Wahrscheinlichkeitstheorie

Universität Ulm
Abteilung Stochastik

Vorlesungsskript
Prof. Dr. Volker Schmidt
Stand: Sommersemester 2005

Ulm, im Juli 2005

Inhaltsverzeichnis

1 Einleitung **5**

1.1 Grundbegriffe . . . 5

1.2 Endlich–dimensionale Verteilungen; Existenzsatz von Kolmogorow . . . 6

1.3 Regularität der Trajektorien . . . 7

1.4 Modifikationen von càdlàg Prozessen . . . 9

1.5 Stationarität und Unabhängigkeit; stationäre bzw. unabhängige Zuwächse . . . 10

2 Prozesse mit stückweise konstanten bzw. stetigen Trajektorien **11**

2.1 Zählprozesse; Erneuerungsprozesse . . . 11

2.1.1 Ergodensatz; Zentraler Grenzwertsatz . . . 11

2.1.2 Erneuerungsfunktion . . . 13

2.1.3 Verzögerte Erneuerungprozesse; stationäre Zuwächse . . . 16

2.2 Zählprozesse vom Poisson–Typ . . . 17

2.2.1 Homogener Poisson–Prozess . . . 17

2.2.2 Zusammengesetzte Poisson–Prozesse . . . 21

2.2.3 Bedingte Erwartung und bedingte Wahrscheinlichkeit . . . 22

2.2.4 Poissonsche Zählmaße; Cox–Prozesse; Simluationsalgorithmus . . . 26

2.3 Markow–Prozesse mit endlich vielen Zuständen . . . 29

2.3.1 Anfangsverteilung und Übergangsfunktion . . . 30

2.3.2 Übergangsintensitäten . . . 31

2.3.3 Kolmogorowsche Differentialgleichungen . . . 34

2.3.4 Eingebettete Markow–Ketten; Simulationsalgorithmus . . . 39

2.3.5 Stationäre Anfangsverteilungen . . . 42

2.4 Wiener–Prozess . . . 44

2.4.1 Vollständige Orthonormalsysteme im L_2; Haar–Funktionen und Schauder–Funktionen . . . 44

2.4.2 Eigenschaften normalverteilter Zufallsvariablen . . . 47

2.4.3 Konstruktion von Wiener–Prozessen; Simulationsalgorithmus . . . 49

2.4.4 Darstellung als Markow–Prozess bzw. Gauß–Prozess . . . 52

2.4.5 Verteilung des Maximums . . . 53

2.4.6 Weitere Verteilungs– und Pfadeigenschaften . . . 57

3 Lévy–Prozesse und Martingale **61**

3.1 Lévy–Prozesse . . . 61

3.1.1 Unbegrenzte Teilbarkeit . . . 61

3.1.2 Lévy–Chintschin–Darstellung . . . 64

3.1.3 Beispiele: Wiener–Prozess, zusammengesetzte Poisson–Prozesse, stabile Lévy–Prozesse . . . 68

3.1.4 Subordinatoren . . . 71
3.2 Martingale . . . 74
3.2.1 Filtrationen und Stoppzeiten . . . 74
3.2.2 Submartingale und Supermartingale; Beispiele . . . 77
3.2.3 Gleichgradige Integrierbarkeit . . . 81
3.2.4 Ungleichung von Doob . . . 83
3.2.5 Gestoppte Martingale . . . 84
3.2.6 Optionales Sampling–Theorem . . . 86
3.3 Anwendungsbeispiele . . . 88
3.3.1 Regeneration von Lévy–Prozessen zu Stoppzeiten . . . 88
3.3.2 Reflexionsprinzip des Wiener–Prozesses; Verteilung des Maximums . . . 90
3.3.3 Wiener–Prozesse mit negativer Drift . . . 92
3.3.4 Subordinatoren als Prozesse von Ersterreichungszeiten . . . 93

4 Stochastische Prozesse und Felder mit allgemeineren Indexmengen 96
4.1 Zählprozesse im $\mathbb{R}^1$ . . . 96
4.1.1 Homogener Poisson–Prozess . . . 97
4.1.2 Allgemeines Konstruktionsprinzip für Zählprozesse mit stationären Zuwächsen . . . 99
4.1.3 Erneuerungsprozesse mit stationären Zuwächsen . . . 102
4.1.4 Semi-Markowsche Zählprozesse; Simulationsalgorithmus . . . 103
4.2 Poissonsche Zählmaße im $\mathbb{R}^d$ . . . 105
4.2.1 Definition und elementare Eigenschaften . . . 105
4.2.2 Messbare Indizierung der Atome . . . 109
4.2.3 Simulationsalgorithmus; Akzeptanz– und Verwerfungsmethode . . . 111
4.2.4 Transformationssätze; radiale Simulation von homogenen Poisson–Prozessen . . . 112

Literatur

[1] Adler, R.J., Taylor, J.E. (2005)
Random Fields and Geometry
Buchmanuskript, Technion/Stanford University

[2] Applebaum, D. (2004)
Lévy Processes and Stochastic Calculus
Cambridge University Press, Cambridge

[3] Breiman, L. (1992)
Probability
SIAM, Philadelphia

[4] Grimmett, G., Stirzaker, D. (2001)
Probability and Random Processes
Oxford University Press, Oxford

[5] Kallenberg, O. (2001)
Foundations of Modern Probability
Springer-Verlag, New York

[6] Kingman, J.F.C. (1993)
Poisson Processes
Oxford University Press, Oxford

[7] Krylov, N.V. (2002)
Introduction to the Theory of Random Processes
AMS-Series: Graduate Studies in Mathematics, Providence

[8] Molchanov, I. (2005)
Theory of Random Sets
Springer-Verlag, Berlin

[9] Resnick, S.I. (1992)
Adventures in Stochastic Processes
Birkhäuser, Boston

[10] Rolski, T., Schmidli, H., Schmidt, V., Teugels, J. (1999)
Stochastic Processes for Insurance and Finance
J. Wiley & Sons, Chichester

[11] Samorodnitsky, G., Taqqu, M.S. (1994)
Stable Non–Gaussian Random Processes
Chapman & Hall, New York

[12] Sato, K.-I. (1999)
Lévy Processes and Infinite Divisibility
Cambridge University Press, Cambridge

[13] Schneider, R., Weil, W. (2000)
Stochastische Geometrie
Teubner-Verlag, Stuttgart

[14] Skorokhod, A.V. (2005)
Basic Principles and Applications of Probability Theory
Springer-Verlag, Berlin/Heidelberg

1 Einleitung

Die Vorlesung gibt eine Einführung in verschiedene Klassen stochastischer Prozesse. Schwerpunkte der Vorlesung sind:

- Zählprozesse (Erneuerungsprozesse, Poisson–Prozesse), Sprungprozesse
- Markow–Prozesse
- Wiener–Prozess (Brownsche Bewegung)
- Lévy–Prozesse
- Martingale
- zufällige Felder
- zufällige Maße

Insbesondere werden analytische, geometrische und asymptotische Eigenschaften dieser Modelle diskutiert, die die Grundlage für statistische Methoden und Simulationsalgorithmen bilden.

Dabei werden wir in dieser Vorlesung vor allem stochastische Prozesse „mit kontinuierlicher Zeit" betrachten, d.h., überabzählbare Familien von Zufallsvariablen.

Die Eigenschaften einzelner Zufallsvariablen bzw. von (endlichen oder abzählbar unendlichen) Folgen von Zufallsvariablen wurden ausführlich in der Vorlesung „Wahrscheinlichkeitsrechnung" behandelt, vgl. das Skript zur Vorlesung „Wahrscheinlichkeitsrechnung" im WS 03/04. Eine elektronische Fassung dieses Skriptes ist im Internet verfügbar:

http://www.mathematik.uni-ulm.de/stochastik/lehre/ws03_04/wr/skript/skript.html

Verweise auf dieses Vorlesungsmanuskript werden wir mit dem Zusatz „WR" vor der Nummer der zitierten Abschnitte, Lemmata, Theoreme, Korollare bzw. Formeln kennzeichnen.

Gelegentlich werden wir auch auf Abschnitte, Lemmata, Theoreme, Korollare bzw. Formeln aus dem Skript zur Vorlesung „Markow–Ketten und Monte–Carlo–Simulation" im SS 2003 verweisen, wobei diese Verweise dann mit dem Zusatz „MK" gekennzeichnet werden.

1.1 Grundbegriffe

Stochastische Prozesse sind Familien $\{X_t,\, t \in I\}$ von Zufallsvariablen $X_t : \Omega \to E$, die über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ definiert sind, wobei $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum sein kann, der oft nicht näher spezifiziert wird.

Die *Indexmenge* I kann eine beliebige Menge sein. Der Bildraum E der Zufallsvariablen X_t kann ebenfalls eine beliebige Menge sein, die lediglich mit einer σ–Algebra $\mathcal{B}(E)$ von Teilmengen von E versehen ist, d.h., im allgemeinen ist $(E, \mathcal{B}(E))$ ein beliebiger Messraum, der sogenannte *Zustandsraum* des stochastischen Prozesses $\{X_t\}$.

Man sagt dann, dass $\{X_t\}$ ein stochastischer Prozess in I mit Werten in $(E, \mathcal{B}(E))$ ist. Dabei unterscheidet man verschiedene Arten von Indexmengen I.

- Falls I eine abzählbare Menge ist, zum Beispiel $I = \{0, 1, 2, \ldots\}$, dann wird $\{X_t\}$ eine *zufällige Folge* bzw. ein stochastischer Prozess mit *diskreter Zeit* genannt.

- Typische Beispiele von (überabzählbaren) Indexmengen $I \subset \mathbb{R}$ für stochastische Prozesse mit *kontinuierlicher Zeit* sind $I = (a, b)$ (beschränktes Interval), $I = [0, \infty)$ (nichtnegative Halbachse) bzw. $I = \mathbb{R}$ (reelle Achse).
- Falls $I \subset \mathbb{R}^d$ eine Teilmenge des d–dimensionalen euklidischen Raumes ist, wobei $d \geq 2$, dann wird $\{X_t\}$ ein *zufälliges Feld* genannt.
- Falls $I \subset \mathcal{B}(\mathbb{R}^d)$ eine Familie von Teilmengen des $\mathbb{R}^d$ ist, dann sagt man, dass $\{X_t\}$ ein *mengen–indizierter Prozess* ist. Wenn $I \subset \mathcal{B}(\mathbb{R}^d)$ eine σ–Algebra ist, $\{X_t\}$ nur nichtnegative Werte annimmt und σ–additiv ist, dann wird $\{X_t\}$ ein *zufälliges Maß* genannt.

Wenn $I \subset \mathbb{R}$ gilt, dann wird für jedes $\omega \in \Omega$ die Funktion $\{X_t(\omega),\, t \in I\}$ eine *Trajektorie* bzw. ein *Pfad* des stochastischen Prozesses $\{X_t\}$ genannt.

Im ersten Teil der Vorlesung werden wir vorwiegend stochastische Prozesse mit $I \subset [0, \infty)$ und $E \subset \mathbb{R}$ betrachten.

1.2 Endlich–dimensionale Verteilungen; Existenzsatz von Kolmogorow

Sei $\{X_t,\, t \in I\}$ ein stochastischer Prozess mit $I = [0, \infty)$ und $E = \mathbb{R}$.

Definition Für jedes $n = 1, 2, \ldots$ und für jedes n–Tupel $t_1, \ldots, t_n \geq 0$ von Indizes wird die Mengenfunktion $P_{t_1,\ldots,t_n} : \mathcal{B}(\mathbb{R}^n) \to [0, 1]$ mit

$$P_{t_1,\ldots,t_n}(B_1 \times \ldots \times B_n) = P(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n)\,, \qquad B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}) \tag{1}$$

endlich–dimensionale Verteilung von $\{X_t\}$ genannt.

Beachte Es ist klar, dass die endlich–dimensionalen Verteilungen von stochastischen Prozessen die folgenden *Konsistenzeigenschaften* besitzen.

- Sei $n = 1, 2, \ldots$ eine beliebige natürliche Zahl, und sei π eine beliebige Permutation von $(1, \ldots, n)$.
- Dann gilt
$$P_{t_1,\ldots,t_n}(B_1 \times \ldots \times B_n) = P_{t_{\pi(1)},\ldots,t_{\pi(n)}}(B_{\pi(1)} \times \ldots \times B_{\pi(n)}) \tag{2}$$
und
$$P_{t_1,\ldots,t_n,t_{n+1}}(B_1 \times \ldots \times B_n \times \mathbb{R}) = P_{t_1,\ldots,t_n}(B_1 \times \ldots \times B_n) \tag{3}$$
für jedes $(n+1)$–Tupel $t_1, \ldots, t_{n+1} \geq 0$ und für beliebige Borel–Mengen $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$.

Wir nehmen nun umgekehrt an, dass $\{P_{t_1,\ldots,t_n},\, t_1, \ldots, t_n \geq 0,\, n = 1, 2, \ldots\}$ eine beliebige Familie von Wahrscheinlichkeitsmaßen ist, die den Bedingungen (2) und (3) genügen.

Der folgende (Existenz-) Satz von Kolmogorow ist eine wichtige Fundamentalaussage in der Theorie stochastischer Prozesse.

Theorem 1.1 *Für jede Familie von Wahrscheinlichkeitsmaßen $\{P_{t_1,\ldots,t_n},\, t_1, \ldots, t_n \geq 0,\, n = 1, 2, \ldots\}$, die den Bedingungen (2) und (3) genügen, gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und einen stochastischen Prozess $\{X_t,\, t \geq 0\}$ über diesem Wahrscheinlichkeitsraum, so dass $\{P_{t_1,\ldots,t_n}\}$ die endlich–dimensionalen Verteilungen von $\{X_t,\, t \geq 0\}$ sind.*

Der *Beweis* von Theorem 1.1 beruht auf Standard–Techniken der Maßtheorie. Wir erwähnen hier nur die Grundidee des Kolmogorowschen Konstruktionsprinzips. Ein detaillierter Beweis kann zum Beispiel in Kallenberg (2001), S. 115–116 nachgelesen werden.

Dabei wird bei der Wahl eines geeigneten Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$ bzw. des zugehörigen stochastischen Prozesses $\{X_t\}$ wie folgt vorgegangen.

- Sei $\Omega = E^I$, d.h., in unserem Fall ist $\Omega = \mathbb{R}^{[0,\infty)}$ die Familie aller Abbildungen $\omega : [0,\infty) \to \mathbb{R}$ von $[0,\infty)$ in die Menge der reellen Zahlen $\mathbb{R}$.
- $\mathcal{F}$ sei die kleinste σ–Algebra von Teilmengen von Ω, die sämtliche *Zylinder–Mengen* der Form

$$A^B_{t_1,\ldots,t_n} = \{\omega \in \Omega : \ (\omega(t_1),\ldots,\omega(t_n)) \in B\} \tag{4}$$

  enthält für jedes beliebige $B \in \mathcal{B}(\mathbb{R}^n)$.
- Mit Techniken der Maßtheorie kann man zeigen, dass es ein (eindeutig bestimmtes) Wahrscheinlichkeitsmaß $P : \mathcal{F} \to [0,1]$ gibt, für das $P(A^B_{t_1,\ldots,t_n}) = P_{t_1,\ldots,t_n}(B)$ für beliebige $n = 1,2,\ldots$ und $t_1,\ldots,t_n \geq 0$ sowie $B \in \mathcal{B}(\mathbb{R}^n)$ gilt.
- Dann stimmen die endlich–dimensionalen Verteilungen des „Koordinaten-Prozesses" $\{X_t,\, t \geq 0\}$, der gegeben ist durch $X_t(\omega) = \omega(t)$ für beliebige $t \geq 0$ und $\omega \in \Omega$, mit den Wahrscheinlichkeitsmaßen $\{P_{t_1,\ldots,t_n}\}$ überein.

Beachte

- Der Existenzsatz von Kolmogorow gilt auch für beliebige Indexmengen I und für beliebige Zustandsräume $(E, \mathcal{B}(E))$, d.h., für Familien von Wahrscheinlichkeitsmaßen $\{P_{t_1,\ldots,t_n},\, t_1,\ldots,t_n \in I\}$ über $(E^n, \mathcal{B}(E^n))$ mit $n = 1,2,\ldots$, die den Bedingungen (2) und (3) genügen.
- Der stochastische Koordinaten–Prozess $\{X_t,\, t \in I\}$ wird dann auf völlig analoge Weise über dem Wahrscheinlichkeitsraum $(E^I, \mathcal{F}(E^I), P)$ konstruiert, wobei $\mathcal{F}(E^I)$ die kleinste σ–Algebra ist, die alle Zylinger–Mengen von E^I enthält.

1.3 Regularität der Trajektorien

Das folgende Beispiel zeigt, dass der gemäß Theorem 1.1 zu vorgegebenen Wahrscheinlichkeitsmaßen $\{P_{t_1,\ldots,t_n}\}$ stets existierende stochastische Prozess, der diese Wahrscheinlichkeitsmaße als endlich–dimensionale Verteilungen hat, im allgemeinen *nicht* eindeutig bestimmt ist.

- Sei U eine (gleichverteilte) Zufallsvariable über dem Wahrscheinlichkeitsraum

$$(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \nu_1)$$

  mit $U(\omega) = \omega$, wobei ν_1 das Lebesgue–Maß bezeichnet.
- Die stochastischen Prozesse $\{X_t,\, t \in [0,1]\}$ und $\{Y_t,\, t \in [0,1]\}$ seien gegeben durch

$$X_t = 0 \quad \text{für jedes } t \in [0,1] \qquad \text{bzw.} \qquad Y_t = \begin{cases} 1\,, & \text{falls } U = t, \\ 0\,, & \text{sonst.} \end{cases}$$

- Es ist klar, dass die Prozesse $\{X_t\}$ und $\{Y_t\}$ die gleichen endlich–dimensionalen Verteilungen besitzen, weil $P(U = t) = 0$ für jedes $t \in [0,1]$ gilt.
- Dennoch sind $\{X_t\}$ und $\{Y_t\}$ zwei verschiedene stochastische Prozesse, denn es gilt beispielsweise

$$P(X_t = 0 \text{ für jedes } t \in [0,1]) = 1 \qquad \text{und} \qquad P(Y_t = 0 \text{ für jedes } t \in [0,1]) = 0\,.$$

Definition

- Stochastische Prozesse $\{X_t\}$ und $\{Y_t\}$, deren endlich–dimensionale Verteilungen übereinstimmen, werden *Versionen* einunddesselben stochastischen Phänomens genannt.

- Wenn für die stochastischen Prozesse $\{X_t\}$ und $\{Y_t\}$

$$P(X_t = Y_t) = 1 \qquad \text{für jedes } t \in I$$

gilt, dann wird $\{Y_t\}$ eine *Modifikation* von $\{X_t\}$ (und umgekehrt) genannt.

Ein weiteres Problem besteht darin, dass Teilmengen von $\Omega = \mathbb{R}^{[0,\infty)}$, die durch Eigenschaften der Trajektorien gegeben sind, im allgemeinen nicht messbar bezüglich der σ–Algebra $\mathcal{F}$ sein müssen, die durch die Zylinder–Mengen in (4) erzeugt wird.

- Ein Beispiel hierfür ist die Teilmenge

$$A = \{\omega \in \Omega : \ X_t(\omega) = 0 \text{ für jedes } t \in [0,1]\}$$

von Ω, die der (überabzählbare) Durchschnitt $A = \bigcap_{t\in[0,1]}\{X_t = 0\}$ der Ereignisse $\{X_t = 0\} \in \mathcal{F}$ ist und die deshalb im allgemeinen nicht messbar sein muss.

- Dieses Problem lässt sich dadurch lösen, dass man nur solche stochastischen Prozesse betrachtet, deren Trajektorien gewisse Regularitätseigenschaften besitzen.
- Falls alle Trajektorien des Prozesses $\{X_t\}$ stetige Funktionen sind, dann wird $\{X_t\}$ ein *stetiger Prozess* genannt.
  - In diesem Fall lässt sich die in oben betrachtete Menge $A \subset \Omega$ als der Durchschnitt von abzählbar vielen Mengen aus $\mathcal{F}$ darstellen.
  - Denn es gilt dann

$$A = \bigcap_{t \in [0,1],\ t \text{ rational}} \{\omega \in \Omega : \ X_t(\omega) = 0\},$$

d.h. $A \in \mathcal{F}$.

Die im folgenden Kapitel 2 betrachteten klassischen Beispiele von stochastischen Prozessen zeigen jedoch, dass es nicht ausreicht, nur Prozesse mit stetigen Trajektorien zu betrachten.

- Deshalb wird im allgemeinen eine größere Klasse von stochastischen Prozessen betrachtet, die alle praktisch relevanten Beispiele umfasst und deren Trajektorien gleichzeitig gute Regularitätseigenschaften besitzen.
- Dabei wird nur gefordert, dass die Trajektorien überall rechtsstetig sind und linksseitige Grenzwerte besitzen. Hinreichende Bedingungen hierfür sind durch das folgende Theorem gegeben.

Theorem 1.2 *Sei $\{X_t,\ t \geq 0\}$ ein beliebiger reellwertiger Prozess, und sei $D \subset [0,\infty)$ eine abzählbare Menge, die dicht in $[0,\infty)$ liegt. Falls*

(i) *$\{X_t\}$ rechtsstetig in Wahrscheinlichkeit ist, d.h., für beliebige $t \geq 0$ und $h \downarrow 0$ gilt $X_{t+h} \xrightarrow{\mathrm{P}} X_t$, wobei $\xrightarrow{\mathrm{P}}$ die Konvergenz in Wahrscheinlichkeit bezeichnet,*

(ii) *die Trajektorien von $\{X_t\}$ für jedes $t \in D$ mit Wahrscheinlichkeit 1 endliche rechts- und linksseitige Grenzwerte besitzen, d.h., für jedes $t \in D$ existieren die Grenzwerte $\lim_{h\downarrow 0} X_{t+h}$ und $\lim_{h\uparrow 0} X_{t+h}$ mit Wahrscheinlichkeit 1,*

dann gibt es eine Version $\{Y_t,\ t \geq 0\}$ von $\{X_t,\ t \geq 0\}$, so dass mit Wahrscheinlichkeit 1

(a) *die Trajektorien von $\{Y_t\}$ rechtsstetig sind und*

(b) *linksseitige Grenzwerte besitzen, d.h., $\lim_{h\uparrow 0} Y_{t+h}$ existiert für jedes $t \geq 0$.*

Der *Beweis* von Theorem 1.2 geht über den Rahmen dieser Vorlesung hinaus und wird deshalb weggelassen; er kann beispielsweise in Breiman (1992), S. 300 nachgelesen werden.

Theorem 1.3 (A.N. Kolmogorow) *Sei $\{X_t,\, t \geq 0\}$ ein beliebiger reellwertiger Prozess, so dass es Konstanten $a, b > 0$ und $c < \infty$ gibt mit*

$$\mathbb{E}\left(|X_t - X_s|^a\right) \leq c\,|t-s|^{1+b} \qquad \forall\, t, s \geq 0\,.$$

Dann gibt es eine stetige Modifikation von $\{X_t\}$.

Einen *Beweis* von Theorem 1.3 kann man zum Beispiel in Krylov (2002), S. 20–21 nachlesen.

1.4 Modifikationen von càdlàg Prozessen

In diesem und in den nachfolgenden Abschnitten setzen wir stets (o.B.d.A.) voraus, dass $(\Omega, \mathcal{F}, P)$ ein *vollständiger Wahrscheinlichkeitsraum* ist, d.h.,

- sämtliche „Nullmengen" des Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$ gehören zu $\mathcal{F}$ bzw. (äquivalent hierzu)
- für $C \subset \Omega$ gilt $C \in \mathcal{F}$, falls es Teilmengen $A, B \in \mathcal{F}$ gibt, so dass $A \subset C \subset B$ und $P(B \setminus A) = 0$.

Beachte

- Zur Erinnerung: Wenn für die stochastischen Prozesse $\{X_t,\, t \geq 0\}$ und $\{Y_t,\, t \geq 0\}$

$$P(X_t = Y_t) = 1 \qquad \text{für jedes } t \geq 0 \tag{5}$$

  gilt, dann wird $\{Y_t\}$ eine *Modifikation* von $\{X_t\}$ (und umgekehrt) genannt.
- Manchmal wird eine verschärfte Version der Bedingung (5) betrachtet. Dabei sagt man, dass die stochastischen Prozesse $\{X_t\}$ und $\{Y_t\}$ *ununterscheidbar* sind, wenn

$$P(X_t = Y_t \text{ für jedes } t \geq 0) = 1\,.$$

- Außerdem sagt man, dass der stochastische Prozess $\{X_t,\, t \geq 0\}$ *càdlàg* ist,
  - wenn fast alle Trajektorien von $\{X_t\}$ rechtsstetige Funktionen sind, die linksseitige Grenzwerte besitzen,
  - wobei „càdlàg" eine Abkürzung der französischen Sprechweise „continu à droite, limites à gauche" ist.
- Alle stochastischen Prozesse mit stückweise konstanten bzw. stetigen Trajektorien, die in Kapitel 2 betrachtet wurden, sind càdlàg.

Theorem 1.4 *Seien $\{X_t,\, t \geq 0\}$ und $\{Y_t,\, t \geq 0\}$ stochastische Prozesse über $(\Omega, \mathcal{F}, P)$, so dass $\{X_t\}$ und $\{Y_t\}$ mit Wahrscheinlichkeit 1 rechtsstetige Trajektorien besitzen. Außerdem sei $\{Y_t\}$ eine Modifikation von $\{X_t\}$. Dann sind die Prozesse $\{X_t\}$ und $\{Y_t\}$ ununterscheidbar.*

Beweis

- Seien $A, A' \subset \Omega$ diejenigen Teilmengen von Ω, für die die Trajektorien von $\{X_t\}$ bzw. $\{Y_t\}$ nicht rechtsstetig sind, wobei dann $P(A) = P(A') = 0$.
- Außerdem sei $C_t = \{\omega \in \Omega :\, X_t(\omega) \neq Y_t(\omega)\}$ für jedes $t \geq 0$ und $C = \bigcup_{t \in \mathbb{Q}^+} C_t$, wobei $\mathbb{Q}^+ = \mathbb{Q} \cap [0, \infty)$ die Menge der rationalen Zahlen in $[0, \infty)$ bezeichnet.

- Weil $\{Y_t\}$ eine Modifikation von $\{X_t\}$ ist, gilt $P(C) = 0$ und deshalb auch
$$P(C') = 0\,, \tag{6}$$
wobei $C' = C \cup A \cup A'$.
- Somit ist klar, dass $X_t(\omega) = Y_t(\omega)$ für beliebige $t \in \mathbb{Q}^+$ und $\omega \notin C'$.

- Sei nun $t \geq 0$ eine beliebige (nichtnegative) Zahl und $\{t_n\}$ eine Folge rationaler Zahlen mit $t_n > t_{n+1} \geq t$ und $t_n \downarrow t$.
  - Weil $X_{t_n}(\omega) = Y_{t_n}(\omega)$ für beliebige $n \geq 1$ und $\omega \notin C'$, ergibt sich nun, dass
$$X_t(\omega) = \lim_{n\to\infty} X_{t_n}(\omega) = \lim_{n\to\infty} Y_{t_n}(\omega) = Y_t(\omega)$$
für beliebige $t \geq 0$ und $\omega \notin C'$.
  - Wegen (6) folgt hieraus, dass die Prozesse $\{X_t\}$ und $\{Y_t\}$ ununterscheidbar sind. □

Aus Theorem 1.4 ergibt sich insbesondere das folgende Resultat.

Korollar 1.1 *Die Prozesse $\{X_t,\, t \geq 0\}$ und $\{Y_t,\, t \geq 0\}$ seien càdlàg, wobei $\{Y_t\}$ eine Modifikation von $\{X_t\}$ sei. Dann sind die Prozesse $\{X_t\}$ und $\{Y_t\}$ ununterscheidbar.*

1.5 Stationarität und Unabhängigkeit; stationäre bzw. unabhängige Zuwächse

Von großer Bedeutung sind stochastische Prozesse, für die die in (1.2.1) eingeführten endlich–dimensionalen Verteilungen gewisse Invarianzeigenschaften besitzen. Zwei wichtige Klassen solcher stochastischen Prozesse sind wie folgt definiert.

Definition Seien $n = 1, 2, \ldots$, $0 \leq t_0 < t_1 < \ldots < t_n$ und $h \geq 0$ beliebige Zahlen.

- Der stochastische Prozess $\{X_t, t \geq 0\}$ heißt *stationär*, wenn die Verteilungen der Zufallsvektoren $(X_{t_1+h}, \ldots, X_{t_n+h})$ nicht von h abhängen.
- Man sagt, dass der stochastische Prozess $\{X_t, t \geq 0\}$ ein *Prozess mit stationären Zuwächsen* ist, wenn die Verteilungen der Zufallsvektoren $(X_{t_1+h} - X_{t_0+h}, \ldots, X_{t_n+h} - X_{t_{n-1}+h})$ nicht von h abhängen.

Eine andere wichtige Klasse von stochastischen Prozessen, die wir in dieser Vorlesung ausführlich behandeln werden, ist wie folgt definiert.

Definition Man sagt, dass der stochastische Prozess $\{X_t, t \geq 0\}$ ein *Prozess mit unabhängigen Zuwächsen* ist, wenn die Zufallsvariablen X_{t_0}, $X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ für alle $n = 1, 2, \ldots$ und $0 \leq t_0 < t_1 < \ldots < t_n$ unabhängig sind.

Gelegentlich werden wir auch den Begriff der Unabhängigkeit von zwei (oder mehreren) stochastischen Prozessen bzw. den Begriff der Unabhängigkeit eines Prozesses von einer σ–Algebra benötigen.

Definition Seien $\{X_t, t \geq 0\}$ und $\{Y_t, t \geq 0\}$ zwei beliebige stochastische Prozesse über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und sei $\mathcal{G} \subset \mathcal{F}$ eine beliebige Teil–σ–Algebra von $\mathcal{F}$.

- Man sagt dann, dass $\{X_t\}$ und $\{Y_t\}$ *unabhängig* sind, wenn die Zufallsvektoren $(X_{t_1}, \ldots, X_{t_n})$ und $(Y_{s_1}, \ldots, Y_{s_m})$ für beliebige (endliche) Folgen von nichtnegativen Zahlen $t_1, \ldots, t_n$ bzw. $s_1, \ldots, s_m$ unabhängig sind.
- Außerdem sagt man, dass $\{X_t\}$ von $\mathcal{G}$ *unabhängig* ist, wenn für jede Folge $t_1, \ldots, t_n$ von nichtnegativen Zahlen und für beliebige $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ und $A \in \mathcal{G}$ die Ereignisse $X_{t_1}^{-1}(B_1) \cap \ldots \cap X_{t_n}^{-1}(B_n)$ und A unabhängig sind, wobei $X_t^{-1}(B) = \{\omega \in \Omega :\, X_t(\omega) \in B\}$.

2 Prozesse mit stückweise konstanten bzw. stetigen Trajektorien

In den ersten drei Abschnitten dieses Kapitels betrachten wir Beispiele von stochastischen Prozessen, deren Pfade stückweise konstante (Sprung–) Funktionen sind. Dagegen ist der in Abschnitt 2.4 betrachtete Wiener–Prozess ein Beispiel von stochastischen Prozessen mit stetigen Trajektorien.

2.1 Zählprozesse; Erneuerungsprozesse

Sei $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum, und sei $S_1, S_2, \ldots : \Omega \to [0, \infty)$ eine beliebige Folge von nichtnegativen Zufallsvariablen mit $0 \leq S_1 \leq S_2 \leq \ldots$.

In der Versicherungsmathematik können die Zufallsvariablen S_n Zeitpunkte modellieren, zu denen Schäden eines bestimmten Typs eintreten. Dies können beispielsweise Schäden sein, die durch Naturkatastrophen wie Erdbeben oder schwere Stürme verursacht werden.

Definition

1. Der stochastische Prozess $\{N_t, t \geq 0\}$ mit
$$N_t = \sum_{k=1}^{\infty} \mathbb{I}(S_k \leq t) \tag{1}$$
wird *Zählprozess* genannt, wobei $\mathbb{I}(A)$ der *Indikator* des Ereignisses $A \in \mathcal{F}$ ist, d.h. $\mathbb{I}(A)(\omega) = 1$, falls $\omega \in A$, und $\mathbb{I}(A)(\omega) = 0$, falls $\omega \notin A$.
2. Sei $T_1, T_2, \ldots : \Omega \to [0, \infty)$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen, die nur nichtnegative Werte annehmen.
   - Dann wird die Folge $\{S_n, n = 0, 1, \ldots\}$ mit $S_0 = 0$ und $S_n = T_1 + \ldots + T_n$ für $n \geq 1$ ein *Erneuerungspunktprozess* genannt, wobei S_n der n-te *Erneuerungszeitpunkt* heißt.
   - Der in (1) gegebene stochastische Prozess $\{N_t, t \geq 0\}$ wird in diesem Fall *Erneuerungszählprozess* bzw. kurz *Erneuerungsprozess* genannt.
   - Dabei wird stets vorausgesetzt, dass die „Zwischenankunftszeiten" T_n nicht mit Wahrscheinlichkeit 1 gleich Null sind, d.h., es gelte $P(T_n > 0) > 0$ für jedes $n \geq 1$.

2.1.1 Ergodensatz; Zentraler Grenzwertsatz

Aus dem starken Gesetz der großen Zahlen (vgl. Theorem WR–5.15) ergibt sich die folgende asymptotische Eigenschaft des Erneuerungsprozesses $\{N_t, t \geq 0\}$; vgl. auch Beispiel 6 in Abschnitt WR–5.2.3. Aussagen dieses Typs werden in der Literatur *individueller Ergodensatz* genannt.

Theorem 2.1 *Sei $0 < \mu < \infty$, wobei $\mu = \mathbb{E}\, T_n$ den Erwartungswert der Zwischenankunftszeiten T_n bezeichnet. Dann gilt mit Wahrscheinlichkeit 1, dass*
$$\lim_{t \to \infty} \frac{N_t}{t} = \frac{1}{\mu}\,. \tag{2}$$

Beweis

- Für jedes $t > 0$ gilt
$$P(N_t < \infty) = 1\,, \tag{3}$$
denn aus Theorem WR–5.15 folgt, dass $\lim_{n \to \infty} S_n = \infty$ mit Wahrscheinlichkeit 1.
- Außerdem ist $N_t(\omega)$ für jedes $\omega \in \Omega$ monoton nichtfallend in $t \geq 0$, d.h., der Grenzwert $\lim_{t \to \infty} N_t(\omega)$ existiert für jedes $\omega \in \Omega$.

- Darüber hinaus gilt mit Wahrscheinlichkeit 1

$$\lim_{t\to\infty} N_t = \infty\,, \tag{4}$$

weil

$$\begin{aligned} P\Big(\lim_{t\to\infty} N_t < \infty\Big) &= \lim_{k\to\infty} P\Big(\lim_{t\to\infty} N_t < k\Big) \\ &= \lim_{k\to\infty}\lim_{t\to\infty} P(N_t < k) \\ &= \lim_{k\to\infty}\lim_{t\to\infty} P(S_k > t) \\ &\le \lim_{k\to\infty}\lim_{t\to\infty}\Big(P(T_1 > \frac{t}{k}) + \ldots + P(T_k > \frac{t}{k})\Big) = 0\,. \end{aligned}$$

- Aus Theorem WR–5.15 folgt also, dass mit Wahrscheinlichkeit 1

$$\lim_{t\to\infty} \frac{S_{N_t}}{N_t} = \mu\,. \tag{5}$$

- Außerdem gilt für beliebige $t \ge 0$ und $n \in \mathbb{N} = \{1, 2, \ldots\}$

$$\{N_t = n\} = \{S_n \le t < S_{n+1}\}\,. \tag{6}$$

- Folglich gilt $S_{N_t} \le t < S_{N_t+1}$ bzw.

$$\frac{S_{N_t}}{N_t} \le \frac{t}{N_t} \le \frac{S_{N_t+1}}{N_t+1}\frac{N_t+1}{N_t} \qquad \forall t \ge T_1\,.$$

- Hieraus und aus (5) ergibt sich nun die Gültigkeit von (2). □

Außerdem gilt der folgende *zentrale Grenzwertsatz*; vgl. auch Beispiel 4 in Abschnitt WR–5.3.2.

Theorem 2.2 *Sei* $0 < \mu = \mathbb{E}\,T_1 < \infty$, $\mathbb{E}\,T_1^2 < \infty$ *und* $\mathrm{Var}\,T_1 > 0$. *Für jedes* $x \in \mathbb{R}$ *gilt dann*

$$\lim_{t\to\infty} P\Big(\frac{N_t - t\mu^{-1}}{\sqrt{ct}} \le x\Big) = \Phi(x)\,, \tag{7}$$

wobei $c = \mathrm{Var}\,T_1/\mu^3$ *und* $\Phi : \mathbb{R} \to [0,1]$ *die Verteilungsfunktion der Standardnormalverteilung ist.*

Beweis

- Aus dem Beweis des zentralen Grenzwertsatzes für Summen von unabhängigen und identisch verteilten Zufallsvariablen (vgl. den Beweis von Theorem WR–5.16) ergibt sich, dass die Konvergenz

$$\lim_{n\to\infty} P\Big(\frac{S_n - n\mu}{\sqrt{n\mathrm{Var}\,T_1}} \le x\Big) = \Phi(x) \tag{8}$$

gleichmäßig in $x \in \mathbb{R}$ erfolgt.

- Sei $\lfloor a \rfloor$ der ganzzahlige Anteil von $a \ge 0$. Mit der Schreibweise $m(t) = \lfloor x\sqrt{ct} + t\mu^{-1} \rfloor$ gilt dann für jedes hinreichend große t

$$\begin{aligned} P\Big(\frac{N_t - t\mu^{-1}}{\sqrt{ct}} \le x\Big) &= P(N_t \le x\sqrt{ct} + t\mu^{-1}) \\ &= P(S_{m(t)+1} > t) \\ &= P\Big(\frac{S_{m(t)+1} - \mu(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}\,T_1}} > \frac{t - \mu(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}\,T_1}}\Big)\,. \end{aligned}$$

- Also gilt

$$\begin{aligned}
&\Big|P\Big(\frac{N_t - t\mu^{-1}}{\sqrt{ct}} \le x\Big) - \Phi(x)\Big| \\
\le\ & \Big|P\Big(\frac{S_{m(t)+1} - \mu(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}\,T_1}} > \frac{t - \mu(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}\,T_1}}\Big) - \Big(1 - \Phi\Big(\frac{t - \mu(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}\,T_1}}\Big)\Big)\Big| \\
&+\Big|\Big(1 - \Phi\Big(\frac{t - \mu(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}\,T_1}}\Big)\Big) - \Phi(x)\Big| .
\end{aligned}$$

- Außerdem gilt für jedes $x \in \mathbb{R}$

$$\lim_{t\to\infty} m(t) = \lim_{t\to\infty} \lfloor x\sqrt{ct} + t\mu^{-1}\rfloor = \infty .$$

- Wegen (8) und wegen der Symmetrieeigenschaft

$$1 - \Phi(-x) = \Phi(x) \qquad \forall x \in \mathbb{R}$$

der Verteilungsfunktion Φ der $\mathrm{N}(0,1)$–Verteilung genügt es somit noch zu zeigen, dass

$$\lim_{t\to\infty} \frac{t - \mu m(t)}{\sqrt{m(t)\mathrm{Var}\,T_1}} = -x . \tag{9}$$

- Weil $m(t) = x\sqrt{ct} + t\mu^{-1} + \varepsilon(t)$ mit $0 \le |\varepsilon(t)| < 1$ für jedes $t \ge 0$ gilt

$$\begin{aligned}
\frac{t - \mu m(t)}{\sqrt{m(t)\mathrm{Var}\,T_1}} &= \frac{t - \mu x\sqrt{ct} - t - \mu\varepsilon(t)}{\sqrt{m(t)\mathrm{Var}\,T_1}} \\
&= -x\,\frac{\sqrt{\mu^{-1}\mathrm{Var}\,T_1 t}}{\sqrt{\mu^{-1}\mathrm{Var}\,T_1 t + x\mathrm{Var}\,T_1\sqrt{ct} + \mathrm{Var}\,T_1\varepsilon(t)}} - \frac{\mu\varepsilon(t)}{\sqrt{m(t)\mathrm{Var}\,T_1}} ,
\end{aligned}$$

wobei der erste Summand gegen $-x$ und der zweite Summand gegen 0 strebt für $t \to \infty$.

- Damit ist die Behauptung (7) bewiesen. □

2.1.2 Erneuerungsfunktion

Definition Die Funktion $H : [0,\infty) \to [0,\infty)$, wobei $H(t) = \mathbb{E}\,N_t$ die erwartete Anzahl von Erneuerungszeitpunkten im Intervall $[0,t]$ bezeichnet, wird *Erneuerungsfunktion* von $\{N_t\}$ genannt. Aus (1) ergibt sich, dass

$$H(t) = \mathbb{E}\sum_{n=1}^{\infty} \mathbb{1}(S_n \le t) = \sum_{n=1}^{\infty} \mathbb{E}\,\mathbb{1}(S_n \le t) = \sum_{n=1}^{\infty} F^{*n}(t) , \tag{10}$$

wobei $F : [0,\infty) \to [0,1]$ die Verteilungsfunktion der Zwischenankunftszeiten T_n und F^{*n} die n-te Faltungspotenz von F bezeichnet. Dabei gilt $F^{*n} = F^{*(n-1)} * F$ und $(F * F)(t) = \int_{-\infty}^{t} F(t-s)\,\mathrm{d}F(s)$ mit $F^{*1}(t) = F(t)$ und $F^{*0} = \mathbb{1}([0,\infty))$, d.h., $F^{*0}(t) = 1$ für $t \ge 0$ und $F^{*0}(t) = 0$ für $t < 0$.

Beachte

1. Es ist klar, dass die Erneuerungsfunktion $H : [0,\infty) \to [0,\infty)$ monoton wachsend ist. Außerdem ist es nicht schwierig zu zeigen, dass $H(t) < \infty$ für jedes $t \ge 0$ gilt.
2. Allerdings ist es nur in wenigen Spezialfällen möglich, einfache geschlossene Formeln für die Erneuerungsfunktion H herzuleiten. Die Laplace–Stieltjes–Transformierte $\hat{l}_H(s) = \int_0^\infty \mathrm{e}^{-sx}\,\mathrm{d}H(x)$ von H lässt sich jedoch immer durch die Laplace–Stieltjes–Transformierte von F ausdrücken.

Theorem 2.3 *Für jedes $s > 0$ gilt*

$$\hat{l}_H(s) = \frac{\hat{l}_F(s)}{1 - \hat{l}_F(s)}\,. \tag{11}$$

Beweis Aus (10) ergibt sich, dass

$$\hat{l}_H(s) = \int_0^\infty \mathrm{e}^{-sx}\,\mathrm{d}(\sum_{n=1}^\infty F^{*n}(x)) = \sum_{n=1}^\infty \int_0^\infty \mathrm{e}^{-sx}\,\mathrm{d}F^{*n}(x) = \sum_{n=1}^\infty \hat{l}_{F^{*n}}(s) = \sum_{n=1}^\infty (\hat{l}_F(s))^n = \frac{\hat{l}_F(s)}{1 - \hat{l}_F(s)}\,,$$

wobei die geometrische Reihe $\sum_{n=1}^\infty (\hat{l}_F(s))^n$ konvergiert, weil $\hat{l}_F(s) < 1$ für jedes $s > 0$ gilt. □

Aus Theorem 2.1 ergibt sich die Vermutung, dass die Erneuerungsfunktion $H(t) = \mathbb{E}\,N_t$ ein ähnliches asymptotisches lineares Verhalten hat, wie es in (2) für den Erneuerungsprozess $\{N_t\}$ gezeigt wurde. Um dies zu zeigen, benötigen wir den folgenden Hilfssatz, der in der Literatur die *Waldsche Identität* für Erneuerungsprozesse genannt wird.

Lemma 2.1 *Sei $g : [0,\infty) \to [0,\infty)$ ein beliebige Borel–messbare Funktion. Dann gilt für jedes $t \geq 0$*

$$\mathbb{E}\Big(\sum_{i=1}^{N_t+1} g(T_i)\Big) = \mathbb{E}\,g(T_1)(\mathbb{E}\,N_t + 1)\,. \tag{12}$$

Beweis

- Um die Gültigkeit von (12) zu zeigen, führen wir die folgenden Hilfsvariablen ein: Sei

$$Y_i = \begin{cases} 0 & \text{falls } i > N_t + 1, \\ 1 & \text{falls } i \leq N_t + 1. \end{cases}$$

- Dann ergibt sich aus dem Satz über die monotone Konvergenz, dass

$$\mathbb{E}\Big(\sum_{i=1}^{N_t+1} g(T_i)\Big) = \mathbb{E}\Big(\sum_{i=1}^{\infty} g(T_i)Y_i\Big) = \sum_{i=1}^{\infty} \mathbb{E}\,(g(T_i)Y_i)\,,$$

  wobei die Erwartungswerte in der letzten Summe genauer bestimmt werden können, weil die Zufallsvariablen $g(T_i)$ und Y_i unabhängig sind.
- Dies gilt, weil die Bernoulli–Variable Y_i genau dann von T_i unabhängig ist, wenn das Ereignis $\{Y_i = 1\}$ von T_i unabhängig ist.
- Es gilt jedoch $\{Y_i = 1\} = \{N_t + 1 \geq i\} = \{S_{i-1} \leq t\}$, und das letzte Ereignis ist unabhängig von T_i.
- Hieraus folgt, dass

$$\begin{aligned} \mathbb{E}\Big(\sum_{i=1}^{N_t+1} g(T_i)\Big) &= \sum_{i=1}^{\infty} \mathbb{E}\,g(T_i)\mathbb{E}\,Y_i = \mathbb{E}\,g(T_1)\sum_{i=1}^{\infty} P(Y_i = 1) \\ &= \mathbb{E}\,g(T_1)\sum_{i=1}^{\infty} P(N_t + 1 \geq i) = \mathbb{E}\,g(T_1)(\mathbb{E}\,N_t + 1)\,. \end{aligned}$$

□

Wir sind nun in der Lage, das folgende Theorem zu beweisen, das in der Literatur der *elementare Erneuerungssatz* genannt wird.

Theorem 2.4 *Sei* $0 < \mu < \infty$. *Dann gilt*

$$\lim_{t\to\infty} \frac{H(t)}{t} = \frac{1}{\mu}\,. \tag{13}$$

Beweis

- Aus Theorem 2.1 ergibt sich mit Hilfe des Lemmas von Fatou, dass

$$\mu^{-1} = \mathbb{E}\liminf_{t\to\infty} t^{-1}N_t \le \liminf_{t\to\infty} t^{-1}\mathbb{E}\,N_t = \liminf_{t\to\infty} t^{-1}H(t)\,.$$

- Andererseits gilt auch

$$\mu^{-1} \ge \limsup_{t\to\infty} t^{-1}H(t)\,. \tag{14}$$

- Um dies zu zeigen, betrachten wir die gestutzten Zwischenankunftszeiten $T'_n = \min\{T_n, b\}$ für ein hinreichend großes $b > 0$, so dass $\mathbb{E}\,T'_n > 0$.
- Sei $\{N'_t, t \ge 0\}$ der entsprechende Zählprozess mit

$$N'_t = \sum_{n=1}^{\infty} \mathbb{I}(S'_n \le t)\,, \qquad S'_n = \sum_{i=1}^{n} T'_i\,, \tag{15}$$

  und $H'(t) = \mathbb{E}\,N'_t$.
- Dann gilt $N'_t \ge N_t$ und somit $H'(t) \ge H(t)$ für jedes $t \ge 0$. Hieraus und aus Lemma 2.1 ergibt sich, dass

$$\begin{aligned}
\limsup_{t\to\infty} t^{-1}H(t) &\le \limsup_{t\to\infty} t^{-1}H'(t) \le \limsup_{t\to\infty} t^{-1}\frac{\mathbb{E}\,S'_{N'_t+1}}{\mathbb{E}\,T'_1} \\
&\le \limsup_{t\to\infty} t^{-1}\mathbb{E}\left(S'_{N'_t} + T'_{N'_t+1}\right)(\mathbb{E}\,T'_1)^{-1} \\
&\le \limsup_{t\to\infty} t^{-1}(t+b)(\mathbb{E}\,T'_1)^{-1} = (\mathbb{E}\,T'_1)^{-1}\,.
\end{aligned}$$

- Damit ist (14) bewiesen, denn es gilt $\lim_{b\to\infty}\mathbb{E}\,T'_1 = \mathbb{E}\,T_1$. □

Außer der in Theorem 2.4 hergeleiteten asymptotischen Linearität lässt sich noch eine wesentliche schärfere Aussage über das asymptotische Verhalten der Erneuerungsfunktion $H(t)$ für $t \to \infty$ herleiten.

Der folgende Grenzwertsatz wird in der Literatur *Erneuerungssatz von Blackwell* bzw. *Haupterneuerungssatz.* genannt. Er besagt, dass sich das *Erneuerungsmaß* $H : \mathcal{B}([0,\infty) \to [0,\infty)$ mit

$$H(B) = \sum_{n=0}^{\infty} \int_B \mathrm{d}F^{*n}(x)\,, \qquad B \in \mathcal{B}([0,\infty)$$

asymptotisch genauso wie das Lebesgue–Maß verhält.

