, 45–63
fast sicher, 45
Funktionale von Zufallsvariablen, 59
Addition, 59
Produkt, 61
Satz von Slutsky, 59, 61
Stetigkeitssatz, 63
Gleichgradige Integrierbarkeit, 63
in L^r, 45, 49
in Verteilung, 46, 50
monotone, 9
schwach, 46
stochastisch, 45, 46
Zusammenhang, 46
Kovarianzfunktion, 117
nicht-zentriert, 126
zentriert, 126
Kumulantenerzeugende Funktion, 44

Laplace-Transformation, 137
Lebesgue
Integral, 2
einfacher Funktionen, 4
Maßtransport, 12
Integral von, 2–6
Lebesgue–Stiltjes–Integral, 3, 12
Lemma von Riemann–Lebesgue, 97
Satz von, 11
Lindeberg
Grenzwertsatz von, *siehe* Grenzwertsätze
Satz von, 90
Satz von Lindeberg–Feller, 91
Ljapunow
Ljapunow–Bedingung, 91, 93
Lomniki
Satz von Lomnicki-Ulam, 111
L^r

Konvergenz in, 45

Maß
signiertes, 16
Marcinkiewicz, Satz von, *siehe* charakteristische Funktion
Markow
schwaches Gesetz der großen Zahlen, *siehe* Grenzwertsätze
Maßtransport
im Lebesgue–Integral, 12
Messraum
Borelsch, 24
mittlerer Zuwachs, 127
Modifikation, 118
Momente
Erwartungswert
absolut stetiger Zufallsvariablen, 12
Additivität, 7
alternative Darstellung, 13
diskreter ZV, 12
Monotonie, 7
Produkt, 14
Momenterzeugende Funktion, 44

Normalverteilung
charakteristische Funktion, *siehe* charakteristische Funktion

Orthogonal Projektion
Bedingte Erwartung, 19

Pólya, Satz von, *siehe* charakteristische Funktion
Poisson–Verteilung
charakteristische Funktion, *siehe* charakteristische Funktion
Poisson-Prozess, 144–154
Inhomogener Poisson-Prozess, 144
Poisson-Zählmaß, 144
zusammengesetzter Poisson-Prozess, 152
Prokhorov
Kolmogorov-Prokhorov, 138

Radon-Nikodým
Satz von, 16
reguläre bedingte Verteilung, 23
reguläre bedingte Verteilungsfunktion, 23
Riemann
Lemma von Riemann–Lebesgue, 97

Schauder-Funktionen, *siehe* Wiener-Prozess
Slutsky
Satz von, *siehe* Konvergenz
Stationarität
im strengen Sinn, 127
im weiten Sinn, 127
intrinsisch, 128
stetig, 119
L^p, 119
fast sicher, 119
stochastisch, 119
Stetigkeitssatz, 63
Stiltjes, *siehe* Lebesgue
stochastische Äquivalenz, 118
stochastische Prozesse
Beispiele, *siehe* Beispiele stochastischer Prozesse
Cox-Prozess, *siehe* Cox-Prozess
Erneuerungsprozesse, *siehe* Erneuerungsprozesse
Pfade
äquivalent, 119
càdlàg, 120
Trajektorie, 110
Poisson-Prozess, *siehe* Poisson-Prozess
separabel, 119
Wiener-Prozess, *siehe* Wiener-Prozess
Zählprozesse, *siehe* Erneuerungsprozesse

Tailfunktion, 13
Trajektorie, *siehe* stochastische Prozesse

Ulam
Satz von Lomnicki-Ulam, 111
Umkehrformel, *siehe* charakteristische Funktion

Verteilung, *siehe* Zufallsvariable
Dirichlet–, 27
endlich dimensionale, *siehe* Zufallsvariable

Waldsche Identität, 138
Weierstrass
  Approximationssatz von
    Warhscheinlichkeitstheoretischer Beweis, 81
Wiener-Prozess, 117, 155–176
  d-dimensionale Brownsche Bewegung, 158
  Differenzierbarkeit des Wiener-Prozess, 174
  Haar-Funktionen, 158
  Maximum des Wiener-Prozess, 170
  Schauder-Funktionen, 160
  Transformierte Wiener-Prozesse, 171

Zentraler Grenzwertsatz, 135
zufällige Funktion, 108–116
  differnzierbar, 125
  Messbarkeit, 115
  stochastischer Prozess, 109
  symmetrisch, 111
  zufällige Folge, 109
  zufälliges Maß, 109
  Zufallsfeld, 109
zufälliges Element, 108
  zufällige abgeschlossene Menge, 108
  Zufallsvariable, 108
  Zufallsvektor, 108
Zufallsvariable, *siehe* Erneuerungsprozesse
  integrierbar, 7
  komplexwertige, 29
  unabhängig von der Zukunft, 138
  Verteilung, 111
    endlich-dimensionale, 111
Zufallsvektor
  mit sphärischen Konturen, 26
zylindrische σ-Algebra, 110