Theorem 2.5 *Sei* $0 < \mu < \infty$, *und die Verteilungsfunktion* F *sei nicht gitterförmig, d.h., die Wachstumspunkte von* F *liegen nicht auf einem regelmäßigen Gitter. Dann gilt*

$$\lim_{x\to\infty}\left(H(x) - H(x-y)\right) = \frac{y}{\mu} \qquad \forall\, y \ge 0\,. \tag{16}$$

Der *Beweis* von Theorem 2.5 geht über den Rahmen dieser Vorlesung hinaus und wird deshalb weggelassen. Er kann beispielsweise in Kallenberg (2001), S. 172–174 nachgelesen werden .

2.1.3 Verzögerte Erneuerungprozesse; stationäre Zuwächse

Wir betrachten nun ein etwas allgemeineres Modell von Erneuerungsprozessen. Dabei setzen wir voraus, dass $T_1, T_2, \ldots$ eine Folge von unabhängigen nichtnegativen Zufallsvariablen ist und dass die Zufallsvariablen $T_2, T_3, \ldots$ identisch verteilt sind mit der Verteilungsfunktion F.

Wir lassen jedoch zu, dass T_1 eine beliebige Verteilungsfunktion F_1 hat, die von F verschieden sein kann.

Definition Die Folge $\{S_n, n \geq 1\}$ mit $S_n = T_1 + \ldots + T_n$ wird dann *verzögerter Erneuerungspunktprozess* genannt. Der Zählprozess $\{N_t,\, t \geq 0\}$, der wiederum so wie in (1) definiert wird, heißt *verzögerter Erneuerungszählprozess* bzw. kurz *verzögerter Erneuerungsprozess.*

Völlig analog zu Theorem 2.4 lässt sich die folgende Aussage beweisen.

Theorem 2.6 *Sei* $0 < \mu < \infty$. *Dann gilt*

$$\lim_{t\to\infty} \frac{H(t)}{t} = \frac{1}{\mu}\,, \tag{17}$$

wobei

$$H(t) = \mathbb{E}\, N_t = \sum_{n=0}^{\infty} (F_1 * F^{*n})(t)\,. \tag{18}$$

Der Fall, wenn F_1 die sogenannte *integrierte Tailverteilungsfunktion* F^{s} von F ist, ist von besonderem Interesse, wobei

$$F^{\mathrm{s}}(x) = \frac{1}{\mu} \int_0^x (1 - F(y))\,\mathrm{d}y \qquad \forall\, x \geq 0\,. \tag{19}$$

Theorem 2.7 *Es gelte* $0 < \mu < \infty$ *und*

$$F_1(x) = F^{\mathrm{s}}(x) \qquad \forall\, x \geq 0\,. \tag{20}$$

Dann ist die Erneuerungsfunktion $H(t)$ *gegeben durch*

$$H(t) = \frac{t}{\mu} \qquad \forall\, t \geq 0\,. \tag{21}$$

Beweis

- Wenn auf beiden Seiten der Gleichung (18) zu den Laplace–Stieltjes–Transformierten übergegangen wird, dann ergibt sich genauso wie im Beweis von Theorem 2.3, dass

$$\hat{l}_H(s) = \frac{\hat{l}_{F_1}(s)}{1 - \hat{l}_F(s)} = \frac{\hat{l}_{F^{\mathrm{s}}}(s)}{1 - \hat{l}_F(s)}\,.$$

- Weil $\hat{l}_{F^{\mathrm{s}}}(s) = (1 - \hat{l}_F(s))(s\mu)^{-1}$ gilt, ergibt sich hieraus, dass $\hat{l}_H(s) = (s\mu)^{-1} = \mu^{-1} \int_0^\infty \mathrm{e}^{-sv}\,\mathrm{d}v$.
- Wegen des eineindeutigen Zusammenhanges zwischen Erneuerungsfunktionen und Laplace–Stieltjes–Transformierten, ist damit die Gültigkeit von (21) bewiesen. □

Für verzögerte Erneuerungsprozesse $\{N_t,\, t \geq 0\}$, die der Bedingung (20) genügen, lässt sich eine wesentlich schärfere Aussage als in Theorem 2.7 herleiten. Hierfür benötigen wir den folgenden Begriff: Die Zufallsvariable $T(t) = S_{N_t+1} - t$ heißt *Exzess* zum Zeitpunkt $t \geq 0$.

Theorem 2.8 *Unter den Voraussetzungen von Theorem* 2.7 *ist der verzögerte Erneuerungsprozess* $\{N_t, t \geq 0\}$ *ein Prozess mit stationären Zuwächsen.*

Beweis

- Weil die Zufallsvariablen $T_1, T_2, \ldots$ unabhängig sind, hängt die gemeinsame Verteilung der Zuwächse $N_{t_1+t} - N_{t_0+t}, \ldots, N_{t_n+t} - N_{t_{n-1}+t}$ nicht von t ab, falls die Verteilung des Exzesses $T(t)$ diese Eigenschaft hat.
- In Anbetracht der in Abschnitt 1.5 eingeführten Definition von Prozessen mit stationären Zuwächsen genügt es zu zeigen, dass $P(T(t) \leq x) = F^{\mathrm{s}}(x)$ für jedes $x \geq 0$ gilt, und zwar unabhängig von t.
- Dies ergibt sich aus den folgenden Überlegungen: Es gilt

$$\begin{aligned} P(T(t) \leq x) &= \sum_{n=1}^{\infty} P(S_{n-1} \leq t < S_n \leq t + x) \\ &= F^{\mathrm{s}}(t+x) - F^{\mathrm{s}}(t) + \sum_{n=1}^{\infty} \int_0^t (F(t+x-y) - F(t-y))\, \mathrm{d}(F^{\mathrm{s}} * F^{*(n-1)})(y) \\ &= F^{\mathrm{s}}(t+x) - F^{\mathrm{s}}(t) + \mu^{-1} \int_0^t (F(t+x-y) - F(t-y))\, \mathrm{d}y = F^{\mathrm{s}}(x)\,, \end{aligned}$$

  wobei sich die dritte Gleichheit aus (18) und (21) ergibt. □

Beachte Weil $\mathbb{E}\, N_t = \mathbb{E}\,(N_t - N_{t'}) + \mathbb{E}\, N_{t'}$ für beliebige $t, t' \geq 0$ mit $t' \leq t$ gilt, ergibt sich unmittelbar aus Theorem 2.8, dass $\mathbb{E}\, N_t$ eine lineare Funktion in t ist. Somit implizieren die Theoreme 2.6 und 2.8 die Aussage von Theorem 2.7.

2.2 Zählprozesse vom Poisson–Typ

2.2.1 Homogener Poisson–Prozess

In diesem Abschnitt betrachten wir (Erneuerungs-) Zählprozesse $\{N_t, t \geq 0\}$ mit

$$N_t = \sum_{k=1}^{\infty} \mathbb{1}(S_k \leq t) \qquad \text{und} \qquad S_n = T_1 + \ldots + T_n\,, \tag{22}$$

wobei $T_1, T_2, \ldots : \Omega \to [0, \infty)$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen ist. Dabei setzen wir zusätzlich voraus, dass die Zwischenankunftszeiten T_n exponentialverteilt sind, d.h.,

$$T_n \sim \mathrm{Exp}(\lambda) \qquad \text{für ein } \lambda > 0. \tag{23}$$

Definition Falls (23) gilt, dann wird der Zählprozess $\{N_t, t \geq 0\}$ ein *homogener Poisson–Prozess* mit der *Intensität* λ genannt.

Beachte

- Eine grundlegende Eigenschaft homogener Poisson–Prozesse besteht darin, dass sie eine Klasse stochastischer Prozesse mit unabhängigen und stationären Zuwächsen bilden.
- Dies werden wir in dem folgenden Theorem genauer analysieren, das verschiedene äquivalente Definitionsmöglichkeiten des homogenen Poisson–Prozesses enthält wobei wir das Adjektiv „homogen" der Kürze wegen weglassen.

Theorem 2.9 *Sei $\{N_t, t \geq 0\}$ ein beliebiger Zählprozess. Dann sind die folgenden Aussagen äquivalent:*

(a) *$\{N_t\}$ ist ein Poisson–Prozess mit Intensität λ.*

(b) *Für beliebige $t \geq 0$ und $n = 1, 2, \ldots$ ist die Zufallsvariable N_t poissonverteilt mit Parameter λt, d.h. $N_t \sim \mathrm{Poi}(\lambda t)$, und unter der Bedingung, dass $\{N_t = n\}$, hat der Zufallsvektor $(S_1, \ldots, S_n)$ die gleiche Verteilung wie die Ordnungsstatistik von n unabhängigen, in $[0, t]$ gleichverteilten Zufallsvariablen.*

(c) *$\{N_t\}$ hat unabhängige Zuwächse mit $\mathbb{E}\, N_1 = \lambda$, und für beliebige $t \geq 0$ und $n = 1, 2, \ldots$ hat der Zufallsvektor $(S_1, \ldots, S_n)$ unter der Bedingung, dass $\{N_t = n\}$, die gleiche Verteilung wie die Ordnungsstatistik von n unabhängigen, in $[0, t]$ gleichverteilten Zufallsvariablen.*

(d) *$\{N_t\}$ hat stationäre und unabhängige Zuwächse, und für $h \downarrow 0$ gilt*

$$P(N_h = 0) = 1 - \lambda h + o(h)\,, \qquad P(N_h = 1) = \lambda h + o(h)\,. \tag{24}$$

(e) *$\{N_t\}$ hat stationäre und unabhängige Zuwächse, und für jedes $t \geq 0$ ist N_t eine $\mathrm{Poi}(\lambda t)$–verteilte Zufallsvariable.*

Beweis Wir führen einen zyklischen Beweis, d.h., wir zeigen, dass (a)⇒(b)⇒(c)⇒(d)⇒(e)⇒(a) gilt.

(a)⇒(b)

- Aus (a) ergibt sich, dass $S_n = \sum_{i=1}^n T_i$ eine Summe von n unabhängigen und $\mathrm{Exp}(\lambda)$–verteilten Zufallsvariablen ist, d.h., S_n ist $\mathrm{Erl}(n, \lambda)$–verteilt, wobei $\mathrm{Erl}(n, \lambda)$ die Erlang–Verteilung mit den Parametern n und λ bezeichnet.
- Hieraus ergibt sich, dass $P(N_t = 0) = P(S_1 > t) = \mathrm{e}^{-\lambda t}$ und

$$\begin{aligned} P(N_t = n) &= P(N_t \geq n) - P(N_t \geq n+1) \\ &= P(S_n \leq t) - P(S_{n+1} \leq t) \\ &= \int_0^t \frac{\lambda^n v^{n-1}}{(n-1)!}\, \mathrm{e}^{-\lambda v}\, \mathrm{d}v - \int_0^t \frac{\lambda^{n+1} v^n}{n!}\, \mathrm{e}^{-\lambda v}\, \mathrm{d}v \\ &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}v}\Big(\frac{(\lambda v)^n}{n!}\, \mathrm{e}^{-\lambda v}\Big)\, \mathrm{d}v = \frac{(\lambda t)^n}{n!}\, \mathrm{e}^{-\lambda t} \end{aligned}$$

  für jedes $n \geq 1$, d.h., N_t ist $\mathrm{Poi}(\lambda t)$–verteilt.
- Weil die gemeinsame Dichte $f_{S_1,\ldots,S_{n+1}}(t_1, \ldots, t_{n+1})$ von $S_1, \ldots, S_{n+1}$ gegeben ist durch

$$f_{S_1,\ldots,S_{n+1}}(t_1, \ldots, t_{n+1}) = \prod_{k=1}^{n+1} \lambda \mathrm{e}^{-\lambda(t_k - t_{k-1})} = \lambda^{n+1} \mathrm{e}^{-\lambda t_{n+1}}$$

  für beliebige $t_0 = 0 \leq t_1 \leq \ldots \leq t_n \leq t_{n+1}$ und $f_{S_1,\ldots,S_{n+1}}(t_1, \ldots, t_{n+1}) = 0$ sonst, gilt für die gemeinsame bedingte Dichte $f_{S_1,\ldots,S_n}(t_1, \ldots, t_n \mid N_t = n)$ von $S_1, \ldots, S_n$ unter der Bedingung, dass $N_t = n$, die Formel

$$\begin{aligned} f_{S_1,\ldots,S_n}(t_1, \ldots, t_n \mid N_t = n) &= f_{S_1,\ldots,S_n}(t_1, \ldots, t_n \mid S_1 \leq t, \ldots, S_n \leq t, S_{n+1} > t) \\ &= \frac{\int_t^\infty \lambda^{n+1} \mathrm{e}^{-\lambda x_{n+1}}\, \mathrm{d}x_{n+1}}{\int_0^t \int_{x_1}^t \cdots \int_{x_{n-1}}^t \int_t^\infty \lambda^{n+1} \mathrm{e}^{-\lambda x_{n+1}}\, \mathrm{d}x_{n+1} \ldots \mathrm{d}x_1} \\ &= \frac{n!}{t^n} \end{aligned}$$

  für $0 \leq t_1 \leq \ldots \leq t_n \leq t$ und $f_{S_1,\ldots,S_n}(t_1, \ldots, t_n \mid N_t = n) = 0$ sonst.
- Das ist die Dichte der Ordnungsstatistik von n unabhängigen, in $[0, t]$ gleichverteilten Zufallsvariablen.

(b)⇒(c)

- Aus (b) ergibt sich ohne weiteres, dass $\mathbb{E}\, N_1 = \lambda$.

- Sei nun $x_1, \ldots, x_n \in \mathbb{N}$ und $t_0 = 0 < t_1 < \ldots < t_n$. Dann ergibt sich aus (b) für $x = x_1 + \ldots + x_n$, dass

$$\begin{aligned} P\Big(\bigcap_{k=1}^n \{N_{t_k} - N_{t_{k-1}} = x_k\}\Big) &= P\Big(\bigcap_{k=1}^n \{N_{t_k} - N_{t_{k-1}} = x_k\} \;\Big|\; N_{t_n} = x\Big) P(N_{t_n} = x) \\ &= \frac{(\lambda t_n)^x}{x!} \mathrm{e}^{-\lambda t_n} \frac{x!}{x_1! \ldots x_n!} \prod_{k=1}^n \left(\frac{t_k - t_{k-1}}{t_n}\right)^{x_k} \\ &= \prod_{k=1}^n \frac{(\lambda(t_k - t_{k-1}))^{x_k}}{x_k!} \mathrm{e}^{-\lambda(t_k - t_{k-1})}, \end{aligned}$$

  d.h., der Zählprozess $\{N_t\}$ hat unabhängige Zuwächse.

(c)$\Rightarrow$(d)

- Wir nehmen an, dass der Zufallsvektor $(S_1, \ldots, S_m)$ unter der Bedingung, dass $N(t_n + h) = m$, die gleiche Verteilung hat wie die Ordnungsstatistik von m unabhängigen, in $[0, t_n + h]$ gleichverteilten Zufallsvariablen.
- Dann gilt für beliebige $x_1, \ldots, x_n \in \mathbb{N}$, $t_0 = 0 < t_1 < \ldots < t_n$ und $h > 0$, dass

$$P\Big(\bigcap_{k=1}^n \{N_{t_k+h} - N_{t_{k-1}+h} = x_k\} \;\Big|\; N_{t_n+h} = m\Big) = P\Big(\bigcap_{k=1}^n \{N_{t_k} - N_{t_{k-1}} = x_k\} \;\Big|\; N_{t_n+h} = m\Big).$$

- Mit Hilfe der Formel der totalen Wahrscheinlichkeit ergibt sich nun, dass der Zählprozess $\{N_t\}$ stationäre Zuwächse hat.
- Außerdem ergibt sich aus der bedingten Gleichverteilungseigenschaft, die in (c) vorausgesetzt wird, dass für $0 < h < 1$

$$\begin{aligned} P(N_h = 0) &= \sum_{k=0}^\infty P(N_h = 0, N_1 - N_h = k) \\ &= \sum_{k=0}^\infty P(N_1 = k)\; P(N_1 - N_h = k \mid N_1 = k) \\ &= \sum_{k=0}^\infty P(N_1 = k)(1-h)^k. \end{aligned}$$

- Somit gilt

$$\begin{aligned} \frac{1}{h}(1 - P(N_h = 0)) &= \frac{1}{h}\Big(1 - \sum_{k=0}^\infty P(N_1 = k)(1-h)^k\Big) \\ &= \sum_{k=1}^\infty P(N_1 = k) \frac{1 - (1-h)^k}{h}. \end{aligned}$$

- Aus der Ungleichung $(1-h)^k \geq 1 - kh$, die für beliebige $0 < h < 1$ und $k = 1, 2, \ldots$ gilt, folgt, dass die Funktionen $g_h(k) = h^{-1}(1 - (1-h)^k)$ in der letzten Summe die gemeinsame Schranke $g(k) = k$ besitzen.
- Diese Schranke ist integrierbar, denn es gilt

$$\sum_{k=1}^\infty k P(N_1 = k) = \mathbb{E}\, N_1 = \lambda < \infty.$$

- Durch Vertauschung von Summation und Grenzwert ergibt sich nun, dass

$$\lim_{h\to 0} \frac{1}{h}\, P(N_h > 0) = \lambda$$

und somit der erste Teil von (24).

- Auf die gleiche Weise erhalten wir, dass

$$\lim_{h\to 0} \frac{1}{h}\, P(N_h = 1) = \lim_{h\to 0} \sum_{k=1}^{\infty} P(N_1 = k)\, k(1-h)^{k-1} = \lambda\,,$$

was äquivalent zum zweiten Teil von (24) ist.

(d)⇒(e)

- Sei $n \in \mathbb{N}$, $t \geq 0$ und $p_n(t) = P(N_t = n)$. Dann gilt für $h > 0$

$$p_0(t+h) = P(N_t = 0, N_{t+h} - N_t = 0) = p_0(t)(1 - \lambda h + o(h)) \tag{25}$$

und für $t \geq h > 0$

$$p_0(t) = p_0(t-h)(1 - \lambda h + o(h))\,. \tag{26}$$

- Hieraus folgt, dass die Funktion $p_0(t)$ stetig in $(0,\infty)$ und rechtsstetig im Punkt $t = 0$ ist.
- Weil $p_0(t-h) = p_0(t) + o(1)$, ergibt sich aus (25) und (26), dass für beliebige $h \geq -t$

$$\frac{p_0(t+h) - p_0(t)}{h} = -\lambda p_0(t) + o(1)\,.$$

- Die Funktion $p_0(t)$ ist somit differenzierbar, und sie genügt der Differentialgleichung

$$p_0^{(1)}(t) = -\lambda p_0(t) \tag{27}$$

für $t > 0$.

- Wegen $p_0(0) = P(N_0 = 0) = 1$ ist die (eindeutig bestimmte) Lösung von (27) gegeben durch

$$p_0(t) = \mathrm{e}^{-\lambda t}, \qquad t \geq 0\,. \tag{28}$$

- Um zu zeigen, dass

$$p_n(t) = \frac{(\lambda t)^n}{n!}\, \mathrm{e}^{-\lambda t}, \qquad t \geq 0 \tag{29}$$

für beliebige $n \in \mathbb{N}$ gilt, kann man genauso wie im Beweis von (28) vorgehen.

- Hierfür genügt es zu beachten, dass (24) die Gültigkeit von $P(N_h > 1) = o(h)$ für $h \downarrow 0$ impliziert.
- Mit Hilfe von (28) ergibt sich nun durch vollständige Induktion nach n, dass (29) gilt.

(e)⇒(a)

- Sei $b_0 = 0 \leq a_1 < b_1 \leq \ldots \leq a_n < b_n$. Dann gilt

$$\begin{aligned} P\Big(\bigcap_{k=1}^{n} \{a_k < S_k \leq b_k\}\Big) \;&=\; P\Big(\bigcap_{k=1}^{n-1} \{N_{a_k} - N_{b_{k-1}} = 0, N_{b_k} - N_{a_k} = 1\} \\ &\qquad \cap \{N_{a_n} - N_{b_{n-1}} = 0, N_{b_n} - N_{a_n} \geq 1\}\Big)\,, \end{aligned}$$

und aus (e) ergibt sich , dass

$$\begin{aligned}
P\Big(\bigcap_{k=1}^{n}\{a_k < S_k \le b_k\}\Big) &= \mathrm{e}^{-\lambda(a_n-b_{n-1})}(1-\mathrm{e}^{-\lambda(b_n-a_n)})\prod_{k=1}^{n-1}\mathrm{e}^{-\lambda(a_k-b_{k-1})}\lambda(b_k-a_k)\mathrm{e}^{-\lambda(b_k-a_k)} \\
&= (\mathrm{e}^{-\lambda a_n}-\mathrm{e}^{-\lambda b_n})\lambda^{n-1}\prod_{k=1}^{n-1}(b_k-a_k) \\
&= \int_{a_1}^{b_1}\dots\int_{a_n}^{b_n}\lambda^n\mathrm{e}^{-\lambda y_n}\,\mathrm{d}y_n\dots\mathrm{d}y_1 \\
&= \int_{a_1}^{b_1}\int_{a_2-x_1}^{b_2-x_1}\dots\int_{a_n-x_1-\dots-x_{n-1}}^{b_n-x_1-\dots-x_{n-1}}\lambda^n\mathrm{e}^{-\lambda(x_1+\dots+x_n)}\,\mathrm{d}x_n\dots\mathrm{d}x_1\,.
\end{aligned}$$

- Die gemeinsame Dichte von $S_1, S_2-S_1, \dots, S_n-S_{n-1}$ ist somit gegeben durch

$$P\Big(\bigcap_{k=1}^{n}\{S_k-S_{k-1}\in \mathrm{d}x_k\}\Big) = \lambda^n\mathrm{e}^{-\lambda(x_1+\dots+x_n)}\,\mathrm{d}x_1\dots\mathrm{d}x_n\,.$$

- Dies bedeutet, dass die Zufallsvariablen $S_1, S_2-S_1, \dots, S_n-S_{n-1}$ unabhängig und Exp(λ)–verteilt sind, d.h., $\{N_t\}$ ist ein Poisson–Prozess mit der Intensität λ. □

2.2.2 Zusammengesetzte Poisson–Prozesse

Wir verallgemeinern das mit Hilfe von (22) eingeführte Modell eines homogenen Poisson–Prozesses $\{N_t, t\ge 0\}$, indem wir nun zulassen, dass die Sprunghöhen beliebige Zufallsvariablen sind, die nicht notwendig gleich Eins sein müssen.

- Außer der Folge der Sprungzeitpunkte $S_1, S_2, \dots$ mit $0 < S_1 < S_2 < \dots$ betrachten wir deshalb zusätzlich eine Folge von Zufallsvariablen $U_1, U_2, \dots : \Omega \to \mathbb{R}$, die die Sprunghöhen des zu konstruierenden (Sprung-) Prozesses sein werden.
- So wie in Abschnitt 2.2.1 nehmen wir an, dass $S_n = T_1 + \dots + T_n$ gilt, wobei $T_1, T_2, \dots : \Omega \to [0,\infty)$ eine Folge von unabhängigen und identisch Exp(λ)–verteilten Zwischenankunftszeiten ist.
- Dabei wird vorausgesetzt, dass $U_1, U_2, \dots$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen ist, die von der Folge der Zwischenankunftszeitpunkte $T_1, T_2, \dots$ unabhängig ist.

Definition Der stochastische Prozess $\{X_t, t\ge 0\}$ mit

$$X_t = \sum_{k=1}^{\infty} U_k\,\mathbb{I}(S_k \le t) = \sum_{k=1}^{N_t} U_k \tag{30}$$

heißt *zusammengesetzter Poisson–Prozess* mit den Charakteristiken (λ, F_U), wobei F_U die Verteilungsfunktion der Sprunghöhen $U_1, U_2, \dots$ bezeichnet.

Theorem 2.10 *Sei $\{X_t\}$ ein zusammengesetzter Poisson–Prozess mit den Charakteristiken (λ, F_U), so dass die momenterzeugende Funktion $\hat{m}_U(s) = \mathbb{E}\,\mathrm{e}^{sU_n}$ von U_n für jedes $s\in\mathbb{R}$ wohldefiniert sei. Dann gilt:*

(a) *$\{X_t\}$ hat stationäre und unabhängige Zuwächse.*

(b) *Für beliebige $s \in \mathbb{R}$ und $t \geq 0$ ist die momenterzeugende Funktion $\hat{m}_{X_t}(s) = \mathbb{E}\, \mathrm{e}^{sX_t}$ von X_t gegeben durch*

$$\hat{m}_{X_t}(s) = \mathrm{e}^{\lambda t(\hat{m}_U(s)-1)}\,, \tag{31}$$

und es gilt

$$\mathbb{E}\, X_t = \lambda t \mu_U\,, \qquad \operatorname{Var} X_t = \lambda t \mu_U^{(2)}\,, \tag{32}$$

wobei $\mu_U = \mathbb{E}\, U_n$ und $\mu_U^{(2)} = \mathbb{E}\, U_n^2$ die ersten beiden Momente der Sprunghöhen U_n sind.

Beweis

- Um die Teilaussage (a) zu beweisen, genügt es zu zeigen, dass für beliebige $n = 1, 2, \ldots, h \geq 0$ und $0 \leq t_0 < t_1 < \ldots < t_n$ die Zufallsvariablen

$$\sum_{i_1=N_{t_0+h}+1}^{N_{t_1+h}} U_{i_1}\,, \ldots\,, \sum_{i_n=N_{t_{n-1}+h}+1}^{N_{t_n+h}} U_{i_n}$$

  unabhängig sind und dass ihre Verteilungen nicht von h abhängen.
- Weil die Folge $\{U_n\}$ aus unabhängigen und identisch verteilten Zufallsvariablen besteht, die unabhängig von $\{T_n\}$ sind, gilt für beliebige $x_1, \ldots, x_n \geq 0$

$$\begin{aligned} &P\Big(\sum_{i_1=N_{t_0+h}+1}^{N_{t_1+h}} U_{i_1} \leq x_1\,, \ldots\,, \sum_{i_n=N_{t_{n-1}+h}+1}^{N_{t_n+h}} U_{i_n} \leq x_n\Big) \\ &= \sum_{k_1,\ldots,k_n \geq 0} \prod_{j=1}^{n} F_U^{*k_j}(x_j) P(N_{t_1+h} - N_{t_0+h} = k_1\,, \ldots\,,\ N_{t_n+h} - N_{t_{n-1}+h} = k_n)\,. \end{aligned}$$

- Es genügt nun zu beachten, dass wir in Theorem 2.9 gezeigt hatten, dass der Poisson–Prozess $\{N_t\}$ unabhängige und stationäre Zuwächse hat.
- Unter Berücksichtigung unserer Unabhängigkeits- und Verteilungsannahmen ergibt sich die Formel (31) in Teilaussage (b) unmittelbar aus der Formel der totalen Wahrscheinlicheit.
- Die Formeln in (32) ergeben sich durch Differenzieren beider Seiten von (31). □

Beachte Die Verteilung einer Zufallsvariablen, deren momenterzeugende Funktion die Form (31) hat, heißt *zusammengesetzte Poisson–Verteilung* mit den Charakteristiken (λ, F_U).

2.2.3 Bedingte Erwartung und bedingte Wahrscheinlichkeit

Um den in Abschnitt 2.2.1 betrachteten Begriff des (homogenen) Poissonschen Zählprozesses $\{N_t, t \geq 0\}$ noch in eine andere Richtung verallgemeinern zu können, benötigen wir als Hilfsmittel die Begriffe der bedingten Erwartung bzw. der bedingten Wahrscheinlichkeit bezüglich einer beliebigen Teil–σ–Algebra $\mathcal{G} \subset \mathcal{F}$ der Ereignis–σ–Algebra $\mathcal{F}$ des zugrundeliegenden Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$.

- Zur Illustration betrachten wir zunächst den folgenden (elementaren) *Spezialfall.*
  - Die σ–Algebra $\mathcal{G} = \sigma(A_1, \ldots, A_n)$ sei eine *endliche* Familie von Teilmengen von Ω, die durch die Mengen $A_1, \ldots, A_n \in \mathcal{F}$ generiert wird, wobei $\bigcup_{i=1}^n A_i = \Omega$ und $A_i \cap A_j = \emptyset$ für beliebige $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ gelte.
  - Außerdem sei $X : \Omega \to \mathbb{R}$ eine beliebige Zufallsvariable über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $\mathbb{E}\,|X| < \infty$.

- Dann ist bedingte Erwartung $\mathbb{E}\,(X \mid \mathcal{G})$ von X bezüglich $\mathcal{G}$ eine Zufallsvariable $\mathbb{E}\,(X \mid \mathcal{G}) : \Omega \to \mathbb{R}$, deren (Funktions–) Werte gegeben sind durch

$$\mathbb{E}\,(X \mid \mathcal{G})(\omega) = \sum_{i=1}^{n} \mathbb{E}\,(X \mid A_i)\ \mathbb{1}(A_i)(\omega) \qquad \forall\, \omega \in \Omega\,, \tag{33}$$

wobei der bedingte Erwartungswert $\mathbb{E}\,(X \mid A_i)$ gegeben ist durch

$$\mathbb{E}\,(X \mid A_i) = \begin{cases} \dfrac{\mathbb{E}\,(X\mathbb{1}(A_i))}{P(A_i)}\,, & \text{falls } P(A_i) > 0, \\ 0\,, & \text{sonst.} \end{cases} \tag{34}$$

Beachte

- Falls $P(A_i) > 0$, dann ist die in (34) gegebene Zahl $\mathbb{E}\,(X|A_i)$ der bedingte Erwartungswert von X unter der Bedingung A_i, d.h., der Erwartungswert bezüglich der *bedingten Verteilung* $\{P_{X|A_i}(B),\ B \in \mathcal{B}(\mathbb{R})\}$, wobei
$$P_{X|A_i}(B) = \frac{P(\{X_i \in B\} \cap A_i)}{P(A_i)} \qquad \forall\, B \in \mathcal{B}(\mathbb{R})\,.$$
- Von besonderer Wichtigkeit ist der Fall, dass die σ–Algebra $\mathcal{G}$ durch die Urbilder einer (diskreten) Zufallsvariablen Y erzeugt wird, die nur endlich viele verschiedene Werte mit positiver Wahrscheinlichkeit annimmt:
  - Sei $Y : \Omega \to \{0, 1, \ldots, n\}$ eine Zufallsvariable, die nur die Werte $0, 1, \ldots, n$ mit positiver Wahrscheinlichkeit annimmt, beispielsweise eine binomialverteilte Zufallsvariable, und
  - sei $\mathcal{G}$ die σ–Algebra $\sigma(Y) = Y^{-1}(\mathcal{B}(\mathbb{R}))$, d.h., die kleinste (endliche) Teil–σ–Algebra von $\mathcal{F}$, die die Urbilder $Y^{-1}(0), Y^{-1}(1), \ldots, Y^{-1}(n)$ enthält,
  - dann wird die bedingte Erwartung $\mathbb{E}\,(X \mid \mathcal{G})$ auch mit dem Symbol $\mathbb{E}\,(X \mid Y)$ bezeichnet (und bedingte Erwartung von X bezüglich Y genannt).
- Sei $\omega \in \Omega$, so dass $Y(\omega) = i$. Dann ist der Funktionswert $\mathbb{E}\,(X|Y)(\omega)$ der Zufallsvariablen $\mathbb{E}\,(X|Y)$ gegeben durch den bedingten Erwartungswert
$$\mathbb{E}\,(X \mid Y)(\omega) = \mathbb{E}\,(X \mid Y = i)\,,$$
wobei $\mathbb{E}\,(X \mid Y = i)$ eine Kurzschreibweise ist für $\mathbb{E}\,(X \mid \{Y = i\})$.
- Die bisher erläuterte Vorgehensweise bei der Definition der bedingten Erwartung $\mathbb{E}\,(X|\mathcal{G})$ bezüglich einer (Teil–) σ–Algebra $\mathcal{G} \subset \mathcal{F}$ kann auch dann beibehalten werden,
  - wenn $\mathcal{G}$ aus *abzählbar* unendlich vielen Teilmengen von Ω besteht,
  - d.h., wenn $\mathcal{G}$ beispielsweise durch die Urbilder einer Poisson–verteilten Zufallsvariablen erzeugt wird,
  - wobei dann lediglich die endliche Summe in (33) durch eine unendliche Summe ersetzt werden muss.
- Die folgenden Eigenschaften der bedingten Erwartung $\mathbb{E}\,(X|Y)$ ergeben sich unmittelbar aus der Definitionsgleichungen (33) und (34):
  - Die Zufallsvariable $\mathbb{E}\,(X|Y) : \Omega \to \mathbb{R}$ ist $(\mathcal{G},\ \mathcal{B}(\mathbb{R}))$–messbar.
  - Für beliebige $A \in \mathcal{G}$ gilt
$$\int_A \mathbb{E}\,(X|Y)(\omega)P(d\omega) = \int_A X(\omega)P(d\omega)\,. \tag{35}$$
  - Insbesondere gilt $\mathbb{E}\,(\mathbb{E}\,(X|Y)) = \mathbb{E}\,X$.

Die folgende (allgemeine) Definition der bedingten Erwartung beruht auf dem Satz von Radon–Nikodym der Maß– und Integrationstheorie. Sie enthält die oben betrachteten Definitionsgleichungen (33) und (34) als Spezialfall.

- Sei $X : \Omega \to \mathbb{R}$ eine beliebige Zufallsvariable über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $\mathbb{E}\,|X| < \infty$,
- und sei $\mathcal{G} \subset \mathcal{F}$ eine beliebige Teil–σ–Algebra von $\mathcal{F}$.
- Dann können wir das (endliche) signierte Maß $Q : \mathcal{G} \to \mathbb{R}$ betrachten, wobei

$$Q(A) = \int_A X(\omega) P(d\omega) \qquad \forall\, A \in \mathcal{G}\,. \tag{36}$$

- Für jedes $A \in \mathcal{G}$ gilt somit die Implikation: $Q(A) = 0$, falls $P(A) = 0$.
- Mit anderen Worten: Das in (36) definierte signierte Maß Q ist *absolutstetig* bezüglich der Einschränkung von P auf die Teil–σ–Algebra $\mathcal{G}$.
- Aus dem *Satz von Radon–Nikodym* folgt nun, dass es eine $(\mathcal{G}, \mathcal{B}(\mathbb{R}))$–messbare Abbildung $Z : \Omega \to \mathbb{R}$ gibt, so dass

$$\int_A Z(\omega) P(d\omega) = \int_A X(\omega) P(d\omega) \qquad \forall\, A \in \mathcal{G}\,, \tag{37}$$

  wobei die Funktionswerte der Abbildung Z mit Wahrscheinlichkeit 1 eindeutig bestimmt sind.

Defintion Jede $(\mathcal{G}, \mathcal{B}(\mathbb{R}))$–messbare Abbildung $Z : \Omega \to \mathbb{R}$, für die (37) gilt, heißt eine Version der *bedingten Erwartung* von X bezüglich $\mathcal{G}$ und wird mit $\mathbb{E}\,(X \mid \mathcal{G})$ bezeichnet.

Aus der Definitionsgleichung (37) und aus den allgemeinen Rechenregeln für das Lebesgue–Integral ergeben sich die folgenden Eigenschaften der bedingten Erwartung, die wir hier lediglich (ohne Beweis) erwähnen.

Theorem 2.11 *Seien $X, Y : \Omega \to \mathbb{R}$ beliebige Zufallsvariablen über $(\Omega, \mathcal{F}, P)$ mit*

$$\mathbb{E}\,|X| < \infty\,, \qquad \mathbb{E}\,|Y| < \infty\,, \qquad \mathbb{E}\,|XY| < \infty\,,$$

und sei $\mathcal{G} \subset \mathcal{F}$ eine beliebige Teil–σ–Algebra von $\mathcal{F}$. Dann gilt

1. $\mathbb{E}\,(X \mid \{\emptyset, \Omega\}) = \mathbb{E}\,X$, $\mathbb{E}\,(X \mid \mathcal{F}) = X$,
2. $\mathbb{E}\,(aX + bY \mid \mathcal{G}) = a\,\mathbb{E}\,(X \mid \mathcal{G}) + b\,\mathbb{E}\,(Y \mid \mathcal{G})$ *für beliebige* $a, b \in \mathbb{R}$,
3. $\mathbb{E}\,(X \mid \mathcal{G}) \le \mathbb{E}\,(Y \mid \mathcal{G})$, *falls* $X \le Y$,
4. $\mathbb{E}\,(XY \mid \mathcal{G}) = Y\mathbb{E}\,(X \mid \mathcal{G})$, *falls Y eine $(\mathcal{G}, \mathcal{B}(\mathbb{R}))$–messbare Zufallsvariable ist,*
5. $\mathbb{E}\,\big(\mathbb{E}\,(X \mid \mathcal{G}_2) \mid \mathcal{G}_1\big) = \mathbb{E}\,(X \mid \mathcal{G}_1)$, *falls $\mathcal{G}_1$ und $\mathcal{G}_2$ Teil–σ–Algebren von $\mathcal{F}$ sind mit $\mathcal{G}_1 \subset \mathcal{G}_2$,*
6. $\mathbb{E}\,(X \mid \mathcal{G}) = \mathbb{E}\,X$, *falls die σ–Algebren $\mathcal{G}$ und $\sigma(X) = X^{-1}(\mathcal{B}(\mathbb{R}))$ unabhängig sind, d.h., falls $P(A \cap A') = P(A)P(A')$ für beliebige $A \in \mathcal{G}$ und $A' \in \sigma(X)$.*
7. $\mathbb{E}\,\big(f(X) \mid \mathcal{G}\big) \ge f\big(\mathbb{E}\,(X \mid \mathcal{G})\big)$, *falls $f : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion ist, so dass $\mathbb{E}\,|f(X)| < \infty$.*

Beachte Außerdem sind die folgenden Eigenschaften bzw. Schreib– und Sprechweisen von Interesse.

- Unmittelbar aus der Definitionsgleichung (37) ergibt sich, dass
  - $\mathbb{E}\,\big(\mathbb{E}\,(X \mid \mathcal{G})\big) = \mathbb{E}\,X$, wenn $A = \Omega$ gesetzt wird,
  - $\mathbb{E}\,(a \mid \mathcal{G}) = a$ für beliebige $a \in \mathbb{R}$ und $\mathcal{G} \subset \mathcal{F}$,
  - $\mathbb{E}\,(Y \mid \mathcal{G}) = Y$ für jede $(\mathcal{G}, \mathcal{B}(\mathbb{R}))$–messbare Zufallsvariable Y.

- Seien $X, Y : \Omega \to \mathbb{R}$ beliebige Zufallsvariablen über $(\Omega, \mathcal{F}, P)$ mit $\mathbb{E}\,|X| < \infty$,
  - und sei $\mathcal{G} = Y^{-1}(\mathcal{B}(\mathbb{R}))$ die Teil–σ–Algebra von $\mathcal{F}$, die durch die Urbilder von Y erzeugt wird.
  - Dann heißt $\mathbb{E}\,(X \mid \mathcal{G})$ die bedingte Erwartung von X bezüglich Y, wobei auch die Schreibweise $\mathbb{E}\,(X \mid Y)$ benutzt wird.
- Wenn es eine Borel–messbare Funktion $g : \mathbb{R} \to \mathbb{R}$ gibt, so dass sich die bedingte Erwartung $\mathbb{E}\,(X \mid Y)$ in der Form $\mathbb{E}\,(X \mid Y) = g(Y)$ darstellen lässt, dann spricht man von *regulärer* bedingter Erwartung.

Beispiele

1. Absolutstetige Zufallsvektoren
   - Sei $(X, Y) : \Omega \to \mathbb{R}^2$ ein beliebiger absolutstetiger Zufallsvektor mit der gemeinsamen Dichte $f_{X,Y} : \mathbb{R}^2 \to [0, \infty]$ und der Randdichte $f_Y : \mathbb{R} \to [0, \infty]$ von Y, wobei $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y)\,\mathrm{d}x$.
   - Die Funktion $g : \mathbb{R} \to [0, \infty)$ mit $g(y) = \int_{\mathbb{R}} |x| f_{X,Y}(x, y)\,\mathrm{d}x$ sei beschränkt.
   - Dann ist eine Version der bedingten Erwartung $\mathbb{E}\,(X \mid Y)$ gegeben durch

$$\mathbb{E}\,(X \mid Y)(\omega) = \begin{cases} \dfrac{\int_{\mathbb{R}} x\, f_{X,Y}(x, Y(\omega))\,\mathrm{d}x}{f_Y(Y(\omega))}\,, & \text{falls } f_Y(Y(\omega)) > 0, \\ 0\,, & \text{falls } f_Y(Y(\omega)) = 0. \end{cases} \tag{38}$$

2. Funktionale unabhängiger Zufallsvariablen
   - Seien $X, Y : \Omega \to \mathbb{R}$ unabhängige Zufallsvariablen, und sei $g : \mathbb{R}^2 \to \mathbb{R}$ eine Borel–messbare Abbildung, so dass die Funktion $g' : \mathbb{R} \to [0, \infty)$ mit $g'(y) = \mathbb{E}\,|g(X, y)|$ beschränkt ist.
   - Dann ist eine Version der bedingten Erwartung $\mathbb{E}\,(g(X, Y) \mid Y)$ gegeben durch

$$\mathbb{E}\,(g(X, Y) \mid Y) = g''(Y)\,, \tag{39}$$

     wobei die Abbildung $g'' : \mathbb{R} \to \mathbb{R}$ gegeben ist durch $g''(y) = \mathbb{E}\,g(X, y)$.
3. Funktionale unabhängiger stochastischer Prozesse
   - Seien $\{X_t,\, t \geq 0\}$ und $\{Y_t,\, t \geq 0\}$ unabhängige stochastische Prozesse, d.h.,
     - $\{X_t\}$ und $\{Y_t\}$ seien unabhängige Familien von $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$–messbaren Abbildungen, und
     - $\mathcal{G} = \sigma\big(\bigcup_{t \geq 0} Y_t^{-1}(\mathcal{B}(\mathbb{R}))\big)$ sei die durch den Prozess $\{Y_t\}$ erzeugte Teil–σ–Algebra von $\mathcal{F}$.
     - Sei $g : (\mathbb{R}^\infty)^2 \to \mathbb{R}$ eine Borel–messbare Abbildung, so dass die Funktion $g' : \mathbb{R}^\infty \to [0, \infty)$ mit $g'(y) = \mathbb{E}\,|g(\{X_t\}, y)|$ beschränkt ist.
   - Für $X = g(\{X_t\}, \{Y_t\})$ gilt dann $\mathbb{E}\,|X| < \infty$, und die bedingte Erwartung $\mathbb{E}\,(X \mid \mathcal{G})$ wird mit dem Symbol $\mathbb{E}\,(X \mid \{Y_t\})$ bezeichnet.
   - Außerdem ist eine Version der bedingten Erwartung $\mathbb{E}\,(X \mid \{Y_t\})$ gegeben durch

$$\mathbb{E}\,(X \mid \{Y_t\}) = g''(\{Y_t\})\,, \tag{40}$$

     wobei die Abbildung $g'' : \mathbb{R}^\infty \to \mathbb{R}$ gegeben ist durch $g''(y) = \mathbb{E}\,g(\{X_t\}, y)$.

Schließlich erwähnen wir noch kurz den Begriff bedingter Wahrscheinlichkeiten bezüglich σ–Algebren.

Definition

- Sei $A \in \mathcal{F}$ ein beliebiges Ereignis, und sei $\mathcal{G} \subset \mathcal{F}$ eine beliebige Teil–σ–Algebra von $\mathcal{F}$.
- Dann ist die *bedingte Wahrscheinlichkeit* $P(A \mid \mathcal{G})$ von A bezüglich $\mathcal{G}$ gegeben durch

$$P(A \mid \mathcal{G}) = \mathbb{E}\,(\mathbb{I}(A) \mid \mathcal{G})\,,$$

  d.h., die Definition bedingter Wahrscheinlichkeiten bezüglich σ–Algebren wird auf den in (37) definierten Begriff der bedingten Erwartung zurückgeführt.

2.2.4 Poissonsche Zählmaße; Cox–Prozesse; Simluationsalgorithmus

Der in Abschnitt 2.2.1 betrachtete Begriff des Poissonschen Zählprozesses $\{N_t, t \geq 0\}$ mit $N_t = \sum_{k=1}^{\infty} \mathbb{1}(S_k \leq t)$ und $S_n = T_1 + \ldots + T_n$, wobei die Zufallsvariablen $T_1, T_2, \ldots : \Omega \to [0, \infty)$ unabhängig und identisch $\mathrm{Exp}(\lambda)$–verteilt sind, lässt sich wie folgt weiter verallgemeinern.

- Anstelle die Anzahl N_t der Sprungpunkte $\{S_n\}$ nur in Intervallen der Form $[0, t]$ für $t \geq 0$ zu betrachten, kann man die Anzahl N_B dieser Punkte in einer *beliebigen* Borel–Menge $B \in \mathcal{B}([0, \infty))$ betrachten, wobei

$$N_B = \sum_{n=1}^{\infty} \mathbb{1}(S_n \in B)\,. \tag{41}$$

- Man kann leicht zeigen, dass der mengen–indizierte Prozess $\{N_B,\, B \in \mathcal{B}([0, \infty))\}$ die Eigenschaften eines zufälligen Zählmaßes hat, d.h., es gilt

$$N_{\bigcup_{n=1}^{\infty} B_n} = \sum_{n=1}^{\infty} N_{B_n} \tag{42}$$

  für beliebige Borel–Mengen $B_1, B_2, \ldots \in \mathcal{B}([0, \infty))$, die paarweise disjunkt sind.

Definition Das in (41) gegebene zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}([0, \infty))\}$, wobei $S_n = T_1 + \ldots + T_n$ und die Zufallsvariablen $T_1, T_2, \ldots : \Omega \to [0, \infty)$ unabhängig und identisch $\mathrm{Exp}(\lambda)$–verteilt sind, wird *homogenes Poissonsches Zählmaß* mit der *Intensität* λ genannt.

Unmittelbar aus Theorem 2.9 ergibt sich die Gültigkeit der folgende beiden Aussagen.

Theorem 2.12 *Sei $\{N_B,\, B \in \mathcal{B}([0, \infty))\}$ ein Poissonsches Zählmaß mit der Intensität λ. Dann gilt:*

- *Die Zufallsvariablen $N_{B_1}, N_{B_2}, \ldots$ sind unabhängig für beliebige beschränkte und paarweise disjunkte Borel–Mengen $B_1, B_2, \ldots \in \mathcal{B}([0, \infty))$.*
- *Für jedes beschränkte $B \in \mathcal{B}([0, \infty))$ gilt $N_B \sim \mathrm{Poi}(\lambda \nu_1(B))$.*

Beachte Mit Hilfe des Satzes über die monotone Konvergenz kann man leicht zeigen, dass die Mengenfunktion $\mu : \mathcal{B}([0, \infty)) \to [0, \infty]$ mit

$$\mu(B) = \mathbb{E}\, N_B\,, \tag{43}$$

die durch das zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}([0, \infty))\}$ induziert wird, σ–additiv ist, d.h., μ ist ein Maß.

Definition Das in (43) gegebene Maß μ wird *Intensitätsmaß* des zufälligen Zählmaßes $\{N_B\}$ genannt.

Beachte

- Für das Intensitätsmaß μ eines homogenen Poissonschen Zählmaßes mit Intensität λ gilt

$$\mu(B) = \lambda \nu_1(B) \qquad \forall\, B \in \mathcal{B}([0, \infty))\,. \tag{44}$$

- Die am Ende von Abschnitt 1.2 erwähnte Variante des Existenzsatzes von Kolmogorow für beliebige Indexmengen I ermöglicht es, Poissonsche Zählmaße mit allgemeineren Intensitätsmaßen als in (44) zu betrachten.

Definition Sei $\mu : \mathcal{B}([0, \infty)) \to [0, \infty]$ ein beliebiges lokal–endliches Maß, d.h., $\mu(B) < \infty$ gilt für jede beschränkte Borel–Menge $B \in \mathcal{B}([0, \infty))$. Das zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}([0, \infty))\}$ wird (inhomogenes) *Poissonsches Zählmaß* mit dem Intensitätsmaß μ genannt, wenn

- die Zufallsvariablen $N_{B_1}, N_{B_2}, \ldots$ unabhängig sind für beliebige beschränkte und paarweise disjunkte Borel–Mengen $B_1, B_2, \ldots \in \mathcal{B}([0,\infty))$ und
- $N_B \sim \mathrm{Poi}(\mu(B))$ für jedes beschränkte $B \in \mathcal{B}([0,\infty))$ gilt.

Man kann sogar noch einen Schritt weiter gehen und sogenannte doppelt–stochastische Poisson–Prozesse betrachten, die in der Literatur auch Cox–Prozesse genannt werden.

Definition Sei $\{\Lambda_b,\, B \in \mathcal{B}([0,\infty))\}$ ein beliebiges zufälliges Maß, das mit Wahrscheinlichkeit 1 lokal–endlich ist.

- Das zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}([0,\infty))\}$ wird *doppelt–stochastisches Poisson–Zählmaß* bzw. *Cox–Zählmaß* mit dem zufälligen Intensitätsmaß Λ genannt, wenn

$$P\Big(\bigcap_{i=1}^{n} \{N_{(a_i,b_i]} = k_i\}\Big) = \mathbb{E}\left(\prod_{i=1}^{n} \frac{\Lambda_{(a_i,b_i]}^{k_i}}{k_i!} \exp\big(-\Lambda_{(a_i,b_i]}\big)\right) \tag{45}$$

  für beliebige $n \in \mathbb{N}$, $k_1, \ldots, k_n \in \{0, 1, \ldots\}$ und $0 \le a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$ gilt.
- Der durch (45) gegebene Zählprozess $\{N_t,\, t \ge 0\}$ mit $N_t = N_{[0,t]}$ heißt *doppelt–stochastischer Poisson–Prozess* bzw. *Cox–Prozess* mit dem zufälligen Intensitätsmaß Λ.

Ein wichtiger Spezialfall liegt dann vor,

- wenn die Realisierungen des zufälligen Intensitätsmaßes Λ mit Wahrscheinlichkeit 1 absolutstetig bezüglich des Lebesgue–Maßes sind, d.h.,
- wenn es einen stochastischen Prozess $\{\lambda_t, t \ge 0\}$ gibt, dessen Trajektorien (mit Wahrscheinlichkeit 1) Borel–messbare und lokal–integrierbare nichtnegative Funktionen sind, so dass für jede beschränkte Borel-Menge $B \in \mathcal{B}([0,\infty))$

$$\Lambda_B = \int_B \lambda_t \,\mathrm{d}t < \infty \tag{46}$$

  mit Wahrscheinlichkeit 1 gilt.
- Der stochastische Prozess $\{\lambda_t\}$ wird dann *Intensitätsprozess* von $\{N_B\}$ genannt.

Im folgenden Theorem wird eine *Zeittransformation* von homogenen Poisson–Prozessen (mit Intensität $\lambda = 1$) betrachtet. Sie liefert einen Algorithmus zur Simulation von Cox–Prozessen mit vorgegebenem Intensitätsprozess.

Theorem 2.13

- *Sei $\{N'_t, t \ge 0\}$ ein (homogener) Poisson–Prozess mit der Intensität $\lambda = 1$, und sei $\{\lambda_t, t \ge 0\}$ ein beliebiger Intensitätsprozess.*
- *Die stochastischen Prozesse $\{N'_t\}$ und $\{\lambda_t\}$ seien unabhängig.*
- *Dann ist der Zählprozess $\{N_t, t \ge 0\}$ mit*

$$N_t = N'_{g_t} \qquad \textit{und} \qquad g_t = \int_0^t \lambda_v \,\mathrm{d}v \tag{47}$$

  ein Cox–Prozess mit dem Intensitätsprozess $\{\lambda_t\}$.

Beweis

- Wir müssen lediglich zeigen, dass (45) gilt.

- Wegen der vorausgesetzten Unabhängigkeit der Prozesse $\{N'_t\}$ und $\{\lambda_t\}$ können wir annehmen, dass der Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, der beide Prozesse „trägt", die Form

$$(\Omega, \mathcal{F}, P) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, P_1 \times P_2)$$

hat, wobei $\{N'_t\}$ über $(\Omega_1, \mathcal{F}_1, P_1)$ und $\{\lambda_t\}$ über $(\Omega_2, \mathcal{F}_2, P_2)$ definiert ist.

- Aus der Definitionsgleichung (47) des Zählprozesses $\{N_t\}$ und aus den in Abschnitt 2.2.3 diskutierten Eigenschaften der bedingten Erwartung ergibt sich, dass

$$\begin{aligned}
P\Big(\bigcap_{i=1}^n \{N_{(a_i,b_i]} = k_i\}\Big) &= P\Big(\bigcap_{i=1}^n \{N_{b_i} - N_{a_i} = k_i\}\Big) \\
&= P\Big(\bigcap_{i=1}^n \{N'_{\int_0^{b_i} \lambda_v \,\mathrm{d}v} - N'_{\int_0^{a_i} \lambda_v \,\mathrm{d}v} = k_i\}\Big) = \mathbb{E}\,\mathbb{1}\big(\bigcap_{i=1}^n \{N'_{\int_0^{b_i} \lambda_v \,\mathrm{d}v} - N'_{\int_0^{a_i} \lambda_v \,\mathrm{d}v} = k_i\}\big) \\
&= \mathbb{E}\Big(\prod_{i=1}^n \mathbb{1}\big(N'_{\int_0^{b_i} \lambda_v \,\mathrm{d}v} - N'_{\int_0^{a_i} \lambda_v \,\mathrm{d}v} = k_i\big)\Big) = \mathbb{E}\Big(\mathbb{E}\,\big(\prod_{i=1}^n \mathbb{1}\big(N'_{\int_0^{b_i} \lambda_v \,\mathrm{d}v} - N'_{\int_0^{a_i} \lambda_v \,\mathrm{d}v} = k_i\big) \mid \{\lambda_t\}\big)\Big) \\
&= \int_{\Omega_2} \mathbb{E}\,\big(\prod_{i=1}^n \mathbb{1}\big(N'_{\int_0^{b_i} \lambda_v(\omega_2) \,\mathrm{d}v} - N'_{\int_0^{a_i} \lambda_v(\omega_2) \,\mathrm{d}v} = k_i\big) P_2(\mathrm{d}\omega_2) \\
&= \mathbb{E}\Big(\prod_{i=1}^n \mathbb{E}\,\big(\mathbb{1}\big(N'_{\int_0^{b_i} \lambda_v \,\mathrm{d}v} - N'_{\int_0^{a_i} \lambda_v \,\mathrm{d}v} = k_i\big) \mid \{\lambda_t\}\big)\Big),
\end{aligned}$$

wobei sich die beiden letzten Gleichheiten aus den Unabhängigkeitseigenschaften der Prozesse $\{N'_t\}$ und $\{\lambda_t\}$ und aus der Darstellungsformel (40) der regulären bedingten Erwartung für Funktionale unabhängiger stochastischer Prozesse ergeben.

- Weil $\{N'_t\}$ ein homogener Poisson–Prozess mit der Intensität $\lambda = 1$ ist, der unabhängig von $\{\lambda_t\}$ ist, ergibt sich durch die erneute Anwendung der Darstellungsformel (40), dass für jedes $i = 1, \ldots, n$

$$\mathbb{E}\,\big(\mathbb{1}\big(N'_{\int_0^{b_i} \lambda_v \,\mathrm{d}v} - N'_{\int_0^{a_i} \lambda_v \,\mathrm{d}v} = k_i\big) \mid \{\lambda_t\}\big) = \frac{\big(\int_{a_i}^{b_i} \lambda_v \,\mathrm{d}v\big)^{k_i}}{k_i!} \exp\big(-\int_{a_i}^{b_i} \lambda_v \,\mathrm{d}v\big).$$

- Hieraus ergibt sich, dass

$$P\Big(\bigcap_{i=1}^n \{N_{(a_i,b_i]} = k_i\}\Big) = \mathbb{E}\,\Big(\prod_{i=1}^n \frac{\big(\int_{a_i}^{b_i} \lambda_v \,\mathrm{d}v\big)^{k_i}}{k_i!} \exp\big(-\int_{a_i}^{b_i} \lambda_v \,\mathrm{d}v\big)\Big),$$

wobei der letzte Ausdruck mit der rechten Seite von (45) übereinstimmt. □

Bespiele

1. Eine spezielle Klasse von Cox–Prozessen sind die *gemischten Poisson–Prozesse*, deren Intensitätsprozess $\{\lambda_t\}$ durch $\lambda_t = Z$ gegeben ist, wobei $Z : \Omega \to [0, \infty)$ eine beliebige nichtnegative Zufallsvariable ist.
2. Sei $\lambda' : [0, \infty) \to [0, \infty)$ eine periodische (und deterministische) Funktion mit der Periode 1, und sei $\{N'_t, t \geq 0\}$ ein (homogener) Poisson–Prozess mit der Intensität $\lambda = 1$.
   - Dann wird der Zählprozess $\{N_t, t \geq 0\}$ mit

$$N_t = N'_{g_t} \qquad \text{und} \qquad g_t = \int_0^t \lambda'_v \,\mathrm{d}v$$

   ein *periodischer Poisson–Prozess* genannt, der im allgemeinen keine stationären Zuwächse hat.

- Es ist jedoch möglich, auf die folgende Weise einen entsprechenden Cox–Prozess mit stationären Zuwächsen zu konstruieren.
  - Sei nämlich $U : \Omega \to [0,1]$ eine $[0,1]$–gleichverteilte Zufallsvariable, die unabhängig von $\{N'_t\}$ ist. Außerdem sei
  $$\lambda_t = \lambda'_{t+U} \qquad \text{und} \qquad \Lambda_t = \int_0^t \lambda_s \, \mathrm{d}s \, .$$
  - Dann ergibt sich aus Theorem 2.13, dass $\{N^*_t, t \geq 0\}$ mit $N^*_t = N'_{\Lambda_t}$ ein Cox–Prozess ist.
  - Außerdem kann man sich leicht überlegen, dass der Cox–Prozess $\{N^*_t\}$ stationäre Zuwächse hat.

2.3 Markow–Prozesse mit endlich vielen Zuständen

Wir beginnen mit einem kurzen Rückblick auf Markow-Prozesse mit diskreter Zeit, die in der Literatur *Markow–Ketten* genannt werden und die beispielsweise in der Vorlesung „Markow–Ketten und Monte–Carlo–Simulation" im SS 2003 ausführlich behandelt wurden.

- Das stochastische Modell der zeitdiskreten Markow-Ketten (mit endlich vielen Zuständen) besteht aus den drei Komponenten: Zustandsraum, Anfangsverteilung und Übergangsmatrix.
  - Den Ausgangspunkt bildet die (endliche) Menge aller möglichen Zustände, die *Zustandsraum* der Markow-Kette genannt wird und die o.B.d.A. mit der Menge $E = \{1, 2, \ldots, \ell\}$ identifiziert werden kann, wobei $\ell \in \{1, 2, \ldots\}$ eine beliebige, jedoch fest vorgegebene natürliche Zahl ist.
  - Für jedes $i \in E$ sei α_i die Wahrscheinlichkeit, dass sich das betrachtete System zum „Zeitpunkt" $n = 0$ im Zustand i befindet, wobei
  $$\alpha_i \in [0,1] \, , \qquad \sum_{i=1}^{\ell} \alpha_i = 1 \tag{1}$$
  vorausgesetzt wird. Der Vektor $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell)^\top$ der (Einzel-) Wahrscheinlichkeiten $\alpha_1, \ldots, \alpha_\ell$ bildet die *Anfangsverteilung* der Markow-Kette.
  - Außerdem betrachten wir für jedes Paar $i, j \in E$ die (bedingte) Wahrscheinlichkeit $p_{ij} \in [0,1]$, dass das betrachtete System in einem Schritt aus dem Zustand i in den Zustand j übergeht.
  - Die $\ell \times \ell$ Matrix $\mathbf{P} = (p_{ij})_{i,j=1,\ldots,\ell}$ der *Übergangswahrscheinlichkeiten* p_{ij} mit
  $$p_{ij} \geq 0 \, , \qquad \sum_{j=1}^{\ell} p_{ij} = 1 \, , \tag{2}$$
  heißt (einstufige) *Übergangsmatrix* der Markow-Kette.
- Für jede Menge $E = \{1, 2, \ldots, \ell\}$, für jeden Vektor $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell)^\top$ bzw. jede Matrix $\mathbf{P} = (p_{ij})$, die den Bedingungen (1) und (2) genügen, kann nun der Begriff der zugehörigen Markow-Kette wie folgt eingeführt werden.

Definition

- Sei $X_0, X_1, \ldots : \Omega \to E$ eine Folge von Zufallsvariablen, die über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ definiert sind und ihre Werte in der Menge $E = \{1, 2, \ldots, \ell\}$ annehmen.
- Dann heißt $X_0, X_1, \ldots$ (homogene) *Markow-Kette* mit der Anfangsverteilung $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell)^\top$ und der Übergangsmatrix $\mathbf{P} = (p_{ij})$, falls
$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \alpha_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} \tag{3}$$
für beliebige $n = 0, 1, \ldots$ und $i_0, i_1, \ldots, i_n \in E$ gilt.

Beachte

- Für beliebige $n \geq 1$ und $i, j \in E$ kann das Produkt $p_{ii_1} p_{i_1 i_2} \dots p_{i_{n-1} j}$ als die Wahrscheinlichkeit des „Pfades" $i \to i_1 \to \dots \to i_{n-1} \to j$ aufgefasst werden.
- Analog hierzu kann die Summe

$$p_{ij}^{(n)} = \sum_{i_1, \dots, i_{n-1} \in E} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} \tag{4}$$

als die Wahrscheinlichkeit des Überganges vom Zustand i in den Zustand j in n Schritten aufgefasst werden, wobei

$$p_{ij}^{(n)} = P(X_n = j \mid X_0 = i)\,, \tag{5}$$

falls $P(X_0 = i) > 0$.

- Die stochastische Matrix $\mathbf{P}^{(n)} = (p_{ij}^{(n)})_{i,j=1,\dots,\ell}$ wird die *n-stufige Übergangsmatrix* der Markow-Kette $\{X_n\}$ genannt.
- Mit der Schreibweise $\mathbf{P}^{(0)} = \mathbf{I}$, wobei $\mathbf{I}$ die $\ell \times \ell$-dimensionale Einheitsmatrix bezeichnet, lässt sich sich die mehrstufigen Übergangsmatrizen $\mathbf{P}^{(n)}$ wie folgt darstellen:
  - Für beliebige $n, m = 0, 1, \dots$ gilt

$$\mathbf{P}^{(n)} = \mathbf{P}^n \tag{6}$$

  - und somit

$$\mathbf{P}^{(n+m)} = \mathbf{P}^{(n)} \mathbf{P}^{(m)}. \tag{7}$$

- Die Matrix-Identität (7) wird in der Literatur *Gleichung von Chapman-Kolmogorov* genannt.

2.3.1 Anfangsverteilung und Übergangsfunktion

Um den Begriff eines Markow–Prozesses $\{X_t,\, t \geq 0\}$ mit stetiger Zeit und mit Werten in der Menge $E = \{1, 2, \dots, \ell\}$ einzuführen, betrachten wir erneut

- einen Vektor von Wahrscheinlichkeiten $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_\ell)^\top$ mit $\sum_{i \in E} \alpha_i = 1$ sowie
- eine Familie von stochastischen Matrizen $\mathbf{P}(h) = (p_{ij}(h))_{i,j \in E}$ für jedes $h \geq 0$.

Beachte

- Im Unterschied zu Markow–Prozessen mit diskreter Zeit muss nun *vorausgesetzt* werden, dass

$$\mathbf{P}(h_1 + h_2) = \mathbf{P}(h_1)\mathbf{P}(h_2) \tag{8}$$

für beliebige $h_1, h_2 \geq 0$ gilt, wobei die Matrix-Identität (8) wiederum die *Gleichung von Chapman-Kolmogorow* genannt wird.

- Außerdem wird vorausgesetzt, dass

$$\lim_{h \downarrow 0} \mathbf{P}(h) = \mathbf{P}(0) = \mathbf{I}\,. \tag{9}$$

- Es ist nicht schwierig zu zeigen, dass die Matrix–Funktion $\mathbf{P}(h)$ dann gleichmäßig stetig in $h \geq 0$ ist.

Definition

1. Eine Familie von stochastischen Matrizen $\{\mathbf{P}(h),\ h > 0\}$, die den Bedingungen (8) und (9) genügt, wird *Matrix–Übergangsfunktion* bzw. kurz *Übergangsfunktion* genannt.

2. Ein stochastischer Prozess $\{X_t, t \geq 0\}$ mit Werten in $E = \{1, 2, \ldots, \ell\}$ heißt *homogener Markow–Prozess*, wenn es eine Übergangsfunktion $\{\mathbf{P}(h), h \geq 0\}$ und eine Wahrscheinlichkeitsfunktion $\boldsymbol{\alpha}$ über E gibt, so dass

$$P(X_0 = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = \alpha_{i_0} p_{i_0 i_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \ldots p_{i_{n-1} i_n}(t_n - t_{n-1}), \tag{10}$$

für beliebige $n = 0, 1, \ldots,\ i_0, i_1, \ldots, i_n \in E,\ 0 \leq t_1 \leq \ldots \leq t_n$ gilt.

3. Die Wahrscheinlichkeitsfunktion $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell)^\top$ heißt *Anfangsverteilung* des Markow–Prozesses $\{X_t\}$.

Ähnlich wie für Markow–Ketten in Theorem MK-2.1 lässt sich die folgende Charakterisierung von homogenen Markow–Prozessen mit stetiger Zeit herleiten, die wir deshalb ohne Beweis angeben. Das Adjektiv „homogen" werden wir im Folgenden der Kürze wegen weglassen.

Theorem 2.14 *Ein stochastischer Prozess $\{X_t, t \geq 0\}$ mit Werten in $E = \{1, 2, \ldots, \ell\}$ ist genau dann ein Markow–Prozess, wenn es eine Übergangsfunktion $\{\mathbf{P}(h), h \geq 0\}$ gibt, so dass*

$$P(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_1} = i_1, X_0 = i_0) = p_{i_{n-1} i_n}(t_n - t_{n-1}) \tag{11}$$

für diejenigen $n \geq 1$, $i_0, i_1, \ldots, i_n \in E$ und $0 \leq t_1 \leq \ldots \leq t_n$, für die $P(X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_1} = i_1, X_0 = i_0) > 0$ gilt.

Analog zu Korollar MK-2.1 gilt auch für Markow–Prozesse mit stetiger Zeit die folgende bedingte Unabhängigkeitseigenschaft.

Korollar 2.1 *Sei $\{X_t\}$ ein Markow–Prozess. Dann gilt*

$$P(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_1} = i_1, X_0 = i_0) = P(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}), \tag{12}$$

falls $P(X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_1} = i_1, X_0 = i_0) > 0$.

2.3.2 Übergangsintensitäten

Mit der Schreibweise $\delta_{ij} = 1$, falls $i = j$, und $\delta_{ij} = 0$, falls $i \neq j$, können wir die folgende grundlegende Differenzierbarkeitseigenschaft von Übergangsfunktionen formulieren.

Theorem 2.15 *Sei $\{\mathbf{P}(h), h \geq 0\}$ eine Übergangsfunktion. Dann existieren die (endlichen) Grenzwerte*

$$q_{ij} = \lim_{h \downarrow 0} \frac{1}{h} \left(p_{ij}(h) - \delta_{ij}\right). \tag{13}$$

Beweis

- Weil in der Behauptung von Theorem 2.15 die Anfangsverteilung nicht näher spezifiziert wird, können wir ohne Einschränkung der Allgemeinheit annehmen,
  - dass $P(X_0 = i) > 0$ für jedes $i \in E$ gilt,
  - wobei wir die Gültigkeit von (13) zuerst für $i \neq j$ zeigen.
- Dabei setzen wir $p_{ii}^{j0}(h) = 1$ und

$$\begin{aligned} p_{ii}^{jv}(h) &= P(X_{vh} = i; X_{kh} \neq j, 1 \leq k < v \mid X_0 = i), \\ f_{ij}^{v}(h) &= P(X_{vh} = j; X_{kh} \neq j, 1 \leq k < v \mid X_0 = i). \end{aligned}$$

- Aus (8) ergibt sich dann, dass

$$p_{ij}(nh) \geq \sum_{v=0}^{n-1} p_{ii}^{jv}(h) p_{ij}(h) p_{jj}((n-v-1)h) \tag{14}$$

und

$$p_{ii}(vh) = p_{ii}^{jv}(h) + \sum_{m=1}^{v-1} f_{ij}^{m}(h) p_{ji}((v-m)h)\,. \tag{15}$$

- Weil $\sum_{m=1}^{v-1} f_{ij}^{m}(h) \leq 1$ gilt, ergibt sich aus (15) die Ungleichung

$$p_{ii}^{jv}(h) \geq p_{ii}(vh) - \max_{1 \leq m < v} p_{ji}((v-m)h)\,. \tag{16}$$

- Mit Hilfe von (9) erhalten wir nun, dass es für beliebige $\varepsilon > 0$ und $i, j \in E$ mit $i \neq j$ ein $h_0 > 0$ gibt, so dass

$$\max_{0 \leq h \leq h_0} p_{ji}(h) < \varepsilon, \quad \min_{0 \leq h \leq h_0} p_{ii}(h) > 1 - \varepsilon, \quad \min_{0 \leq h \leq h_0} p_{jj}(h) > 1 - \varepsilon. \tag{17}$$

  - Für $nh < h_0$ und $v \leq n$ ergibt sich somit aus (16), dass $p_{ii}^{jv}(h) > 1 - 2\varepsilon$.
  - Wenn man dies in (14) einsetzt, dann ergibt sich die Ungleichung

$$p_{ij}(nh) \geq (1-2\varepsilon) \sum_{v=0}^{n-1} p_{ij}(h)(1-\varepsilon) \geq (1-3\varepsilon) n p_{ij}(h)$$

    bzw.

$$\frac{p_{ij}(nh)}{nh} \geq (1-3\varepsilon) \frac{p_{ij}(h)}{h} \tag{18}$$

    für $nh < h_0$.

- Mit der Schreibweise $a_{ij} = \liminf_{h \to 0} p_{ij}(h)/h$ ergibt dies, dass $a_{ij} < \infty$.
  - Denn, wenn $a_{ij} = \infty$ gelten würde, dann würde es ein beliebig kleines h geben, so dass $p_{ij}(h)/h$ beliebig groß ist.
  - Wegen (18), würde auch $p_{ij}(nh)/nh$ beliebig groß werden.
  - Wenn andererseits n so gewählt wird, dass $h_0/2 \leq nh < h_0$, dann ergibt (17), dass

$$\frac{p_{ij}(nh)}{nh} < \frac{\varepsilon}{nh} < h_0^{-1} 2\varepsilon\,.$$

- Somit gilt $a_{ij} < \infty$, und es genügt zu zeigen, dass

$$\limsup_{h \to 0} \frac{p_{ij}(h)}{h} \leq a_{ij}\,. \tag{19}$$

  - Aus der Definition von a_{ij} folgt, dass es ein $h_1 < h_0$ gibt mit $h_1^{-1} p_{ij}(h_1) < a_{ij} + \varepsilon$.
  - Weil $p_{ij}(h)$ stetig ist, gilt für jedes hinreichend kleine t_0 mit $h_1 + t_0 < h_0$ die Ungleichung

$$\frac{p_{ij}(t)}{t} < a_{ij} + \varepsilon$$

    für $h_1 - t_0 < t < h_1 + t_0$.
  - Aus (18) ergibt sich nun, dass man für jedes $h < t_0$ eine ganze Zahl n_h finden kann, so dass

$$h_1 - t_0 < n_h h < h_1 + t_0 \qquad \text{und} \qquad (1-3\varepsilon)\,\frac{p_{ij}(h)}{h} \leq \frac{p_{ij}(n_h h)}{n_h h} < a_{ij} + \varepsilon\,.$$

  - Weil $\varepsilon > 0$ beliebig ist, ergibt sich hieraus die Gültigkeit von (19).

- Damit ist die Existenz der Grenzwerte q_{ij} in (13) für $i \neq j$ bewiesen.
- Weil der Zustandsraum E endlich ist und $\mathbf{P}(h)$ eine stochastische Matrix ist, gilt außerdem

$$\lim_{h\downarrow 0} \frac{p_{ii}(h)-1}{h} = -\lim_{h\downarrow 0} \sum_{j\neq i} \frac{p_{ij}(h)}{h} = -\sum_{j\neq i} \lim_{h\downarrow 0} \frac{p_{ij}(h)}{h} = -\sum_{j\neq i} q_{ij}\,. \tag{20}$$

Damit ist die Behauptung vollständig bewiesen. □

Definition Die Matrix $\mathbf{Q} = (q_{ij})_{i,j=1,\ldots,\ell}$ heißt *Intensitätsmatrix* des Markow–Prozesses $\{X_t\}$. Die Eintragungen q_{ij} von $\mathbf{Q}$ werden *Übergangsintensitäten* genannt.

Korollar 2.2 *Für beliebige $i \neq j$ gilt $q_{ij} \geq 0$ und $q_{ii} \leq 0$. Außerdem gilt $\mathbf{Qe} = \mathbf{0}$ bzw. (äquivalent hierzu)*

$$\sum_{j\in E} q_{ij} = 0 \qquad \forall\, i \in E\,, \tag{21}$$

wobei $\mathbf{e} = (1,\ldots,1)^\top$ bzw. $\mathbf{0} = (0,\ldots,0)^\top$ den ℓ–dimensionalen Einheits– bzw. Nullvektor bezeichnet.

Beweis

- Aus der Definition (13) der Übergangsintensitäten q_{ij} ergibt sich unmittelbar, dass $q_{ij} \geq 0$ und $q_{ii} \leq 0$ für beliebige $i \neq j$ gilt.
- Die Behauptung (21) folgt somit aus (20). □

Beachte

- Die Definition von Markow–Prozessen und ihre Charakterisierung in Theorem 2.14 lassen sich völlig analog für Markow–Prozesse mit einem (abzählbar) unendlichen Zustandsraum formulieren, beispielsweise für $E = \mathbb{N}$.
- Die in Theorem 2.15 hergeleitete Differenzierbarkeitseigenschaft der Übergangsfunktion $\{\mathbf{P}(h), h \geq 0\}$ bleibt ebenfalls (in einer etwas modifizierten Form) gültig.
  - Im Beweis von Theorem 2.15 wird nämlich die Endlichkeit des Zustandsraumes E *nicht* benötigt, um die Existenz und Endlichkeit der Übergangsintensitäten q_{ij} für $i \neq j$ zu zeigen.
  - Außerdem kann man auch zeigen, dass die Grenzwerte q_{ii} in (13) existieren, wenn der Zustandsraum E unendlich ist; sie müssen aber nicht notwendig endlich sein.
  - Im allgemeinen kann man dann (anstelle von (20)) nur zeigen, dass $-q_{ii} \geq \sum_{j\neq i} q_{ij}$ für jedes $i \in E$ gilt, wobei der Fall der Gleichheit von besonderer Bedeutung ist.

Definition Wenn der Zustandsraum E unendlich ist, dann sagt man, dass die Übergangsfunktion $\{\mathbf{P}(h), h \geq 0\}$ *konservativ* ist, falls

$$\sum_{j\neq i} q_{ij} = -q_{ii} < \infty \qquad \forall\, i \in E\,. \tag{22}$$

Beachte Die meisten Ergebnisse, die wir in dieser Vorlesung für Markow–Prozesse mit endlichem Zustandsraum herleiten, bleiben für konservative Übergangsfunktionen in (abzählbar) unendlichen Zustandsräumen gültig. Die Beweise sind dann jedoch oft wesentlich aufwändiger.

Beispiel

- Man kann sich leicht überlegen, dass jeder stochastische Prozess $\{X_t,\, t \geq 0\}$ mit Werten in $E = \{0, 1, \ldots\}$, der unabhängige und stationäre Zuwächse hat, ein Markow–Prozess ist.
- Insbesondere ist somit der (homogene) Poisson–Prozess mit Intensität λ ein Markow–Prozess.

– In diesem Fall gilt $\alpha_0 = 1$ und

$$p_{ij}(h) = \begin{cases} \mathrm{e}^{-\lambda h}\dfrac{(\lambda h)^{j-i}}{(j-i)!}\,, & \text{falls } j \geq i, \\ 0, & \text{sonst.} \end{cases} \tag{23}$$

– Hieraus folgt, dass

$$q_{ij} = \begin{cases} \lambda\,, & \text{falls } j = i+1, \\ -\lambda\,, & \text{falls } j = i, \\ 0, & \text{sonst.} \end{cases}$$

2.3.3 Kolmogorowsche Differentialgleichungen

- In diesem Abschnitt zeigen wir, dass es einen eineindeutigen Zusammenhang zwischen (Matrix-) Übergangsfunktionen und ihren Intensitätsmatrizen gibt.
- Dabei zeigen wir zunächst in Verallgemeinerung von Theorem 2.15, dass die Übergangsfunktionen $p_{ij}(h)$ für *jedes* $h \geq 0$ differenzierbar sind.

Theorem 2.16 *Für beliebige $i, j \in E$ und $h \geq 0$ sind die Übergangsfunktionen $p_{ij}(h)$ differenzierbar und genügen dem folgenden System von Differentialgleichungen:*

$$p_{ij}^{(1)}(h) = \sum_{k \in E} q_{ik} p_{kj}(h) \qquad \forall\, i, j \in E,\ h \geq 0\,. \tag{24}$$

Beweis

- Sei $h, h' > 0$ mit $h' < h$. Aus der Gleichung von Chapman–Kolmogorow, d.h. aus (8), ergibt sich dann, dass

$$\begin{aligned} p_{ij}(h+h') - p_{ij}(h) &= \sum_{k \in E} p_{ik}(h') p_{kj}(h) - p_{ij}(h) \\ &= \sum_{k \neq i} p_{ik}(h') p_{kj}(h) + \big(p_{ii}(h') - 1\big) p_{ij}(h)\,. \end{aligned}$$

- Auf die gleiche Weise erhalten wir die Gleichungen:

$$\begin{aligned} p_{ij}(h-h') - p_{ij}(h) &= p_{ij}(h-h') - \sum_{k \in E} p_{ik}(h') p_{kj}(h-h') \\ &= -\sum_{k \neq i} p_{ik}(h') p_{kj}(h-h') - \big(p_{ii}(h') - 1\big) p_{ij}(h-h')\,. \end{aligned}$$

- Wenn wir nun beide Seiten durch h' dividieren, den Grenzwert $h' \downarrow 0$ bilden und die Stetigkeit von $p_{ij}(h)$ beachten, dann erhalten wir (24).
- Dabei nutzen wir die beiden folgenden Gleichungen, die sich aus Theorem 2.15 ergeben:

$$\lim_{h' \downarrow 0} \frac{1}{h'} \sum_{k \neq i} p_{ik}(h') p_{kj}(h) = \sum_{k \neq i} q_{ik} p_{kj}(h)$$

und

$$\lim_{h' \downarrow 0} \frac{1}{h'} \sum_{k \neq i} p_{ik}(h') p_{kj}(h-h') = \sum_{k \neq i} q_{ik} p_{kj}(h)\,. \qquad \square$$

Beachte

- Die Differentialgleichungen in (24) werden *Kolmogorowsche Rückwärtsgleichungen* genannt.
- In Matrix–Schreibweise nimmt (24) die folgende Form an:

$$\mathbf{P}^{(1)}(h) = \mathbf{Q}\mathbf{P}(h) \qquad \forall\, h \geq 0\,. \tag{25}$$

- Auf die gleiche Weise wie im Beweis von (24) lässt sich zeigen, dass

$$\mathbf{P}^{(1)}(h) = \mathbf{P}(h)\mathbf{Q} \qquad \forall\, h \geq 0\,. \tag{26}$$

- Dies ist die Matrix–Schreibweise der *Kolmogorowschen Vorwärtsgleichung.*
- Die Anfangsbedingung für beide Kolmogorowschen Differentialgleichungen ist $\mathbf{P}(0) = \mathbf{I}$.

Zur Lösung der Differentialgleichungen (25) und (26) benötigen wir Hilfsmittel aus der Matrizenrechnung.

- Dabei setzen wir in diesem Abschnitt stets voraus, dass die betrachteten Matrizen die Dimension $\ell \times \ell$ haben, wobei Vektoren die Dimension $1 \times \ell$ haben.
- Die Konvergenz von Matrizen bzw. Vektoren erfolgt eintragungsweise.
- Wenn beispielsweise $\{\mathbf{A}_n\}$ eine Folge von Matrizen ist, dann bedeutet $\mathbf{A}_n \to \mathbf{0}$ für $n \to \infty$, dass $(\mathbf{A}_n)_{ij} \to 0$ für beliebige $i, j = 1, \ldots, \ell$.
- Die Matrix–Norm definieren wir so, dass die oben beschriebene Topologie induziert wird:
  - Für jeden Vektor $\mathbf{x} = (x_1, \ldots, x_\ell)^\top$ sei $\|\mathbf{x}\| = \sum_{i=1}^{\ell} |x_i|$, und
  - für jede $\ell \times \ell$–Matrix $\mathbf{A} = (a_{ij})$ sei $\|\mathbf{A}\| = \sum_{i,j=1,\ldots,\ell} |a_{ij}|$.

Beachte

- Für jedes $h \in \mathbb{R}$ gilt $\|h\mathbf{A}\| = |h|\,\|\mathbf{A}\|$, wobei $\|\mathbf{A}\| = 0$ genau dann, wenn $\mathbf{A} = \mathbf{0}$.
- Es ist klar, dass $\mathbf{A}_n \to \mathbf{0}$ genau dann, wenn $\|\mathbf{A}_n\| \to 0$.
- Außerdem gilt für beliebige $\ell \times \ell$–Matrizen $\mathbf{A}, \mathbf{B}$

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|\,, \qquad \|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\|\,\|\mathbf{B}\|\,. \tag{27}$$

Lemma 2.2 *Für jedes $h_0 > 0$ konvergiert die Reihe $\sum_{n=0}^{\infty}(h\mathbf{A})^n/(n!)$ gleichmäßig in $h \in [-h_0, h_0]$.*

Beweis Sei $h \in [-h_0, h_0]$, $\varepsilon > 0$ und $m \in \mathbb{N}$. Dann ergibt sich aus (27), dass

$$\Big\|\sum_{n=0}^{\infty}\frac{(h\mathbf{A})^n}{n!} - \sum_{n=0}^{m}\frac{(h\mathbf{A})^n}{n!}\Big\| = \Big\|\sum_{n=m+1}^{\infty}\frac{(h\mathbf{A})^n}{n!}\Big\| \leq \sum_{n=m+1}^{\infty}\frac{|h|^n\|\mathbf{A}\|^n}{n!} \leq \sum_{n=m+1}^{\infty}\frac{h_0^n\|\mathbf{A}\|^n}{n!} < \varepsilon$$

für jedes hinreichend große m, und zwar gleichmäßig in $h \in [-h_0, h_0]$. □

Definition

1. Die Reihe $\sum_{n=0}^{\infty}(h\mathbf{A})^n/(n!)$ ist somit eine wohldefinierte Matrix–Funktion, die stetig für jedes $h \in \mathbb{R}$ ist. Sie wird *Matrix–Exponentialfunktion* genannt und mit

$$\exp(h\mathbf{A}) = \mathbf{I} + h\mathbf{A} + \ldots + \frac{(h\mathbf{A})^n}{n!} + \ldots \tag{28}$$

bezeichnet.

2. Sei $\mathbf{A}(h)$ eine beliebige Matrix–Funktion, so dass sämtliche Eintragungen differenzierbar in h sind. Dann ist die *Matrix–Ableitung* $\mathbf{A}^{(1)}(h)$ gegeben durch

$$\mathbf{A}^{(1)}(h) = \lim_{h' \to 0} \frac{1}{h'} \left(\mathbf{A}(h+h') - \mathbf{A}(h)\right).$$

Lemma 2.3 *Die Matrix–Exponentialfunktion* $\exp(h\mathbf{A})$ *ist für jedes* $h \in \mathbb{R}$ *differenzierbar, und es gilt*

$$\frac{\mathrm{d}\exp(h\mathbf{A})}{\mathrm{d}h} = \mathbf{A}\exp(h\mathbf{A}) = \exp(h\mathbf{A})\mathbf{A}. \tag{29}$$

Beweis

- Es gilt

$$\frac{\exp((h+h')\mathbf{A}) - \exp(h\mathbf{A})}{h'} = \sum_{n=1}^{\infty} \frac{(h+h')^n - h^n}{h'} \frac{\mathbf{A}^n}{n!} = \sum_{n=1}^{\infty} nh^{n-1}\frac{\mathbf{A}^n}{n!} + h' \sum_{n=1}^{\infty} r_n(h,h')\frac{\mathbf{A}^n}{n!},$$

  wobei für jedes $|h'| \le |h|$

$$0 \le |r_n(h,h')| \le n(n-1)(2|h|)^n. \tag{30}$$

- Die Schranke (30) ergibt sich dabei durch Taylor-Reihenentwicklung der Funktion $g(x) = (h+x)^n - h^n$.
- Für ein $\theta \in [0,1]$ gilt also

$$g(x) = xnh^{n-1} + \frac{x^2}{2}n(n-1)(h+\theta x)^{n-2}.$$

- Für $h' \to 0$ ergibt sich hieraus, dass

$$\frac{\mathrm{d}\exp(h\mathbf{A})}{\mathrm{d}h} = \sum_{n=1}^{\infty} nh^{n-1}\frac{\mathbf{A}^n}{n!}.$$

- Außerdem gilt offenbar

$$\sum_{n=1}^{\infty} nh^{n-1}\frac{\mathbf{A}^n}{n!} = \mathbf{A}\sum_{n=0}^{\infty}\frac{(h\mathbf{A})^n}{n!} = \sum_{n=0}^{\infty}\frac{(h\mathbf{A})^n}{n!}\mathbf{A}.$$

□

Beachte Die $\ell \times \ell$–Matrizen $\mathbf{A}, \mathbf{A}'$ heißen *kommutativ*, wenn $\mathbf{A}\mathbf{A}' = \mathbf{A}'\mathbf{A}$ gilt. Es ist nicht schwierig zu zeigen, dass in diesem Fall

$$\exp(\mathbf{A} + \mathbf{A}') = \exp(\mathbf{A})\exp(\mathbf{A}'). \tag{31}$$

Lemma 2.4

- *Sei* $\mathbf{Q}$ *eine beliebige* $\ell \times \ell$*–Matrix, so dass* $q_{ij} \ge 0$ *für* $i \ne j$ *und* $\mathbf{Qe} = \mathbf{0}$ *gilt.*
- *Dann ist die Matrix–Exponentialfunktion* $\{\exp(h\mathbf{Q}), h \ge 0\}$ *eine Übergangsfunktion, die den Kolmogorowschen Differentialgleichungen* (25) *und* (26) *genügt, d.h.,*
- *die Intensitätsmatrix der Übergangsfunktion* $\{\exp(h\mathbf{Q}), h \ge 0\}$ *ist gleich* $\mathbf{Q}$.

Beweis

- Wenn $\mathbf{Q} = \mathbf{0}$, dann ist die Aussage offensichtlich. Es gelte also nun $\mathbf{Q} \ne \mathbf{0}$.
- Wir überprüfen zunächst, dass $\exp(h\mathbf{Q})$ für jedes $h \ge 0$ eine stochastische Matrix ist.
  - Weil $\mathbf{Qe} = \mathbf{0}$, gilt $\mathbf{Q}^n\mathbf{e} = \mathbf{0}$ für jedes $n = 1, 2, \ldots$

- und somit $\exp(h\mathbf{Q})\mathbf{e} = \mathbf{e}$.

- Um zu zeigen, dass sämtliche Eintragungen von $\exp(h\mathbf{Q})$ nichtnegativ sind, setzen wir

$$a = \max\{-q_{ii} : i = 1, \ldots, \ell\}\,.$$

  - Dann ist die Matrix $\tilde{\mathbf{P}}$ mit

$$\tilde{\mathbf{P}} = a^{-1}\mathbf{Q} + \mathbf{I} \tag{32}$$

    nichtnegativ, und wegen $\tilde{\mathbf{P}}\mathbf{e} = a^{-1}\mathbf{Q}\mathbf{e} + \mathbf{I}\mathbf{e} = \mathbf{e}$ ist $\tilde{\mathbf{P}}$ auch stochastisch.
  - Aus der Darstellungsformel

$$\exp(h\mathbf{Q}) = \exp(ah(\tilde{\mathbf{P}} - \mathbf{I})) = \sum_{n=0}^{\infty} \frac{(ah)^n(\tilde{\mathbf{P}})^n}{n!}\,\mathrm{e}^{-ah} \tag{33}$$

    ergibt sich nun, dass die Eintragungen von $\exp(h\mathbf{Q})$ nichtnegativ sind.
  - Dabei wird bei der Herleitung von (33) die Identität (31) benutzt zusammen mit der Tatsache, dass $\exp(-ah\mathbf{I}) = \mathrm{e}^{-ah}\mathbf{I}$.

- Außerdem ergibt sich aus (31), dass $\exp(h\mathbf{Q})$ der Chapman–Kolmogorow–Gleichung (8) genügt.
- Mit Hilfe von (29) ergibt sich nun, dass $\mathbf{Q}$ die Intensitätsmatrix der Übergangsfunktion $\exp(h\mathbf{Q})$ ist, d.h., $\exp(h\mathbf{Q})$ ist eine Lösung von (25) und (26). □

Theorem 2.17 *Jede Übergangsfunktion $\{\mathbf{P}(h), h \geq 0\}$ lässt sich wie folgt durch ihre Intensitätsmatrix $\mathbf{Q}$ ausdrücken: Es gilt*

$$\mathbf{P}(h) = \exp(h\mathbf{Q}) \qquad \forall\, h \geq 0\,. \tag{34}$$

Beweis

- Aus Lemma 2.4 folgt, dass die Matrix–Funktion $\{\mathbf{P}'(h)\}$ mit $\mathbf{P}'(h) = \exp(h\mathbf{Q})$ eine Lösung der Kolmogorowschen Rückwärtsgleichung (25) ist, die der Anfangsbedingung $\mathbf{P}'(0) = \mathbf{I}$ genügt.
- Aus der Theorie der gewöhnlichen linearen Differentialgleichungssysteme ergibt sich dann, dass diese Lösung eindeutig bestimmt ist.
- Somit gilt $\mathbf{P}'(h) = \mathbf{P}(h)$ für jedes $h \geq 0$. □

Beachte

- Zusammen mit Theorem 2.17 ergibt sich aus (32) und (33)
  - die folgende Darstellungsformel der Übergangsfunktion $\{\mathbf{P}(h), h \geq 0\}$:

$$\mathbf{P}(h) = \sum_{n=0}^{\infty} \frac{(ah)^n(\tilde{\mathbf{P}})^n}{n!}\,\mathrm{e}^{-ah} \qquad \forall\, h \geq 0\,, \tag{35}$$

    wobei $\tilde{\mathbf{P}} = a^{-1}\mathbf{Q} + \mathbf{I}$ eine stochastische Matrix ist und $a = \max\{-q_{ii} : i = 1, \ldots, \ell\}$.
  - Aus (35) folgt insbesondere, dass $p_{ii}(h) > 0$ für beliebige $i \in E$ und $h \geq 0$.
- Ähnlich wie bei Markow–Ketten mit diskreter Zeit (vgl. Abschnitt MK-2.1.4) kann die Spektraldarstellung von $\mathbf{Q}^n$ zur Berechnung der Matrix-Exponentialfunktion $\mathbf{P}(h) = \exp(h\mathbf{Q})$ verwendet werden, falls die Eigenwerte $\theta_1, \ldots, \theta_\ell \neq 0$ von $\mathbf{Q}$ paarweise verschieden sind.
- Zur Erinnerung
  - Sei $\mathbf{A}$ eine beliebige $\ell \times \ell$ Matrix, seien $\boldsymbol{\phi}, \boldsymbol{\psi} \neq \mathbf{0}$ zwei ℓ-dimensionale (Spalten-) Vektoren, so daß jeweils mindestens eine ihrer Komponenten von 0 verschieden ist, und sei θ eine beliebige (reelle oder komplexe) Zahl.

- Falls
$$\mathbf{A}\boldsymbol{\phi} = \theta\boldsymbol{\phi} \qquad \text{bzw.} \qquad \boldsymbol{\psi}^\top\mathbf{A} = \theta\boldsymbol{\psi}^\top\,, \tag{36}$$
dann ist θ ein *Eigenwert* von $\mathbf{A}$, und $\boldsymbol{\phi}$ bzw. $\boldsymbol{\psi}$ ist ein rechter bzw. linker (zu θ gehörender) *Eigenvektor*.
- Weil (36) und
$$(\mathbf{A} - \theta\mathbf{I})\boldsymbol{\phi} = \mathbf{0} \qquad \text{bzw.} \qquad \boldsymbol{\psi}^\top(\mathbf{A} - \theta\mathbf{I}) = \mathbf{0}^\top\,,$$
äquivalent sind, ist θ genau dann ein Eigenwert von $\mathbf{A}$, wenn θ eine Lösung der sogenannten *charakteristischen Gleichung* ist:
$$\det(\mathbf{A} - \theta\mathbf{I}) = 0\,. \tag{37}$$
- Weil (37) eine algebraische Gleichung der Ordnung ℓ ist, besitzt sie somit ℓ Lösungen $\theta_1, \ldots, \theta_\ell$, die komplex sein können und die nicht alle voneinander verschieden sein müssen.
- Wir nehmen o.B.d.A. an, dass die Eigenwerte $\theta_1, \ldots, \theta_\ell$ so numeriert sind, dass
$$|\theta_1| \geq |\theta_2| \geq \ldots \geq |\theta_\ell|\,.$$
- Für jeden Eigenwert θ_i existieren (jeweils von $\mathbf{0}$ verschiedene) rechte bzw. linke Eigenvektoren $\boldsymbol{\phi}_i$ bzw. $\boldsymbol{\psi}_i$.
- Sei $\boldsymbol{\Phi} = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_\ell)$ die $\ell \times \ell$ Matrix, die aus den rechten Eigenvektoren $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_\ell$ besteht, und sei
$$\boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\psi}_1^\top \\ \vdots \\ \boldsymbol{\psi}_\ell^\top \end{pmatrix}$$
die $\ell \times \ell$ Matrix, die aus den linken Eigenvektoren $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_\ell$ gebildet wird.
- Mit dieser Schreibweise ergibt sich, dass
$$\mathbf{A}\boldsymbol{\Phi} = \boldsymbol{\Phi}\,\operatorname{diag}(\boldsymbol{\theta})\,, \tag{38}$$
wobei $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_\ell)^\top$ und $\operatorname{diag}(\boldsymbol{\theta})$ die Diagonalmatrix mit den Diagonalelementen $\theta_1, \ldots, \theta_\ell$ bezeichnet.

- Falls die Eigenwerte $\theta_1, \ldots, \theta_\ell \neq 0$ paarweise verschieden sind, dann sind die Eigenvektoren $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_\ell$ linear unabhängig (vgl. beispielsweise Lemma MK-2.2).
  - Somit existiert die inverse Matrix $\boldsymbol{\Phi}^{-1}$, und wir können $\boldsymbol{\Psi} = \boldsymbol{\Phi}^{-1}$ setzen.
  - Außerdem ergibt sich in diesem Fall aus (38), dass
$$\mathbf{A} = \boldsymbol{\Phi}\,\operatorname{diag}(\boldsymbol{\theta})\boldsymbol{\Phi}^{-1} = \boldsymbol{\Phi}\,\operatorname{diag}(\boldsymbol{\theta})\boldsymbol{\Psi}$$
bzw.
$$\mathbf{A}^n = \boldsymbol{\Phi}(\operatorname{diag}(\boldsymbol{\theta}))^n\boldsymbol{\Phi}^{-1} = \boldsymbol{\Phi}(\operatorname{diag}(\boldsymbol{\theta}))^n\boldsymbol{\Psi}\,.$$
  - Hieraus ergibt sich die *Spektraldarstellung* von $\mathbf{A}^n$ mit
$$\mathbf{A}^n = \sum_{i=1}^{\ell} \theta_i^n \boldsymbol{\phi}_i \boldsymbol{\psi}_i^\top\,. \tag{39}$$

- Falls die Eigenwerte $\theta_1, \ldots, \theta_\ell \neq 0$ der Intensitätsmatrix $\mathbf{Q}$ paarweise verschieden sind, dann gilt also
$$\mathbf{P}(h) = \sum_{i=1}^{\ell} \mathrm{e}^{h\theta_i} \boldsymbol{\phi}_i \boldsymbol{\psi}_i^\top \tag{40}$$

für jedes $h \geq 0$, wobei $\boldsymbol{\phi}_i, \boldsymbol{\psi}_i$ (rechte bzw. linke) Eigenvektoren des Eigenwertes θ_i sind.

2.3.4 Eingebettete Markow–Ketten; Simulationsalgorithmus

- Sei $E = \{1, 2, \ldots, \ell\}$ und sei $\mathbf{Q} = (q_{ij})_{i,j \in E}$ eine beliebige Intensitätsmatrix, d.h., es gelte

$$q_{ij} \geq 0 \text{ für beliebige } i \neq j \qquad \text{und} \qquad \sum_{j \in E} q_{ij} = 0 \text{ für jedes } i \in E.$$

- Wir diskutieren nun eine Methode zur Konstruktion von Markow–Prozessen mit vorgegebener Intensitätsmatrix $\mathbf{Q}$, wobei wir zunächst die folgende *eingebettete Markow–Kette* $\{\tilde{X}_n, n \geq 0\}$ (mit diskreter Zeit) betrachten.
  - Sei $\boldsymbol{\alpha}$ eine beliebige Anfangsverteilung, und sei $q(i) = \sum_{j \neq i} q_{ij}$ für jedes $i \in E$.
  - Außerdem sei die stochastische Matrix $\widetilde{\mathbf{P}} = (\widetilde{p}_{ij})$ gegeben durch

$$\widetilde{p}_{ij} = \begin{cases} \dfrac{q_{ij}}{q(i)}\,, & \text{falls } i \neq j, \\ 0\,, & \text{falls } i = j. \end{cases} \tag{41}$$

    für beliebige $i, j \in E$ mit $q(i) > 0$.
  - Falls $q(i) = 0$, dann wird die i–te Zeile von $\widetilde{\mathbf{P}}$ gleich $\mathbf{e}_i^\top = (0, \ldots, 0, 1, 0, \ldots, 0)$ gesetzt, wobei die i–te Komponente von $\mathbf{e}_i$ gleich 1 ist.
  - Sei $\{\widetilde{X}_n, n \geq 0\}$ eine Markow–Kette mit der Anfangsverteilung $\boldsymbol{\alpha}$ und der Übergangsmatrix $\widetilde{\mathbf{P}}$.
- Außerdem sei $\{Z_n, n \geq 0\}$ eine Folge von unabhängigen und Exp(1)–verteilten Zufallsvariablen, die von $\{\widetilde{X}_n\}$ unabhängig sind.
- Mit Hilfe der Folgen $\{Z_n, n \geq 0\}$ und $\{\widetilde{X}_n, n \geq 0\}$ konstruieren wir nun einen Markow–Prozess $\{X_t, t > 0\}$, dessen Trajektorien die Form

$$X_t = \sum_{n=0}^{\infty} \widetilde{X}_n \mathbb{1}(S_n \leq t < S_{n+1}) \tag{42}$$

haben, wobei $0 = S_0 < S_1 < \ldots$ eine Folge von zufälligen (Sprung–) Zeitpunkten ist:

Schritt 1 Sei $S_0 = 0$ und $Z_0' = Z_0/q(\widetilde{X}_0)$, wobei $Z_0' = \infty$ gesetzt wird, falls $q(\widetilde{X}_0) = 0$.

- Die Zufallsvariable Z_0' ist die *Aufenthaltsdauer* des zu konstruierenden Markow–Prozesses $\{X_t\}$ im (zufälligen) Anfangszustand $\widetilde{X}_0$, der zum Zeitpunkt $S_0 = 0$ gewählt wird.
- Dabei gilt $P(Z_0' > x \mid \widetilde{X}_0 = i) = \mathrm{e}^{-q(i)x}$ für jedes $i \in E$ mit $P(\widetilde{X}_0 = i) > 0$; $x \geq 0$.

Schritt 2 Sei $S_1 = S_0 + Z_0'$ und $X_t = \widetilde{X}_0$ für $S_0 = 0 \leq t < S_1$, wodurch die Trajektorie von $\{X_t\}$ bis zum ersten Sprungzeitpunkt S_1 gegeben ist.

Schritt 3 (analog zu Schritt 1) Sei $Z_1' = Z_1/q(\widetilde{X}_1)$, wodurch

- die Aufenthaltsdauer von $\{X_t\}$ im Zustand $\widetilde{X}_1$ gegeben ist, der zum Zeitpunkt S_1 gewählt wird.
- Dabei gilt $P(Z_1' > x \mid \widetilde{X}_1 = i) = \mathrm{e}^{-q(i)x}$, falls $P(\widetilde{X}_1 = i) > 0$.

Schritt 4 Sei $S_2 = S_1 + Z_1'$ und $X_t = \widetilde{X}_1$ für $S_1 \leq t < S_2$.

$\vdots$

Schritt 2n+1 Wir nehmen nun an, dass die Größen $Z_0', Z_1', \ldots, Z_{n-1}', S_0, S_1, \ldots, S_n$ und $\{X_t, t \in [0, S_n)\}$ für ein $n \geq 1$ bereits definiert worden sind, und setzen $Z_n' = Z_n/q(\widetilde{X}_n)$.

Schritt 2(n + 1) Sei $S_{n+1} = S_n + Z'_n$ und $X_t = \widetilde{X}_n$ für $S_n \leq t < S_{n+1}$.

Beachte

- Auf diese Weise lassen sich die Trajektorien von $\{X_t, t \geq 0\}$ für beliebige $t \geq 0$ konstruieren, weil man sich leicht überlegen kann, dass $P(\lim_{n\to\infty} S_n = \infty) = 1$.
- Somit ist durch das oben beschriebene Konstruktionsprinzip ein Algorithmus zur *Simulation* des zeitstetigen Prozesses $\{X_t, t \geq 0\}$ gegeben, wobei vorausgesetzt wird, dass Algorithmen
  - zur Simulation der Folge $\{Z_n, n \geq 0\}$ von unabhängigen und Exp(1)–verteilten Zufallsvariablen sowie
  - der eingebetteten Marlow–Kette $\{\widetilde{X}_n, n \geq 0\}$ (vgl. Übungsaufgabe 3.7)

  vorhanden sind.
- Außerdem kann man zeigen, dass die endlich–dimensionalen Verteilungen des Prozesses $\{X_t, t \geq 0\}$ die Faktorisierungeigenschaft besitzen, die in der Definitionsgleichung (10) von Markow–Prozessen postuliert wird.

Theorem 2.18 *Der in* (42) *gegebene stochastische Prozess* $\{X_t, t \geq 0\}$ *ist ein homogener Markow–Prozess.*

Beweis

- Der Einfachheit wegen geben wir hier lediglich eine Beweisskizze an.
  - Einen ausführlicheren Beweis kann man beispielsweise in Resnick (1992), S. 378–379 finden.
  - Dabei betrachten wir hier nur den folgenden Spezialfall: Sei $\alpha_i > 0$ und $q_{ij} > 0$ für beliebige $i, j \in \{1, \ldots, \ell\}$ mit $i \neq j$.
- Sei $\{N_t,\ t \geq 0\}$ der Zählprozess der Sprungzeitpunkte $S_1, S_2, \ldots$ mit $N_t = \max\{n : S_n \leq t\}$.
- Dann ergibt sich unmittelbar aus dem Konstruktionsprinzip des stochastischen Prozesses $\{X_t,\ t \geq 0\}$, dass

$$\begin{aligned}
&P(X_0 = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) \\
&\quad = \sum_{m_1,\ldots,m_n \geq 0} P(\tilde{X}_0 = i_0, N_{t_1} = m_1, \tilde{X}_{m_1} = i_1, N_{t_2} = m_1 + m_2, \tilde{X}_{m_1+m_2} = i_2, \\
&\qquad\qquad \ldots, N_{t_n} = m_1 + \ldots + m_n, \tilde{X}_{m_1+\ldots+m_n} = i_n) > 0
\end{aligned}$$

  für beliebige $n = 0, 1, \ldots,\ i_0, i_1, \ldots, i_n \in E,\ 0 \leq t_1 \leq \ldots \leq t_n$.
- Hieraus und aus der „Gedächtnislosigkeit" der Exponentialverteilung ergibt sich, dass

$$\begin{aligned}
&P(X_0 = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) \\
&\quad = \sum_{m_1,\ldots,m_n \geq 0} P(\tilde{X}_0 = i_0) P(N_{t_1} = m_1, \tilde{X}_{m_1} = i_1 \mid \tilde{X}_0 = i_0) \\
&\qquad P(N_{t_2} = m_1 + m_2, \tilde{X}_{m_1+m_2} = i_2 \mid N_{t_1} = m_1, \tilde{X}_{m_1} = i_1) \\
&\qquad \vdots \\
&\qquad P(N_{t_n} = m_1 + \ldots + m_n, \tilde{X}_{m_1+\ldots+m_n} = i_n \mid N_{t_{n-1}} = m_1 + \ldots + m_{n-1}, \tilde{X}_{m_1+\ldots+m_{n-1}} = i_{n-1}) \\
&\quad = \sum_{m_1,\ldots,m_n \geq 0} P(\tilde{X}_0 = i_0) P(N_{t_1} = m_1, \tilde{X}_{m_1} = i_1 \mid \tilde{X}_0 = i_0) \\
&\qquad P(N_{t_2-t_1} = m_2, \tilde{X}_{m_2} = i_2 \mid \tilde{X}_0 = i_1) \ \ldots\ P(N_{t_n-t_{n-1}} = m_n, \tilde{X}_{m_n} = i_n \mid \tilde{X}_0 = i_{n-1}) \\
&\quad = P(\tilde{X}_0 = i_0) \sum_{m_1=0}^{\infty} P(N_{t_1} = m_1, \tilde{X}_{m_1} = i_1 \mid \tilde{X}_0 = i_0) \\
&\qquad \sum_{m_2=0}^{\infty} P(N_{t_2-t_1} = m_2, \tilde{X}_{m_2} = i_2 \mid \tilde{X}_0 = i_1) \ \ldots\ \sum_{m_n=0}^{\infty} P(N_{t_n-t_{n-1}} = m_n, \tilde{X}_{m_n} = i_n \mid \tilde{X}_0 = i_{n-1}) .
\end{aligned}$$

- Es gilt also

$$P(X_0 = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = \alpha_{i_0} p_{i_0 i_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \ \ldots \ p_{i_{n-1} i_n}(t_n - t_{n-1}), \tag{43}$$

  wobei $\alpha_i = P(\widetilde{X}_0 = i) = P(X_0 = i)$ und

$$p_{ij}(h) = \sum_{m=0}^{\infty} P(N_h = m, \tilde{X}_m = j \mid \tilde{X}_0 = i) \qquad \Big(= P(X_h = j \mid X_0 = i)\Big).$$

- Aus (43) ergibt sich insbesondere, dass $p_{ij}(h_1 + h_2) = \sum_{k \in E} p_{ik}(h_1)\, p_{kj}(h_2)$ für beliebige $i, j \in E$ und $h_1, h_2 \geq 0$.
- Damit ist gezeigt, dass die endlich–dimensionalen Verteilungen des Prozesses $\{X_t, t \geq 0\}$ die Faktorisierungeigenschaft (10) besitzen. □

Wir zeigen nun, dass der in diesem Abschnitt konstruierte stochastische Prozess $\{X_t,\, t \geq 0\}$ der „richtige" Markow–Prozess ist, d.h., seine Intensitätsmatrix ist gleich der vorgegebenen Matrix $\mathbf{Q}$.

- Hierfür muss gezeigt werden, dass sich die Übergangswahrscheinlichkeiten $p_{ij}(h)$ des Markow–Prozesses $\{X_t,\, t \geq 0\}$ durch die „lokalen" Charakteristiken $\{q(i)\}_{i \in E}$ und $\widetilde{\mathbf{P}} = (\widetilde{p}_{ij})_{i,j \in E}$ ausdrücken lassen.
- Dabei sagt man, dass $i \in E$ ein *absorbierender Zustand* ist, falls $q(i) = 0$.

Theorem 2.19 *Für beliebige $i, j \in E$ und $h \geq 0$ gilt*

$$p_{ij}(h) = \delta_{ij} \mathrm{e}^{-q(i)h} + \int_0^h q(i) \mathrm{e}^{-q(i)t} \sum_{k \neq i} \widetilde{p}_{ik} p_{kj}(h - t)\, \mathrm{d}t. \tag{44}$$

Insbesondere gilt $p_{ij}(h) = \delta_{ij}$ für beliebige $j \in E$ und $h \geq 0$, wenn $i \in E$ ein absorbierender Zustand ist.

Beweis

- So wie im Beweis von Theorem 2.15 können wir ohne Einschränkung der Allgemeinheit annehmen, dass $P(X_0 = i) > 0$ für jedes $i \in E$.
- Es ist klar, dass sich die Übergangswahrscheinlichkeiten $p_{ij}(h)$ als Summe $p_{ij}(h) = I_{ij}(h) + I'_{ij}(h)$ darstellen lassen, wobei

$$I_{ij}(h) = P(X_h = j, S_1 > h \mid X_0 = i) \qquad \text{und} \qquad I'_{ij}(h) = P(X_h = j, S_1 \leq h \mid X_0 = i).$$

- Für die Zerlegungskomponenten $I_{ij}(h)$ und $I'_{ij}(h)$ gilt dann

$$I_{ij}(h) = \begin{cases} P(Z'_0 > h \mid \widetilde{X}_0 = i) = \mathrm{e}^{-q(i)h}, & \text{falls } i = j, \\ 0, & \text{falls } i \neq j, \end{cases}$$

  und

$$\begin{aligned} I'_{ij}(h) &= \sum_{k \neq i} \int_0^h \frac{P(X_h = j, S_1 \in \mathrm{d}t, \widetilde{X}_1 = k, \widetilde{X}_0 = i)}{P(\widetilde{X}_0 = i)} \\ &= \sum_{k \neq i} \int_0^h P(X_h = j \mid X_t = k) P\Big(\frac{Z_0}{q(i)} \in \mathrm{d}t\Big) P(\widetilde{X}_1 = k \mid \widetilde{X}_0 = i) \\ &= \int_0^h q(i) \mathrm{e}^{-q(i)t} \sum_{k \neq i} \widetilde{p}_{ik} p_{kj}(h - t)\, \mathrm{d}t. \end{aligned}$$

- Falls $i \in E$ ein absorbierender Zustand ist, dann verbleibt der in diesem Abschnitt konstruierte Prozess $\{X_t, t \geq 0\}$ für immer im Zustand i, sobald er diesen Zustand zum ersten Mal erreicht. Somit gilt $p_{ij}(h) = \delta_{ij}$ für beliebige $j \in E$ und $h \geq 0$. □

Korollar 2.3 *Für beliebige $i, j \in E$ gilt*

$$p_{ij}^{(1)}(0+) = q_{ij} = \begin{cases} -q(i)\,, & \textit{falls } i = j, \\ \widetilde{p}_{ij} q(i) & \textit{falls } i \neq j. \end{cases} \tag{45}$$

Beweis Um (45) zu beweisen, genügt es, in (44) die Ableitung bezüglich h zu bilden und danach den Grenzwert für $h \downarrow 0$ zu betrachten. □

Beachte Wenn man die Formeln (13) und (45) miteinander vergleicht, dann erkennt man, dass die Intensitätsmatrix des in diesem Abschnitt konstruierten Markow–Prozesses $\{X_t, t \geq 0\}$ mit der vorgegebenen Matrix $\mathbf{Q}$ übereinstimmt.

Beispiel

- Der in diesem Abschnitt konstruierte Markow–Prozess $\{X_t, t \geq 0\}$ heißt *Geburts– und Todesprozess*, wenn $\widetilde{p}_{i,i-1} + \widetilde{p}_{i,i+1} = 1$ für beliebige $1 < i < \ell$ und $\widetilde{p}_{12} = \widetilde{p}_{\ell,\ell-1} = 1$ gilt.
- Die Produkte $\widetilde{p}_{i,i+1} q(i)$ und $\widetilde{p}_{i,i-1} q(i)$ heißen *Geburtsrate* bzw. *Sterberate*.
- Diese Begriffsbildungen lassen sich damit begründen, dass die Größen $\widetilde{p}_{i,i+1} q(i)$ und $\widetilde{p}_{i,i-1} q(i)$ gemäß Korollar 2.3 mit den Übergangsintensitäten $q_{i,i+1}$ und $q_{i,i-1}$ für die Übergänge $i \to i+1$ bzw. $i \to i-1$ im Sinne der Definitionsgleichung (13) übereinstimmen.

2.3.5 Stationäre Anfangsverteilungen

Ähnlich wie bei Markow–Ketten mit diskreter Zeit (vgl. Abschnitt MK-2.2.4) betrachten wir nun eine spezielle Klasse von Anfangsverteilungen, die invariant bezüglich der Übergangsfunktion von Markow–Prozessen sind.

Definition Die Wahrscheinlichkeitsfunktion $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell)^\top$ heißt *stationäre Anfangsverteilung* bezüglich der Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$, falls

$$\boldsymbol{\alpha}^\top \mathbf{P}(h) = \boldsymbol{\alpha}^\top \qquad \forall\, h \geq 0\,. \tag{46}$$

Beachte

- Wegen der Endlichkeit des Zustandsraumes E gilt für jede stationäre Anfangsverteilung $\boldsymbol{\alpha}$:

$$\mathbf{0} = \lim_{h \to 0} \frac{\boldsymbol{\alpha}^\top (\mathbf{P}(h) - \mathbf{I})}{h} = \boldsymbol{\alpha}^\top \mathbf{Q}\,.$$

- Es gilt auch die umgekehrte Implikation: Aus $\boldsymbol{\alpha}^\top \mathbf{Q} = \mathbf{0}$ folgt die Gültigkeit von $\boldsymbol{\alpha}^\top \mathbf{Q}^n = \mathbf{0}$ und somit auch

$$\boldsymbol{\alpha}^\top \mathbf{P}(h) = \boldsymbol{\alpha}^\top \sum_{n=0}^{\infty} \frac{(h\mathbf{Q})^n}{n!} = \boldsymbol{\alpha}^\top \qquad \forall\, h \geq 0\,.$$

- Im allgemeinen muss jedoch die Existenz einer Wahrscheinlichkeitsfunktion $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell)^\top$, die der Gleichung (46) genügt, bzw. die eindeutige Lösbarkeit von (46) nicht immer gewährleistet sein.
- Andererseits lassen sich, ähnlich wie bei (zeitdiskreten) aperiodischen und irreduziblen Markow–Ketten, stationäre Anfangsverteilungen von Markow–Prozessen auch als Grenzverteilungen charakterisieren.

- Dabei ist die Situation im zeitstetigen Fall sogar einfacher, denn aus der Darstellungsformel (35) der Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ folgt, dass $p_{ii}(h) > 0$ für beliebige $i \in E$ und $h \geq 0$.
- Die bei Markow–Ketten betrachtete Bedingung der Aperiodizität muss also bei Markow–Prozessen nicht extra vorausgesetzt werden, sondern wir benötigen lediglich die Irreduzibilität.

Definition Die Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ heißt *irreduzibel*, falls

$$p_{ij}(h) > 0 \qquad \forall\, i \neq j,\ h > 0\,. \tag{47}$$

Beachte Man kann leicht zeigen, dass

- die Bedingung (47) mit der Irreduzibilität der Intensitätsmatrix $\mathbf{Q}$ äquivalent ist,
- d.h., für beliebige $i, j \in E$ mit $i \neq j$ gibt es eine Folge von Zuständen $i_1, \ldots, i_{n-1} \in E$ mit $i_k \neq i_l$, so dass $q_{ii_1} q_{i_1 i_2} \ldots q_{i_{n-1} j} > 0$.

Das folgende Theorem zeigt darüber hinaus: Die Irreduzibilität der Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ bzw. der Intensitätsmatrix $\mathbf{Q}$ impliziert, dass es eine (eindeutig bestimmte) stationäre Anfangsverteilung bezüglich der Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ gibt.

Theorem 2.20 *Wenn die Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ irreduzibel ist, dann existiert der Grenzwert*

$$\lim_{t \to \infty} P(X_t = i) = \pi_i \tag{48}$$

für jedes $i \in E$, wobei $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_\ell)^\top$ die (eindeutig bestimmte) stationäre Anfangsverteilung bezüglich $\{\mathbf{P}(h),\ h \geq 0\}$ ist.

Beweis

- Die Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ sei irreduzibel.
- Wegen (35) gilt dann $p_{ij}(h) > 0$ für beliebige $i, j \in E$ und $h > 0$, d.h., für jedes $h > 0$ sind sämtliche Eintragungen der Übergangsmatrix $\mathbf{P}(h)$ positiv.
- Aus Theorem MK-2.4 ergibt sich nun, dass die eingebettete Markow–Kette $\{X_{nh},\ n \geq 0\}$ mit der (einstufigen) Übergangsmatrix $\mathbf{P}(h)$ für jedes $h > 0$ ergodisch ist, d.h., es gilt

  $$\lim_{n \to \infty} \mathbf{P}(nh) = \mathbf{\Pi} \qquad \forall\, h > 0$$

  wobei $\mathbf{\Pi}$ eine $\ell \times \ell$–Matrix ist, deren Zeilen gleich $\boldsymbol{\pi}^\top$ sind.
- Weil die Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ gleichmäßig stetig in $h \geq 0$ ist, gibt es für jedes $\varepsilon > 0$ eine (kleine) Zahl $h_0 > 0$, so dass es für jedes hinreichend große $t > 0$ eine natürliche Zahl n gibt mit

  $$\|\mathbf{P}(t) - \mathbf{\Pi}\| \leq \|\mathbf{P}(t) - \mathbf{P}(nh_0)\| + \|\mathbf{P}(nh_0) - \mathbf{\Pi}\| \leq 2\varepsilon\,. \qquad \square$$

Beachte

- Sei $\{X_t,\ t \geq 0\}$ ein Markow–Prozess mit der Übergangsfunktion $\{\mathbf{P}(h),\ h \geq 0\}$ und der stationären Anfangsverteilung $\boldsymbol{\alpha}$.
- Dann kann man leicht zeigen, dass $\{X_t,\ t \geq 0\}$ ein *stationärer Prozess* ist, d.h., für die endlich–dimensionalen Verteilungen von $\{X_t\}$ gilt

  $$P(X_0 = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = P(X_h = i_0, X_{t_1+h} = i_1, \ldots, X_{t_n+h} = i_n)$$

  für beliebige $n = 0, 1, \ldots,\ i_0, i_1, \ldots, i_n \in E,\ 0 \leq t_1 \leq \ldots \leq t_n,\ h > 0$.
- Insbesondere ist dann $\{X_t,\ t \geq 0\}$ ein Prozess mit stationären Zuwächsen, die jedoch im allgemeinen *nicht* unabhängig sein müssen.

2.4 Wiener–Prozess

Der Wiener–Prozess, der in der Literatur auch *Brownsche Bewegung* genannt wird, ist ein weiteres grundlegendes Modell in der Theorie stochastischer Prozesse. Dabei nehmen wir für die Indexmenge I an, dass entweder $I = [0, b]$ für eine (endliche) Zahl $b \in (0, \infty)$ oder $I = [0, \infty)$ gilt.

Definition Ein stochastischer Prozess $\{X_t,\, t \in I\}$ über einem (im allgemeinen nicht näher spezifizierten) Wahrscheinlichkeitsraum $(\Omega.\mathcal{F}, P)$ heisst *Wiener–Prozess* (im Zeitintervall I), wenn

- $\{X_t,\, t \in I\}$ unabhängige Zuwächse hat,
- $X_{t_2} - X_{t_1} \sim \mathrm{N}(0, t_2 - t_1)$ für beliebige $t_1, t_2 \in I$ mit $t_1 < t_2$,
- $X_0 = 0$,
- die Trajektorie $t \to X_t(\omega)$, $t \in I$, für jedes $\omega \in \Omega$ eine stetige Funktion ist.

Beachte

- Aus dem Existenzsatz von Kolmogorow, d.h. aus Theorem 1.1, folgt, dass es einen Wahrscheinlichkeitsraum und einen stochastischen Prozess $\{X_t,\, t \in I\}$ über diesem Wahrscheinlichkeitsraum gibt, so dass die ersten drei Bedingungen in der Definition des Wiener–Prozesses erfüllt sind.
- Aus dem Stetigkeitssatz von Kolmogorow, d.h. aus Theorem 1.3, folgt außerdem, dass zu jedem Prozess $\{X_t,\, t \geq 0\}$, der den ersten drei Bedingungen in der Definition des Wiener–Prozesses genügt, eine stetige Modifikation existiert.
- Weniger offensichtlich ist, ob bzw. wie man einen solchen Prozess (mit unabhängigen und normalverteilten Zuwächsen) explizit konstruieren kann, dessen Trajektorien stetige Funktionen sind.
- In den drei folgenden Abschnitten erläutern wir eine Methode zur Konstruktion von Wiener–Prozessen (mit stetigen Trajektorien), die gleichzeitig zu einem Algorithmus zur Simulation von Wiener–Prozessen führt.

2.4.1 Vollständige Orthonormalsysteme im L_2; Haar–Funktionen und Schauder–Funktionen

Wir betrachten zunächst den Fall, dass $I = [0, 1]$ das Einheitsintervall ist. Dabei benötigen wir zur Konstruktion von Wiener–Prozessen in $[0, 1]$ einige analytische Hilfsmittel.

- Insbesondere betrachten wir den *Hilbert–Raum* $L_2 = L_2([0, 1])$
  - aller auf dem Einheitsintervall $I = [0, 1]$ definierten (Borel–messbaren) Funktionen $f : [0, 1] \to \mathbb{R}$, die bezüglich des Lebesgue–Maßes quadratisch integrierbar sind, d.h., für die $\int_0^1 f^2(s)\,\mathrm{d}s < \infty$ gilt.
  - Für beliebige $f, g \in L_2$ sei $(f, g) = \int_0^1 f(s)g(s)\,\mathrm{d}s$ das zugehörige *Skalarprodukt*.
- Eine Folge $\{f_n,\, n \geq 1\}$ von Funktionen $f_n \in L_2$ wird ein vollständiges *Orthonormalsystem* in L_2 genannt,
  - wenn $(f_n, f_n) = 1$ und $(f_n, f_m) = 0$ für beliebige $n \neq m$ und
  - wenn für jedes $g \in L_2$ die Gültigkeit von $(g, f_n) = 0$ für jedes $n \geq 1$ impliziert, dass $g(s) = 0$ für fast jedes $s \in [0, 1]$.

Die Konstruktion von Wiener–Prozessen in $I = [0, 1]$ beruht auf der Idee,

- ein speziell gewähltes Orthonormalsystem $\{H_n,\, n \geq 1\}$ in L_2 (die sogenannten Haar–Funktionen) sowie
- eine Folge $\{Y_n,\, n \geq 1\}$ von unabhängigen und $\mathrm{N}(0, 1)$–verteilten Zufallsvariablen über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ zu betrachten und zu zeigen, dass

- der stochastische Prozess $\{X_t,\, t \in [0,1]\}$ mit

$$X_t = \sum_{n=1}^{\infty} Y_n \int_0^t H_n(s)\,\mathrm{d}s \qquad \forall\, t \in [0,1] \tag{1}$$

über einer gewissen Einschränkung des Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$ wohldefiniert ist und den Bedingungen eines Wiener–Prozesses in $[0,1]$ genügt.

Definition Die durch den Ansatz

$$H_1(s) = 1 \qquad \text{für jedes } s \in [0,1], \tag{2}$$

$$H_{2^{m+1}}(s) = \begin{cases} 2^{\frac{m}{2}} & \text{für } s \in \left[\frac{2^m-1}{2^m}\,,\, \frac{2^{m+1}-1}{2^{m+1}}\right), \\ -2^{\frac{m}{2}} & \text{für } s \in \left[\frac{2^{m+1}-1}{2^{m+1}}\,,\, 1\right], \\ 0 & \text{sonst}, \end{cases} \tag{3}$$

$$H_{2^m+k}(s) = \begin{cases} 2^{\frac{m}{2}} & \text{für } s \in \left[\frac{k-1}{2^m}\,,\, \frac{2k-1}{2^{m+1}}\right), \\ -2^{\frac{m}{2}} & \text{für } s \in \left[\frac{2k-1}{2^{m+1}}\,,\, \frac{k}{2^m}\right), \\ 0 & \text{sonst} \end{cases} \tag{4}$$

für $k = 1, 2, \ldots, 2^m - 1$ und $m = 0, 1, 2, \ldots$ gegebenen Funktionen $H_n : [0,1] \to \mathbb{R}$ heißen *Haar–Funktionen.*

Lemma 2.5

- *Die in* (2)–(4) *definierten Haar–Funktionen* $H_n : [0,1] \to \mathbb{R}$ *mit* $n = 1, 2, \ldots$ *bilden ein vollständiges Orthonormalsystem in* L_2.
- *Insbesondere gilt die sogenannte* Parseval–Identität

$$(f,g) = \sum_{n=1}^{\infty} (f, H_n)(g, H_n) \qquad \forall\, f, g \in L_2\,. \tag{5}$$

Beweis

- Unmittelbar aus der Definition der Haar–Funktionen folgt, dass $(H_n, H_n) = 1$ und $(H_n, H_m) = 0$ für beliebige $n, m \geq 1$ mit $n \neq m$ gilt.
- Außerdem kann man leicht zeigen, dass das Orthonormalsystem $\{H_n,\, n \geq 1\}$ vollständig in L_2 ist.
  - Sei $g \in L_2$ eine quadratisch integrierbare Funktion mit

$$\int_0^1 g(s)\, H_n(s)\,\mathrm{d}s = 0 \qquad \forall\, n \geq 1\,.$$

  - Hieraus folgt, dass $\int_{k/2^m}^{(k+1)/2^m} g(s)\,\mathrm{d}s = 0$ für beliebige $m = 0, 1, \ldots$ und $k = 1, \ldots, 2^m - 1$.
  - Insbesondere gilt also für die Funktion $G : [0,1] \to \mathbb{R}$ mit

$$G(t) = \int_0^t g(s)\,\mathrm{d}s\,,$$

dass $G(k/2^m) = 0$ für beliebige $m = 0, 1, \ldots$ und $k = 1, \ldots, 2^m - 1$.

- Weil G eine stetige Funktion ist, folgt hieraus, dass $G(t) = 0$ für jedes $t \in [0,1]$ und somit $g(s) = 0$ für fast jedes $s \in [0,1]$.

- Die Gültigkeit von (5) ergibt sich aus der Vollständigkeit des Orthonormalsystems $\{H_n,\, n \geq 1\}$, vgl. beispielsweise Satz 5.3.6 in W. Arendt (2002) Funktionalanalysis (Vorlesungskript, Universität Ulm) bzw. Satz V.4.9 in D. Werner (1997) Funktionalanalysis (Springer, Berlin).

Lemma 2.6

- *Für jedes $n = 1, 2, \ldots$ sei die Funktion $S_n : [0,1] \to \mathbb{R}$ gegeben durch*

$$S_n(t) = \int_0^t H_n(s)\,\mathrm{d}s \qquad \forall\, t \in [0,1]\,. \tag{6}$$

- *Die in (6) definierten Funktionen $S_1, S_2, \ldots : [0,1] \to \mathbb{R}$, die* Schauder–Funktionen *genannt werden, besitzen die folgenden Eigenschaften:*

  1 *Für beliebige $n \geq 1$ und $t \in [0,1]$ gilt $S_n(t) \geq 0$.*

  2 *Für beliebige $m = 1, 2, \ldots$ und $t \in [0,1]$ gilt die Ungleichung*

$$\sum_{k=1}^{2^m} S_{2^m+k}(t) \leq \frac{1}{2}\, 2^{-m/2}\,. \tag{7}$$

  3 *Sei $a_1, a_2, \ldots \in \mathbb{R}$ eine beliebige Folge reeller Zahlen mit $a_n = \mathrm{O}((\log n)^{1/2})$ für $n \to \infty$. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n\, S_n(t)$ absolut und gleichmäßig in $t \in [0,1]$.*

Beweis

- Die Tatsache, dass die Schauder–Funktionen S_n nur nichtnegative Werte haben, ergibt sich unmittelbar aus den Definitionsgleichungen (2)–(4) bzw. (6).

- Um die Gültigkeit der Ungleichung (7) zu zeigen, genügt es zu beachten, dass die Funktionen S_{2^m+k} für $k = 1, \ldots, 2^m$ disjunkte Träger haben und dass

$$S_{2^m+k}(t) \leq S_{2^m+k}\Big(\frac{2k-1}{2^{m+1}}\Big) = 2^{m/2}\,\frac{1}{2^{m+1}} \;=\; \frac{1}{2}\,2^{-m/2} \qquad \forall\, t \in [0,1]\,.$$

- Sei nun $c > 0$ eine Konstante, so dass $|a_n| \leq c\,(\log n)^{1/2}$ für jedes $n \geq 1$.
  - Dann ergibt sich aus (7), dass für $\ell \to \infty$

$$\sum_{n=2^\ell+1}^{\infty} |a_n| S_n(t) \leq |a_1| + \sum_{m=\ell}^{\infty} \sum_{k=2^m+1}^{2^{m+1}} |a_k| S_k(t) \leq |a_1| + \frac{c}{2} \sum_{m=\ell}^{\infty} \frac{\big(\log 2^{m+1}\big)^{1/2}}{2^{m/2}} \longrightarrow 0$$

    gleichmäßig in $t \in [0,1]$,
  - weil $m + 1 < c' 2^{\varepsilon m}$ für jedes $m \geq 0$ und für gewisse Konstanten $c' < \infty$ bzw. $\varepsilon < 1$
  - und weil somit

$$\sum_{m=\ell}^{\infty} \frac{\big(\log 2^{m+1}\big)^{1/2}}{2^{m/2}} = (\log 2)^{1/2} \sum_{m=\ell}^{\infty} \Big(\frac{m+1}{2^m}\Big)^{1/2} \leq (c' \log 2)^{1/2} \sum_{m=\ell}^{\infty} \Big(\frac{1}{2^{(1-\varepsilon)m}}\Big)^{1/2} \longrightarrow 0\,.$$

  - Damit ist auch die dritte Teilaussage bewiesen. □

2.4.2 Eigenschaften normalverteilter Zufallsvariablen

Bei der Konstruktion von Wiener–Prozessen benötigen wir noch die folgenden Eigenschaften von normalverteilten Zufallsvariablen.

Lemma 2.7 *Sei $\{Y_n,\, n \geq 1\}$ eine Folge von (nicht notwendig unabhängigen) $\mathrm{N}(0,1)$–verteilten Zufallsvariablen über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Dann gilt mit Wahrscheinlichkeit 1, dass*

$$|Y_n| = \mathrm{O}((\log n)^{1/2}) \qquad \text{für } n \to \infty. \tag{8}$$

Beweis

- Sei X eine beliebige $\mathrm{N}(0,1)$–verteilte Zufallsvariable. Dann gilt für jedes $x > 0$, dass

$$(2\pi)^{-1/2}\Big(x + \frac{1}{x}\Big)^{-1} \mathrm{e}^{-x^2/2} \leq P(X > x) \leq (2\pi)^{-1/2} x^{-1}\, \mathrm{e}^{-x^2/2}. \tag{9}$$

  - Denn mit partieller Integration ergibt sich, dass

$$\int_x^\infty \frac{1}{y^2}\, \mathrm{e}^{-y^2/2}\, \mathrm{d}y = \Big[-\frac{1}{y}\, \mathrm{e}^{-y^2/2}\Big]_x^\infty - \int_x^\infty \mathrm{e}^{-y^2/2}\, \mathrm{d}y$$

  und somit

$$\int_x^\infty \Big(1 + \frac{1}{y^2}\Big)\, \mathrm{e}^{-y^2/2}\, \mathrm{d}y = \frac{1}{x}\, \mathrm{e}^{-x^2/2}.$$

  - Hieraus folgt, dass

$$\frac{1}{x}\, \mathrm{e}^{-x^2/2} > \int_x^\infty \mathrm{e}^{-y^2/2}\, \mathrm{d}y > \frac{1}{x}\, \mathrm{e}^{-x^2/2} \Big(1 + \frac{1}{x^2}\Big)^{-1}.$$

  - Wenn sämtliche Ausdrücke der letzten Ungleichungskette durch $\sqrt{2\pi}$ geteilt werden, dann ergibt sich (9), weil $P(X > x) = (2\pi)^{-1/2} \int_x^\infty \mathrm{e}^{-y^2/2}\, \mathrm{d}y$.

- Aus (9) ergibt sich, dass für jedes $c > \sqrt{2}$

$$\sum_{n=3}^\infty P(|Y_n| > c\sqrt{\log n}) \leq \sqrt{\frac{2}{\pi}} \sum_{n=3}^\infty \frac{1}{c}\, (\log n)^{-1/2}\, \exp\Big(-\frac{c^2}{2}\, \log n\Big) \leq \sqrt{\frac{1}{\pi}} \sum_{n=3}^\infty n^{-\frac{c^2}{2}} < \infty.$$

- Aus dem Lemma von Borel–Cantelli (vgl. Lemma WR-2.3) ergibt sich nun die Behauptung. □

Lemma 2.8 *Die Zufallsvariablen $Y_1, Y_n, \ldots$ seien unabhängig und $\mathrm{N}(0,1)$–normalverteilt. Außerdem seien $a_i, b_i \in \mathbb{R}$ beliebige Konstanten, so dass für jedes $m = 1, 2, \ldots$*

$$\sum_{k=1}^{2^m} |a_{2^m+k}| \leq 2^{-\frac{m}{2}} \qquad \text{und} \qquad \sum_{k=1}^{2^m} |b_{2^m+k}| \leq 2^{-\frac{m}{2}}. \tag{10}$$

Dann existieren die Grenzwerte $U = \sum_{i=1}^\infty a_i Y_i$ und $V = \sum_{i=1}^\infty b_i Y_i$ mit Wahrscheinlichkeit 1, und es gilt

$$U \sim \mathrm{N}(0, \sum_{i=1}^\infty a_i^2) \qquad \text{und} \qquad V \sim \mathrm{N}(0, \sum_{i=1}^\infty b_i^2), \tag{11}$$

wobei

$$\mathrm{Cov}\,(U, V) = \sum_{i=1}^\infty a_i b_i. \tag{12}$$

Außerdem sind die Zufallsvariablen U und V sind genau dann unabhängig, wenn $\mathrm{Cov}\,(U, V) = 0$.

Beweis

- Mit Hilfe von Lemma 2.7 ergibt sich auf die gleiche Weise wie im Beweis von Lemma 2.6, dass die Grenzwerte U und V mit Wahrscheinlichkeit 1 existieren und somit wohldefinierte Zufallsvariablen sind.
- Für jedes $m = 1, 2, \ldots$ ergibt sich die Normalverteilheit der Zufallsvariablen $U^{(m)} = \sum_{i=1}^m a_i Y_i$ und $V^{(m)} = \sum_{i=1}^m b_i Y_i$ unmittelbar aus der Faltungsstabilität der Normalverteilung (vgl. Korollar WR-3.2), wobei
$$U^{(m)} \sim \mathrm{N}(0, \sum_{i=1}^m a_i^2) \qquad \text{und} \qquad V^{(m)} \sim \mathrm{N}(0, \sum_{i=1}^m b_i^2)\,.$$
- Hieraus folgt (11), weil $U^{(m)} \xrightarrow{\mathrm{d}} U$ bzw. $V^{(m)} \xrightarrow{\mathrm{d}} V$ für $m \to \infty$.
- Die Formel (12) ergibt sich aus der Unabhängigkeit der Zufallsvariablen $Y_1, \ldots, Y_n$ und aus den allgemeinen Rechenregeln für die Kovarianz, und zwar gilt
$$\begin{aligned}
\mathrm{Cov}\,(U,V) &= \lim_{m\to\infty} \mathrm{Cov}\Big(\sum_{i=1}^m a_i Y_i, \sum_{j=1}^m b_j Y_j\Big) = \lim_{m\to\infty} \sum_{i=1}^m a_i \mathrm{Cov}\,\Big(Y_i, \sum_{j=1}^m b_j Y_j\Big) \\
&= \lim_{m\to\infty} \sum_{i=1}^m \sum_{j=1}^m a_i b_j \mathrm{Cov}\,(Y_i, Y_j) = \lim_{m\to\infty} \sum_{i=1}^m a_i b_i \mathrm{Cov}\,(Y_i, Y_i) \\
&= \sum_{i=1}^\infty a_i b_i\,.
\end{aligned}$$
- Die Notwendigkeit der Bedingung in Teilaussage 2 ergibt sich aus der Multiplikationsformel für den Erwartungswert des Produktes von unabhängigen Zufallsvariablen; vgl. Korollar WR-4.5.
- Um die Hinlänglichkeit der Bedingung in Teilaussage 2 zu beweisen, betrachten wir die (gemeinsame) charakteristische Funktion $\varphi_{U,V}(t,s)$ des Zufallsvektors (U,V).
  - Weil $(U^{(m)}, V^{(m)}) \xrightarrow{\mathrm{d}} (U,V)$ für $m \to \infty$, gilt dann für beliebige $t, s \in \mathbb{R}$
$$\begin{aligned}
\varphi_{U,V}(t,s) &= \lim_{m\to\infty} \mathbb{E}\,\exp\Big(\mathrm{i}\,\Big(t\sum_{i=1}^m a_i Y_i + s\sum_{j=1}^m b_j Y_j\Big)\Big) = \lim_{m\to\infty} \mathbb{E}\,\exp\Big(\mathrm{i}\sum_{i=1}^m (ta_i + sb_i) Y_i\Big) \\
&= \lim_{m\to\infty} \prod_{i=1}^m \mathbb{E}\,\exp\Big(\mathrm{i}\,(ta_i + sb_i) Y_i\Big) = \lim_{m\to\infty} \prod_{i=1}^m \exp\Big(-\frac{(ta_i + sb_i)^2}{2}\Big) \\
&= \exp\Big(-\frac{\sum_{i=1}^\infty (ta_i + sb_i)^2}{2}\Big) \overset{(*)}{=} \exp\Big(-\frac{\sum_{i=1}^\infty (ta_i)^2 + \sum_{j=1}^m (sb_j)^2}{2}\Big) \\
&= \exp\Big(-\frac{t^2 \sum_{i=1}^\infty a_i^2}{2}\Big)\ \exp\Big(-\frac{s^2 \sum_{j=1}^\infty b_j^2}{2}\Big) \\
&= \varphi_U(t)\ \varphi_V(s)\,,
\end{aligned}$$
  - wobei sich die Gleichheit $(*)$ aus der Annahme ergibt, dass die Zufallsvariablen U und V unkorreliert sind und dass deshalb $\mathrm{Cov}\,(U,V) = \sum_{i=1}^\infty a_i b_i = 0$ gilt.
- Die Hinlänglichkeit der Bedingung in Teilaussage 2 ergibt sich nun
  - aus dem Eindeutigkeitssatz für charakteristische Funktionen von Zufallsvektoren (der eine Verallgemeinerung des in Korollar WR-5.5 betrachteten eindimensionalen Falles ist),
  - weil das Produkt der charakteristischen Funktionen von unabhängigen Zufallsvektoren Z und Z' gleich der charakteristischen Funktion des Zufallsvektors (Z, Z') ist. □

2.4.3 Konstruktion von Wiener–Prozessen; Simulationsalgorithmus

Aus Lemma 2.7 ergibt sich, dass man

- von dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, über dem die Folge $\{Y_n,\ n \geq 1\}$ von $\mathrm{N}(0,1)$–verteilten Zufallsvariablen gegeben ist,
- zu dem folgenden (eingeschränkten) Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', P')$ übergehen kann, wobei

$$\begin{aligned} \Omega' &= \{\omega \in \Omega : |X_n(\omega)| = \mathrm{O}(\sqrt{\log n}) \text{ für } n \to \infty\}, \\ \mathcal{F}' &= \mathcal{F} \cap \Omega', \qquad \text{d.h., } \mathcal{F}' = \{A \cap \Omega' : A \in \mathcal{F}\}, \\ P' &= P_{|\mathcal{F}'}, \qquad \text{d.h., } P'(A \cap \Omega') = P(A) \text{ für jedes } A \in \mathcal{F}. \end{aligned}$$

Beachte

- Weil vorausgesetzt wird, dass die $\mathrm{N}(0,1)$–verteilten Zufallsvariablen $Y_1, Y_2, \ldots$ unabhängig sind, kann $(\Omega, \mathcal{F}, P)$ beispielsweise der sogenannte *kanonische Wahrscheinlichkeitsraum* sein mit

$$(\Omega, \mathcal{F}, P) = \big(\mathbb{R} \times \mathbb{R} \times \ldots,\ \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \otimes \ldots,\ \mathrm{N}(0,1) \times \mathrm{N}(0,1) \times \ldots\big).$$

- Weil man sich leicht überlegen kann, dass $\Omega' \in \mathcal{F}$ gilt und weil $P(\Omega') = 1$ wegen Lemma 2.7, ist somit $(\Omega', \mathcal{F}', P')$ ein wohldefinierter Wahrscheinlichkeitsraum.

Insgesamt ergibt sich nun der folgende Ansatz zur Konstruktion von Wiener–Prozessen in $[0,1]$.

Theorem 2.21

- *Sei $\{Y_n,\ n \geq 1\}$ eine Folge von unabhängigen und $N(0,1)$–verteilten Zufallsvariablen über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, und sei $S_1, S_2, \ldots : [0,1] \to [0,\infty)$ das in (2)–(4) bzw. (6) gegebene System von Schauder–Funktionen.*
- *Dann ist der stochastische Prozess $\{X_t,\ t \in [0,1]\}$ mit*

$$X_t = \sum_{n=1}^{\infty} Y_n\, S_n(t) \qquad \forall\, t \in [0,1] \tag{13}$$

ein Prozess mit unabhängigen und normalverteilten Zuwächsen über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und ein Wiener–Prozess in $[0,1]$ über dem eingeschränkten Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', P')$.

Beweis

- Mit Hilfe von (7) ergibt sich aus der ersten Teilaussage von Lemma 2.8, dass die in (13) eingeführte Größe X_t für jedes $t \in [0,1]$ eine wohldefinierte (normalverteilte) Zufallsvariable ist, wobei

$$X_t \sim \mathrm{N}(0,t) \qquad \forall\, t \in [0,1]. \tag{14}$$

  - Denn es gilt

$$\sum_{n=1}^{\infty} S_n^2(t) = \sum_{n=1}^{\infty} \Big(\int_0^1 H_n(s) \mathbb{1}([0,t])(s)\, \mathrm{d}s\Big)^2 = \sum_{n=1}^{\infty} \big(H_n, \mathbb{1}([0,t])\big)^2 = \big(\mathbb{1}([0,t]),\ \mathbb{1}([0,t])\big) = t,$$

  - Dabei ergibt sich die vorletzte Gleichheit aus Lemma 2.5, weil $\mathbb{1}([0,t]) \in L_2$ und weil somit aus (5) folgt, dass

$$\big(\mathbb{1}([0,t]),\ \mathbb{1}([0,t])\big) = \sum_{n=1}^{\infty} \big(H_n, \mathbb{1}([0,t])\big)^2.$$

- Außerdem gilt $X_{t_2} - X_{t_1} = \sum_{n=1}^{\infty} Y_n \left(S_n(t_2) - S_n(t_1)\right)$ für beliebige $t_1, t_2 \geq 0$ mit $t_1 < t_2$,
  - und ähnlich wie beim Beweis von (14) kann man zeigen, dass
  $$X_{t_2} - X_{t_1} \sim \mathrm{N}(0, t_2 - t_1)\,,$$
  - weil
  $$\sum_{n=1}^{\infty} \left(S_n(t_2) - S_n(t_1)\right)^2 = t_2 - t_1\,,$$
  - d.h. insbesondere, dass die Zuwächse $X_{t_2} - X_{t_1}$ des stochastischen Prozesses $\{X_t\}$ normalverteilt sind.
- Wir zeigen nun, dass $\{X_t\}$ ein Prozess mit unabhängigen Zuwächsen ist.
  - Aus (12) ergibt sich, dass für beliebige $t_1, \ldots, t_4 \geq 0$ mit $t_1 < t_2 \leq t_3 < t_4$
  $$\mathrm{Cov}\left(X_{t_2} - X_{t_1},\, X_{t_4} - X_{t_3}\right) = \sum_{n=1}^{\infty} \left(S_n(t_2) - S_n(t_1)\right)\left(S_n(t_4) - S_n(t_3)\right).$$
  - Dabei ist
  $$\begin{aligned} &\sum_{n=1}^{\infty} \left(S_n(t_2) - S_n(t_1)\right)\left(S_n(t_4) - S_n(t_3)\right) \\ &= \sum_{n=1}^{\infty} \int_0^1 H_n(s) \mathbb{1}([t_1, t_2])(s)\,\mathrm{d}s \int_0^1 H_n(s) \mathbb{1}([t_3, t_4])(s)\,\mathrm{d}s = \sum_{n=1}^{m} \left(H_n,\, \mathbb{1}([t_1, t_2])\right)\left(H_n,\, \mathbb{1}([t_3, t_4])\right) \\ &= \left(\mathbb{1}([t_1, t_2])\, \mathbb{1}([t_3, t_4])\right) = 0\,, \end{aligned}$$
  wobei sich die vorletzte Gleichheit erneut aus Formel (5) in Lemma 2.5 ergibt.
  - Somit gilt $\mathrm{Cov}\left(X_{t_2} - X_{t_1},\, X_{t_4} - X_{t_3}\right) = 0$, und aus der zweiten Teilaussage von Lemma 2.8 ergibt sich nun, dass $\{X_t\}$ unabhängige Zuwächse hat.
- Schließlich ergibt sich aus der dritten Teilaussage von Lemma 2.6 und aus Lemma 2.7, dass sämtliche Trajektorien des Prozess $\{X_t,\, t \in [0,1]\}$ über dem eingeschränkten Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', P')$ stetige Funktionen sind, weil
  - die Summe $\sum_{n=1}^{m} Y_n\, S_n(t)$ für jedes $m \geq 1$ eine stetige Funktion in t ist und
  - die Konvergenz $\sum_{n=1}^{m} Y_n\, S_n(t) \longrightarrow_{m\to\infty} \sum_{n=1}^{\infty} Y_n\, S_n(t)$ gleichmäßig in t erfolgt.
- Weil offenbar auch $X_0 = 0$ gilt, ist damit die Behauptung vollständig bewiesen. □

Beachte

- Die Aussage von Theorem 2.21 rechtfertigt den folgenden Algorithmus zur (näherungsweisen) Simulation des in (13) gegebenen Wiener–Prozesses $\{X_t,\, t \in [0,1]\}$.
- Um für ein $k = 2^m$ mit $m \geq 0$ die Approximation $\{X_t^{(k)},\, t \in [0,1]\}$ mit $X_t^{(k)} = \sum_{n=1}^{k} Y_n\, S_n(t)$ des Wiener–Prozesses $\{X_t\}$ zu simulieren,
  - generiert man zunächst die Realisierungen $y_1, \ldots, y_k$ der unabhängigen und $\mathrm{N}(0,1)$–verteilten Zufallsvariablen $Y_1, \ldots, Y_k$ und
  - berechnet dann *schrittweise* die 2^n Segmente der Polygonzüge $\{X_t^{(2^n)}, t \in [0,1]\}$ für $n = 0, 1, \ldots, m$.
- Dabei kann man zum Beispiel für $m = 2$ und $n = 0, 1, 2$ wie folgt vorgehen:

  Schritt 0 Für die Realisierung y_1 von Y_1 ergibt sich der Endpunkt $(1, x_1^{(1)}) = (1, y_1)$ des Segmentes $(0,0) \to (1, x_1^{(1)})$, das die entsprechende Trajektorie von $\{X_t^{(1)},\, t \in [0,1]\}$ ist.

Schritt 1 Berechne den Mittelpunkt des Segmentes $(0,0) \to (1, x_1^{(1)})$ und addiere den Wert $y_2 S_2(1/2)$ $(= y_2/2)$ zu dessen zweiter Komponente. Dies ergibt den Wert

$$x_{1/2}^{(2)} = y_1 S_1(1/2) + y_2 S_2(1/2)\,,$$

wobei die Trajektorie von $\{X_t^{(2)},\, t \in [0,1]\}$ aus den zwei Segmenten

$$(0,0) \to (1/2, x_{1/2}^{(2)}) \qquad \text{und} \qquad (1/2, x_{1/2}^{(2)}) \to (1, x_1^{(1)})$$

besteht.

Schritt 2 Berechne die Mittelpunkte der Segmente

$$(0,0) \to (1/2, x_{1/2}^{(2)}) \qquad \text{und} \qquad (1/2, x_{1/2}^{(2)}) \to (1, x_1^{(1)})$$

und addiere jeweils die Werte $y_3 S_3(1/4)$ $(= y_3/2^{3/2})$ bzw. $y_4 S_4(3/4)$ $(= y_4/2^{3/2})$ zu deren zweiten Komponenten. Dies ergibt die Werte

$$x_{1/4}^{(4)} = y_1 S_1(1/4) + y_2 S_2(1/4) + y_3 S_3(1/4) \quad \text{und} \quad x_{3/4}^{(4)} = y_1 S_1(3/4) + y_2 S_2(3/4) + y_4 S_4(3/4)\,,$$

wobei die Trajektorie von $\{X_t^{(4)},\, t \in [0,1]\}$ nun aus vier Segmenten besteht:

$$(0,0) \to (1/4, x_{1/4}^{(4)})\,, \qquad (1/4, x_{1/4}^{(4)}) \to (1/2, x_{1/2}^{(2)})$$

und

$$(1/2, x_{1/2}^{(2)}) \to (3/4, x_{3/4}^{(4)})\,, \qquad (3/4, x_{3/4}^{(4)}) \to (1, x_1^{(1)})\,.$$

Beachte

- Die in Abschnitt 2.4.3 enthaltenen Darlegungen lassen sich auf völlig analoge Weise
  - zur Konstruktion und Simulation von Wiener–Prozessen in beliebigen beschränkten Intervallen der Form $I = [0,c]$ verwenden, wobei $c < \infty$ eine beliebige positive Zahl ist.
  - Denn man kann sich leicht überlegen, dass $\{Y_t,\, t \in [0,c]\}$ mit $Y_t = \sqrt{c}\, X_{t/c}$ ein Wiener–Prozess in $[0,c]$ ist, wenn $\{X_t,\, t \in [0,1]\}$ ein Wiener–Prozess in $[0,1]$ ist; vgl. auch Theorem 2.24 in Abschnitt 2.4.6.
- Wir erwähnen nun noch eine Möglichkeit, wie Wiener–Prozesse in dem unbeschränkten Intervall $I = [0,\infty)$ auf einfache Weise konstruiert werden können.
  - Hierfür genügt es, eine Doppelfolge $\{Y_{nm},\, n,m \geq 1\}$ von unabhängigen und $\mathrm{N}([0,1])$–verteilten Zufallsvariablen zu betrachten, die über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ gegeben sind.
  - Gemäß Theorem 2.21 lässt sich über dem eingeschränkten Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', P')$ eine Folge $\{X_t^{(1)},\, t \in [0,1]\}$, $\{X_t^{(2)},\, t \in [0,1]\}$, ... von unabhängigen Wiener–Prozessen jeweils im Intervall $[0,1]$ konstruieren.
  - Man kann sich leicht überlegen, dass dann durch den Ansatz

$$X_t = \begin{cases} X_t^{(1)} & \text{für } t \in [0,1], \\ X_1^{(1)} + X_{t-1}^{(2)} & \text{für } t \in [1,2], \\ X_1^{(1)} + X_1^{(2)} + X_{t-2}^{(3)} & \text{für } t \in [2,3], \\ \vdots & \\ \sum_{k=1}^{n-1} X_1^{(k)} + X_{t-(n-1)}^{(n)} & \text{für } t \in [n-1,n], \\ \vdots & \end{cases}$$

ein stochastischer Prozess $\{X_t,\, t \geq 0\}$ gegeben ist, der ein Wiener–Prozess in $[0,\infty)$ über dem eingeschränkten Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', P')$ ist.

2.4.4 Darstellung als Markow–Prozess bzw. Gauß–Prozess

Wir zeigen nun, dass der Wiener–Prozess ein Markow–Prozess im Sinne der folgenden (verallgemeinerten) Definition dieses Begriffes ist, die ähnlich zu der in Abschnitt 2.3.1 betrachteten Definition von Markow–Prozessen mit endlichem Zustandsraum ist.

Sei $\mathcal{P}(\mathbb{R})$ die Familie sämtlicher Wahrscheinlichkeitsmaße auf der Borel–σ–Algebra $\mathcal{B}(\mathbb{R})$. Um den Begriff eines Markow–Prozesses $\{X_t,\, t \geq 0\}$ mit stetiger Zeit und mit Werten in dem (überabzählbaren) Zustandsraum $E = \mathbb{R}$ einzuführen, betrachten wir

- ein Wahrscheinlichkeitsmaß $\boldsymbol{\alpha} : \mathcal{B}(\mathbb{R}) \to [0,1]$ sowie
- einen stochastischen Kern $\mathbf{P} : [0,\infty) \times \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1]$, so dass

$$\mathbf{P}(h, x, \cdot) \in \mathcal{P}(\mathbb{R})\,, \tag{15}$$

$$\mathbf{P}(0, x, \{x\}) = 1\,, \tag{16}$$

$$\mathbf{P}(\cdot, \cdot, B) \text{ ist eine messbare Funktion auf } [0,\infty) \times \mathbb{R}, \tag{17}$$

$$\mathbf{P}(h_1 + h_2, x, B) = \int_{\mathbb{R}} \mathbf{P}(h_2, y, B)\, \mathbf{P}(h_1, x, \mathrm{d}y) \tag{18}$$

für beliebige $h, h_1, h_2 \geq 0$, $x \in \mathbb{R}$ und $B \in \mathcal{B}(\mathbb{R})$.

Definition

1. Ein stochastischer Kern $\mathbf{P}$, der den Bedingungen (15)–(18) genügt, wird *Übergangskern* genannt.
2. Ein stochastischer Prozess $\{X_t, t \geq 0\}$ mit Werten in $E = \mathbb{R}$ heißt *homogener Markow–Prozess*, wenn es einen Übergangskern $\mathbf{P}$ und ein Wahrscheinlichkeitsmaß $\boldsymbol{\alpha}$ über $\mathcal{B}(\mathbb{R})$ gibt, so dass

$$\begin{aligned} &P(X_0 \in B_0, X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n) \\ &= \int_{B_0} \int_{B_1} \ldots \int_{B_n} \mathbf{P}(t_n - t_{n-1}, x_{n-1}, \mathrm{d}x_n) \ldots \mathbf{P}(t_1, x_0, \mathrm{d}x_1)\, \boldsymbol{\alpha}(\mathrm{d}x_0)\,, \end{aligned} \tag{19}$$

für beliebige $n = 0, 1, \ldots$, $B_0, B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$, $t_0 = 0 \leq t_1 \leq \ldots \leq t_n$.

Beachte

- Das Wahrscheinlichkeitsmaß $\boldsymbol{\alpha}$ in (19) wird *Anfangsverteilung* genannt.
- Außerdem wird $\mathbf{P}(h, x, B)$ als die Wahrscheinlichkeit interpretiert, dass der Prozess $\{X_t\}$ in h Zeiteinheiten vom Zustand x in einen Zustand aus B übergeht.

Theorem 2.22 *Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess. Dann ist $\{X_t\}$ ein Markow–Prozess mit $\boldsymbol{\alpha} = \delta_0$, d.h., $\boldsymbol{\alpha}(\{0\}) = 1$, und*

$$\mathbf{P}(h, x, B) = \sqrt{\frac{1}{2\pi h}} \int_B \exp\Big(-\frac{(y-x)^2}{2h}\Big)\, \mathrm{d}y \tag{20}$$

für beliebige $h > 0$, $x \in \mathbb{R}$ und $B \in \mathcal{B}(\mathbb{R})$.

Beweis

- Wegen $X_0 = 0$ genügt es, anstelle von (19), zu zeigen, dass die folgende Desintegrationsgleichung

$$P(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n) = \int_{B_1} \ldots \int_{B_n} \mathbf{P}(t_n - t_{n-1}, x_{n-1}, \mathrm{d}x_n) \ldots \mathbf{P}(t_1, 0, \mathrm{d}x_1)\,, \tag{21}$$

für beliebige $n = 1, 2 \ldots$, $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ und $0 < t_1 < \ldots < t_n$ gilt.

- Aus der Unabhängigkeit der Zuwächse des Wiener–Prozesses $\{X_t\}$ ergibt sich, dass

$$\begin{aligned}
&P(X_{t_1} \in B_1, X_{t_2} \in B_2, \ldots, X_{t_n} \in B_n) \\
&= P(X_{t_1} \in B_1, X_{t_2} - X_{t_1} \in B_2 - X_{t_1}, \ldots, X_{t_n} - X_{t_1} \in B_n - X_{t_1}) \\
&= \int_{B_1} P(X_{t_2} - X_{t_1} \in B_2 - x_1, \ldots, X_{t_n} - X_{t_1} \in B_n - x_1) P(X_{t_1} \in \mathrm{d}x_1) \\
&= \int_{B_1} \int_{B_2 - x_1} P(X_{t_3} - X_{t_2} \in B_3 - x_1 - x_2, \ldots, X_{t_n} - X_{t_2} \in B_n - x_1 - x_2) \\
&\qquad\qquad P(X_{t_2} - X_{t_1} \in \mathrm{d}x_2)\, P(X_{t_1} \in \mathrm{d}x_1) \\
&\;\vdots \\
&= \int_{B_1} \int_{B_2 - x_1} \cdots \int_{B_n - x_1 - \ldots - x_{n-1}} P(X_{t_n} - X_{t_{n-1}} \in \mathrm{d}x_n) \ldots P(X_{t_2} - X_{t_1} \in \mathrm{d}x_2)\, P(X_{t_1} \in \mathrm{d}x_1)
\end{aligned}$$

- Somit ergibt sich aus der Normalverteiltheit der Zuwächse von $\{X_t\}$, dass

$$\begin{aligned}
&P(X_{t_1} \in B_1, X_{t_2} \in B_2, \ldots, X_{t_n} \in B_n) \\
&= \int_{B_1} \int_{B_2 - x_1} \cdots \int_{B_n - x_1 - \ldots - x_{n-1}} \sqrt{\frac{1}{2\pi(t_n - t_{n-1})}}\, \exp\Big(-\frac{x_n^2}{2(t_n - t_{n-1})}\Big)\, \mathrm{d}x_n \\
&\qquad \cdots \sqrt{\frac{1}{2\pi(t_2 - t_1)}}\, \exp\Big(-\frac{x_2^2}{2(t_2 - t_1)}\Big)\, \mathrm{d}x_2 \sqrt{\frac{1}{2\pi t_1}}\, \exp\Big(-\frac{x_1^2}{2t_1}\Big)\, \mathrm{d}x_1 \\
&= \int_{B_1} \int_{B_2} \cdots \int_{B_n} \sqrt{\frac{1}{2\pi(t_n - t_{n-1})}}\, \exp\Big(-\frac{(x_n - x_{n-1})^2}{2(t_n - t_{n-1})}\Big)\, \mathrm{d}x_n \\
&\qquad \cdots \sqrt{\frac{1}{2\pi(t_2 - t_1)}}\, \exp\Big(-\frac{(x_2 - x_1)^2}{2(t_2 - t_1)}\Big)\, \mathrm{d}x_2 \sqrt{\frac{1}{2\pi t_1}}\, \exp\Big(-\frac{x_1^2}{2t_1}\Big)\, \mathrm{d}x_1 .
\end{aligned}$$

- Damit ist die Gültigkeit von (21) gezeigt, wobei der Übergangskern $\mathbf{P}$ durch (20) gegeben ist. □

Beachte

- Aus dem Beweis von Theorem 2.22 ergibt sich insbesondere, dass $(X_{t_1}, \ldots, X_{t_n})$ ein absolutstetiger Zufallsvektor ist, dessen (gemeinsame) Dichte $f_{t_1,\ldots,t_n} : \mathbb{R}^n \to [0, \infty)$ für $0 < t_1 < \ldots < t_n$ gegeben ist durch

$$f_{t_1,\ldots,t_n}(x_1, \ldots, x_n) = (2\pi)^{-n/2} \big(t_1(t_2 - t_1) \ldots (t_n - t_{n-1})\big)^{-1/2} \exp\Big(-\frac{1}{2} \sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}\Big), \tag{22}$$

  wobei $t_0 = x_0 = 0$.
- Aus (22) folgt, dass der Zufallsvektor $(X_{t_1}, \ldots, X_{t_n})$ eine (multivariate) Normalverteilung hat, wobei $\mathrm{Cov}\,(X_s, X_t) = \min\{s, t\}$ für beliebige $s, t \geq 0$.
- Ein stochastischer Prozess $\{X_t,\, t \geq 0\}$, dessen endlich–dimensionale Verteilungen (multivariate) Normalverteilungen sind, wird *Gauß–Prozess* genannt.

2.4.5 Verteilung des Maximums

In diesem Abschnitt setzen wir die Diskussion von Eigenschaften des Wiener–Prozesses fort, die mit relativ elementaren Hilfsmitteln gezeigt werden können. Dabei leiten wir zunächst eine obere Schranke für die Tailfunktion des Maximums $\max_{t \in [0,1]} X_t$ von Wiener–Prozessen in $[0, 1]$ her.

Später zeigen wir in Kapitel 3 mit den Techniken von Martingalen bzw. Lévy–Prozessen, dass diese Schranke „optimal" ist, d.h., mit der Tailfunktion von $\max_{t \in [0,1]} X_t$ übereinstimmt.

Theorem 2.23

- *Sei* $\{X_t,\, t \in [0,1]\}$ *ein Wiener–Prozess über einem gewissen Wahrscheinlichkeitsraum* $(\Omega, \mathcal{F}, P)$.
- *Dann gilt*

$$P\big(\max_{t\in[0,1]} X_t > x\big) \le \sqrt{\frac{2}{\pi}} \int_x^\infty \mathrm{e}^{-y^2/2}\,\mathrm{d}y \qquad \forall\, x \ge 0\,. \tag{23}$$

Beachte

- Die in (23) betrachtete Abbildung $\max_{t\in[0,1]} X_t : \Omega \to [0,\infty)$ ist eine wohldefinierte Zufallsvariable, weil die Trajektorien von $\{X_t\}$ stetige Funktionen sind und somit

$$\big(\max_{t\in[0,1]} X_t\big)(\omega) = \lim_{k\to\infty} \big(\max_{i=1,\ldots,k} X_{i/k}\big)(\omega) \qquad \forall\, \omega \in \Omega\,. \tag{24}$$

- Aus (23) folgt, dass die Zufallsvariable $\max_{t\in[0,1]} X_t$ einen exponentiell beschränkten Tail hat. Hieraus ergibt sich insbesondere, dass sämtliche Momente von $\max_{t\in[0,1]} X_t$ existieren (und endlich sind).
- Im Beweis von Theorem 2.23 benötigen wir Bezeichnungen und Hilfsmittel, die teilweise bereits in der Vorlesung „Wahrscheinlichkeitsrechnung" bereitgestellt wurden.
  - Sei $\{X_t,\, t \ge 0\}$ ein Wiener–Prozess, und sei $Z_1, Z_2, \ldots$ eine Folge von unabhängigen Zufallsvariablen mit $P(Z_i = 1) = P(Z_i = -1) = 1/2$ für jedes $i \ge 1$.
  - Außerdem betrachten wir für jedes $n \ge 1$ den stochastischen Prozess $\{\widetilde{X}_t^{(n)},\, t \in [0,1]\}$ mit

$$\widetilde{X}_t^{(n)} = S_{\lfloor nt \rfloor}/\sqrt{n} + (nt - \lfloor nt \rfloor)\, Z_{\lfloor nt \rfloor + 1}/\sqrt{n}\,, \tag{25}$$

    wobei $S_i = Z_1 + \ldots + Z_i$ für jedes $i \ge 1$ und $\lfloor t \rfloor$ den ganzzahligen Teil von $t \ge 0$ bezeichnet; $S_0 = 0$.
- Das folgende Lemma zeigt, dass $\{\widetilde{X}_t^{(n)}\}$ als Approximation des Wiener–Prozesses $\{X_t\}$ aufgefasst werden kann.
  - Allerdings konvergiert der in (25) gegebene Prozess $\{\widetilde{X}_t^{(n)}\}$ nur „in Verteilung" für $n \to \infty$ gegen $\{X_t\}$,
  - während die in Abschnitt 2.4.3 betrachtete Approximation $\{X_t^{(n)}\}$ mit $X_t^{(n)} = \sum_{k=1}^n Y_k\; S_k(t)$ „pfadweise" für $n \to \infty$ gegen den Wiener–Prozess $\{X_t\}$ konvergiert.

Lemma 2.9 *Für jedes* $k \ge 1$ *und für beliebige* $t_1, \ldots, t_k \in [0,1]$ *gilt*

$$\big(\widetilde{X}_{t_1}^{(n)}, \ldots, \widetilde{X}_{t_k}^{(n)}\big) \xrightarrow{\mathrm{d}} \big(X_{t_1}, \ldots, X_{t_k}\big) \qquad \textit{für } n \to \infty. \tag{26}$$

Beweis

- Wir zeigen die Behauptung nur für den Spezialfall $k = 2$; für beliebige $k \ge 2$ verläuft der Beweis analog. Außerdem nehmen wir o.B.d.A. an, dass $t_1 < t_2$.
- Dann gilt für beliebige $s_1, s_2 \in \mathbb{R}$, dass

$$\begin{aligned} s_1 \widetilde{X}_{t_1}^{(n)} + s_2 \widetilde{X}_{t_2}^{(n)} \quad = \quad & (s_1 + s_2) S_{\lfloor nt_1 \rfloor}/\sqrt{n} + s_2\big(S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor + 1}\big)/\sqrt{n} \\ & + Z_{\lfloor nt_1 \rfloor + 1}\big((nt_1 - \lfloor nt_1 \rfloor) s_1/\sqrt{n} + s_2/\sqrt{n}\big) + Z_{\lfloor nt_2 \rfloor + 1}(nt_2 - \lfloor nt_2 \rfloor) s_2/\sqrt{n}\,, \end{aligned}$$

  - wobei die vier Summanden auf der rechten Seite der letzten Gleichheit unabhängige Zufallsvariablen sind,
  - so dass die letzten beiden Summanden jeweils (in Verteilung) gegen Null konvergieren.

- Aus der Charakterisierung der Verteilungskonvergenz mittels charakteristischer Funktionen (siehe Theorem WR-5.20) ergibt sich somit, dass die charakteristischen Funktionen der beiden letzten Summanden punktweise gegen 1 konvergieren.
- Insgesamt gilt also

$$\lim_{n\to\infty} \mathbb{E}\ \mathrm{e}^{\mathrm{i}(s_1\widetilde{X}_{t_1}^{(n)}+s_2\widetilde{X}_{t_2}^{(n)})} = \lim_{n\to\infty} \big(\varphi(s_1/\sqrt{n}+s_2/\sqrt{n})\big)^{\lfloor nt_1\rfloor}\ \big(\varphi(s_2/\sqrt{n})\big)^{\lfloor nt_2\rfloor-\lfloor nt_1\rfloor-1}$$

  wobei $\varphi(s) = \mathbb{E}\ \exp(\mathrm{i}sZ_1)$ für $s \in \mathbb{R}$.
- Weil $\varphi^n(s/\sqrt{n})$ die charakteristische Funktion der Zufallsvariablen $(Z_1+\ldots+Z_n)/\sqrt{n}$ ist,
  - ergibt sich aus dem zentralen Grenzwertsatz, dass

$$\lim_{n\to\infty} \varphi^n(s/\sqrt{n}) = \mathrm{e}^{-s^2/2}$$

  - und somit

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\ \mathrm{e}^{\mathrm{i}(s_1\widetilde{X}_{t_1}^{(n)}+s_2\widetilde{X}_{t_2}^{(n)})} &= \exp\big(-((s_1+s_2)^2t_1+s_2^2(t_2-t_1))/2\big)\\ &= \exp\big(-(s_1^2t_1+2s_1s_2\min\{t_1,t_2\}+s_2^2t_2)/2\big)\,.\end{aligned}$$

- Weil der letzte Ausdruck die charakteristische Funktion des Zufallsvektors (X_{t_1}, X_{t_2}) ist, ergibt sich nun die Behauptung
  - durch die erneute Anwendung der Charakterisierung der Verteilungskonvergenz mittels charakteristischer Funktionen
  - und aus dem eineindeutigen Zusammenhang zwischen charakteristischer Funktion und Verteilung (vgl. Korollar WR-5.5 für den eindimensionalen Fall). □

Lemma 2.10 *Sei* $\widetilde{X}^{(n)} = \max_{t\in[0,1]} \widetilde{X}_t^{(n)}$. *Dann gilt*

$$\widetilde{X}^{(n)} = \frac{1}{\sqrt{n}} \max_{k=1,\ldots,n} S_k \qquad \forall\, n = 1,2,\ldots \tag{27}$$

und

$$\lim_{n\to\infty} P(\widetilde{X}^{(n)} \le x) = \sqrt{\frac{2}{\pi}} \int_0^x \mathrm{e}^{-y^2/2}\,\mathrm{d}y \qquad \forall\, x \ge 0\,. \tag{28}$$

Beweis

- Die Gleichung (27) folgt unmittelbar aus der in (25) gegebenen Definition von $\widetilde{X}_t^{(n)}$.
- Die Gleichung (28) beruht auf dem sogenannten
  - *Reflexionsprinzip* der maximalen Gewinnsume bei n–maligem Münzwurf
  - und wurde bereits in Abschnitt WR-5.3.2 hergeleitet (vgl. Formel (68) in Abschnitt WR-5.3.2). □

Beweis von Theorem 2.23

- Aus Lemma 2.9 ergibt sich, dass für beliebige $x \ge 0$, $k \ge 1$ und $t_1,\ldots,t_k \in [0,1]$

$$\lim_{n\to\infty} P\big(\max_{t\in\{t_1,\ldots,t_k\}} \widetilde{X}_t^{(n)} > x\big) = P\big(\max_{t\in\{t_1,\ldots,t_k\}} X_t > x\big)\,.$$

- Hieraus folgt, dass

$$\lim_{n\to\infty} P\big(\max_{t\in[0,1]} \widetilde{X}_t^{(n)} > x\big) \ge P\big(\max_{t\in\{t_1,\ldots,t_k\}} X_t > x\big)\,.$$

- Mit $\{t_1, \ldots, t_k\} = \{1/k, 2/k, \ldots, k/k\}$ ergibt sich nun unter Beachtung von (24), dass
$$\lim_{n\to\infty} P(\max_{t\in[0,1]} \widetilde{X}_t^{(n)} > x) \geq P(\max_{t\in[0,1]} X_t > x)\,.$$
- Wegen Lemma 2.10 ist damit die Ungleichung (23) bewiesen. □

Beachte Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess über $(\Omega, \mathcal{F}, P)$. Dabei setzen wir (wie stets in diesem Skript) voraus, dass $(\Omega, \mathcal{F}, P)$ ein vollständiger Wahrscheinlichkeitsraum ist, d.h.,

- dass sämtliche „Nullmengen" des Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$, über dem der Wiener–Prozess $\{X_t\}$ gegeben ist, zu $\mathcal{F}$ gehören, bzw. (äquivalent hierzu)
- für $C \subset \Omega$ gilt $C \in \mathcal{F}$, falls es Teilmengen $A, B \in \mathcal{F}$ gibt, so dass $A \subset C \subset B$ und $P(B \setminus A) = 0$.

Aus Theorem 2.23 ergibt sich nun die folgende asymptotische Eigenschaft des Wiener–Prrozesses.

Korollar 2.4 *Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess. Dann gilt mit Wahrscheinlichkeit* 1
$$\lim_{t\to\infty} \frac{X_t}{t} = 0\,. \tag{29}$$

Beweis

- Für beliebige $n = 1, 2, \ldots$ und $t \in [n, n+1)$ gilt
$$\begin{aligned}\left|\frac{X_t}{t} - \frac{X_n}{n}\right| &\leq \left|\frac{X_t}{t} - \frac{X_n}{t}\right| + \left|\frac{X_n}{t} - \frac{X_n}{n}\right| \\ &\leq |X_n|\left|\frac{1}{t} - \frac{1}{n}\right| + \frac{1}{n}\sup_{t\in[n,n+1]} |X_t - X_n| \\ &\leq \frac{|X_n|}{n^2} + \frac{Z_n}{n}\,,\end{aligned}$$
wobei $Z_n = \sup_{t\in[0,1]} |X_{n+t} - X_n|$ unabhängige und identisch verteilte Zufallsvariablen sind; $n \geq 0$.
- Hieraus und aus der Vollständigkeit des Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$ ergibt sich,
  - dass $\lim_{t\to\infty} X_t/t$ eine wohldefinierte Zufallsvariable mit der Eigenschaft (29) ist,
  - wenn wir zeigen, dass mit Wahrscheinlichkeit 1
$$\lim_{n\to\infty} \frac{|X_n|}{n^2} = 0 \qquad \text{und} \qquad \lim_{n\to\infty} \frac{Z_n}{n} = 0\,. \tag{30}$$
- Die erste Nullkonvergenz in (30) ergibt sich unmittelbar aus dem starken Gesetz der großen Zahlen für Summen unabhängiger und identisch verteilter Zufallsvariablen, denn es gilt
$$\frac{|X_n|}{n} = \left|\frac{\sum_{i=1}^n (X_i - X_{i-1})}{n}\right| \longrightarrow |\mathbb{E}\, X_1| = 0$$
und somit auch $|X_n|/n^2 \to 0$ für $n \to \infty$.
- Um auch die Gültigkeit der zweiten Nullkonvergenz in (30) mit dem starken Gesetz der großen Zahlen begründen zu können, muss noch gezeigt werden, dass $\mathbb{E}\, Z_1 < \infty$.
  - Man kann sich leicht überlegen, dass $\{-X_t,\, t \geq 0\}$ ebenfalls ein Wiener–Prozess ist (vgl. auch Theorem 2.24). Deshalb gilt
$$P(Z_1 > x) \leq P\big(\max_{t\in[0,1]} X_t > x\big) + P\big(\max_{t\in[0,1]} (-X_t) > x\big) = 2\, P\big(\max_{t\in[0,1]} X_t > x\big)\,.$$
  - Mit Hilfe von (23) ergibt sich somit, dass
$$P(Z_1 > x) \leq 2\sqrt{\frac{2}{\pi}} \int_x^\infty \mathrm{e}^{-y^2/2}\,\mathrm{d}y \qquad \forall\, x \geq 0\,.$$
  - Hieraus folgt, dass $\mathbb{E}\, Z_1 = \int_0^\infty P(Z_1 > x)\,\mathrm{d}x < \infty$. □

2.4.6 Weitere Verteilungs– und Pfadeigenschaften

Wir zeigen nun einige *Invarianzeigenschaften* des Wiener–Prozesses in $[0, \infty)$, womit die Tatsache gemeint ist, dass bestimmte Transformationen des Wiener–Prozesses erneut zu einem Wiener–Prozess führen.

Theorem 2.24

- *Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess über einem (vollständigen) Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$.*
- *Die stochastischen Prozesse $\{Y_t^{(i)},\, t \geq 0\}$, $i = 1, \ldots, 4$ mit*

$$Y_t^{(1)} = -X_t\,, \qquad \text{(Symmetrie)} \tag{31}$$

$$Y_t^{(2)} = X_{t+t_0} - X_{t_0} \quad \textit{für ein } t_0 > 0 \qquad \text{(Verschiebung des Nullpunktes)} \tag{32}$$

$$Y_t^{(3)} = \sqrt{c} X_{t/c} \quad \textit{für ein } c > 0, \qquad \text{(Skalierung)} \tag{33}$$

$$Y_t^{(4)} = \begin{cases} t\, X_{1/t} & \textit{für } t > 0, \\ 0 & \textit{für } t = 0, \end{cases} \qquad \text{(Spiegelung bei } t = \infty) \tag{34}$$

sind dann ebenfalls Wiener–Prozesse, wobei $\{Y_t^{(i)},\, t \geq 0\}$, $i = 1, 2, 3$ über dem gleichen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ wie $\{X_t\}$ gegeben sind, während $\{Y_t^{(4)},\, t \geq 0\}$ über einem eingeschränkten Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', P')$ gegeben ist.

Beweis

- Wir zeigen, dass die Prozesse $\{Y_t^{(i)},\, t \geq 0\}$ für $i = 1, \ldots, 4$ die vier Bedingungen in der Definition des Wiener–Prozesses erfüllen.
  - Man kann sich leicht überlegen, dass die Prozesse $\{Y_t^{(i)},\, t \geq 0\}$ unabhängige Zuwächse besitzen mit $Y_{t_2}^{(i)} - Y_{t_1}^{(i)} \sim \mathrm{N}(0, t_2 - t_1)$ für $0 \leq t_1 < t_2$ und $i = 1, \ldots, 4$,
  - Offenbar gilt auch $Y_0^{(i)} = 0$ für $i = 1, \ldots, 4$.
  - Außerdem ist klar, dass die Prozesse $\{Y_t^{(i)},\, t \geq 0\}$ für $i = 1, 2, 3$ stetige Trajektorien besitzen.
- Es bleibt also noch zu zeigen,
  - dass mit Wahrscheinlichkeit 1
  $$\lim_{t \downarrow 0} Y_t^{(4)} = 0 \tag{35}$$
  - bzw. dass es für fast jedes $\omega \in \Omega$ und für jedes $n > 0$ eine Zahl $m > 0$ gibt, so dass
  $$\sup_{t \in (0, 1/m]} |Y_t^{(4)}(\omega)| < \frac{1}{n}\,,$$
  - weil $\{Y_t^{(4)},\, t \geq 0\}$ dann stetige Trajektorien über dem Wahrscheinlichkeitrsraum $(\Omega', \mathcal{F}', P')$ hat, wobei $\Omega' = \{\omega \in \Omega : \lim_{t \downarrow 0} Y_t^{(4)}(\omega) = 0\}$, $\mathcal{F}' = \mathcal{F} \cap \Omega'$ und P' die Einschränkung von P auf $\mathcal{F}'$ bezeichnet.
- Weil die Trajektorien von $\{Y_t^{(4)},\, t \geq 0\}$ für jedes $t > 0$ stetig sind, gilt
$$\sup_{t \in (0, 1/m]} |Y_t^{(4)}| = \sup_{t \in \mathbb{Q} \cap (0, 1/m]} |Y_t^{(4)}|\,,$$
wobei $\mathbb{Q} \subset (0, \infty)$ die (abzählbar unendliche) Menge der positiven rationalen Zahlen ist.

- Hieraus folgt, dass

$$\begin{aligned} P\big(\lim_{t\downarrow 0} Y_t^{(4)} = 0\big) &= P\Big(\bigcap_{n>0}\bigcup_{m>0}\big\{\sup_{t\in(0,1/m]} |Y_t^{(4)}| < 1/n\big\}\Big) \\ &= P\Big(\bigcap_{n>0}\bigcup_{m>0}\sup_{t\in\mathbb{Q}\cap(0,1/m]}\big\{|Y_t^{(4)}| < 1/n\big\}\Big) \\ &= \lim_{n',m',k=\infty} P\Big(\bigcap_{n=1}^{n'}\bigcup_{m=1}^{m'}\sup_{t\in\mathbb{Q}_k\cap(0,1/m]}\big\{|Y_t^{(4)}| < 1/n\big\}\Big) \\ &= \lim_{n',m',k\to\infty} P\Big(\bigcap_{n=1}^{n'}\bigcup_{m=1}^{m'}\sup_{t\in\mathbb{Q}_k\cap(0,1/m]}\big\{|X_t| < 1/n\big\}\Big)\,, \end{aligned}$$

  - wobei $\mathbb{Q}_k \subset \mathbb{Q}$ die (endliche) Menge der „ersten" k rationalen Zahlen ist
  - und sich die letzte Gleichheit aus der Tatsache ergibt, dass die (endlich–dimensionalen) Zufallsvektoren $\{Y_t^{(4)},\, t \in \mathbb{Q}_k\}$ und $\{X_t,\, t \in \mathbb{Q}_k\}$ identisch verteilt sind.
- Weil die Trajektorien des Wiener–Prozesses $\{X_t\, t \in [0,1]\}$ rechtsseitig stetig in $t = 0$ sind, ist der letzte Grenzwert gleich 1. Damit ist (35) bewiesen. □

Beachte Die Gültigkeit von (35) kann auch direkt aus Korollar 2.4 gefolgert werden, denn es gilt

$$\lim_{t\downarrow 0} Y_t^{(4)} = \lim_{t\downarrow 0} t\, X_{1/t} = \lim_{t\to\infty} \frac{X_t}{t} = 0\,.$$

Korollar 2.5 *Sei $\{X_t, \geq 0\}$ ein Wiener–Prozess. Dann gilt*

$$P\big(\sup_{t\geq 0} X_t = \infty\big) = P\big(\inf_{t\geq 0} X_t = -\infty\big) = 1\,. \tag{36}$$

Beweis

- Aus der Skalierungseigenschaft (33) in Theorem 2.24 ergibt sich,
  - dass für beliebige $c, x > 0$

$$P\big(\sup_{t\geq 0} X_t > x\big) = P\big(\sup_{t\geq 0} \sqrt{c} X_{t/c} > x\big) = P\big(\sup_{t\geq 0} X_t > x/\sqrt{c}\big)\,.$$

  - d.h. insbesondere, dass die Wahrscheinlichkeit $P\big(\sup_{t\geq 0} X_t > x\big)$ nicht von x abhängt.
- Hieraus folgt, dass

$$P\big(\{\sup_{t\geq 0} X_t = 0\} \cup \{\sup_{t\geq 0} X_t = \infty\}\big) = 1\,. \tag{37}$$

- Weil der Wiener–Prozess $\{X_t\}$ unabhängige und stationäre Zuwächse hat, gilt

$$\begin{aligned} &P\big(\sup_{t\geq 0} X_t = 0\big) = P\big(\sup_{t\geq 0} X_t \leq 0\big) \leq P\big(X_1 \leq 0, \sup_{t\geq 1} X_t \leq 0\big) = P\big(X_1 \leq 0, \sup_{t\geq 1}(X_t - X_1) \leq -X_1\big) \\ &= \int_{-\infty}^{0} P\big(\sup_{t\geq 1}(X_t - X_1) \leq -x\big)\, P(X_1 \in \mathrm{d}x) = \int_{-\infty}^{0} P\big(\sup_{t\geq 0} X_t \leq -x\big)\, P(X_1 \in \mathrm{d}x) \\ &\overset{(37)}{=} \int_{-\infty}^{0} P\big(\sup_{t\geq 0} X_t = 0\big)\, P(X_1 \in \mathrm{d}x) = P(X_1 \leq 0)\; P\big(\sup_{t\geq 0} X_t = 0\big) \\ &= \frac{1}{2}\, P\big(\sup_{t\geq 0} X_t = 0\big)\,, \end{aligned}$$

  wobei sich die dritte Gleichheit aus der Invarianz–Eigenschaft von $\{X_t\}$ bezüglich der Verschiebung des Nullpunktes ergibt; vgl. die Formel (32).

- Hieraus folgt, dass $P\big(\sup_{t\geq 0} X_t = 0\big) = 0$ und somit $P\big(\sup_{t\geq 0} X_t = \infty\big) = 1$.
- Die Gültigkeit von $P\big(\inf_{t\geq 0} X_t = -\infty\big) = 1$ ergibt sich auf analoge Weise. □

Beachte

- Die Aussage von Korollar 2.5 ist äquivalent mit $P\big(\sup_{t\geq 0} X_t = \infty,\ \inf_{t\geq 0} X_t = -\infty\big) = 1$.
- Hieraus folgt insbesondere, dass fast alle Pfade des Wiener–Prozesses $\{X_t,\ t \geq 0\}$ in dem unbeschränkten Intervall $[0, \infty)$ unendlich oft zwischen positiven und negativen Werten oszillieren.
- Wir zeigen nun, dass die Pfade des Wiener–Prozesses $\{X_t\}$
  - auch in beschränkten Intervallen „wild" oszillieren,
  - denn es stellt sich heraus, dass sie zwar stetige, jedoch mit Wahrscheinlichkeit 1 nirgendwo differenzierbare Funktionen sind.

Theorem 2.25 *Sei $\{X_t, \geq 0\}$ ein Wiener–Prozess. Dann gilt*

$$P(\omega \in \Omega :\ X.(\omega) \textit{ ist nirgendwo differenzierbar in } [0,\infty)\,) = 1\,. \tag{38}$$

Beweis

- Wegen der Identität

$$\begin{aligned} &\{\omega \in \Omega :\ X.(\omega) \text{ ist nirgendwo differenzierbar in } [0,\infty)\} \\ &= \bigcap_{i\geq 0} \{\omega \in \Omega :\ X.(\omega) \text{ ist nirgendwo differenzierbar in } [i, i+1]\} \end{aligned}$$

  genügt es zu zeigen, dass $P(\omega \in \Omega :\ X.(\omega)$ ist nirgendwo differenzierbar in $[0,1]\,) = 1$ bzw., äquivalent hierzu, dass

$$P(\omega \in \Omega :\ X.(\omega) \text{ ist differenzierbar für ein } t_0 = t_0(\omega) \in [0,1]\,) = 0\,. \tag{39}$$

- Um die Gültigkeit von (39) zu beweisen, setzen wir für beliebige natürliche Zahlen $n, m \geq 1$

$$A_{nm} = \big\{\omega \in \Omega :\ \text{es gibt ein } t_0 = t_0(\omega) \in [0,1] \text{ mit } |X_{t_0(\omega)+h}(\omega) - X_{t_0(\omega)}(\omega)| \leq mh,\ \forall\, h \in [0, 4/n]\big\}\,.$$

  - Dann gilt

$$\{\omega \in \Omega :\ X.(\omega) \text{ ist differenzierbar für ein } t_0 = t_0(\omega)\} \subset \bigcup_{m\geq 1} \bigcup_{n\geq 1} A_{nm}\,,$$

  - und es genügt somit zu zeigen, dass

$$P(A_{nm}) = 0 \qquad \forall\, n, m \geq 1\,. \tag{40}$$

- Sei $k_0(\omega) = \min\big\{k \in \{0,1,\ldots\} :\ k/n \geq t_0(\omega)\big\}$. Dann gilt für jedes $\omega \in A_{nm}$ und für $j = 0,1,2$

$$\begin{aligned} &\big|X_{(k_0(\omega)+j+1)/n}(\omega) - X_{(k_0(\omega)+j)/n}(\omega)\big| \\ &\leq \big|X_{(k_0(\omega)+j+1)/n}(\omega) - X_{t_0(\omega)}(\omega)\big| + \big|X_{(k_0(\omega)+j)/n}(\omega) - X_{t_0(\omega)}(\omega)\big| \\ &\leq \frac{8m}{n} \end{aligned}$$

- Mit der Schreibweise $\Delta_n(k) = X_{(k+1)/n} - X_{k/n}$ gilt also

$$A_{nm} \subset \bigcup_{k=0}^{n} \bigcap_{j=0}^{2} \{|\Delta_n(k+j)| \leq \frac{8m}{n}\}$$

und somit

$$\begin{aligned} P(A_{nm}) &\leq P\Big(\bigcup_{k=0}^{n}\bigcap_{j=0}^{2}\{|\Delta_n(k+j)|\leq \frac{8m}{n}\}\Big) \\ &\leq \sum_{k=0}^{n} P\Big(\bigcap_{j=0}^{2}\{|\Delta_n(k+j)|\leq \frac{8m}{n}\}\Big) \\ &= (n+1)P\Big(|\Delta_n(0)|\leq \frac{8m}{n}\Big)^3 \\ &\leq (n+1)\Big(\frac{16\,m}{\sqrt{2\pi n}}\Big)^3 \longrightarrow 0 \qquad \text{für } n\to\infty\,. \end{aligned}$$

- Weil $A_{nm}\subset A_{n+1,m}$ für jedes $n\geq 1$, ergibt sich hieraus die Gültigkeit von (40). □

Korollar 2.6 *Mit Wahrscheinlichkeit* 1 *gilt*

$$P\big(\sup_{n\geq 1}\ \sup_{0\leq t_0<\ldots<t_n\leq 1}\ \sum_{i=1}^{n}|X_{t_i}-X_{t_{i-1}}|=\infty\big)=1\,, \tag{41}$$

d.h., fast alle Trajektorien des Wiener–Prozesses $\{X_t,\ t\in[0,1]\}$ *sind Funktionen mit unbeschränkter Variation.*

Beweis Weil jede stetige Funktion $g:[0,1]\to\mathbb{R}$ mit beschränkter Variation fast überall differenzierbar ist, ergibt sich die Behauptung unmittelbar aus Theorem 2.25. □

Beachte

- Ein *direkter* Beweis von Korollar 2.6 kann wie folgt geführt werden: Es genügt zu zeigen, dass mit Wahrscheinlichkeit 1
$$\lim_{n\to\infty}\sum_{i=1}^{2^n}\big|X_{it/2^n}-X_{(i-1)t/2^n}\big|=\infty \qquad \forall\, t\in(0,1]\,. \tag{42}$$
- Um (42) zu beweisen, setzen wir
$$Z_n=\sum_{i=1}^{2^n}\big(X_{it/2^n}-X_{(i-1)t/2^n}\big)^2-t\,.$$
- Dann gilt $\mathbb{E}\,Z_n=0$, und außerdem kann man leicht zeigen, dass $\mathbb{E}\,(Z_n^2)=t^2 2^{-n+1}$.
- Aus der Tschebyschew–Ungleichung (vgl. Abschnitt WR-4.4.3) ergibt sich, dass für jedes $\varepsilon>0$
$$P(|Z_n|\geq\varepsilon)\leq\varepsilon^{-2}\mathbb{E}\,(Z_n^2)=(t/\varepsilon)^2 2^{-n+1}$$
und somit $\sum_{n=1}^{\infty}P(|Z_n|\geq\varepsilon)<\infty$.
- Aus dem Lemma von Borel–Cantelli (vgl. Korollar WR-2.3) ergibt sich nun, dass $\lim_{n\to\infty}Z_n=0$ mit Wahrscheinlichkeit 1.
- Hieraus folgt, dass mit Wahrscheinlichkeit 1
$$0<t\leq\liminf_{n\to\infty}\max_{1\leq k\leq 2^n}\big|X_{kt/2^n}-X_{(k-1)t/2^n}\big|\ \sum_{i=1}^{2^n}\big|X_{it/2^n}-X_{(i-1)t/2^n}\big|\,.$$
- Damit ist die Gültigkeit von (42) bewiesen, weil der Wiener–Prozess $\{X_t\}$ stetige Trajektorien hat und weil deshalb
$$\lim_{n\to\infty}\max_{1\leq k\leq 2^n}\big|X_{kt/2^n}-X_{(k-1)t/2^n}\big|=0\,.$$

3 Lévy–Prozesse und Martingale

3.1 Lévy–Prozesse

In Abschnitt 2.2.2 hatten wir zusammengesetzte Poisson–Prozesse betrachtet und gezeigt, dass sie unabhängige und stationäre Zuwächse besitzen (vgl. Theorem 2.10). Der Wiener–Prozess besitzt ebenfalls diese beiden Eigenschaften, was sich unmittelbar aus seiner Definition ergibt (vgl. Abschnitt 2.4).

Wir betrachten nun eine allgemeine Klasse von stochastischen Prozessen mit unabhängigen und stationären Zuwächsen, die sowohl zusammengesetzte Poisson–Prozesse als auch den Wiener–Prozess als Spezialfälle umfasst.

Definition Ein stochastischer Prozess $\{X_t,\, t \geq 0\}$ über einem (im allgemeinen nicht näher spezifizierten) Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ heißt *Lévy–Prozess*, wenn

- $X_0 = 0$,
- $\{X_t\}$ unabhängige und stationäre Zuwächse hat,
- $\{X_t\}$ *stochastisch stetig* ist, d.h., für beliebige $\varepsilon > 0$ und $t_0 \geq 0$ gilt $\lim_{t\to t_0} P(|X_t - X_{t_0}| > \varepsilon) = 0$.

Beachte

- Man kann sich leicht überlegen, dass für zusammengesetzte Poisson–Prozesse die dritte Bedingung ebenfalls erfüllt ist. Denn es gilt dann für jedes $\varepsilon > 0$ und für $t \to t_0$

$$P(|X_t - X_{t_0}| > \varepsilon) \leq P(|X_t - X_{t_0}| > 0) \leq 1 - \mathrm{e}^{-\lambda|t-t_0|} \longrightarrow 0\,,$$

  wobei $\lambda > 0$ die Intensität der Sprungzeitpunkte bezeichnet.
- Darüber hinaus gilt auch im Fall des Wiener–Prozesses für jedes $\varepsilon > 0$ und für $t \to t_0$

$$P(|X_t - X_{t_0}| > \varepsilon) = \sqrt{\frac{2}{\pi|t-t_0|}} \int_\varepsilon^\infty \exp\Big(-\frac{y^2}{2|t-t_0|}\Big)\,\mathrm{d}y = \sqrt{\frac{2}{\pi}} \int_{\varepsilon/\sqrt{|t-t_0|}}^\infty \mathrm{e}^{-y^2/2}\,\mathrm{d}y \longrightarrow 0\,.$$

3.1.1 Unbegrenzte Teilbarkeit

Wir zeigen zunächst, dass für jeden Lévy–Prozess $\{X_t\}$ und für jedes $t \geq 0$ die Zufallsvariable X_t unbegrenzt teilbar ist, und geben dann in Abschnitt 3.1.2 eine Darstellungsformel für die charakteristische Funktion von X_t an.

Definition Sei $X : \Omega \to \mathbb{R}$ eine beliebige Zufallsvariable über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Man sagt, dass X bzw. die Verteilung P_X von X *unbegrenzt teilbar* ist, wenn es für jedes $n \geq 1$ unabhängige und identisch verteilte Zufallsvariablen $Y_1^{(n)}, \ldots, Y_n^{(n)} : \Omega \to \mathbb{R}$ gibt mit $X \stackrel{\mathrm{d}}{=} Y_1^{(n)} + \ldots + Y_n^{(n)}$.

Theorem 3.1 *Sei $\{X_t,\, t \geq 0\}$ ein Lévy–Prozess. Dann ist die Zufallsvariable X_t für jedes $t \geq 0$ unbegrenzt teilbar.*

Beweis

- Für beliebige $t \geq 0$ und $n \geq 1$ gilt offenbar, dass

$$X_t = X_{t/n} + (X_{2t/n} - X_{t/n}) + \ldots + (X_{nt/n} - X_{(n-1)t/n})\,.$$

- Weil $\{X_t\}$ unabhängige und stationäre Zuwächse hat, sind die Summanden auf der rechten Seite der letzten Gleichheit unabhängige und identische verteilte Zufallsvariablen. □

Das folgende Lemma enthält ein einfaches (jedoch nützliches) Hilfsmittel zur Untersuchung von unbegrenzt teilbaren Zufallsvariablen.

Lemma 3.1 *Die Zufallsvariable $X : \Omega \to \mathbb{R}$ ist genau dann unbegrenzt teilbar, wenn sich die charakteristische Funktion φ_X von X für jedes $n \geq 1$ darstellen lässt in der Form*

$$\varphi_X(s) = \big(\varphi_n(s)\big)^n \qquad \forall\, s \in \mathbb{R}, \tag{1}$$

wobei φ_n die charakteristische Funktion einer Zufallsvariablen ist.

Beweis

- Die Notwendigkeit der Bedingung (1) ergibt sich unmittelbar
  - aus der Definition der unbegrenzten Teilbarkeit und aus der Tatsache, dass
  - die charakteristische Funktion der Summe von unabhängigen Zufallsvariablen gleich dem Produkt der charakteristischen Funktionen der Summanden ist (vgl. Theorem WR-5.18).
- Es gelte nun umgekehrt (1), und $Y_1^{(n)}, \ldots, Y_n^{(n)}$ seien unabhängige Zufallsvariablen mit der charakteristischen Funktion φ_n.
  - Dann ergibt sich durch die erneute Anwendung von Theorem WR-5.18, dass $\big(\varphi_n(s)\big)^n$ die charakteristische Funktion der Summe $Y_1^{(n)} + \ldots + Y_n^{(n)}$ ist.
  - Wegen (1) stimmen also die charakteristischen Funktionen von X und $Y_1^{(n)} + \ldots + Y_n^{(n)}$ überein.
  - Aus dem Eindeutigkeitssatz für charakteristische Funktionen (vgl. Korollar WR-5.5) ergibt sich nun, dass auch die Verteilungen von X und $Y_1^{(n)} + \ldots + Y_n^{(n)}$ übereinstimen. □

Wir benötigen noch einen Stetigkeitssatz für charakteristische Funktionen, der eine (teilweise) Verschärfung von Theorem WR-5.20 ist und den wir hier ohne Beweis angeben.

Lemma 3.2

- *Sei $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ eine beliebige Folge von Zufallsvariablen, und seien $\varphi_{X_1}, \varphi_{X_2}, \ldots : \mathbb{R} \to \mathbb{C}$ ihre charakteristischen Funktionen.*
- *Falls es eine Funktion $\varphi : \mathbb{R} \to \mathbb{C}$ gibt, so dass $\varphi(s)$ stetig im Punkt $s = 0$ ist und $\lim_{n\to\infty} \varphi_n(s) = \varphi(s)$ für jedes $s \in \mathbb{R}$ gilt,*
  - *dann ist φ die charakteristische Funktion einer Zufallsvariablen X,*
  - *und es gilt $X_n \xrightarrow{\mathrm{d}} X$.*

Außerdem spielt der Begriff des Lévy–Maßes eine wichtige Rolle bei der Untersuchung von unbegrenzt teilbaren Verteilungen.

Definition Sei ν ein Maß über dem Messraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Man sagt, dass ν ein *Lévy–Maß* ist, wenn $\nu(\{0\}) = 0$ und

$$\int_{\mathbb{R}} \min\{y^2, 1\}\, \nu(\mathrm{d}y) < \infty\,. \tag{2}$$

Beachte

- Man kann sich leicht überlegen, dass jedes Lévy–Maß ν ein σ–endliches Maß ist und dass

$$\nu((-\varepsilon, \varepsilon)^c) < \infty \qquad \forall\, \varepsilon > 0\,, \tag{3}$$

  wobei $(-\varepsilon, \varepsilon)^c = \mathbb{R} \setminus (-\varepsilon, \varepsilon)$.

- Insbesondere ist jedes endliche Maß ν über $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ein Lévy–Maß, falls $\nu(\{0\}) = 0$.
- Eine zu (2) äquivalente Bedingung ist

$$\int_{\mathbb{R}} \frac{y^2}{1+y^2}\, \nu(\mathrm{d}y) < \infty\,, \tag{4}$$

denn es gilt

$$\frac{y^2}{1+y^2} \le \min\{y^2, 1\} \le 2\, \frac{y^2}{1+y^2} \qquad \forall\, y \in \mathbb{R}\,.$$

Der folgende Ansatz zur Konstruktion charakteristischer Funktionen von unbegrenzt teilbaren Verteilungen wird in der Literatur die *Lévy–Chintschin–Formel* genannt.

Theorem 3.2 *Seien $a \in \mathbb{R}$ und $b \ge 0$ beliebige Konstanten, und sei ν ein beliebiges Lévy–Maß. Dann ist durch die Funktion $\varphi : \mathbb{R} \to \mathbb{C}$ mit*

$$\varphi(s) = \exp\Big(\mathrm{i}as - \frac{bs^2}{2} + \int_{\mathbb{R}} \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{1}_{(-1,1)}(y)\big)\, \nu(\mathrm{d}y)\Big) \qquad \forall\, s \in \mathbb{R} \tag{5}$$

die charakteristische Funktion einer unbegrenzt teilbaren Zufallsvariablen gegeben.

Beweis

- Weil es für jedes $s \in \mathbb{R}$ eine Konstante $c < \infty$ gibt, so dass

$$|\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy| \le cy^2 \qquad \forall\, y \in (-1,1)\,, \tag{6}$$

folgt aus (2) und (3), dass die in (5) gegebene Funktion $\varphi : \mathbb{R} \to \mathbb{C}$ wohldefiniert ist.
- Sei nun $\{c_n\}$ eine beliebige Zahlenfolge mit $c_n > c_{n+1} > 0$ und $\lim_{n\to\infty} c_n = 0$.
  - Dann ist die Funktion $\varphi_n : \mathbb{R} \to \mathbb{C}$ mit

$$\begin{aligned}\varphi_n(s) &= \exp\Big(\mathrm{i}s\big(a - \int_{[-c_n,c_n]^c\cap(-1,1)} y\,\nu(\mathrm{d}y)\big) - \frac{bs^2}{2} + \int_{[-c_n,c_n]^c} \big(\mathrm{e}^{\mathrm{i}sy} - 1\big)\,\nu(\mathrm{d}y)\Big) \\ &= \exp\Big(\mathrm{i}s\big(a - \int_{[-c_n,c_n]^c\cap(-1,1)} y\,\nu(\mathrm{d}y)\big) - \frac{bs^2}{2}\Big)\, \exp\Big(\int_{[-c_n,c_n]^c} \big(\mathrm{e}^{\mathrm{i}sy} - 1\big)\,\nu(\mathrm{d}y)\Big) \end{aligned} \tag{7}$$

    die charakteristische Funktion der Summe $Z_1^{(n)} + Z_2^{(n)}$ von zwei unabhängigen Zufallsvariablen $Z_1^{(n)}$ und $Z_2^{(n)}$,
  - denn der erste Faktor in (7) ist die charakteristische Funktion der Normalverteilung mit Erwartungswert $a - \int_{[-c_n,c_n]^c\cap(-1,1)} y\,\nu(\mathrm{d}y)$ und Varianz b,
  - während der zweite Faktor in (7) die charakteristische Funktion der zusammengesetzten Poisson–Verteilung mit den Charakteristiken $\lambda = \nu([-c_n, c_n]^c)$ und $P_U = \nu(\cdot \cap [-c_n, c_n]^c)/\nu([-c_n, c_n]^c)$ ist; vgl. die Teilaussage (b) von Theorem 2.10.
- Außerdem gilt

$$\lim_{n\to\infty} \varphi_n(s) = \varphi(s) \qquad \forall\, s \in \mathbb{R}\,,$$

wobei man sich leicht überlegen kann, dass die in (5) gegebene Funktion $\varphi : \mathbb{R} \to \mathbb{C}$ stetig im Punkt $s = 0$ ist.
  - Denn für die Funktion $\psi : \mathbb{R} \to \mathbb{C}$ im Exponent von (5) mit

$$\psi(s) = \int_{\mathbb{R}} \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{1}_{(-1,1)}(y)\big)\, \nu(\mathrm{d}y) \qquad \forall\, s \in \mathbb{R}$$

    gilt wegen (6), dass

$$|\psi(s)| \le c \int_{(-1,1)} y^2\, \nu(dy) + \int_{(-1,1)^c} \big|\mathrm{e}^{\mathrm{i}sy} - 1\big|\, \nu(\mathrm{d}y).$$

- Hieraus und aus (3) folgt nun mit Hilfe des Satzes von Lebesgue über die majorisierte Konvergenz, dass $\lim_{s\to 0}\psi(s) = 0$.

- Aus Lemma 3.2 ergibt sich somit, dass die in (5) gegebene Funktion $\varphi : \mathbb{R} \to \mathbb{C}$ die charakteristische Funktion einer Zufallsvariablen ist.
- Die unbegrenzte Teilbarkeit dieser Zufallsvariablen folgt schließlich aus Lemma 3.1 und aus der Tatsache, dass für jedes $n \geq 1$ mit ν auch $\nu(\cdot)/n$ ein Lévy–Maß ist und dass

$$\varphi(s) = \Big(\exp\Big(\mathrm{i}\,\frac{a}{n}\,s - \frac{\frac{b}{n}\,s^2}{2} + \int_{\mathbb{R}} \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{I}_{(-1,1)}(y)\big)\,(\nu/n)(\mathrm{d}y)\Big)\Big)^n \qquad \forall\, s \in \mathbb{R}. \qquad \square$$

Beachte

- Das Tripel (a, b, ν), das in der Lévy–Chintschin–Formel (5) auftritt, wird *Lévy–Charakteristik* der zugehörigen unbegrenzt teilbaren Verteilung genannt.
- Die Abbildung $\eta : \mathbb{R} \to \mathbb{C}$ mit

$$\eta(s) = \mathrm{i}as - \frac{bs^2}{2} + \int_{\mathbb{R}} \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{I}_{(-1,1)}(y)\big)\,\nu(\mathrm{d}y) \qquad \forall\, s \in \mathbb{R} \tag{8}$$

  im Exponent von (5) heißt *Lévy–Exponent* dieser unbegrenzt teilbaren Verteilung.

3.1.2 Lévy–Chintschin–Darstellung

In diesem Abschnitt zeigen wir, dass für jeden Lévy–Prozess $\{X_t\}$ und für jedes $t \geq 0$ die charakteristische Funktion der unbegrenzt teilbaren Zufallsvariablen X_t durch die Lévy–Chintschin–Formel (5) dargestellt werden kann.

Dabei zeigen wir zunächst, dass die charakteristische Funktion von X_t auf einfache Weise durch die charakteristische Funktion von X_1 ausgedrückt werden kann. Hierfür benötigen wir den folgenden Hilfssatz.

Lemma 3.3 *Sei $\{X_t,\, t \geq 0\}$ ein stochastisch stetiger Prozess, d.h. für beliebige $x > 0$ und $t_0 \geq 0$ gelte*

$$\lim_{t\to t_0} P(|X_t - X_{t_0}| > x) = 0\,.$$

Dann ist für jedes $s \in \mathbb{R}$ durch $t \to \varphi_{X_t}(s)$ eine stetige Abbildung von $[0, \infty)$ nach $\mathbb{C}$ gegeben.

Beweis

- Für jedes $s \in \mathbb{R}$ ist die Abbildung $y \to \mathrm{e}^{\mathrm{i}sy}$ offenbar stetig im Punkt $y = 0$, d.h., für jedes $\varepsilon > 0$ gibt es ein $\delta_1 > 0$, so dass

$$\sup_{y\in(-\delta_1,\delta_1)} |\mathrm{e}^{\mathrm{i}sy} - 1| < \frac{\varepsilon}{2}\,.$$

- Weil $\{X_t\}$ stochastisch stetig ist, gibt es außerdem ein $\delta_2 > 0$, so dass für jedes $t_0 \geq 0$

$$\sup_{t\geq 0,\, |t-t_0|<\delta_2} P(|X_t - X_{t_0}| > \delta_1) < \frac{\varepsilon}{4}\,.$$

- Hieraus ergibt sich, dass für beliebige $s \in \mathbb{R}$ und $t \geq 0$ mit $|t - t_0| < \delta_2$

$$\begin{aligned}
|\varphi_{X_t}(s) - \varphi_{X_{t_0}}(s)| &= \left| \int_\Omega \mathrm{e}^{\mathrm{i}sX_{t_0}(\omega)} \left(\mathrm{e}^{\mathrm{i}s(X_t - X_{t_0})(\omega)} - 1\right) P(\mathrm{d}\omega) \right| \\
&\leq \int_{\mathbb{R}} \left|\mathrm{e}^{\mathrm{i}sy} - 1\right| P_{X_t - X_{t_0}}(\mathrm{d}y) \\
&= \int_{(-\delta_1,\delta_1)} \left|\mathrm{e}^{\mathrm{i}sy} - 1\right| P_{X_t - X_{t_0}}(\mathrm{d}y) + \int_{(-\delta_1,\delta_1)^c} \left|\mathrm{e}^{\mathrm{i}sy} - 1\right| P_{X_t - X_{t_0}}(\mathrm{d}y) \\
&\leq \sup_{y \in (-\delta_1,\delta_1)} |\mathrm{e}^{\mathrm{i}sy} - 1| + 2\, P(|X_t - X_{t_0}| > \delta_1) \\
&\leq \varepsilon\,.
\end{aligned}$$

- Damit ist die Behauptung bewiesen. □

Theorem 3.3 *Sei $\{X_t,\, t \geq 0\}$ ein Lévy–Prozess. Dann ist für jedes $t \geq 0$ die charakteristische Funktion φ_{X_t} von X_t gegeben durch*

$$\varphi_{X_t}(s) = \mathrm{e}^{t\,\eta(s)} \qquad \forall\, s \in \mathbb{R}\,, \tag{9}$$

wobei $\eta : \mathbb{R} \to \mathbb{C}$ eine stetige Funktion ist. Insbesondere gilt somit $\varphi_{X_t}(s) = (\varphi_{X_1}(s))^t$ für beliebige $s \in \mathbb{R}$, $t \geq 0$.

Beweis

- Weil $\{X_t\}$ unabhängige und stationäre Zuwächse hat, gilt für beliebige $s \in \mathbb{R}$ und $t, t' \geq 0$

$$\varphi_{X_{t+t'}}(s) = \mathbb{E}\,\mathrm{e}^{\mathrm{i}sX_{t+t'}} = \mathbb{E}\left(\mathrm{e}^{\mathrm{i}sX_t}\,\mathrm{e}^{\mathrm{i}s(X_{t+t'} - X_t)}\right) = \mathbb{E}\,\mathrm{e}^{\mathrm{i}sX_t}\,\mathbb{E}\,\mathrm{e}^{\mathrm{i}sX_{t'}} = \varphi_{X_t}(s)\varphi_{X_{t'}}(s)\,.$$

- Für jedes $s \in \mathbb{R}$ gilt also für die Funktion $g_s : [0,\infty) \to \mathbb{C}$ mit $g_s(t) = \varphi_{X_t}(s)$, dass

$$g_s(t + t') = g_s(t) g_s(t') \qquad \forall\, t, t' \geq 0\,. \tag{10}$$

- Weil $X_0 = 0$, gilt außerdem für jedes $s \in \mathbb{R}$

$$g_s(0) = 1\,, \tag{11}$$

  und in Lemma 3.3 hatten wir gezeigt, dass die Funktion $g_s : [0,\infty) \to \mathbb{C}$ für jedes $s \in \mathbb{R}$ stetig ist.
- Man kann sich leicht überlegen, dass es für jede Familie von stetigen Funktionen $g_s : [0,\infty) \to \mathbb{C}$, die den Bedingungen (10) und (11) genügen, eine Funktion $\eta : \mathbb{R} \to \mathbb{C}$ gibt, so dass

$$g_s(t) = \mathrm{e}^{t\,\eta(s)} \qquad \forall\, s \in \mathbb{R},\ t \geq 0\,. \tag{12}$$

- Es genügt nun, $t = 1$ zu setzen, um zu erkennen, dass $\varphi_{X_1}(s) = \mathrm{e}^{\eta(s)}$ für jedes $s \in \mathbb{R}$ gilt und dass somit $\eta : \mathbb{R} \to \mathbb{C}$ eine stetige Funktion ist. □

Wir zeigen nun, dass für jeden Lévy–Prozess $\{X_t\}$ die charakteristische Funktion von X_1 durch die Lévy–Chintschin–Formel (5) dargestellt werden kann. Hierfür benötigen wir den folgenden Hilfssatz zur Charakterisierung der relativen Kompaktheit von Familien (gleichmäßig beschränkter) endlicher Maße.

Lemma 3.4 *Sei $\mu_1, \mu_2, \ldots$ eine Folge von endlichen Maßen, die über der Borel–σ–Algebra $\mathcal{B}(\mathbb{R})$ definiert sind. Falls $\sup_{n\geq 1} \mu_n(\mathbb{R}) < c$ für ein $c < \infty$ und falls es für jedes $\varepsilon > 0$ eine kompakte Menge $B_\varepsilon \in \mathcal{B}(\mathbb{R})$ gibt, so dass*

$$\sup_{n \geq 1} \mu_n(B_\varepsilon^c) \leq \varepsilon\,, \tag{13}$$

dann gibt es eine Teilfolge $\mu_{n_1}, \mu_{n_2}, \ldots$ und ein endliches Maß μ über $\mathcal{B}(\mathbb{R})$, so dass für jede stetige und beschränkte Funktion $f : \mathbb{R} \to \mathbb{C}$

$$\lim_{k\to\infty} \int_{\mathbb{R}} f(y)\, \mu_{n_k}(\mathrm{d}y) = \int_{\mathbb{R}} f(y)\, \mu(\mathrm{d}y)\,. \tag{14}$$

Der *Beweis* von Lemma 3.4 geht über den Rahmen dieser Vorlesung hinaus und wird deshalb weggelassen; er kann beispielsweise im Buch von A.V. Skorokhod (2005), S. 122–123 nachgelesen werden.

Theorem 3.4 *Sei $\{X_t,\, t \geq 0\}$ ein Lévy–Prozess. Dann gibt es Konstanten $a \in \mathbb{R}$, $b \geq 0$ und ein Lévy–Maß ν, so dass*

$$\varphi_{X_1}(s) = \exp\Big(\mathrm{i}as - \frac{bs^2}{2} + \int_{\mathbb{R}} \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{I}_{(-1,1)}(y)\big)\,\nu(\mathrm{d}y)\Big) \qquad \forall\, s \in \mathbb{R}\,. \tag{15}$$

Beweis

- Aus Theorem 3.3 ergibt sich, dass für jede Nullfolge $t_1, t_2, \ldots$ positiver Zahlen

$$\eta(s) = \lim_{n\to\infty} \frac{\mathrm{e}^{t_n\,\eta(s)} - 1}{t_n} = \lim_{n\to\infty} \frac{\varphi_{X_{t_n}}(s) - 1}{t_n}\,. \tag{16}$$

- Weil $\eta : \mathbb{R} \to \mathbb{C}$ eine stetige Funktion ist, ergibt sich durch Taylor–Reihenentwicklung der Funktion $\mathrm{e}^{t\,\eta(s)}$ nach t, dass für jedes $s_0 < \infty$ die Konvergenz in (16) gleichmäßig in $s \in [-s_0, s_0]$ erfolgt.
- Wir setzen nun $t_n = 1/n$ und bezeichnen die Verteilung von $X_{1/n}$ mit P_n. Mit dieser Schreibweise gilt

$$\lim_{n\to\infty} n \int_{\mathbb{R}} \big(\mathrm{e}^{\mathrm{i}sy} - 1\big)\, P_n(\mathrm{d}y) = \eta(s) \tag{17}$$

  gleichmäßig in $s \in [-s_0, s_0]$.
  - Wenn auf beiden Seiten dieser Gleichung über $s \in [-s_0, s_0]$ integriert wird,
  - dann ergibt sich aus dem Satz von Lebesgue über die majorisierte Konvergenz (und nach Vertauschung der Integrationsreihenfolge auf der linken Seite), dass für jedes $s_0 < \infty$

$$\lim_{n\to\infty} n \int_{\mathbb{R}} \Big(1 - \frac{\sin(s_0 y)}{s_0 y}\Big)\, P_n(\mathrm{d}y) = -\,\frac{1}{2s_0} \int_{-s_0}^{s_0} \eta(s)\,\mathrm{d}s\,. \tag{18}$$

- Weil $\eta : \mathbb{R} \to \mathbb{C}$ eine stetige Funktion ist mit $\eta(0) = 0$, gibt es für jedes $\varepsilon > 0$ ein $s_0 > 0$, so dass

$$\Big|\, \frac{1}{2s_0} \int_{-s_0}^{s_0} \eta(s)\,\mathrm{d}s \Big| < \varepsilon\,.$$

- Weil außerdem $1 - \sin(s_0 y)/(s_0 y) \geq 1/2$ für $|s_0 y| \geq 2$, ergibt sich somit aus (18),
  - dass es für jedes $\varepsilon > 0$ ein $s_0 > 0$ gibt, so dass

$$\limsup_{n\to\infty} \frac{n}{2} \int_{\{y:\,|y|\geq 2/s_0\}} P_n(\mathrm{d}y) \leq \varepsilon\,,$$

  - bzw. dass es für jedes $\varepsilon > 0$ ein $s_0 > 0$ und ein $n_0 > 0$ gibt, so dass

$$n \int_{\{y:\,|y|\geq 2/s_0\}} P_n(\mathrm{d}y) \leq 4\,\varepsilon \qquad \forall\, n \geq n_0\,. \tag{19}$$

- Wir verkleinern nun $s_0 > 0$ weiter (falls erforderlich), so dass die Ungleichung in (19) auch für jedes $n \in \{1, \ldots, n_0\}$ gilt.
- Insgesamt haben wir somit gezeigt, dass es für jedes $\varepsilon > 0$ ein $s_0 > 0$ gibt, so dass

$$\sup_{n\geq 1}\, n \int_{\{y:\,|y|\geq 2/s_0\}} P_n(\mathrm{d}y) \leq 4\,\varepsilon\,. \tag{20}$$

- Weil es ein $c < \infty$ gibt, so dass

$$\frac{y^2}{1+y^2} \le c\,(1 - \frac{\sin y}{y}) \qquad \forall\, y \neq 0\,,$$

ergibt sich aus (18), dass es eine Konstante $c' < \infty$ gibt, so dass

$$\sup_{n\ge 1} n \int_{\mathbb{R}} \frac{y^2}{1+y^2}\, P_n(\mathrm{d}y) \le c'\,. \tag{21}$$

- Für jedes $n \ge 1$ sei das (endliche) Maß $\mu_n : \mathcal{B}(\mathbb{R}) \to [0,\infty)$ gegeben durch

$$\mu_n(B) = n \int_B \frac{y^2}{1+y^2}\, P_n(\mathrm{d}y) \qquad \forall\, B \in \mathcal{B}(\mathbb{R})\,.$$

  - Wegen (21) sind die Maße $\{\mu_n,\, n \ge 1\}$ gleichmäßig beschränkt, und wegen

$$\frac{y^2}{1+y^2} \le 1 \qquad \forall\, y \in \mathbb{R}$$

    ergibt sich aus (20), dass die Folge $\{\mu_n,\, n \ge 1\}$ relativ kompakt ist, d.h. der Bedingung (13) genügt.
  - Aus Lemma 3.4 folgt also, dass es eine Teilfolge $\mu_{n_1}, \mu_{n_2}, \ldots$ und ein endliches Maß μ über $\mathcal{B}(\mathbb{R})$ gibt, so dass für jede stetige und beschränkte Funktion $f : \mathbb{R} \to \mathbb{C}$

$$\lim_{k\to\infty} \int_{\mathbb{R}} f(y)\, \mu_{n_k}(\mathrm{d}y) = \int_{\mathbb{R}} f(y)\, \mu(\mathrm{d}y)\,. \tag{22}$$

- Weil für jedes $s \in \mathbb{R}$ die Funktion $f_s : \mathbb{R} \to \mathbb{C}$ mit

$$f_s(y) = \begin{cases} \left(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}s\sin y\right) \dfrac{1+y^2}{y^2} & \text{für } y \neq 0, \\ -\dfrac{s^2}{2} & \text{für } y = 0 \end{cases}$$

stetig und beschränkt ist, ergibt sich aus (17) und (22), dass

$$\begin{aligned} \eta(s) &= \lim_{n\to\infty} n \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}sy} - 1\right) P_n(\mathrm{d}y) = \lim_{n\to\infty} \Big(\int_{\mathbb{R}} f_s(y)\, \mu_n(\mathrm{d}y) + \mathrm{i}sn \int_{\mathbb{R}} \sin y P_n(\mathrm{d}y) \Big) \\ &= \lim_{k\to\infty} \Big(\int_{\mathbb{R}} f_s(y)\, \mu_{n_k}(\mathrm{d}y) + \mathrm{i}sn_k \int_{\mathbb{R}} \sin y P_{n_k}(\mathrm{d}y) \Big) \\ &= \int_{\mathbb{R}} f_s(y)\, \mu(\mathrm{d}y) + \mathrm{i}s \lim_{k\to\infty} n_k \int_{\mathbb{R}} \sin y P_{n_k}(\mathrm{d}y)\,, \end{aligned}$$

wobei sich die Existenz und Endlichkeit des letzten Grenzwertes aus der Tatsache ergibt, das alle übrigen Grenzwerte in dieser Gleichungskette existieren und endlich sind.

- Hieraus folgt, dass

$$\eta(s) = \mathrm{i}a's - \frac{bs^2}{2} + \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}s\sin y\right) \nu(\mathrm{d}y) \qquad \forall\, s \in \mathbb{R}, \tag{23}$$

wobei

$$a' = \lim_{k\to\infty} n_k \int_{\mathbb{R}} \sin y P_{n_k}(\mathrm{d}y)\,, \qquad b = \mu(\{0\})$$

und das Lévy–Maß $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ gegeben ist durch

$$\nu(\mathrm{d}y) = \begin{cases} \dfrac{1+y^2}{y^2}\, \mu(\mathrm{d}y) & \text{für } y \neq 0, \\ 0 & \text{für } y = 0. \end{cases}$$

- Außerdem gilt
$$\int_{\mathbb{R}} |y \mathbb{1}_{(-1,1)}(y) - \sin y| \, \nu(\mathrm{d}y) < \infty \,,$$
weil es eine Konstante $c'' < \infty$ gibt, so dass
$$|y \mathbb{1}_{(-1,1)}(y) - \sin y| \, \frac{1+y^2}{y^2} < c'' \qquad \forall\, y \neq 0 \,.$$
- Aus (23) ergibt sich somit, dass
$$\eta(s) = \mathrm{i}as - \frac{bs^2}{2} + \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy \mathbb{1}_{(-1,1)}(y) \right) \nu(\mathrm{d}y) \qquad \forall\, s \in \mathbb{R} \,,$$
wobei $a = a' + \int_{\mathbb{R}} \left(y \mathbb{1}_{(-1,1)}(y) - \sin y \right) \nu(\mathrm{d}y)$. □

Korollar 3.1 *Sei $\{X_t,\, t \geq 0\}$ ein Lévy–Prozess. Dann sind sämtliche endlich–dimensionalen Verteilungen von $\{X_t\}$ eindeutig durch die Lévy–Charakteristik (a, b, ν) von X_1 bestimmt.*

Beweis

- Aus den Theoremen 3.3 und 3.4 ergibt sich zusammen mit dem Eindeutigkeitssatz für charakteristische Funktionen (vgl. Korollar WR-5.5), dass die Verteilung von X_t für jedes $t \geq 0$ eindeutig durch die Lévy–Charakteristik (a, b, ν) von X_1 bestimmt ist.
- Weil $\{X_t\}$ unabhängige und stationäre Zuwächse hat, sind damit auch sämtliche endlich–dimensionalen Verteilungen von $\{X_t\}$ eindeutig durch die Lévy–Charakteristik von X_1 bestimmt. □

3.1.3 Beispiele: Wiener–Prozess, zusammengesetzte Poisson–Prozesse, stabile Lévy–Prozesse

Bereits am Anfang von Abschnitt 3.1 hatten wir gezeigt, dass sowohl der Wiener–Prozess als auch beliebige zusammengesetzte Poisson–Prozesse den drei Bedingungen in der Definition von Lévy–Prozessen genügen.

Wir bestimmen nun die Lévy–Charakteristik (a, b, ν) für diese und weitere Beispiele von Lévy–Prozessen.

1. *Wiener–Prozess (mit Drift)*
   - Sei $\{X_t,\, t \geq 0\}$ ein (Standard–) Wiener–Prozess, d.h., insbesondere, dass $X_1 \sim \mathrm{N}(0,1)$.
   - Dann gilt also
   $$\varphi_{X_1}(s) = \mathrm{e}^{-s^2/2} \qquad \forall\, s \in \mathbb{R} \,.$$
   - Hieraus und aus der Lévy–Chintschin–Formel (5) ergibt sich nun, dass die Lévy–Charakteristik (a, b, ν) des Wiener–Prozesse $\{X_t\}$ gegeben ist durch $a = 0$, $b = 1$ und $\nu = 0$.
   - Beachte: Für beliebige $\mu \in \mathbb{R}$ und $c \neq 0$ wird der stochastische Prozess $\{Y_t,\, t \geq 0\}$ mit $Y_t = c\, X_t + \mu\, t$ ein *Wiener–Prozess mit Drift* genannt.
     - Man kann sich leicht überlegen, dass $\{Y_t\}$ ebenfalls ein Lévy–Prozess ist und
     - und dass seine Charakteristik (a, b, ν) gegeben ist durch $a = \mu$, $b = c^2$ und $\nu = 0$.
2. *zusammengesetzte Poisson–Prozesse*
   - Sei $\{X_t,\, t \geq 0\}$ ein zusammengesetzter Poisson–Prozess mit den Charakteristiken (λ, P_U), wobei $\lambda > 0$ die Intensität der Sprungzeitpunkte und P_U die Verteilung der Sprunghöhen bezeichnet.

- Dann gilt
$$\varphi_{X_1}(s) = \exp\Big(\lambda \int_{\mathbb{R}} \big(\mathrm{e}^{\mathrm{i}sy} - 1\big)\, P_U(\mathrm{d}y)\Big) \qquad \forall\, s \in \mathbb{R}\,,$$
vgl. die Teilaussage (b) von Theorem 2.10.
- Hieraus und aus (5) ergibt sich, dass die Lévy–Charakteristik (a, b, ν) von $\{X_t\}$ gegeben ist durch
$$a = \lambda \int_{-1}^{1} y\, P_U(\mathrm{d}y)\,, \qquad b = 0\,, \qquad \nu = \lambda\, P_U\,. \tag{24}$$
- Beachte: Für das in (24) betrachtete Lévy–Maß ν ist $\nu(B)$ die erwartete Anzahl von Sprüngen je Zeiteinheit mit Sprunghöhen in der Menge $B \in \mathcal{B}(\mathbb{R})$.

3. *Lévy–Prozesse vom Gauß–Poisson–Typ*

- Sei $\{X_t^{(1)},\, t \geq 0\}$ ein zusammengesetzter Poisson–Prozess (mit den Charakteristiken (λ, P_U)), und sei $\{X_t^{(2)},\, t \geq 0\}$ ein Wiener–Prozess mit Drift (mit den Charakteristiken (μ, c)).
- Außerdem seien die stochastischen Prozesse $\{X_t^{(1)}\}$ und $\{X_t^{(2)}\}$ unabhängig.
- Dann kann man sich leicht überlegen, dass der „überlagerte" Prozess $\{X_t,\, t \geq 0\}$ mit $X_t = X_t^{(1)} + X_t^{(2)}$ ein Lévy–Prozess ist, dessen Charakteristik (a, b, ν) gegeben ist durch
$$a = \mu + \lambda \int_{-1}^{1} y\, P_U(\mathrm{d}y)\,, \qquad b = c^2\,, \qquad \nu = \lambda\, P_U\,.$$

4. *stabile Lévy–Prozesse*

- Eine wichtige Eigenschaft der Normalverteilung ist ihre Faltungsstabilität (vgl. Korollar WR-3.2), die sich wie folgt charakterisieren lässt:
  - Sei $X \sim \mathrm{N}(\mu, \sigma^2)$ mit $\mu \in \mathbb{R}$ und $\sigma^2 > 0$, und $X_1, \ldots, X_n$ seien unabhängige Zufallsvariablen mit $X_k \sim \mathrm{N}(\mu, \sigma^2)$ für jedes $k = 1, \ldots, n$.
  - Dann gilt
$$X_1 + \ldots + X_n \stackrel{\mathrm{d}}{=} \sqrt{n}\, X + \mu\big(n - \sqrt{n}\big) \qquad \forall\, n \geq 1\,. \tag{25}$$
- Die Faltungsgleichung (25) lässt sich auf die folgende Weise verallgemeinern.
  - Die unbegrenzt teilbare Zufallsvariable X bzw. ihre Verteilung P_X heißt *stabil*, falls es für jedes $n \geq 1$ eine Folge von unabhängigen Zufallsvariablen $X_1, \ldots, X_n$ gibt,
  - so dass $X_k \stackrel{\mathrm{d}}{=} X$ für jedes $k = 1, \ldots, n$ und
$$X_1 + \ldots + X_n \stackrel{\mathrm{d}}{=} n^{1/\alpha}\, X + d_n \qquad \forall\, n \geq 1\,, \tag{26}$$
    wobei $\alpha \in (0, 2]$ und $d_n \in \mathbb{R}$ für jedes $n \geq 1$ gewisse Konstanten sind.
  - Dabei heißt $\alpha \in (0, 2]$ der *Stabilitätsindex* von X.
- Man kann zeigen, dass für jedes $\alpha \in (0, 2]$ eine Lösung der (verallgemeinerten) Faltungsgleichung (26) existiert, vgl. Samorodnitsky und Taqqu (1994) bzw. K.-I. Sato (1999). Dabei gilt für ein $\mu \in \mathbb{R}$
$$d_n = \begin{cases} \mu(n - n^{1/\alpha})\,, & \text{falls } \alpha \neq 1, \\ \mu\, n \log n\,, & \text{falls } \alpha = 1. \end{cases}$$
- Neben dem klassischen Fall $\alpha = 2$, der zur Normalverteilung führt, gibt es weitere wohlbekannte Beispiele stabiler Verteilungen, für die man zeigen kann, dass sie der Gleichung (26) genügen. Und zwar

- für $\alpha = 1$: die *Cauchy–Verteilung* mit den Parametern $\mu \in \mathbb{R}$, $\sigma^2 > 0$, deren Dichte $f : \mathbb{R} \to [0, \infty)$ gegeben ist durch
$$f(x) = \frac{\sigma}{\pi} \frac{1}{(x-\mu)^2 + \sigma^2} \qquad \forall\, x \in \mathbb{R}\,,$$
- für $\alpha = 1/2$: die *Lévy–Verteilung* mit den Parametern $\mu \in \mathbb{R}$, $\sigma^2 > 0$, deren Dichte $f : \mathbb{R} \to [0, \infty)$ gegeben ist durch
$$f(x) = \begin{cases} \left(\frac{\sigma}{2\pi}\right)^{1/2} \frac{1}{(x-\mu)^{3/2}} \exp\left(-\frac{\sigma}{2(x-\mu)}\right), & \text{falls } x > \mu, \\ 0\,, & \text{sonst.} \end{cases}$$

- Es ist klar, dass die Faltungsgleichung (26) genau dann gilt, wenn
$$\big(\varphi_X(s)\big)^n = \mathrm{e}^{\mathrm{i} d_n s} \varphi_X(n^{1/\alpha} s) \qquad \forall\, s \in \mathbb{R},\ n \geq 1\,. \tag{27}$$
- Wenn $X \stackrel{\mathrm{d}}{=} -X$ (wie beispielsweise bei der Cauchy–Verteilung mit $\mu = 0$),
  - dann gilt $d_n = 0$ für jedes $n \geq 1$,
  - und aus (27) folgt, dass es eine Konstante $c > 0$ gibt, so dass
$$\varphi_X(s) = \mathrm{e}^{-c|s|^\alpha} \qquad \forall\, s \in \mathbb{R}\,. \tag{28}$$
- Außerdem kann man zeigen, dass für jede stabile Zufallsvariable X die Charakteristik (a, b, ν) in der Lévy–Chintschin–Darstellung (5) von φ_X gegeben ist durch eine (beliebige) reelle Zahl $a \in \mathbb{R}$,
$$b = \begin{cases} c^2\,, & \text{falls } \alpha = 2, \\ 0\,, & \text{falls } \alpha < 2, \end{cases}$$
und
$$\nu(\mathrm{d}y) = \begin{cases} 0\,, & \text{falls } \alpha = 2, \\ \frac{c_1}{y^{1+\alpha}} \mathbb{1}_{(0,\infty)}(y)\,\mathrm{d}y + \frac{c_2}{|y|^{1+\alpha}} \mathbb{1}_{(-\infty,0)}(y)\,\mathrm{d}y\,, & \text{falls } \alpha < 2, \end{cases} \tag{29}$$
wobei $c \neq 0$ und $c_1, c_2 \geq 0$ gewisse Konstanten sind mit $c_1 + c_2 > 0$.
- Die Tails von stabilen Zufallsvariablen können entweder exponentiell beschränkt sein (für $\alpha = 2$) oder polynomial abklingen (für $\alpha < 2$):
  - Wenn $X \sim \mathrm{N}(\mu, \sigma^2)$, d.h., X hat eine stabile Verteilung mit Stabilitätsindex $\alpha = 2$, dann gilt
$$\lim_{x\to\infty} \frac{P(|X| > x)}{\mathrm{e}^{-x^2/(2\sigma^2)}} = 0\,,$$
d.h., X hat sogenannte „leichte" (exponentiell beschränkte) Tails.
  - Wenn X dagegen eine stabile Zufallsvariable mit Stabilitätsindex $\alpha < 2$ ist, dann kann man zeigen, dass
$$\lim_{x\to\infty} \frac{P(|X| > x)}{x^{-\alpha}} = c_\alpha$$
für ein $c_\alpha < \infty$, d.h., X hat „schwere" (polynomial abklingende) Tails.
- Beachte: Der Lévy–Prozess $\{X_t,\, t \geq 0\}$ heißt *stabiler Lévy–Prozess*, falls X_1 eine stabile Verteilung hat.

3.1.4 Subordinatoren

Eine weitere Klasse von Lévy–Prozessen, die insbesondere die zusammengesetzten Poisson–Prozesse mit positiven Sprunghöhen als Spezialfall umfassen, sind die sogenannten Subordinatoren.

Definition Ein Lévy–Prozess $\{X_t,\, t \geq 0\}$ heißt *Subordinator*, wenn mit Wahrscheinlichkeit 1

$$X_{t_1} \leq X_{t_2} \qquad \forall\, t_1, t_2 \geq 0 \text{ mit } t_1 \leq t_2, \tag{30}$$

wobei sich aus (30) wegen $X_0 = 0$ sofort ergibt, dass $X_t \geq 0$ für jedes $t \geq 0$.

Bevor wir einige konkrete Beispiele von Subordinatoren näher diskutieren, zeigen wir zunächst, wie sich die Lévy–Chintschin–Darstellung (15) von φ_{X_1} spezifizieren lässt, wenn $\{X_t\}$ ein Subordinator ist.

Theorem 3.5 *Der Lévy–Prozess $\{X_t,\, t \geq 0\}$ ist genau dann ein Subordinator, wenn sich der Lévy–Exponent $\eta(s)$ von X_1 darstellen lässt in der Form*

$$\eta(s) = \mathrm{i}as + \int_0^\infty \left(\mathrm{e}^{\mathrm{i}sy} - 1\right) \nu(\mathrm{d}y) \qquad \forall\, s \in \mathbb{R}, \tag{31}$$

wobei $a \geq 0$ und für das Lévy–Maß ν zusätzlich gilt, dass

$$\nu((-\infty, 0)) = 0 \qquad \textit{und} \qquad \int_0^\infty \min\{y, 1\}\, \nu(\mathrm{d}y) < \infty\,. \tag{32}$$

Beweis Wir zeigen zuerst die Hinlänglichkeit der Bedingungen (31) und (32).

- Falls $\nu = 0$, dann gilt $X_1 = a \geq 0$,
  - und aus der Stationarität der Zuwächse sowie der stochastischen Stetigkeit des Lévy–Prozesses $\{X_t\}$ ergibt sich, dass $X_t = at$ für jedes $t \geq 0$ gilt,
  - d.h., $\{X_t\}$ ist ein Subordinator.
- Sei nun $\nu(1/n, \infty) > 0$ für jedes hinreichend große $n \geq n_0$. Aus (31) und (32) ergibt sich dann, dass für jedes $s \in \mathbb{R}$

$$\varphi_{X_1}(s) = \mathrm{e}^{\mathrm{i}as}\, \exp\Big(\int_0^\infty \left(\mathrm{e}^{\mathrm{i}sy} - 1\right) \nu(\mathrm{d}y)\Big) = \mathrm{e}^{\mathrm{i}as} \lim_{n\to\infty} \exp\Big(\int_{1/n}^\infty \left(\mathrm{e}^{\mathrm{i}sy} - 1\right) \nu(\mathrm{d}y)\Big)\,.$$

- Für jedes $n \geq n_0$ ist dabei $\varphi_n : \mathbb{R} \to \mathbb{C}$ mit

$$\varphi_n(s) = \exp\Big(\int_{1/n}^\infty \left(\mathrm{e}^{\mathrm{i}sy} - 1\right) \nu(\mathrm{d}y)\Big) \qquad \forall\, s \in \mathbb{R}$$

  die charakteristische Funktion einer nichtnegativen Zufallsvariablen Z_n,
  - die eine zusammengesetzte Poisson–Verteilung mit den Charakteristiken (λ, P_U) hat
  - und zwar mit der „Poisson–Intensität" $\lambda = \nu(1/n, \infty)$, wobei $0 < \lambda < \infty$,
  - und mit der „Sprunghöhen-Verteilung" $P_U(\cdot) = \nu(\cdot \cap (1/n, \infty))/\nu(1/n, \infty)$, vgl. Abschnitt 2.2.2.
- Es gilt also

$$X_1 \stackrel{\mathrm{d}}{=} a + \lim_{n\to\infty} Z_n\,, \tag{33}$$

  wobei $a \geq 0$ und $\{Z_n\}$ eine Folge von nichtnegativen Zufallsvariablen ist.
  - Aus (33) folgt insbesondere, dass

$$P(X_1 \geq 0) = 1\,. \tag{34}$$

- Außerdem gilt für jedes $n \geq 1$

$$X_1 \stackrel{\mathrm{d}}{=} X_{1/n} + \big(X_{2/n} - X_{1/n}\big) + \ldots + \big(X_{n/n} - X_{(n-1)/n}\big)\,,$$

  wobei die Zufallsvariablen $X_{1/n}$, $X_{2/n} - X_{1/n}$, $\ldots$, $X_{n/n} - X_{(n-1)/n}$ unabhängig und identisch verteilt sind.
- Hieraus und aus (34) folgt, dass mit Wahrscheinlichkeit 1

$$X_{k/n} - X_{(k-1)/n} \geq 0 \qquad \forall\, n \geq 1,\, k = 1, \ldots, n\,.$$

- Aus der Stationarität der Zuwächse des Lévy–Prozesses $\{X_t\}$ ergibt sich nun, dass

$$X_{k/n} - X_{(k-1)/n} \geq 0 \qquad \forall\, k, n \geq 1$$

  bzw.

$$X_{q_2} - X_{q_1} \geq 0 \qquad \forall\, q_1, q_2 \in \mathbb{Q} \text{ mit } 0 \leq q_1 \leq q_2. \tag{35}$$

- Für $t_1, t_2 \geq 0$ mit $t_1 < t_2$ seien $\{q_n^{(1)}\}$ und $\{q_n^{(2)}\}$ zwei Folgen rationaler Zahlen, so dass $q_n^{(1)} < q_n^{(2)}$ für jedes $n \geq 1$ und $q_n^{(1)} \downarrow t_1$ bzw. $q_n^{(2)} \uparrow t_2$.
  - Dann ergibt sich aus (35), dass für jedes $\varepsilon > 0$

$$\begin{aligned} P(X_{t_2} - X_{t_1} < -\varepsilon) &= P\big(X_{t_2} - X_{q_n^{(2)}} + \big(X_{q_n^{(2)}} - X_{q_n^{(1)}}\big) + X_{q_n^{(1)}} - X_{t_1} < -\varepsilon\big) \\ &\leq P\big(X_{t_2} - X_{q_n^{(2)}} + X_{q_n^{(1)}} - X_{t_1} < -\varepsilon\big) \longrightarrow_{n\to\infty} 0\,, \end{aligned}$$

  - wobei sich der letzte Grenzwert aus der Tatsache ergibt, dass wegen der stochastischen Stetigkeit von $\{X_t\}$

$$X_{t_2} - X_{q_n^{(2)}} \xrightarrow{\mathrm{P}} 0 \qquad \text{und} \qquad X_{q_n^{(1)}} - X_{t_1} \xrightarrow{\mathrm{P}} 0\,.$$

- Damit ist gezeigt, dass $P(X_{t_2} - X_{t_1} < 0) = \lim_{\varepsilon\downarrow 0} P(X_{t_2} - X_{t_1} < -\varepsilon) = 0$ bzw. $P(X_{t_2} - X_{t_1} \geq 0) = 1$, d.h., $\{X_t\}$ ist ein Subordinator.

Die Notwendigkeit der Bedingungen (31) und (32) kann man sich wie folgt überlegen.

- Der Lévy–Prozess $\{X_t,\, t \geq 0\}$ sei ein Subordinator, d.h. insbesondere, dass $X_1 \geq 0$.
  - Dann ergibt sich unmittelbar aus dem Beweis von Theorem 3.4, dass die Lévy–Charakteristik (a, b, ν) von X_1 so gewählt werden kann, dass $\nu((-\infty, 0)) = 0$.
  - Gemäß Theorem 3.4 gilt somit für die charakteristische Funktion φ_{X_1} von X_1, dass für jedes $s \in \mathbb{R}$

$$\begin{aligned} \varphi_{X_1}(s) &= \exp\Big(\mathrm{i}as - \frac{bs^2}{2} + \int_0^\infty \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{I}_{(0,1)}(y)\big)\,\nu(\mathrm{d}y)\Big) \\ &= \exp\Big(\mathrm{i}as - \frac{bs^2}{2}\Big) \exp\Big(\int_0^\infty \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{I}_{(0,1)}(y)\big)\,\nu(\mathrm{d}y)\Big)\,. \end{aligned}$$

  - Hieraus ergibt sich mit Hilfe von Theorem 3.2 und dem Eindeutigkeitssatz für charakteristische Funktionen (vgl. Korollar WR-5.5), dass $X_1 \stackrel{\mathrm{d}}{=} Y_1 + Y_2$, wobei Y_1, Y_2 zwei unabhängige und unbegrenzt teilbare Zufallsvariablen sind mit $Y_1 \sim \mathrm{N}(a, b)$ sind.
  - Weil andererseits $X_1 \geq 0$ gilt, ist dies nur möglich, wenn $b = 0$.
- Wir haben somit gezeigt, dass für beliebige $s \in \mathbb{R}$ und $\varepsilon \in (0, 1)$

$$\begin{aligned} \varphi_{X_1}(s) &= \mathrm{e}^{\mathrm{i}as} \exp\Big(\int_0^\infty \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\mathbb{I}_{(0,1)}(y)\big)\,\nu(\mathrm{d}y)\Big) \\ &= \exp\Big(\mathrm{i}s\big(a - \int_\varepsilon^1 y\,\nu(\mathrm{d}y)\big)\Big) \exp\Big(\int_0^\varepsilon \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\big)\,\nu(\mathrm{d}y)\Big) \exp\Big(\int_\varepsilon^\infty \big(\mathrm{e}^{\mathrm{i}sy} - 1\big)\,\nu(\mathrm{d}y)\Big) \\ &= \exp\Big(\mathrm{i}s\big(a - \int_\varepsilon^1 y\,\nu(\mathrm{d}y)\big)\Big)\, \varphi_1(s)\, \varphi_2(s)\,, \end{aligned}$$

wobei

$$\varphi_1(s) = \exp\Big(\int_0^\varepsilon \big(\mathrm{e}^{\mathrm{i}sy} - 1 - \mathrm{i}sy\big)\,\nu(\mathrm{d}y)\Big)\,, \qquad \varphi_2(s) = \exp\Big(\int_\varepsilon^\infty \big(\mathrm{e}^{\mathrm{i}sy} - 1\big)\,\nu(\mathrm{d}y)\Big)\,.$$

- Dabei sind φ_1 und φ_2 die charakteristischen Funktionen von gewissen (unabhängigen) Zufallsvariablen, die wir mit Z_1 bzw. Z_2 bezeichnen, wobei

$$X_1 \stackrel{\mathrm{d}}{=} a - \int_\varepsilon^1 y\,\nu(\mathrm{d}y) + Z_1 + Z_2\,. \tag{36}$$

  - Weil die charakteristische Funktion $\varphi_1(s)$ von Z_1 zweimal an der Stelle $s = 0$ differenzierbar ist, existiert der Erwartungswert von Z_1, wobei $\mathbb{E}\,Z_1 = \varphi_1^{(1)}(0)/\mathrm{i}$; vgl. beispielsweise S. 183 im Buch von Grimmett und Stirzaker (2001).
  - Außerdem kann man sich leicht überlegen, dass $\varphi_1^{(1)}(0) = 0$ und somit $\mathbb{E}\,Z_1 = 0$, d.h. insbesondere $P(Z_1 \le 0) > 0$.
  - Andererseits gilt auch $P(Z_2 \le 0) > 0$, weil Z_2 für jedes $\varepsilon \in (0,1)$ eine zusammengesetzte Poisson–Verteilung hat.
  - Wegen der (angenommenen) Unabhängigkeit der Zufallsvariablen Z_1 und Z_2 folgt hieraus, dass

$$P(Z_1 + Z_2 \le 0) \ge P(Z_1 \le 0)\,P(Z_2 \le 0) > 0\,. \tag{37}$$

- Wegen (36) und (37) kann also $X_1 \ge 0$ nur dann mit Wahrscheinlichkeit 1 gelten, wenn

$$a - \int_\varepsilon^1 y\,\nu(\mathrm{d}y) \ge 0 \qquad \forall\,\varepsilon \in (0,1)\,.$$

- Hieraus folgt, dass $a \ge 0$ und $\int_0^\infty \min\{y,1\}\,\nu(\mathrm{d}y) < \infty$. □

Beachte

- Das Paar (a,ν) in (31) wird *Lévy–Charakteristik* des Subordinators $\{X_t\}$ genannt.
- Es ist klar, dass die in (32) betrachtete (verschärfte) Integrierbarkeitseigenschaft des Lévy–Maßes ν äquivalent ist mit der Bedingung, dass

$$\int_0^\infty \frac{y}{1+y}\,\nu(\mathrm{d}y) < \infty\,. \tag{38}$$

- Außerdem kann man sich leicht überlegen,
  - dass für jedes $t \ge 0$ die Abbildung $s \to \mathbb{E}\,\mathrm{e}^{\mathrm{i}sX_t}$ analytisch in das Teilgebiet $\{\mathrm{i}u : u \ge 0\}$ der komplexen Ebene fortgesetzt werden kann und
  - dass sich dabei die folgende Formel für die *Laplace–Transformierte* $\mathbb{E}\,\mathrm{e}^{-uX_t}$ von X_t ergibt: Für beliebige $t, u \ge 0$ gilt

$$\mathbb{E}\,\mathrm{e}^{-uX_t} = \mathrm{e}^{-t\,\xi(u)}\,, \qquad \text{wobei} \qquad \xi(u) = au + \int_0^\infty \big(1 - \mathrm{e}^{-uy}\big)\,\nu(\mathrm{d}y)\,. \tag{39}$$

- Die Funktion $\xi : [0,\infty) \to [0,\infty)$ in (39) wird *Laplace–Exponent* des Subordinators $\{X_t\}$ genannt.

Außer den zusammengesetzten Poisson–Prozessen mit positiven Sprunghöhen, die wir bereits am Anfang dieses Abschnittes erwähnt haben und deren Lévy–Maß ν endlich ist, gibt es noch weitere Klassen von Subordinatoren (mit unendlichem Lévy–Maß).

Beispiel (α–stabile Subordinatoren)

- Sei $\{X_t,\, t \geq 0\}$ ein Subordinator mit $a = 0$, dessen Lévy–Maß ν gegeben ist durch

$$\nu(\mathrm{d}y) = \begin{cases} \dfrac{\alpha}{\Gamma(1-\alpha)} \dfrac{1}{y^{1+\alpha}}\, \mathrm{d}y & \text{für } y > 0, \\ 0 & \text{für } y \leq 0, \end{cases} \tag{40}$$

  wobei $\alpha \in (0,1)$ und $\Gamma : (0,\infty) \to (0,\infty)$ die *Gammafunktion* bezeichnet mit

$$\Gamma(p) = \int_0^\infty \mathrm{e}^{-y}\, y^{p-1}\, \mathrm{d}y\,, \qquad \forall\, p > 0\,.$$

- Beachte: Das in (40) gegebene Lévy–Maß ν hat die Form (29).
- Außerdem kann man in diesem Fall leicht zeigen, dass (28) gilt bzw. (äquivalent hierzu) dass

$$\mathbb{E}\, \mathrm{e}^{-uX_t} = \mathrm{e}^{-tu^\alpha} \qquad \forall\, t, u \geq 0\,. \tag{41}$$

  - Denn für beliebige $u \geq 0$ und $\alpha \in (0,1)$ gilt die Identität

$$u^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (1 - \mathrm{e}^{-uy})\, \frac{\mathrm{d}y}{y^{1+\alpha}}\,.$$

  - Dies ergibt sich aus den folgenden Überlegungen: Es gilt

$$\begin{aligned} \int_0^\infty (1 - \mathrm{e}^{-uy})\, y^{-1-\alpha}\, \mathrm{d}y &= \int_0^\infty \Big(\int_0^y u\mathrm{e}^{-uz}\, \mathrm{d}z \Big)\, y^{-1-\alpha}\, \mathrm{d}y \\ &= \int_0^\infty \Big(\int_z^\infty y^{-1-\alpha}\, \mathrm{d}y \Big)\, u\mathrm{e}^{-uz}\, \mathrm{d}z \\ &= \frac{u}{\alpha} \int_0^\infty \mathrm{e}^{-uz} z^{-\alpha}\, \mathrm{d}z = \frac{u^\alpha}{\alpha} \int_0^\infty \mathrm{e}^{-x} x^{-\alpha}\, \mathrm{d}x \\ &= \frac{u^\alpha}{\alpha}\, \Gamma(1-\alpha)\,. \end{aligned}$$

  - Aus (39) und (40) ergibt sich nun, dass die Laplace–Transformierte $\mathbb{E}\, \mathrm{e}^{-uX_t}$ von X_t durch (41) gegeben ist.
  - Die Zufallsvariable X_t hat also für jedes $t \geq 0$ eine stabile Verteilung mit dem Stabilitätsindex $\alpha \in (0,1)$, d.h., der Subordinator $\{X_t\}$ mit dem in (40) gegebenen (unendlichen) Lévy–Maß ν ist ein α–stabiler Lévy–Prozess.

3.2 Martingale

In diesem Abschnitt diskutieren wir einige grundlegende Begriffe und Ergebnisse der Martingaltheorie, die wir dann in Abschnitt 3.3 bei der weiteren Untersuchung von Lévy–Prozessen anwenden werden.

3.2.1 Filtrationen und Stoppzeiten

Definition Sei $(\Omega, \mathcal{F}, P)$ ein vollständiger Wahrscheinlichkeitsraum.

- Eine Familie $\{\mathcal{F}_t,\, t \geq 0\}$ von Teil–σ–Algebren von $\mathcal{F}$ heißt *Filtration*, wenn $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ für beliebige $s, t \geq 0$ mit $s \leq t$ gilt.

- Die Filtration $\{\mathcal{F}_t,\, t \geq 0\}$ heißt *vollständig*, wenn $\mathcal{F}_0$ (und damit auch $\mathcal{F}_t$ für jedes $t \geq 0$) sämtliche Nullmengen enthält.
- Die Filtration $\{\mathcal{F}_t,\, t \geq 0\}$ heißt *rechtsstetig*, wenn $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u$ für jedes $t \geq 0$.

- Sei $\{X_t,\, t \geq 0\}$ ein stochastischer Prozess über $(\Omega, \mathcal{F}, P)$, und für jedes $t \geq 0$ sei $\mathcal{F}_t^X$ die kleinste σ–Algebra, für die
$$\{\omega \in \Omega :\, (X_{t_1}(\omega), \ldots, X_{t_n}(\omega)) \in B\} \in \mathcal{F}_t^X$$
für beliebige $n \geq 1$, $B \in \mathcal{B}(\mathbb{R}^n)$ und $t_1, \ldots, t_n \in [0, t]$ gilt. Dann heißt $\{\mathcal{F}_t^X,\, t \geq 0\}$ die *natürliche Filtration*, die von $\{X_t\}$ erzeugt wird.
- Sei $\{\mathcal{F}_t,\, t \geq 0\}$ eine beliebige Filtration, und sei $T : \Omega \to [0, \infty]$ eine beliebige Zufallsvariable über $(\Omega, \mathcal{F}, P)$. Man sagt, dass T eine *Stoppzeit* (bezüglich der Filtration $\{\mathcal{F}_t\}$) ist, wenn
$$\{\omega \in \Omega :\, T(\omega) \leq t\} \in \mathcal{F}_t \qquad \forall\, t \in [0, \infty)\,. \tag{1}$$

Beachte Wir werden stets (o.B.d.A.) voraussetzen, dass die jeweils betrachtete Filtration $\{\mathcal{F}_t\}$ vollständig ist.

Lemma 3.5 *Sei $\{\mathcal{F}_t,\, t \geq 0\}$ eine rechtsstetige Filtration. Die Zufallsvariable $T : \Omega \to [0, \infty]$ ist genau dann eine Stoppzeit, wenn*
$$\{T < t\} \in \mathcal{F}_t \qquad \forall\, t \in [0, \infty)\,. \tag{2}$$

Beweis

- Für jedes $\varepsilon > 0$ gilt offenbar, dass $\{T \leq t\} = \bigcap_{u \in (t, t+\varepsilon)} \{T < u\}$.
- Aus (2) und aus der Rechtsstetigkeit der Filtration $\{\mathcal{F}_t\}$ ergibt sich somit, dass
$$\{T \leq t\} \in \bigcap_{u>t} \mathcal{F}_u = \mathcal{F}_t \qquad \forall\, t \in [0, \infty)\,.$$
- Sei nun T eine Stoppzeit. Dann gilt
$$\{T < t\} = \bigcup_{\varepsilon \in (0,t)} \{T \leq t - \varepsilon\} \in \bigcup_{\varepsilon \in (0,t)} \mathcal{F}_{t-\varepsilon} \subset \mathcal{F}_t \qquad \forall\, t \in [0, \infty)\,.$$

□

Definition Sei $\{X_t,\, t \geq 0\}$ ein stochastischer Prozess über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit der Filtration $\{\mathcal{F}_t,\, t \geq 0\}$. Man sagt, dass der Prozess $\{X_t\}$ *adaptiert* (bezüglich der Filtration $\{\mathcal{F}_t\}$) ist, wenn
$$\{X_t \in B\} \in \mathcal{F}_t \qquad \forall\, t \geq 0,\, B \in \mathcal{B}(\mathbb{R})\,. \tag{3}$$
Außerdem heißt die Abbildung $T_B : \Omega \to [0, \infty]$ mit $T_B(\omega) = \inf\{t \geq 0 :\, X_t(\omega) \in B\}$ die *Ersterreichungszeit* der Borel–Menge $B \in \mathcal{B}(\mathbb{R})$ durch den Prozess $\{X_t\}$.

Für „càdlàg" Prozesse, deren Trajektorien mit Wahrscheinlichkeit 1 rechtsstetige Funktionen mit linksseitigen Grenzwerten sind (vgl. Abschnitt 1.4), diskutieren wir nun einige Beispiele von Stoppzeiten.

Theorem 3.6 *Sei $\{\mathcal{F}_t,\, t \geq 0\}$ eine rechtsstetige Filtration und sei $\{X_t,\, t \geq 0\}$ ein adaptierter càdlàg Prozess. Für jede offene Menge $B \in \mathcal{B}(\mathbb{R})$ ist dann die Ersterreichungszeit T_B eine Stoppzeit. Wenn $B \in \mathcal{B}(\mathbb{R})$ eine abgeschlossene Menge ist, dann ist die Abbildung $\widetilde{T}_B : \Omega \to [0, \infty]$ mit*
$$\widetilde{T}_B(\omega) = \inf\{t \geq 0 :\, X_t(\omega) \in B \;\; \textit{oder} \;\; X_{t-}(\omega) \in B\}$$
eine Stoppzeit, wobei $X_{t-} = \lim_{s \uparrow t} X_s$.

Beweis

- Sei $B \in \mathcal{B}(\mathbb{R})$ zunächst eine offene Menge.
  - Wegen Lemma 3.5 genügt es zu zeigen, dass $\{T_B < t\} \in \mathcal{F}_t$ für jedes $t \in [0, \infty)$ gilt.
  - Andererseits gilt
$$\{T_B < t\} = \bigcup_{s \in \mathbb{Q} \cap [0,t)} \{X_s \in B\},$$
    weil B eine offene Menge ist und $\{X_t\}$ rechtsstetige Trajektorien hat.
  - Hieraus ergibt sich der erste Teil der Behauptung, weil $\{X_s \in B\} \in \mathcal{F}_s \subset \mathcal{F}_t$ für jedes $s \leq t$.
- Sei nun $B \in \mathcal{B}(\mathbb{R})$ eine abgeschlossene Menge.
  - Für jedes $\varepsilon > 0$ bezeichnen wir mit $B_\varepsilon = \{x \in \mathbb{R} : d(x, B) < \varepsilon\}$ die sogenannte ε–„Parallelmenge" der Menge B, wobei $d(x, B) = \min\{|x - y| : y \in B\}$ der kleinste euklidische Abstand des Punktes $x \in \mathbb{R}$ zu einem Punkt von B ist.
  - Dann ist B_ε eine offene Menge, und man kann sich leicht überlegen, dass
$$\{\widetilde{T}_B \leq t\} = \{X_t \in B\} \cup \{X_{t-} \in B\} \cup \bigcap_{n \geq 1} \bigcup_{s \in \mathbb{Q} \cap [0,t)} \{X_s \in B_{1/n}\}.$$
  - Weil $\{X_t\}$ adaptiert ist, gehören sämtliche Ereignisse auf der rechten Seite der letzten Gleichheit zu $\mathcal{F}_t$. □

Theorem 3.7 *Seien $T, T' : \Omega \to [0, \infty]$ beliebige Stoppzeiten. Dann sind auch $\min\{T, T'\}$, $\max\{T, T'\}$, $T + T'$ bzw. αT für jedes $\alpha > 1$ Stoppzeiten.*

Beweis

- Weil T und T' Stoppzeiten sind, gilt $\{T \leq t\} \in \mathcal{F}_t$ bzw. $\{T' \leq t\} \in \mathcal{F}_t$ und somit
$$\{\min\{T, T'\} \leq t\} = \{T \leq t\} \cup \{T' \leq t\} \in \mathcal{F}_t,$$
  d.h., $\min\{T, T'\}$ ist eine Stoppzeit.
- Völlig analog ergibt sich, dass $\max\{T, T'\}$ ein Stoppzeit ist, denn aus $\{T \leq t\} \in \mathcal{F}_t$ und $\{T' \leq t\} \in \mathcal{F}_t$ folgt, dass
$$\{\max\{T, T'\} \leq t\} = \{T \leq t\} \cap \{T' \leq t\} \in \mathcal{F}_t.$$
- Um zu zeigen, dass auch $T + T'$ eine Stoppzeit ist, genügt es zu zeigen, dass $\{T + T' > t\} \in \mathcal{F}_t$ für jedes $t \geq 0$.
  - Dies ergibt sich aus der Identität
$$\{T + T' > t\} = \{T > t\} \cup \{T' > t\} \cup \{T \geq t, T' > 0\} \cup \{0 < T < t, T + T' > t\}.$$
  - Denn es ist klar, dass die ersten drei Ereignisse auf der rechten Seite der letzten Gleichheit zu $\mathcal{F}_t$ gehören,
  - und für das vierte Ereignis gilt
$$\{0 < T < t, T + T' > t\} = \bigcup_{s \in \mathbb{Q} \cap [0,t)} \{s < T < t, T' > t - s\} \in \mathcal{F}_t.$$
- Außerdem gilt für jedes $\alpha > 1$, dass $\{\alpha T \leq t\} = \{T \leq t/\alpha\} \in \mathcal{F}_{t/\alpha} \subset \mathcal{F}_t$. □

Manchmal ist es nützlich, beliebige (nichtnotwendig diskrete) Stoppzeiten durch monotone Folgen von Stoppzeiten zu approximieren, die jeweils nur abzählbar unendlich viele Werte mit positiver Wahrscheinlichkeit annehmen können. Aus dem Beweis des folgenden Theorems ergibt sich, wie eine solche monotone Folge von diskreten Stoppzeiten konstruiert werden kann.

Theorem 3.8 *Sei $T : \Omega \to (0, \infty]$ eine beliebige endliche Stoppzeit über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, d.h., es gelte $P(T < \infty) = 1$. Dann gibt es eine Folge $\{T^{(n)}, n \geq 1\}$ von diskreten Stoppzeiten, so dass mit Wahrscheinlichkeit* 1

$$T^{(1)} \geq T^{(2)} \geq \dots \qquad \textit{und} \qquad \lim_{n\to\infty} T^{(n)} = T\,. \tag{4}$$

Beweis

- Für jedes $n \geq 1$ sei die Zufallsvariable $T^{(n)} : \Omega \to [0, \infty)$ gegeben durch

$$T^{(n)} = \begin{cases} 0\,, & \text{falls } T = 0, \\ \dfrac{k+1}{2^n}\,, & \text{falls } \dfrac{k}{2^n} < T \leq \dfrac{k+1}{2^n} \text{ für ein } k = 0, 1, \dots. \end{cases} \tag{5}$$

- Für $k/2^n \leq t < (k+1)/2^n$ gilt dann

$$\{T^{(n)} \leq t\} = \{T^{(n)} \leq \frac{k}{2^n}\} = \{T \leq \frac{k}{2^n}\} \in \mathcal{F}_{k/2^n} \subset \mathcal{F}_t\,. \tag{6}$$

- Hieraus folgt, dass $T^{(n)}$ eine Stoppzeit ist, wobei $T^{(1)} \geq T^{(2)} \geq \dots$ und $\lim_{n\to\infty} T^{(n)} = T$. □

Korollar 3.2 *Sei $T : \Omega \to [0, \infty]$ eine beliebige endliche Stoppzeit. Der Prozess $\{X_t, t \geq 0\}$ sei càdlàg und über dem gleichen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ wie T gegeben. Dann gilt $\{\omega \in \Omega : X_{T(\omega)}(\omega) \in B\} \in \mathcal{F}$ für jedes $B \in \mathcal{B}(\mathbb{R})$, d.h., die Abbildung $X_T : \Omega \to \mathbb{R}$ mit*

$$X_T(\omega) = X_{T(\omega)}(\omega) \tag{7}$$

ist $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$–messbar.

Beweis

- Weil $\{X(t)\}$ càdlàg ist, ergibt sich aus (4) und (6), dass $X_T = \lim_{n\to\infty} X_{T^{(n)}}$.
- Es genügt nun zu beachten, dass $X_{T^{(n)}} : \Omega \to \mathbb{R}$ eine $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$–messbare Abbildung ist, denn für jedes $B \in \mathcal{B}(\mathbb{R})$ gilt

$$\{X_{T^{(n)}} \in B\} = \bigcup_{k=0}^{\infty} \{X_{k/2^n} \in B\} \cap \{T^{(n)} = \frac{k}{2^n}\} \in \mathcal{F}\,.$$ □

3.2.2 Submartingale und Supermartingale; Beispiele

Definition Sei $\{X_t, t \geq 0\}$ ein stochastischer Prozess über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit der Filtration $\{\mathcal{F}_t, t \geq 0\}$.

- Der Prozess $\{X_t\}$ sei adaptiert, und es gelte $\mathbb{E}\,|X_t| < \infty$ für jedes $t \geq 0$.
- Man sagt, dass $\{X_t\}$ ein *Martingal* ist, wenn $\mathbb{E}\,(X_t \mid \mathcal{F}_s) = X_s$ für beliebige $s, t \geq 0$ mit $s \leq t$, wobei $\mathbb{E}\,(X_t \mid \mathcal{F}_s)$ die bedingte Erwartung von X_t bezüglich der σ–Algebra $\mathcal{F}_s$ bezeichnet; vgl. Abschnitt 2.2.3.
- Außerdem sagt man, dass $\{X_t\}$ ein *Submartingal* bzw. *Supermartingal* ist, wenn $\mathbb{E}\,(X_t \mid \mathcal{F}_s) \geq X_s$ bzw. $\mathbb{E}\,(X_t \mid \mathcal{F}_s) \leq X_s$ für beliebige $s, t \geq 0$ mit $s \leq t$.

Beispiele

1. *Poisson–Prozess*
   - Sei $\{X_t,\, t \geq 0\}$ ein (homogener) Poisson–Prozess mit der Intensität $\lambda > 0$.
   - Dann kann man sich leicht überlegen, dass der stochastische Prozess $\{Y_t,\, t \geq 0\}$ mit $Y_t = X_t - t\lambda$ ein Martingal bezüglich der (natürlichen) Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$ ist.
     - Aus der Unabhängigkeit und Stationarität der Zuwächse von $\{X_t\}$ und aus den Teilaussagen 4 und 6 von Theorem 2.11 ergibt sich nämlich, dass
     $$\begin{aligned}\mathbb{E}\,(X_t \mid \mathcal{F}_s^X) &= \mathbb{E}\,(X_s + (X_t - X_s) \mid \mathcal{F}_s^X) = \mathbb{E}\,(X_s \mid \mathcal{F}_s^X) + \mathbb{E}\,(X_t - X_s \mid \mathcal{F}_s^X) \\ &= X_s + \mathbb{E}\,(X_t - X_s) = X_s + (t-s)\lambda\end{aligned}$$
     für beliebige $s, t \geq 0$ mit $s \leq t$,
     - d.h.,
     $$\mathbb{E}\,(X_t - t\lambda \mid \mathcal{F}_s^X) = X_s - s\lambda \qquad \forall\, s \leq t\,.$$
   - Außerdem kann man zeigen, dass der stochastische Prozess $\{Y_t',\, t \geq 0\}$ mit $Y_t' = (X_t - t\lambda)^2 - t\lambda$ ein Martingal ist.
     - Sei $\widetilde{X}_t = X_t - t\lambda$.
     - Dann ergibt sich aus der Unabhängigkeit der Zuwächse von $\{\widetilde{X}_t\}$ und aus den Eigenschaften der bedingten Erwartung (vgl. Theorem 2.11), dass für beliebige $s, t \geq 0$ mit $s \leq t$
     $$\begin{aligned}\mathbb{E}\,(Y_t' \mid \mathcal{F}_s^X) &= \mathbb{E}\,(\widetilde{X}_t^2 - t\lambda \mid \mathcal{F}_s^X) \\ &= \mathbb{E}\,\big(\big(\widetilde{X}_s + (\widetilde{X}_t - \widetilde{X}_s)\big)^2 - t\lambda \mid \mathcal{F}_s^X\big) \\ &= \mathbb{E}\,\big(\widetilde{X}_s^2 + 2\widetilde{X}_s(\widetilde{X}_t - \widetilde{X}_s) + (\widetilde{X}_t - \widetilde{X}_s)^2 - t\lambda \mid \mathcal{F}_s^X\big) \\ &= \widetilde{X}_s^2 - s\lambda + 2\widetilde{X}_s\mathbb{E}\,\big(\widetilde{X}_t - \widetilde{X}_s\big) + \mathbb{E}\,\big((\widetilde{X}_t - \widetilde{X}_s)^2\big) - (t-s)\lambda \\ &= \widetilde{X}_s^2 - s\lambda = Y_s'\,,\end{aligned}$$
     - wobei in der vorletzten Gleichheit die Tatsache genutzt wurde, dass $\mathrm{Var}\,(\widetilde{X}_t - \widetilde{X}_s) = (t-s)\lambda$.

2. *zusammengesetzte Poisson–Prozesse*
   - Sei $\{X_t,\, t \geq 0\}$ ein zusammengesetzter Poisson–Prozess mit den Charakteristiken (λ, P_U), wobei $\lambda > 0$ die Intensität der Sprungzeitpunkte und P_U die Verteilung der Sprunghöhen bezeichnet.
   - Wenn die zusätzliche Integrierbarkeitsbedingung $\mathbb{E}\,|U| < \infty$ erfüllt ist, dann ist der stochastische Prozess $\{Y_t,\, t \geq 0\}$ mit $Y_t = X_t - t\lambda\mathbb{E}\,U$ ein Martingal.
   - Denn genauso wie in dem oben diskutierten Fall des (nicht zusammengesetzten) Poisson–Prozesses ergibt sich, dass
   $$\mathbb{E}\,(X_t - t\lambda\mathbb{E}\,U \mid \mathcal{F}_s^X) = X_s - s\lambda\mathbb{E}\,U \qquad \forall\, s \leq t\,.$$

3. *Wiener–Prozess*
   - Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess. Dann kann man auf die gleiche Weise wie bei den beiden vorhergehenden Beispielen zeigen, dass $\{X_t\}$ ein Martingal bezüglich der natürlichen Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$ dieses Prozesses ist.
   - Außerdem kann man zeigen, dass auch die stochstischen Prozesse $\{Y_t',\, t \geq 0\}$ und $\{\widetilde{Y}_t,\, t \geq 0\}$ Martingale sind, wobei
   $$Y_t' = X_t^2 - t \qquad \text{bzw.} \qquad \widetilde{Y}_t = \mathrm{e}^{uX_t - u^2 t/2}$$
   und $u \in \mathbb{R}$ eine beliebige (jedoch fixierte) Zahl ist.
   - Denn genauso wie im Poisson–Fall ergibt sich, dass für beliebige $s, t \geq 0$ mit $s \leq t$
   $$\mathbb{E}\,(Y_t' \mid \mathcal{F}_s^X) = \ldots = X_s^2 - s + 2X_s\mathbb{E}\,(X_t - X_s) + \mathbb{E}\,\big((X_t - X_s)^2\big) - (t-s) = X_s^2 - s = Y_s'\,,$$
   weil $\mathbb{E}\,(X_t - X_s) = 0$ und $\mathbb{E}\,\big((X_t - X_s)^2\big) = t - s$, d.h., $\{Y_t'\}$ ist ein Martingal.

- Auf ähnliche Weise ergibt sich, dass für beliebige $s, t \geq 0$ mit $s \leq t$

$$\mathbb{E}\,(\mathrm{e}^{uX_t - u^2 t/2} \mid \mathcal{F}_s^X) = \mathbb{E}\,\mathrm{e}^{u(X_t - X_s)}\,\mathrm{e}^{uX_s - u^2 t/2}\,,$$

  woaus folgt, dass $\{\widetilde{Y}_t\}$ ein Martingale ist, weil $\mathbb{E}\,\mathrm{e}^{u(X_t - X_s)} = \mathrm{e}^{u^2(t-s)/2}$.

4. *Lévy–Prozesse*
   - Genauso wie bei den ersten drei Beispielen lässt sich für jeden Lévy–Prozess $\{X_t,\, t \geq 0\}$ mit $\mathbb{E}\,|X_1| < \infty$ ein Martingal konstruieren.
     - Denn für beliebige $s, t \geq 0$ mit $s \leq t$ gilt

$$\mathbb{E}\,(X_t \mid \mathcal{F}_s^X) = X_s + (t-s)\,\mathbb{E}\,X_1\,,$$

     - d.h., $\{Y_t,\, t \geq 0\}$ mit $Y_t = X_t - t\,\mathbb{E}\,X_1$ ist ein Martingal bezüglich der Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$.
   - Im allgemeinen, d.h., wenn $\mathbb{E}\,|X_1| < \infty$ nicht vorausgesetzt wird, ist für jedes $u \in \mathbb{R}$ ein (komplexwertiges) Martingal $\{\widetilde{Y}_t,\, t \geq 0\}$ gegeben durch

$$\widetilde{Y}_t = \mathrm{e}^{\mathrm{i}uX_t - t\eta(u)}\,,$$

     wobei $\eta : \mathbb{R} \to \mathbb{C}$ der Lévy–Exponent von $\{X_t\}$ ist; vgl. die Formel (8) in Abschnitt 3.1.
     - Denn für beliebige $u \in \mathbb{R}$ und $t \geq 0$ gilt $\mathbb{E}\,|\widetilde{Y}_t| = \mathrm{e}^{-t\eta(u)} < \infty$.
     - Außerdem gilt für beliebige $u \in \mathbb{R}$ und $s, t \geq 0$ mit $s \leq t$

$$\begin{aligned}
\mathbb{E}\,(\widetilde{Y}_t \mid \mathcal{F}_s^X) &= \mathbb{E}\left(\mathrm{e}^{\mathrm{i}uX_t - t\eta(u)} \mid \mathcal{F}_s^X\right) \\
&= \mathbb{E}\left(\mathrm{e}^{\mathrm{i}uX_s - s\eta(u)}\mathrm{e}^{\mathrm{i}u(X_t - X_s) - (t-s)\eta(u)} \mid \mathcal{F}_s^X\right) \\
&= \mathrm{e}^{\mathrm{i}uX_s - s\eta(u)}\,\mathbb{E}\left(\mathrm{e}^{\mathrm{i}u(X_t - X_s) - (t-s)\eta(u)} \mid \mathcal{F}_s^X\right) \\
&= \mathrm{e}^{\mathrm{i}uX_s - s\eta(u)}\,\mathbb{E}\,\mathrm{e}^{\mathrm{i}u(X_t - X_s) - (t-s)\eta(u)} \\
&= \mathrm{e}^{\mathrm{i}uX_s - s\eta(u)}\,\mathbb{E}\,\mathrm{e}^{\mathrm{i}uX_{t-s} - (t-s)\eta(u)} = \mathrm{e}^{\mathrm{i}uX_s - s\eta(u)} = \widetilde{Y}_s\,,
\end{aligned}$$

       wobei sich die dritte Gleichheit aus Teilaussage 4 in Theorem 2.11 ergibt, während die vorletzte Gleichheit aus Theorem 3.3 folgt.

5. *Markow–Prozesse*
   - Wir zeigen nun, wie Martingale für Funktionen von Markow–Prozessen (mit endlich vielen Zuständen) konstruiert werden können.
     - Sei also $\{X_t,\, t \geq 0\}$ ein Markow–Prozess mit Werten in der Menge $E = \{1, 2, \ldots, \ell\}$,
     - und sei $\mathbf{Q}$ die in Abschnitt 2.3.2 eingeführte Intensitätsmatrix von $\{X_t\}$.
   - Für jeden Vektor $\mathbf{b} = (b_1, \ldots, b_\ell)^\top \in \mathbb{R}^\ell$ ist dann der stochastische Prozess $\{Y_t\,, t \geq 0\}$ mit

$$Y_t = b_{X_t} - b_{X_0} - \int_0^t (\mathbf{Qb})_{X_v}\,\mathrm{d}v \qquad \forall\, t \geq 0 \tag{8}$$

     ein Martingal, wobei das Integral in (8) pfadweise gebildet wird.
     - Weil $\{X_t\}$ ein homogener Markow–Prozess ist, gilt für beliebige $i \in E$ und $s, t \geq 0$ mit $s \leq t$

$$\begin{aligned}
&\mathbb{E}\left(b_{X_t} - b_{X_s} - \int_s^t (\mathbf{Qb})_{X_v}\,\mathrm{d}v \;\middle|\; X_s = i\right) \\
&\quad = \mathbb{E}\left(b_{X_{t-s}} - b_{X_0} - \int_0^{t-s} (\mathbf{Qb})_{X_v}\,\mathrm{d}v \;\middle|\; X_0 = i\right).
\end{aligned}$$

– Um zu zeigen, dass der in (8) gegebene Prozess $\{Y_t\}$ ein Martingal ist, genügt es also zu zeigen, dass für beliebige $i \in E$ und $h \geq 0$

$$\mathbb{E}\,(b_{X_h} \mid X_0 = i) - b_i = \int_0^h \mathbb{E}\,((\mathbf{Qb})_{X_v} \mid X_0 = i)\,\mathrm{d}v\,, \tag{9}$$

denn man kann sich leicht überlegen, dass

$$\mathbb{E}\,\Big(\int_0^h (\mathbf{Qb})_{X_v}\,\mathrm{d}v \mid X_0 = i\Big) = \int_0^h \mathbb{E}\,((\mathbf{Qb})_{X_v} \mid X_0 = i)\,\mathrm{d}v\,.$$

– In Theorem 2.17 hatten wir gezeigt, dass die Übergangsfunktion $\{\mathbf{P}(h), h \geq 0\}$ des Markow–Prozesse $\{X_t\}$ für jedes $h \geq 0$ durch $\mathbf{P}(h) = \exp(h\mathbf{Q})$ gegeben ist.

– Hieraus folgt, dass

$$\mathbb{E}\,(b_{X_h} \mid X_0 = i) = \mathbf{e}_i^\top \exp(\mathbf{Q}h)\mathbf{b} \tag{10}$$

und

$$\mathbb{E}\,\big((\mathbf{Qb})_{X_v} \mid X_0 = i\big) = \mathbf{e}_i^\top \exp(\mathbf{Q}v)\mathbf{Qb}\,, \tag{11}$$

wobei $\mathbf{e}_i^\top$ ein ℓ-dimensionaler (Zeilen–) Vektor ist, für den sämtliche Komponenten gleich 0 sind, bis auf die i–te Komponente, die gleich 1 ist.

– Aus (10) und (11) ergibt sich nun, dass (9) äquivalent ist mit

$$\mathbf{e}_i^\top \exp(\mathbf{Q}h)\mathbf{b} - b_i = \int_0^h \mathbf{e}_i^\top \exp(\mathbf{Q}v)\mathbf{Qb}\,\mathrm{d}v\,, \tag{12}$$

wobei sich die Gültigkeit dieser Gleichung aus Lemma 2.3 ergibt, wenn auf beiden Seiten von (12) die Ableitung nach h gebildet wird.

6. *abgeschlossene Martingale*

- Sei $X : \Omega \to \mathbb{R}$ eine beliebige Zufallsvariable mit $\mathbb{E}\,|X| < \infty$, und sei $\{\mathcal{F}_t,\, t \geq 0\}$ eine beliebige Filtration.
- Dann kann man sich leicht überlegen, dass der stochastische Prozess $\{Y_t,\, t \geq 0\}$ mit

$$Y_t = \mathbb{E}\,(X \mid \mathcal{F}_t) \tag{13}$$

ein Martingal ist, denn aus Teilaussage 6 von Theorem 2.11 ergibt sich, dass für beliebige $s, t \geq 0$ mit $s \leq t$

$$\mathbb{E}\,\big(\mathbb{E}\,(X \mid \mathcal{F}_t) \mid \mathcal{F}_s\big) = \mathbb{E}\,(X \mid \mathcal{F}_s)\,.$$

- Manchmal sagt man, dass der in (13) gegebene stochastische Prozess ein *abgeschlossenes Martingal* ist.

Beachte

- Für die Martingale $\{Y_t\}$ und $\{Y_t'\}$, die in den Beispielen 1 – 4 betrachtet wurden, gilt $\mathbb{E}\,Y_t = \mathbb{E}\,Y_t' = 0$ für jedes $t \geq 0$. Ein Martingal mit dieser Eigenschaft wird ein *zentriertes Martingal* genannt.
- Für die Martingale $\{\widetilde{Y}_t\}$, die in den Beispielen 3 und 4 betrachtet wurden, gilt dagegen $\mathbb{E}\,\widetilde{Y}_t = 1$ für jedes $t \geq 0$.
- Die Definitionsgleichung (8) des Martingals $\{Y_t\}$ in Beispiel 5 wird *Dynkin–Formel* für Markow–Prozesse genannt.

Aus den Monotonieeigenschaften der bedingten Erwartung (vgl. die Teilaussage 3 von Theorem 2.11) ergibt sich, dass jeder adaptierte Prozess $\{X_t,\, t \geq 0\}$ mit nichtfallenden Trajektorien und mit $\mathbb{E}\,|X_t| < \infty$ für jedes $t \geq 0$ ein Submartingal ist.

Insbesondere ist jeder „integrierbare" Subordinator ein Submartingal, d.h. jeder (nichtnegative) Lévy–Prozess $\{X_t,\, t \geq 0\}$ mit nichtfallenden Trajektorien, so dass $\mathbb{E}\,|X_1| < \infty$.

Theorem 3.9 *Der Prozess $\{X_t,\, t \geq 0\}$ sei adaptiert und $f : \mathbb{R} \to \mathbb{R}$ sei eine konvexe Funktion, so dass $\mathbb{E}\,|f(X_t)| < \infty$ für jedes $t \geq 0$. Dann ist $\{f(X_t),\, t \geq 0\}$ ein Submartingal, wenn*

- *$\{X_t\}$ ein Martingal oder*
- *$\{X_t\}$ ein Submartingal und $f : \mathbb{R} \to \mathbb{R}$ eine nichtfallende konvexe Funktion ist.*

Beweis Aus der Jensen–Ungleichung für bedingte Erwartungen (vgl. die Teilaussage 7 von Theorem 2.11) ergibt sich in beiden Fällen, dass für beliebige $s, t \geq 0$ mit $s \leq t$

$$\mathbb{E}\left(f(X_t) \mid \mathcal{F}_s\right) \geq f\big(\mathbb{E}\left(X_t \mid \mathcal{F}_s\right)\big) \geq f(X_s)\,. \qquad \square$$

3.2.3 Gleichgradige Integrierbarkeit

Sei T eine endliche Stoppzeit, und der stochastische Prozess $\{X_t,\, t \geq 0\}$ sei càdlàg mit $\mathbb{E}\,|X_t| < \infty$ für jedes $t \in \mathbb{R}$. In Korollar 3.2 hatten wir gezeigt, dass dann X_T eine wohldefinierte Zufallsvariable ist.

Zur Herleitung von Bedingungen, so dass auch $\mathbb{E}\,|X_T| < \infty$ gilt, benötigen wir den Begriff der gleichgradigen Integrierbarkeit von Zufallsvariablen.

Hierfür benutzen wir die folgende Schreibweise: Sei $X : \Omega \to \mathbb{R}$ eine beliebige Zufallsvariable mit $\mathbb{E}\,|X| < \infty$. Für jedes $A \in \mathcal{F}$ setzen wir dann $\mathbb{E}\,[X; A] = \mathbb{E}\,(X \mathbb{1}(A))$.

Definition Die Folge $X_1, X_2, \ldots : \Omega :\to \mathbb{R}$ von Zufallsvariablen heißt *gleichgradig integrierbar*, wenn $\mathbb{E}\,|X_n| < \infty$ für jedes $n \geq 1$ und

$$\lim_{x\to\infty} \big(\sup_{n\geq 1} \mathbb{E}\,[|X_n|; |X_n| > x]\big) = 0\,. \tag{14}$$

Lemma 3.6 *Die Folge $\{X_n,\, n \geq 1\}$ ist genau dann gleichgradig integrierbar, wenn*

(i) $\sup_{n\geq 1} \mathbb{E}\,|X_n| < \infty$ *und*

(ii) *wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass $\mathbb{E}\,[|X_n|; A] < \varepsilon$ für jedes $n \geq 1$ und für jedes $A \in \mathcal{F}$ mit $P(A) < \delta$.*

Beweis

- Wir zeigen zuerst, dass (14) gilt, wenn die Bedingungen (i) und (ii) erfüllt sind.
  - Offenbar gilt $xP(|X_n| > x) \leq \mathbb{E}\,|X_n|$ für beliebige $n \geq 1$ und $x > 0$. Hieraus und aus (i) folgt, dass für jedes $\delta > 0$ und für jedes hinreichend große $x > 0$

$$\sup_{n\geq 1} P(|X_n| > x) \leq \frac{1}{x} \sup_{n\geq 1} \mathbb{E}\,|X_n| < \delta\,.$$

  - Wegen (ii) gilt somit $\limsup_{x\to\infty} \big(\sup_{n\geq 1} \mathbb{E}\,[|X_n|; |X_n| > x]\big) \leq \varepsilon$ für jedes $\varepsilon > 0$.
  - Hieraus ergibt sich die Gültigkeit von (14), weil $\varepsilon > 0$ beliebig klein gewählt werden kann.
- Sei nun $\{X_n,\, n \geq 1\}$ gleichgradig integrierbar.
  - Dann gilt für jedes hinreichend große $x > 0$

$$\sup_{n\geq 1} \mathbb{E}\,|X_n| \leq \sup_{n\geq 1} \big\{\mathbb{E}\,[|X_n|; |X_n| > x] + x\big\} < \infty\,,$$

    d.h., die Bedingung (i) ist erfüllt.

- Für jedes $\varepsilon > 0$ sei nun $x > 0$ so gewählt, dass $\sup_{n\geq 1} \mathbb{E}\left[|X_n|; |X_n| > x\right] < \varepsilon/2$.
- Für jedes $\delta > 0$ mit $\delta < \varepsilon/(2x)$ und für jedes $A \in \mathcal{F}$ mit $P(A) < \delta$ gilt dann

$$\mathbb{E}\left[|X_n|; A\right] \leq \mathbb{E}\left[|X_n|; |X_n| > x\right] + xP(A) \leq \varepsilon\,,$$

d.h., die Bedingung (ii) ist ebenfalls erfüllt. □

Die Bedeutung der gleichgradigen Integrierbarkeit für die Konvergenz von Zufallsvariablen wird durch das folgende Lemma deutlich.

Lemma 3.7 *Sei $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ eine beliebige Folge von Zufallsvariablen mit $\mathbb{E}\,|X_n| < \infty$ für jedes $n \geq 1$ und $\lim_{n\to\infty} X_n = X$ mit Wahrscheinlichkeit 1 für eine gewisse Zufallsvariable $X : \Omega \to \mathbb{R}$. Dann sind die beiden folgenden Aussagen äquivalent:* (a) *$\{X_n,\, n \geq 1\}$ ist gleichgradig integrierbar.* (b) *Es gilt $\mathbb{E}\,|X| < \infty$ und $\lim_{n\to\infty} \mathbb{E}\,|X_n - X| = 0$.*

Beweis

- Sei $\{X_n,\, n \geq 1\}$ gleichgradig integrierbar.
  - Weil $\sup_{n\geq 1} \mathbb{E}\,|X_n| < \infty$ wegen Teilaussage (i) in Lemma 3.6 und weil $\lim_{n\to\infty} X_n = X$ vorausgesetzt wird, ergibt sich aus dem Lemma von Fatou, dass $\mathbb{E}\,|X| < \infty$.
  - Andererseits kann man sich leicht überlegen (vgl. Korollar WR-5.1), dass aus der Konvergenz $\lim_{n\to\infty} X_n = X$ mit Wahrscheinlichkeit 1 folgt, dass für jedes $\varepsilon > 0$

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0\,. \tag{15}$$

  - Außerdem gilt für jedes $\varepsilon > 0$

$$\begin{aligned} \mathbb{E}\,|X_n - X| &\leq \mathbb{E}\left[|X_n - X|; |X_n - X| \leq \varepsilon\right] + \mathbb{E}\left[|X_n - X|; |X_n - X| > \varepsilon\right] \\ &\leq \varepsilon + \mathbb{E}\left[|X_n|; |X_n - X| > \varepsilon\right] + \mathbb{E}\left[|X|; |X_n - X| > \varepsilon\right]. \end{aligned}$$

  - Wegen (15) ergibt sich dann aus der Teilaussage (ii) in Lemma 3.6 mit $A = \{|X_n - X| > \varepsilon\}$, dass der zweite Summand des letzten Ausdruckes gegen 0 strebt.
  - Der ditte Summand konvergiert ebenfalls gegen 0, weil $\mathbb{E}\,|X| < \infty$.
  - Hieraus folgt, dass $\lim_{n\to\infty} \mathbb{E}\,|X_n - X| = 0$, weil $\varepsilon > 0$ beliebig klein gewählt werden kann.
- Wir nehmen nun umgekehrt an, dass (b) gilt.
  - Dann gilt

$$\sup_{n\geq 1} \mathbb{E}\,|X_n| \leq \sup_{n\geq 1} \mathbb{E}\,|X_n - X| + \mathbb{E}\,|X| < \infty\,. \tag{16}$$

  - Für jedes $\varepsilon > 0$ sei $n_0 \geq 1$ so gewählt, dass für beliebige $n \geq n_0$ und $A \in \mathcal{F}$

$$\mathbb{E}\left[|X_n - X|; A\right] \leq \mathbb{E}\,|X_n - X| \leq \frac{\varepsilon}{2}\,.$$

  - Sei nun $\delta > 0$ so gewählt, dass $P(A) < \delta$ die Gültigkeit der folgenden Ungleichung impliziert:

$$\max_{0\leq n\leq n_0} \mathbb{E}\left[|X_n - X|; A\right] + \mathbb{E}\left[|X|; A\right] \leq \frac{\varepsilon}{2}\,.$$

  - Hieraus folgt, dass es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für jedes $A \in \mathcal{F}$ mit $P(A) < \delta$

$$\sup_{n\geq 1} \mathbb{E}\left[|X_n|; A\right] \leq \sup_{n\geq 1} \mathbb{E}\left[|X_n - X|; A\right] + \mathbb{E}\left[|X|; A\right] \leq \varepsilon\,.$$

- Wegen Lemma 3.6 ergibt sich hieraus und aus (16), dass $\{X_n, n \geq 1\}$ gleichgradig integrierbar ist. □

Korollar 3.3 *Sei $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ eine Folge von Zufallsvariablen, so dass $\mathbb{E}\,|X_n| < \infty$ für jedes $n \geq 1$ und $\lim_{n\to\infty} X_n = X$ mit Wahrscheinlichkeit 1. Wenn $\{X_n, n \geq 1\}$ gleichgradig integrierbar ist, dann gilt $\mathbb{E}\,|X| < \infty$ und $\lim_{n\to\infty} \mathbb{E}\,X_n = \mathbb{E}\,X$.*

Beweis Wegen $|\mathbb{E}\,X_n - \mathbb{E}\,X| \leq \mathbb{E}\,|X_n - X|$ ergibt sich die Behauptung unmittelbar aus Lemma 3.7. □

3.2.4 Ungleichung von Doob

Ein wichtiges (beweistechnisches) Hilfsmittel ist die folgende *Ungleichung von Doob* für Submartingale. Dabei benutzen wir die Schreibweise $x_+ = \max\{x, 0\}$ für jedes $x \in \mathbb{R}$.

Theorem 3.10 *Der Prozess $\{X_t,\, t \geq 0\}$ sei adaptiert und càdlàg. Wenn $\{X_t\}$ ein Submartingal ist, dann gilt für beliebige $x > 0$ und $t \geq 0$*

$$P\big(\sup_{0\leq v\leq t} X_v > x\big) \leq \frac{\mathbb{E}\,(X_t)_+}{x}\,. \tag{17}$$

Beweis

- Wir können o.B.d.A. annehmen, dass $P(X_t \geq 0) = 1$ für jedes $t \geq 0$, denn für jedes $x \geq 0$ gilt

  $$P\big(\sup_{0\leq v\leq t} X_v > x\big) = P\big(\sup_{0\leq v\leq t} \max\{X_v, 0\} > x\big)$$

  und aus Theorem 3.9 ergibt sich, dass der stochastische Prozess $\{\widetilde{X}_t,\, t \geq 0\}$ mit $\widetilde{X}_t = \max\{X_t, 0\}$ ebenfalls ein Submartingal ist.
- Sei nun $\{X_t,\, t \geq 0\}$ ein nichtnegatives Submartingal, und sei $A = \{\max_{1\leq k\leq n} X_{t_k} > x\}$ für beliebige $t_1, \ldots, t_n \in [0, t]$ mit $t_1 \leq \ldots \leq t_n$.
  - Man kann sich leicht überlegen, dass sich das Ereignis A darstellen lässt als die Vereinigung $A = A_1 \cup \ldots \cup A_n$ einer Folge von paarweise disjunkten Mengen $A_1, \ldots, A_n$, wobei

    $$A_1 = \{X_{t_1} > x\} \in \mathcal{F}_{t_1}\,, \qquad A_k = \{X_{t_1} \leq x, \ldots, X_{t_{k-1}} \leq x, X_{t_k} > x\} \in \mathcal{F}_{t_k} \qquad \forall\, k \in \{2, \ldots, n\}\,.$$
  - Weil $\{X_t,\, t \geq 0\}$ ein nichtnegatives Submartingal ist, gilt

    $$\mathbb{E}\,[X_{t_n}\,;\, A_k] \geq \mathbb{E}\,[X_{t_k}\,;\, A_k] \geq x\,P(A_k)\,.$$
  - Hieraus folgt, dass

    $$\mathbb{E}\,X_{t_n} \geq \mathbb{E}\,[X_{t_n}\,;\, A] = \sum_{k=1}^{n} \mathbb{E}\,[X_{t_n}\,;\, A_k] \geq \sum_{k=1}^{n} x\,P(A_k) = xP(A)\,.$$
- Sei nun B eine beliebige endliche Teilmenge des Intervalls $[0, t]$ so, dass $0 \in B$ und $t \in B$.
  - Dann gilt also

    $$xP\big(\max_{v\in B} X_v > x\big) \leq \mathbb{E}\,X_t\,. \tag{18}$$
  - Wegen der Monotonie von Wahrscheinlichkeitsmaßen gilt außerdem für jede monoton wachsende Folge $B_1, B_2, \ldots \subset [0, t]$ von endlichen Mengen, dass

    $$\lim_{n\to\infty} P\big(\max_{v\in B_n} X_v > x\big) = P\big(\bigcup_{n\geq 1} \{\max_{v\in B_n} X_v > x\}\big) = P\big(\{\sup_{v\in\bigcup_{n\geq 1} B_n} X_v > x\}\big)\,.$$
  - Hieraus und aus (18) folgt, dass für $B = \bigcup_{n\geq 1} B_n = ([0, t) \cap \mathbb{Q}) \cup \{t\}$

    $$xP\big(\sup_{v\in B} X_v > x\big) \leq \mathbb{E}\,X_t\,.$$

- Weil $\{X(t)\}$ rechtsstetige Trajektorien hat, ergibt sich hieraus die Gültigkeit von (17). □

Beachte

- Sei $\{X_t,\, t \geq 0\}$ ein (Standard–) Wiener–Prozess, und für $\mu > 0$ sei $\{Y_t,\, t \geq 0\}$ mit $Y_t = X_t - t\mu$ ein Wiener–Prozess mit negativer Drift.
  - In Beispiel 3 von Abschnitt 3.2.2 hatten wir gezeigt, dass $\{\mathrm{e}^{u(Y_t+t\mu)-u^2t/2},\, t \geq 0\}$ für jedes $u \in \mathbb{R}$ ein Martingal bezüglich der (natürlichen) Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$ ist.
  - Für $u = 2\mu$ ergibt also insbesondere, dass $\{\mathrm{e}^{2\mu Y_t},\, t \geq 0\}$ ein Martingal ist.
  - Weil die Ungleichung $\sup_{v\in[0,t]} Y_v > x$ genau dann gilt, wenn $\sup_{v\in[0,t]} \mathrm{e}^{2\mu Y_v} > \mathrm{e}^{2\mu x}$, ergibt sich aus Theorem 3.10, dass

$$P\big(\sup_{t\geq 0} Y_t > x\big) = \lim_{t\to\infty} P\big(\sup_{v\in[0,t]} Y_v > x\big) = \lim_{t\to\infty} P\big(\sup_{v\in[0,t]} \mathrm{e}^{2\mu Y_v} > \mathrm{e}^{2\mu x}\big) \leq \frac{\lim_{t\to\infty} \mathbb{E}\,\mathrm{e}^{2\mu Y_t}}{\mathrm{e}^{2\mu x}}\,.$$

  - Weil $\mathbb{E}\,\mathrm{e}^{2\mu Y_t} = 1$ für jedes $t \geq 0$, ergibt sich hieraus, dass

$$P\big(\sup_{t\geq 0} Y_t > x\big) \leq \mathrm{e}^{-2\mu x} \qquad \forall\, x \geq 0\,. \tag{19}$$

- Um zu zeigen, dass in (19) sogar die Gleichheit gilt, benötigen wir ein *optionales Sampling–Theorem*, für dessen Herleitung wir zunächst sogenannte „gestoppte Martingale" betrachten.

3.2.5 Gestoppte Martingale

Lemma 3.8

- *Der Prozess $\{X_t,\, t \geq 0\}$ sei adaptiert und càdlàg. Außerdem sei $T : \Omega \to [0,\infty)$ eine endliche Stoppzeit und sei $\{T^{(n)},\, n \geq 1\}$ die in (5) eingeführte Folge von diskreten Stoppzeiten mit $T^{(n)} \geq T^{(n+1)}$ und $\lim_{n\to\infty} T^{(n)} = T$.*
- *Wenn $\{X_t\}$ ein Martingal ist, dann ist die Folge $X_{T^{(1)}\wedge t}, X_{T^{(2)}\wedge t}, \ldots$ gleichgradig integrierbar für jedes $t \geq 0$, wobei $s \wedge t = \min\{s,t\}$.*

Beweis

- Weil $\mathbb{E}\,|X_{T^{(n)}\wedge t}| \leq \sum_{\{k:\, k<2^n t\}} \mathbb{E}\,\big|X_{k/2^n}\big| + \mathbb{E}\,|X_t| < \infty$ für beliebige $n \geq 1$ und $t \geq 0$, genügt es zu zeigen, dass die Bedingung (14) in der Definition der gleichgradigen Integrierbarkeit erfüllt ist.
  - Aus Theorem 3.9 folgt, dass $\{|X_t|,\, t \geq 0\}$ ein Submartingal ist. Somit gilt

$$\begin{aligned}
&\sup_{n\geq 1} \mathbb{E}\,[|X_{T^{(n)}\wedge t}|; |X_{T^{(n)}\wedge t}| > x] \\
&\quad = \sup_{n\geq 1}\Big(\sum_{\{k:\, k<2^n t\}} \mathbb{E}\,[|X_{k/2^n}|; \{T^{(n)} = \frac{k}{2^n}\} \cap \{|X_{T^{(n)}\wedge t}| > x\}] \\
&\qquad\qquad + \mathbb{E}\,[|X_t|; \{T^{(n)} \geq t\} \cap \{|X_{T^{(n)}\wedge t}| > x\}]\Big) \\
&\quad \leq \sup_{n\geq 1}\Big(\sum_{\{k:\, k<2^n t\}} \mathbb{E}\,[|X_t|; \{T^{(n)} = \frac{k}{2^n}\} \cap \{|X_{T^{(n)}\wedge t}| > x\}] \\
&\qquad\qquad + \mathbb{E}\,[|X_t|; \{T^{(n)} \geq t\} \cap \{|X_{T^{(n)}\wedge t}| > x\}]\Big) \\
&\quad = \sup_{n\geq 1} \mathbb{E}\,[|X_t|; |X_{T^{(n)}\wedge t}| > x] \leq \sup_{n\geq 1} \mathbb{E}\,[|X_t|\,; Y > x] = \mathbb{E}\,[|X_t|; Y > x]\,,
\end{aligned}$$

  wobei $Y = \sup_{n\geq 1} |X_{T^{(n)}\wedge t}|$.

- Andererseits ergibt die Anwendung von Theorem 3.10 auf das Submartingal $\{|X_t|,\, t \geq 0\}$, dass

$$P(Y > x) \leq P(\sup_{0 \leq v \leq t} |X_v| > x) \leq \frac{\mathbb{E}\,|X_t|}{x} \qquad \forall\, t, x > 0\,.$$

- Hieraus folgt insbesondere, dass $\lim_{x\to\infty} P(Y > x) = 0$.

- Aus dem Satz von Lebesgue über die majorisierte Konvergenz ergibt sich nun insgesamt, dass

$$\lim_{x\to\infty} \sup_{n\geq 1} \mathbb{E}\,[|X_{T^{(n)}\wedge t}|; |X_{T^{(n)}\wedge t}| > x] \leq \lim_{x\to\infty} \mathbb{E}\,[|X_t|; Y \geq x] = 0\,. \qquad \square$$

Theorem 3.11 *Der Prozess $\{X_t,\, t \geq 0\}$ sei adaptiert und càdlàg. Wenn $\{X_t\}$ ein Martingal und $T : \Omega \to [0, \infty)$ eine endliche Stoppzeit ist, dann ist auch der stochastische Prozess $\{X_{T\wedge t}, t \geq 0\}$ ein Martingal.*

Beweis

- Ähnlich wie im Beweis von Korollar 3.2 kann man zeigen, dass $X_{T\wedge t}$ für jedes $t \geq 0$ eine $\mathcal{F}_t$–messbare Zufallsvariable ist.
- Es muss also lediglich noch gezeigt werden, dass für beliebige $t \geq 0$, $v \in [0, t)$ und $A \in \mathcal{F}_v$

$$\mathbb{E}\,|X_{T\wedge t}| < \infty \qquad \text{und} \qquad \mathbb{E}\,[X_{T\wedge t}; A] = \mathbb{E}\,[X_{T\wedge v}; A]\,. \tag{20}$$

- Dabei betrachten wir zunächst die in (5) eingeführte Folge $\{T^{(n)},\, n \geq 1\}$ von diskreten Stoppzeiten und zeigen, dass für jedes $n \geq 1$ und für beliebige $t \geq 0$, $v \in [0, t)$ und $A \in \mathcal{F}_v$

$$\mathbb{E}\,|X_{T^{(n)}\wedge t}| < \infty \qquad \text{und} \qquad \mathbb{E}\,[X_{T^{(n)}\wedge t}; A] = \mathbb{E}\,[X_{T^{(n)}\wedge v}; A]\,. \tag{21}$$

  - Genauso wie im Beweis von Lemma 3.8 ergibt sich, dass $\mathbb{E}\,|X_{T^{(n)}\wedge t}| < \infty$.
  - Seien nun $t_1, \ldots, t_k \in (v, t)$ die Werte, die die Stoppzeit $T^{(n)}$ mit positiver Wahrscheinlichkeit zwischen v und t annehmen kann.
  - Dann ergibt sich aus den Teilaussagen 4 und 5 von Theorem 2.11, dass

$$\begin{aligned}
\mathbb{E}\,(X_{T^{(n)}\wedge t} \mid \mathcal{F}_v) &= \mathbb{E}\,\big(\mathbb{E}\,(X_{T^{(n)}\wedge t} \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_v\big) \\
&= \mathbb{E}\,(\mathbb{E}\,(\mathbb{1}(T^{(n)} \leq t_k) X_{T^{(n)}\wedge t} \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_v) + \mathbb{E}\,(\mathbb{E}\,(\mathbb{1}(T^{(n)} > t_k) X_{T^{(n)}\wedge t} \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_v) \\
&= \mathbb{E}\,(\mathbb{1}(T^{(n)} \leq t_k) \mathbb{E}\,(X_{T^{(n)}\wedge t_k} \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_v) + \mathbb{E}\,(\mathbb{1}(T^{(n)} > t_k) \mathbb{E}\,(X_t \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_v) \\
&= \mathbb{E}\,(\mathbb{1}(T^{(n)} \leq t_k) X_{T^{(n)}\wedge t_k} \mid \mathcal{F}_v) + \mathbb{E}\,(\mathbb{1}(T^{(n)} > t_k) X_{t_k} \mid \mathcal{F}_v) \\
&= \mathbb{E}\,(\mathbb{1}(T^{(n)} \leq t_k) X_{T^{(n)}\wedge t_k} \mid \mathcal{F}_v) + \mathbb{E}\,(\mathbb{1}(T^{(n)} > t_k) X_{T^{(n)}\wedge t_k} \mid \mathcal{F}_v) \\
&= \mathbb{E}\,(X_{T^{(n)}\wedge t_k} \mid \mathcal{F}_v) \\
&\;\;\vdots \\
&= \mathbb{E}\,(X_{T^{(n)}\wedge t_1} \mid \mathcal{F}_v) = \mathbb{E}\,(X_{T^{(n)}\wedge v} \mid \mathcal{F}_v) = X_{T^{(n)}\wedge v}\,.
\end{aligned}$$

  - Damit ist (21) bewiesen.

- Weil $\{X_t\}$ is càdlàg ist, gilt $\lim_{n\to\infty} X_{T^{(n)}\wedge t} = X_{T\wedge t}$ und $\lim_{n\to\infty} X_{T^{(n)}\wedge v} = X_{T\wedge v}$.
- Gemäß Lemma 3.8 sind die Folgen $\{X_{T^{(n)}\wedge v},\ n \geq 1\}$ und $\{X_{T^{(n)}\wedge t},\ n \geq 1\}$ außerdem gleichgradig integrierbar.
- Wegen (21) ergibt sich somit die Gültigkeit von (20) mit Hilfe von Korollar 3.3. $\square$

3.2.6 Optionales Sampling–Theorem

Für jede endliche Stoppzeit $T : \Omega \to [0, \infty)$ betrachten wir die σ–Algebra $\mathcal{F}_T$, die gegeben ist durch

$$\mathcal{F}_T = \{A \in \mathcal{F} : \, A \cap \{T \leq t\} \in \mathcal{F}_t \text{ für jedes } t \geq 0\}\,, \tag{22}$$

wobei die folgenden Eigenschaften der „gestoppten" σ–Algebra $\mathcal{F}_T$ nützlich sind.

Lemma 3.9 *Für beliebige endliche Stoppzeiten $S, T : \Omega \to [0, \infty)$ mit $P(S \leq T) = 1$ gilt $\mathcal{F}_S \subset \mathcal{F}_T$. Außerdem gilt für jeden adaptierten càdlàg Prozess $\{X_t,\, t \geq 0\}$*

$$\{\omega \in \Omega : \, X_T(\omega) \in B\} \in \mathcal{F}_T \qquad \forall\, B \in \mathcal{B}(\mathbb{R})\,, \tag{23}$$

d.h., X_T ist eine $\big(\mathcal{F}_T, \mathcal{B}(\mathbb{R})\big)$–messbare Zufallsvariable.

Beweis

- Wir zeigen zunächst, dass $\mathcal{F}_S \subset \mathcal{F}_T$, wenn $P(S \leq T) = 1$.
  - Sei $A \in \mathcal{F}_S$, d.h., für jedes $t \geq 0$ gelte $A \cap \{S \leq t\} \in \mathcal{F}_t$.
  - Hieraus folgt, dass $A \cap \{T \leq t\} = A \cap \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t$ für jedes $t \geq 0$, weil

$$\{T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \qquad \text{und} \qquad \{T \leq t\} \in \mathcal{F}_t \qquad \forall\, t \geq 0\,.$$

- Wir zeigen nun, dass $\{X_T \in B\} \cap \{T \leq t\} \in \mathcal{F}_t$ für beliebige $B \in \mathcal{B}(\mathbb{R})$ und $t \geq 0$.
  - Für jedes $t \geq 0$ fassen wir dabei den (eingeschränkten) Prozess $\{X_s,\, s \in [0, t]\}$ als Abbildung $X : [0, t] \times \Omega \to \mathbb{R}$ auf, die gegeben ist durch

$$(s, \omega) \mapsto X_s(\omega)\,. \tag{24}$$

  - Weil der Prozess $\{X_t\}$ càdlàg ist, gilt für beliebige $(s, \omega) \in [0, t] \times \Omega$

$$X_s(\omega) = X_0(\omega)\mathbb{1}_{\{0\}}(s) + \lim_{n\to\infty} \sum_{k=1}^{2^n} \mathbb{1}_{((k-1)t/2^n\,,\, kt/2^n]}(s) X_{kt/2^n}(\omega)\,.$$

  - Weil $\{X_t\}$ adaptiert ist, sind sämtliche Summanden auf der rechten Seite des letzten Ausdruckes $\big(\mathcal{B}([0, t]) \otimes \mathcal{F}_t,\, \mathcal{B}(\mathbb{R})\big)$–messbar.
  - Damit besitzt auch ihr Grenzwert diese Messbarkeitseigenschaft, d.h., die in (24) gegebene Abbildung ist $\big(\mathcal{B}([0, t]) \otimes \mathcal{F}_t,\, \mathcal{B}(\mathbb{R})\big)$–messbar.
  - Außerdem betrachten wir die Abbildung $g : \Omega \cap \{T \leq t\} \to [0, t] \times \Omega$ mit

$$g(\omega) = \big(T(\omega), \omega\big)\,, \tag{25}$$

    wobei diese Abbildung $\big(\mathcal{F}_t \cap \{T \leq t\},\, \mathcal{B}([0, t]) \otimes \mathcal{F}_t\big)$–messbar ist.
  - Hieraus folgt, dass die Superposition $X \circ g : \Omega \cap \{T \leq t\} \to \mathbb{R}$ der in (24) bzw. (25) gegebenen Abbildungen $\big(\mathcal{F}_t \cap \{T \leq t\},\, \mathcal{B}(\mathbb{R})\big)$–messbar ist. □

Wir kommen nun zum sogenannten *optionalen Sampling–Theorem* für Martingale bzw. für gestoppte Martingale.

Theorem 3.12 *Der Prozess $\{X_t,\, t \geq 0\}$ sei adaptiert und càdlàg. Wenn $\{X_t\}$ ein Martingal und $T : \Omega \to [0, \infty)$ eine endliche Stoppzeit ist, dann gilt*

$$\mathbb{E}\,(X_t \mid \mathcal{F}_T) = X_{T\wedge t} \qquad \forall\, t \geq 0\,. \tag{26}$$

Beweis

- Wir zeigen zunächst, dass

$$\mathbb{E}\left(X_t \mid \mathcal{F}_{T^{(n)}}\right) = X_{T^{(n)}\wedge t} \qquad \forall\, n \geq 1,\ t \geq 0\,, \tag{27}$$

  wobei $\{T^{(n)},\, n \geq 1\}$ die in (5) eingeführte Folge von diskreten Stoppzeiten ist.
  - Hierfür sei $t_k = t$ und $t_1 \leq \ldots \leq t_{k-1} \leq t$ bezeichne die Werte, die die Zufallsvariable $T^{(n)} \wedge t$ mit positiver Wahrscheinlichkeit annehmen kann.
  - Dann gilt für jedes $A \in \mathcal{F}_{T^{(n)}}$

$$\begin{aligned}(X_t - X_{T^{(n)}\wedge t})\mathbb{1}(A) &= \Big(\sum_{i=1}^{k-1}(X_t - X_{t_i})\mathbb{1}(T^{(n)} \wedge t = t_i)\Big)\mathbb{1}(A)\\ &= \Big(\sum_{i=2}^{k}(X_{t_i} - X_{t_{i-1}})\mathbb{1}(T^{(n)} \wedge t < t_i)\Big)\mathbb{1}(A)\,.\end{aligned}$$

  - Hieraus folgt, dass

$$\begin{aligned}\mathbb{E}\Big((X_t - X_{T^{(n)}\wedge t})\mathbb{1}(A)\Big) &= \sum_{i=2}^{k}\mathbb{E}\Big((X_{t_i} - X_{t_{i-1}})\mathbb{1}(T^{(n)} \wedge t < t_i)\mathbb{1}(A)\Big)\\ &= \sum_{i=2}^{k}\mathbb{E}\Big(\mathbb{E}\Big((X_{t_i} - X_{t_{i-1}})\mathbb{1}(T^{(n)} \wedge t < t_i)\mathbb{1}(A) \mid \mathcal{F}_{t_{i-1}}\Big)\Big)\\ &= \sum_{i=2}^{k}\mathbb{E}\Big(\mathbb{1}(T^{(n)} \wedge t < t_i)\mathbb{1}(A)\,\mathbb{E}\left(X_{t_i} - X_{t_{i-1}} \mid \mathcal{F}_{t_{i-1}}\right)\Big)\\ &= 0\,,\end{aligned}$$

    wobei sich die vorletzte Gleichheit aus der Definitionsgleichung (22) der σ–Algebra $\mathcal{F}_{T^{(n)}}$ und aus Teilaussage 4 von Theorem 2.11 ergibt.
  - Somit gilt

$$\mathbb{E}\left(X_t \mid \mathcal{F}_{T^{(n)}}\right) = \mathbb{E}\left(X_{T^{(n)}\wedge t} \mid \mathcal{F}_{T^{(n)}}\right) = X_{T^{(n)}\wedge t}\,,$$

    weil $X_{T^{(n)}\wedge t}$ eine $\left(\mathcal{F}_{T^{(n)}}, \mathcal{B}(\mathbb{R})\right)$–messbare Zufallsvariable ist; vgl. Lemma 3.9.
- Aus Lemma 3.9 folgt außerdem, dass $\mathcal{F}_T \subset \mathcal{F}_{T^{(n)}}$, weil $T \leq T^{(n)}$.
- Mit Hilfe der Teilaussage 5 von Theorem 2.11 ergibt sich nun hieraus und aus (27), dass

$$\mathbb{E}\left(X_t \mid \mathcal{F}_T\right) = \mathbb{E}\left(\mathbb{E}\left(X_t \mid \mathcal{F}_{T^{(n)}}\right) \mid \mathcal{F}_T\right) = \mathbb{E}\left(X_{T^{(n)}\wedge t} \mid \mathcal{F}_T\right).$$

- Weil $\lim_{n\to\infty} T^{(n)} \downarrow T$, weil der Prozess $\{X_t,\, t \geq 0\}$ càdlàg ist und weil die Folge $\{X_{T^{(n)}\wedge t},\ n \geq 1\}$ gemäß Lemma 3.8 gleichgradig integrierbar ist, gilt also

$$\mathbb{E}\left(X_t \mid \mathcal{F}_T\right) = \lim_{n\to\infty}\mathbb{E}\left(X_{T^{(n)}\wedge t} \mid \mathcal{F}_T\right) = \mathbb{E}\left(X_{T\wedge t} \mid \mathcal{F}_T\right) = X_{T\wedge t}\,,$$

  wobei sich die letzte Gleichheit aus Lemma 3.9 ergibt. □

Korollar 3.4 *Der Prozess $\{X_t,\, t \geq 0\}$ sei adaptiert und càdlàg. Wenn $\{X_t\}$ ein Martingal und $S, T : \Omega \to [0, \infty)$ beliebige endliche Stoppzeiten mit $P(S \leq T) = 1$ sind, dann gilt*

$$\mathbb{E}\left(X_{T\wedge t} \mid \mathcal{F}_S\right) = X_{S\wedge t} \qquad \forall\, t \geq 0\,. \tag{28}$$

Insbesondere gilt

$$\mathbb{E}\, X_{T\wedge t} = \mathbb{E}\, X_0 \qquad \forall\, t \geq 0\,. \tag{29}$$

Beweis

- Sei $\{X_t\}$ ein Martingal. Aus Theorem 3.11 folgt dann, dass der stochastische Prozess $\{X_{T\wedge t}, t \geq 0\}$ ebenfalls ein Martingal ist.
  - Es ist klar, dass mit $\{X_t\}$ auch $\{X_{T\wedge t}\}$ càdlàg ist. Außerdem ergibt sich aus Lemma 3.9, dass $\{X_{T\wedge t}\}$ adaptiert ist.
  - Wir können also Theorem 3.12 auf das Martingal $\{X_{T\wedge t}, t \geq 0\}$ anwenden und erhalten, dass

$$\mathbb{E}\left(X_{T\wedge t} \mid \mathcal{F}_S\right) = X_{T\wedge(S\wedge t)} = X_{S\wedge t} \qquad \forall\, t \geq 0\,,$$

  - wobei sich die letzte Gleichheit aus der Annahme ergibt, dass $S \leq T$.
- Wenn wir die „Stoppzeit" $S = 0$ in (28) einsetzen, dann ergibt sich, dass $\mathbb{E}\left(X_{T\wedge t} \mid \mathcal{F}_0\right) = X_0$ bzw. $\mathbb{E}\, X_{T\wedge t} = \mathbb{E}\, X_0$ für jedes $t \geq 0$. □

3.3 Anwendungsbeispiele

Um die in Abschnitt 3.2 dargelegten Begriffe und Ergebnisse der Martingaltheorie bei der Untersuchung von Lévy–Prozessen anwenden zu können, benötigen wir noch zwei weitere grundlegende Eigenschaften von Lévy–Prozessen, die wir hier ohne Beweis angeben.

Theorem 3.13 *Sei $\{X_t,\, t \geq 0\}$ ein Lévy–Prozess mit den Charakteristiken (a, b, ν).*

- *Dann gibt es eine càdlàg Modifikation $\{\widetilde{X}_t,\, t \geq 0\}$ von $\{X_t\}$, so dass $\{\widetilde{X}_t\}$ ebenfalls ein Lévy–Prozess ist und zwar mit den gleichen Charakteristiken (a, b, ν) wie $\{X_t\}$.*
- *Wenn der Lévy–Prozess $\{X_t,\, t \geq 0\}$ càdlàg ist, dann ist die (durch $\{X_t\}$ erzeugte) natürliche Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$ rechtsstetig.*

Ein *Beweis* von Theorem 3.13 kann zum Beispiel in Applebaum (2004), S. 74–77 nachgelesen werden.

Wegen Theorem 3.13 werden wir in den weiteren Abschnitten dieses Skriptes (o.B.d.A.) voraussetzen, das die jeweils betrachteten Lévy–Prozesse càdlàg sind.

3.3.1 Regeneration von Lévy–Prozessen zu Stoppzeiten

In Theorem 2.23 hatten wir für Fall, dass $\{X_t,\, t \geq 0\}$ ein (Standard–) Wiener–Prozess ist, ein Schranke für die Tailfunktion von $\max_{t\in[0,1]} X_t$ hergeleitet.

Im nachfolgenden Abschnitt 3.3.2 werden wir zeigen, dass diese Schranke „optimal" ist, d.h. mit der Tailfunktion von $\max_{t\in[0,1]} X_t$ übereinstimmt. Hierfür zeigen wir zunächst, dass Lévy–Prozesse zu endlichen Stoppzeiten die folgende *Regenerationseigenschaft* besitzen.

Theorem 3.14

- *Sei $\{X_t,\, t \geq 0\}$ ein Lévy–Prozess über $(\Omega, \mathcal{F}, P)$ und sei $T : \Omega \to [0, \infty)$ eine endliche Stoppzeit bezüglich der natürlichen Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$ von $\{X_t\}$.*
- *Dann ist der Prozess $\{Y_t,\, t \geq 0\}$ mit $Y_t = X_{T+t} - X_T$ ebenfalls ein Lévy–Prozess, der adaptiert ist bezüglich der Filtration $\{\mathcal{G}_t,\, t \geq 0\}$ mit $\mathcal{G}_t = \mathcal{F}_{T+t}^X$, wobei*
  - *der Lévy–Prozess $\{Y_t\}$ unabhängig von $\mathcal{F}_T^X$ ist und*
  - *$\{Y_t\}$ die gleiche Lévy–Charakteristik wie $\{X_t\}$ hat.*

Beweis

- Wir betrachten zunächst den Fall, dass $T : \Omega \to [0, \infty)$ eine beschränkte Stoppzeit ist, d.h., es gelte $P(T < c) = 1$ für ein $c < \infty$.
  - Für beliebige $n \geq 1$ und $u_1, \ldots, u_n \in \mathbb{R}$ sind dann durch den Ansatz

$$\widetilde{Y}_t^{(j)} = \mathrm{e}^{\mathrm{i} u_j X_t - t\eta(u_j)}$$

    (komplexwertige) Martingale $\{\widetilde{Y}_t^{(j)}, t \geq 0\}$ für jedes $j = 1, \ldots, n$ gegeben, wobei $\eta : \mathbb{R} \to \mathbb{C}$ der Lévy–Exponent von $\{X_t\}$ ist; vgl. Beispiel 4 in Abschnitt 3.2.2.
  - Dann gilt für beliebige $A \in \mathcal{F}_T^X$ und $t_0, t_1, \ldots, t_n \geq 0$ mit $t_0 < t_1 < \ldots < t_n$

$$\begin{aligned}
\mathbb{E}\left(\mathbb{1}(A)\, \exp\big(\mathrm{i} \sum_{j=1}^n u_j (X_{T+t_j} - X_{T+t_{j-1}})\big)\right) &= \mathbb{E}\left(\mathbb{1}(A) \prod_{j=1}^n \frac{\widetilde{Y}_{T+t_j}^{(j)}}{\widetilde{Y}_{T+t_{j-1}}^{(j)}} \frac{\mathrm{e}^{(T+t_j)\eta(u_j)}}{\mathrm{e}^{(T+t_{j-1})\eta(u_j)}}\right) \\
&= \mathbb{E}\left(\mathbb{1}(A) \prod_{j=1}^n \frac{\widetilde{Y}_{T+t_j}^{(j)}}{\widetilde{Y}_{T+t_{j-1}}^{(j)}} \mathrm{e}^{(t_j - t_{j-1})\eta(u_j)}\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}(A) \prod_{j=1}^n \frac{\widetilde{Y}_{T+t_j}^{(j)}}{\widetilde{Y}_{T+t_{j-1}}^{(j)}} \mathrm{e}^{(t_j - t_{j-1})\eta(u_j)} \mid \mathcal{F}_{T+t_{n-1}}^X\right)\right),
\end{aligned}$$

    wobei sich die letzte Gleichheit aus Teilaussage 5 von Theorem 2.11 ergibt.
  - Aus Lemma 3.9 und aus Teilaussage 4 von Theorem 2.11 ergibt sich nun, dass

$$\begin{aligned}
&\mathbb{E}\left(\mathbb{E}\left(\mathbb{1}(A) \prod_{j=1}^n \frac{\widetilde{Y}_{T+t_j}^{(j)}}{\widetilde{Y}_{T+t_{j-1}}^{(j)}} \mathrm{e}^{(t_j - t_{j-1})\eta(u_j)} \mid \mathcal{F}_{T+t_{n-1}}^X\right)\right) \\
&= \mathbb{E}\left(\mathbb{1}(A) \left(\prod_{j=1}^{n-1} \frac{\widetilde{Y}_{T+t_j}^{(j)}}{\widetilde{Y}_{T+t_{j-1}}^{(j)}} \mathrm{e}^{(t_j - t_{j-1})\eta(u_j)}\right) \frac{\mathrm{e}^{(t_n - t_{n-1})\eta(u_n)}}{\widetilde{Y}_{T+t_{n-1}}^{(n)}} \mathbb{E}\left(\widetilde{Y}_{T+t_n}^{(n)} \mid \mathcal{F}_{T+t_{n-1}}^X\right)\right) \\
&= \mathbb{E}\left(\mathbb{1}(A) \left(\prod_{j=1}^{n-1} \frac{\widetilde{Y}_{T+t_j}^{(j)}}{\widetilde{Y}_{T+t_{j-1}}^{(j)}} \mathrm{e}^{(t_j - t_{j-1})\eta(u_j)}\right) \mathrm{e}^{(t_n - t_{n-1})\eta(u_n)}\right),
\end{aligned}$$

    weil aus dem optionalen Sampling–Theorem in Korollar 3.4 folgt, dass

$$\mathbb{E}\left(\widetilde{Y}_{T+t_n}^{(n)} \mid \mathcal{F}_{T+t_{n-1}}^X\right) = \widetilde{Y}_{T+t_{n-1}}^{(n)},$$

    wobei der Zeitpunkt $t \geq 0$ in der Formel (28) des Korollars 3.4 so zu wählen ist, dass $c + t_n \leq t$.
  - Durch Iteration dieser Überlegungen ergibt sich dann insgesamt, dass

$$\begin{aligned}
\mathbb{E}\left(\mathbb{1}(A)\, \exp\big(\mathrm{i} \sum_{j=1}^n u_j (Y_{t_j} - Y_{t_{j-1}})\big)\right) &= \mathbb{E}\left(\mathbb{1}(A)\, \exp\big(\mathrm{i} \sum_{j=1}^n u_j (X_{T+t_j} - X_{T+t_{j-1}})\big)\right) \\
&= P(A) \prod_{j=1}^n \mathrm{e}^{(t_j - t_{j-1})\eta(u_j)} = P(A)\, \mathbb{E}\left(\exp\big(\mathrm{i} \sum_{j=1}^n u_j (X_{t_j} - X_{t_{j-1}})\big)\right),
\end{aligned}$$

    d.h., für beliebige $A \in \mathcal{F}_T^X$ und $t_0, t_1, \ldots, t_n \geq 0$ mit $t_0 < t_1 < \ldots < t_n$ gilt

$$\mathbb{E}\left(\mathbb{1}(A)\, \exp\big(\mathrm{i} \sum_{j=1}^n u_j (Y_{t_j} - Y_{t_{j-1}})\big)\right) = P(A)\, \mathbb{E}\left(\exp\big(\mathrm{i} \sum_{j=1}^n u_j (X_{t_j} - X_{t_{j-1}})\big)\right). \tag{1}$$

- Wir zeigen nun, dass (1) auch für endliche (nicht notwendig beschränkte) Stoppzeiten gilt.

- Sei also $T : \Omega \to [0, \infty)$ eine beliebige endliche Stoppzeit.
- Man kann sich leicht überlegen, dass dann $A \cap \{T \le k\} \in \mathcal{F}^X_{T\wedge k}$ für beliebige $k \ge 1$ und $A \in \mathcal{F}^X_T$.
- Somit ergibt sich aus (1), dass

$$\mathbb{E}\left(\mathbb{1}(A_k)\, \exp\big(\mathrm{i}\sum_{j=1}^n u_j(Y_{k,t_j} - Y_{k,t_{j-1}})\big)\right) = P(A_k)\, \mathbb{E}\left(\exp\big(\mathrm{i}\sum_{j=1}^n u_j(X_{t_j} - X_{t_{j-1}})\big)\right), \tag{2}$$

  wobei $A_k = A \cap \{T \le k\}$ und $Y_{k,t} = X_{(T\wedge k)+t} - X_{T\wedge k}$.
- Durch Grenzübergang für $k \to \infty$ auf beiden Seiten von (2) ergibt sich nun mit Hilfe des Satzes von Lebesgue über die majorisierte Konvergenz, dass (1) auch für beliebige endliche Stoppzeiten gilt.

- Für $A = \Omega$ folgt aus (1), dass $\{Y_t\}$ ein Lévy–Prozess mit der gleichen Verteilung wie $\{X_t\}$ ist.
- Außerdem ergibt sich aus Lemma 3.9, dass der Lévy–Prozess $\{Y_t\}$ adaptiert ist bezüglich der Filtration $\{\mathcal{G}_t,\, t \ge 0\}$ mit $\mathcal{G}_t = \mathcal{F}^X_{T+t}$.
- Es bleibt also noch zu zeigen, dass $\{Y_t\}$ unabhängig von $\mathcal{F}^X_T$ ist, d.h., dass
  - für jede Folge $t_1, \ldots, t_n$ von nichtnegativen Zahlen und für beliebige $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ und $A \in \mathcal{F}^X_T$
  - die Ereignisse $Y_{t_1}^{-1}(B_1) \cap \ldots \cap Y_{t_n}^{-1}(B_n)$ und A unabhängig sind.
- Hierfür betrachten wir die bedingte Verteilung $P_{Y_{t_1},\ldots,Y_{t_n}|\mathcal{F}^X_T} : \mathcal{B}(\mathbb{R}^n) \times \Omega \to [0,1]$ mit

$$P_{Y_{t_1},\ldots,Y_{t_n}|\mathcal{F}^X_T}(B,\omega) = P\big((Y_{t_1},\ldots,Y_{t_n})^{-1}(B) \mid \mathcal{F}^X_T\big)(\omega) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^n),\ \omega \in \Omega\,, \tag{3}$$

  wobei $P(A \mid \mathcal{F}^X_T) = \mathbb{E}\,(\mathbb{1}(A) \mid \mathcal{F}^X_T)$.
  - Man kann sich leicht überlegen, dass (3) impliziert, dass mit Wahrscheinlichkeit 1

$$\mathbb{E}\left(g(Y_{t_1},\ldots,Y_{t_n}) \mid \mathcal{F}^X_T\right)(\omega) = \int_{\mathbb{R}^n} g(y)\, P_{Y_{t_1},\ldots,Y_{t_n}|\mathcal{F}^X_T}(\mathrm{d}y,\omega) \tag{4}$$

    für jede Borel–messbare Funktion $g : \mathbb{R}^n \to \mathbb{C}$ mit $\mathbb{E}\left|g(Y_{t_1},\ldots,Y_{t_n})\right| < \infty$.
  - Außerdem folgt aus (1), dass für beliebige $A \in \mathcal{F}^X_T$ und $t_1, \ldots, t_n \ge 0$

$$\mathbb{E}\left(\mathbb{1}(A)\, \exp\big(\mathrm{i}\sum_{j=1}^n u_j Y_{t_j}\big)\right) = P(A)\, \mathbb{E}\left(\exp\big(\mathrm{i}\sum_{j=1}^n u_j Y_{t_j}\big)\right).$$

  - Hieraus und aus (4) ergibt sich, dass die charakteristischen Funktionen von $P_{Y_{t_1},\ldots,Y_{t_n}|\mathcal{F}^X_T}$ und $P_{Y_{t_1},\ldots,Y_{t_n}}$ mit Wahrscheinlichkeit 1 übereinstimmen.
  - Aus dem Eindeutigkeitssatz für charakteristische Funktionen folgt nun, dass auch die Verteilungen $P_{Y_{t_1},\ldots,Y_{t_n}|\mathcal{F}^X_T}$ und $P_{Y_{t_1},\ldots,Y_{t_n}}$ mit Wahrscheinlichkeit 1 übereinstimmen.
  - Dies ist gleichbedeutend damit, dass für beliebige $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ und $A \in \mathcal{F}^X_T$ die Ereignisse $Y_{t_1}^{-1}(B_1) \cap \ldots \cap Y_{t_n}^{-1}(B_n)$ und A unabhängig sind. □

3.3.2 Reflexionsprinzip des Wiener–Prozesses; Verteilung des Maximums

Mit Hilfe von Theorem 3.14 leiten wir nun das folgende *Reflexionsprinzip* des Wiener–Prozesses her. Es betrifft eine weitere Invarianzeigenschaft des Wiener–Prozesses (zusätzlich zu den bereits in Theorem 2.24 hergeleiteten Eigenschaften dieses Typs).

Theorem 3.15

- *Sei $\{X_t,\, t \geq 0\}$ ein (Standard–) Wiener–Prozess über $(\Omega, \mathcal{F}, P)$ und sei $T : \Omega \to [0, \infty)$ eine endliche Stoppzeit bezüglich der natürlichen Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$ von $\{X_t\}$.*
- *Dann hat $\{X_t\}$ die gleiche Verteilung wie der „reflektierte Prozess" $\{Y_t,\, t \geq 0\}$ mit*

$$Y_t = X_{T \wedge t} - (X_t - X_{T \wedge t})\,, \tag{5}$$

d.h., $\{Y_t\}$ ist ebenfalls ein (Standard–) Wiener–Prozess.

Beweis

- Die stochastischen Prozesse $\{X'_t,\, t \geq 0\}$ und $\{\widetilde{X}_t,\, t \geq 0\}$ seien gegeben durch

$$X'_t = X_{T \wedge t} \qquad \text{bzw.} \qquad \widetilde{X}_t = X_{T+t} - X_T\,.$$

- Aus Theorem 3.14 folgt dann, dass
  - $\{\widetilde{X}_t\}$ ein Wiener–Prozess ist, der unabhängig von $(T, \{X'_t\})$ ist.
  - Außerdem ergibt sich aus Teilaussage 1 von Theorem 2.24, dass $\{\widetilde{X}_t\} \stackrel{\mathrm{d}}{=} \{-\widetilde{X}_t\}$.
  - Insgesamt gilt also

$$\big(T, \{X'_t\}, \{\widetilde{X}_t\}\big) \stackrel{\mathrm{d}}{=} \big(T, \{X'_t\}, -\{\widetilde{X}_t\}\big)\,. \tag{6}$$

- Es genügt nun zu beachten,
  - dass $X_t = X'_t + \widetilde{X}_{(t-T)_+}$ und $Y_t = X'_t - \widetilde{X}_{(t-T)_+}$ für jedes $t \geq 0$, wobei $x_+ = \max\{x, 0\}$.
  - D.h., X_t bzw. Y_t sind messbare Abbildungen von $\big(T, \{X'_t\}, \{\widetilde{X}_t\}\big)$ bzw. $\big(T, \{X'_t\}, -\{\widetilde{X}_t\}\big)$.
  - Wegen (6) gilt somit $\{X_t\} \stackrel{\mathrm{d}}{=} \{Y_t\}$. □

Beachte

- Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess. Die natürliche Filtration $\{\mathcal{F}_t^X,\, t \geq 0\}$ von $\{X_t\}$ ist dann gemäß Theorem 3.13 rechtsstetig.
- Aus Theorem 3.6 folgt somit, dass für jedes $z > 0$ die Ersterreichungszeit $T_{\{z\}}^X = \min\{t \geq 0 :\, X_t = z\}$ eine Stoppzeit bezüglich $\{\mathcal{F}_t^X\}$ ist.
- Außerdem ergibt sich aus Korollar 2.5, dass $P(T_{\{z\}}^X < \infty) = 1$, d.h., $T_{\{z\}}^X$ ist eine endliche Stoppzeit.

Die Anwendung von Theorem 3.15 auf die Ersterreichungszeit $T_{\{z\}}^X$ führt nun zu der folgenden Identität.

Korollar 3.5 *Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess und sei $\{M_t,\, t \geq 0\}$ der Maximum–Prozess mit*

$$M_t = \max_{s \in [0,t]} X_s \qquad \forall\, t \geq 0\,. \tag{7}$$

Dann gilt für beliebige $y \geq 0$ und $z > 0$

$$P(X_t < z - y,\, M_t \geq z) = P(X_t > y + z)\,. \tag{8}$$

Beweis

- Die in (7) betrachtete Abbildung $M_t : \Omega \to [0, \infty)$ ist eine wohldefinierte Zufallsvariable, weil die Trajektorien von $\{X_t\}$ stetige Funktionen sind, vgl. auch die Formel (24) in Abschnitt 2.4.5.

– Für die endliche Stoppzeit $T = T^X_{\{z\}} = \min\{t \geq 0 : X_t = z\}$ ergibt sich dann aus Theorem 3.15, dass

$$\{X_t\} \stackrel{\mathrm{d}}{=} \{Y_t\} \qquad \text{bzw.} \qquad \big(T^X_{\{z\}},\, \{X_t\}\big) \stackrel{\mathrm{d}}{=} \big(T^Y_{\{z\}},\, \{Y_t\}\big)\,,$$

– wobei $T^Y_{\{z\}} = \min\{t \geq 0 : Y_t = z\}$ die Ersterreichungszeit des Niveaus z durch den in (5) eingeführten Wiener–Prozess $\{Y_t\}$ ist.

– Hieraus folgt, dass

$$P(T^X_{\{z\}} \leq t,\, X_t < z - y) = P(T^Y_{\{z\}} \leq t,\, Y_t < z - y)\,, \tag{9}$$

- Weil außerdem $T^X_{\{z\}}(\omega) = T^Y_{\{z\}}(\omega)$ für jedes $\omega \in \Omega$, ergibt sich aus (5), dass

$$\{T^Y_{\{z\}} \leq t\} \cap \{Y_t < z - y\} = \{T^X_{\{z\}} \leq t\} \cap \{2z - X_t < z - y\}\,.$$

– Hieraus und aus (9) folgt nun, dass

$$P(T^X_{\{z\}} \leq t,\, X_t < z - y) = P(T^X_{\{z\}} \leq t,\, 2z - X_t < z - y) = P(T^X_{\{z\}} \leq t,\, X_t > z + y)\,.$$

– Damit ist die Behauptung (8) beweisen, weil offenbar

$$P(T^X_{\{z\}} \leq t,\, X_t > z + y) = P(X_t > z + y)$$

und

$$P(T^X_{\{z\}} \leq t,\, X_t < z - y) = P(M_t \geq z,\, X_t < z - y)\,. \qquad \square$$

Mit Hilfe von Korollar 3.5 können wir nun zeigen, dass die in Theorem 2.23 hergeleitete Schranke für die Tailfunktion von $M_t = \max_{s\in[0,t]} X_s$ „optimal" ist, d.h. mit der Tailfunktion von M_t übereinstimmt.

Theorem 3.16 *Sei $\{X_t,\, t \in [0,1]\}$ ein Wiener–Prozess. Dann gilt für beliebige $t > 0$ und $x \geq 0$*

$$P(M_t > x) = \sqrt{\frac{2}{\pi t}} \int_x^\infty \mathrm{e}^{-ty^2/2}\, \mathrm{d}y\,. \tag{10}$$

Beweis

- Für $y = 0$ ergibt sich aus Korollar 3.5, dass $P(X_t < z,\, M_t \geq z) = P(X_t > z)$.
- Wenn wir nun auf beiden Seiten dieser Gleichung die Wahrscheinlichkeit $P(X_t > x)$ addieren und dabei berücksichtigen, dass $\{X_t > z\} = \{X_t > z,\, M_t \geq z\}$, ergibt sich

$$P(M_t \geq z) = 2\,P(X_t > z) \quad \text{bzw.} \qquad P(M_t > z) = 2\,P(X_t > z) \qquad \forall\, z \geq 0\,,$$

weil $P(X_t = z) = 0$ für die (normalverteilte) Zufallsvariable X_t und somit auch $P(M_t = z) = 0$ für jedes $z \geq 0$. $\square$

3.3.3 Wiener–Prozesse mit negativer Drift

Sei $\{X_t,\, t \geq 0\}$ ein (Standard–) Wiener–Prozess und für $\mu > 0$ sei $\{Y_t,\, t \geq 0\}$ mit $Y_t = X_t - t\mu$ ein Wiener–Prozess mit negativer Drift.

Wir beweisen nun eine Formel für die Tailfunktion des Supremums $\sup_{t\geq 0} Y_t$, wobei $P(\sup_{t\geq 0} Y_t \geq x)$ für jedes $x \geq 0$ als „Ruinwahrscheinlichkeit" über dem unendlichen Zeithorizont $[0, \infty)$ aufgefasst werden kann.

Dabei zeigen wir insbesondere, dass die Schranke für $P(\sup_{t\geq 0} Y_t \geq x)$ „exakt" ist, die wir bereits in Formel (19) des Abschnittes 3.2.4 hergeleitet hatten.

Theorem 3.17 *Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess. Dann gilt für beliebige $\mu > 0$ und $x \geq 0$*

$$P\big(\sup_{t\geq 0} Y_t > x\big) = \mathrm{e}^{-2\mu x}\,, \qquad \textit{wobei } Y_t = X_t - t\mu. \tag{11}$$

Beweis

- Für beliebige $u, x \geq 0$ betrachten wir das Martingal $\{\mathrm{e}^{uY_t - t(u^2/2-\mu u)},\, t \geq 0\}$ und die (endliche) Stoppzeit
$$T_{\{x\}} \wedge t = \inf\{s \geq 0 :\, Y_s = x\} \wedge t \qquad \forall\, t \geq 0\,.$$
- Aus Korollar 3.4 ergibt sich dann, dass für beliebige $u, t, x \geq 0$
$$\mathbb{E}\,\exp\big(uY_{T_{\{x\}}\wedge t} - (T_{\{x\}} \wedge t)(u^2/2 - \mu u)\big) = \mathbb{E}\,\mathrm{e}^{uY_0} = 1$$
bzw.
$$\mathbb{E}\,\big[\exp\big(ux - T_{\{x\}}(u^2/2-\mu u)\big);\, T_{\{x\}} < t\big] + \mathbb{E}\,\big[\exp\big(uY_t - t(u^2/2-\mu u)\big);\, T_{\{x\}} \geq t\big] = 1\,. \tag{12}$$
- Weil aus Korollar 2.4 folgt, dass $\lim_{t\to\infty} Y_t = -\infty$ mit Wahrscheinlichkeit 1, gilt für jedes $u \geq 2\mu$
$$\lim_{t\to\infty}\Big(\exp\big(uY_t - t(u^2/2-\mu u)\big)\; \mathbb{1}\big(T_{\{x\}} \geq t\big)\Big) = 0$$
mit Wahrscheinlichkeit 1.
- Außerdem gilt
$$\exp\big(uY_t - t(u^2/2-\mu u)\big)\; \mathbb{1}\big(T_{\{x\}} \geq t\big) \leq \exp\big(ux - t(u^2/2-\mu u)\big)\; \mathbb{1}\big(T_{\{x\}} \geq t\big)\,,$$
wobei der letzte Ausdruck eine (bezüglich P) integrierbare Majorante ist.
- Aus dem Satz von Lebesgue über die majorisierte Konvergenz folgt also, dass für jedes $u \geq 2\mu$
$$\lim_{t\to\infty} \mathbb{E}\,\big[\exp\big(uY_t - t(u^2/2-\mu u)\big);\, T_{\{x\}} \geq t\big] = 0\,.$$
- Hieraus und aus (12) ergibt sich nun, dass
$$\mathrm{e}^{-ux} = \mathbb{E}\,\big[\exp\big(-T_{\{x\}}(u^2/2-\mu u)\big);\, T_{\{x\}} < \infty\big] \qquad \forall\, u \geq 2\mu\,.$$
- Insbesondere ergibt sich für $u = 2\mu$, dass für jedes $x \geq 0$
$$\mathrm{e}^{-2\mu x} = P(T_{\{x\}} < \infty) = P\big(\sup_{t\geq 0} Y_t \geq x\big)\,.$$
- Damit gilt auch $P\big(\sup_{t\geq 0} Y_t > x\big) = \mathrm{e}^{-2\mu x}$ für jedes $x \geq 0$. □

3.3.4 Subordinatoren als Prozesse von Ersterreichungszeiten

In diesem Abschnitt diskutieren wir zwei weitere Beispiele, bei dem Theorem 3.15 auf die Ersterreichungszeit $T^X_{\{z\}} = \min\{t \geq 0 :\, X_t = z\}$ des (Standard–) Wiener–Prozesses $\{X_t,\, t \geq 0\}$ angewendet wird; $z \geq 0$.

Zur Erinnerung: Ein Subordinator $\{Y_t,\, t \geq 0\}$ heißt *α–stabiler Lévy–Prozess* mit Stabilitätsindex $\alpha \in (0,1)$, wenn $a = b = 0$ und das Lévy–Maß ν gegeben ist durch

$$\nu(\mathrm{d}y) = \begin{cases} \dfrac{\alpha}{\Gamma(1-\alpha)}\,\dfrac{1}{y^{1+\alpha}}\,\mathrm{d}y & \text{für } y > 0, \\ 0 & \text{für } y \leq 0, \end{cases} \tag{13}$$

wobei $\alpha \in (0,1)$ und $\Gamma : (0,\infty) \to (0,\infty)$ die *Gammafunktion* bezeichnet mit

$$\Gamma(p) = \int_0^\infty \mathrm{e}^{-y}\, y^{p-1}\,\mathrm{d}y\,, \qquad \forall\, p > 0\,.$$

Beachte

- In Formel (41) des Abschnittes 3.1.4 hatten wir gezeigt, dass in diesem Fall

$$\mathbb{E}\,\mathrm{e}^{-uY_t} = \mathrm{e}^{-tu^\alpha} \qquad \forall\, t, u \geq 0\,. \tag{14}$$

- Der α–stabile Lévy–Prozess $\{Y_t\}$ mit $\alpha = 1/2$ wird manchmal *Lévy–Subordinator* genannt.

Theorem 3.18 *Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess. Dann ist der Prozess der Ersterreichungszeiten $\{T_t,\, t \geq 0\}$ mit*

$$T_t = \min\{s \geq 0 :\, X_s = t/\sqrt{2}\} \tag{15}$$

ein Lévy–Prozess, der die gleiche Verteilung wie der Lévy–Subordinator hat.

Beweis

- Wir zeigen zunächst, dass der in (15) gegebene Prozess $\{T_t\}$ ein Lévy–Prozess ist.
  - Offenbar gilt $T_0 = 0$.
  - Aus Theorem 3.14 folgt, dass $\{T_t\}$ unabhängige und stationäre Zuwächse hat.
  - Außerdem ergibt sich aus Theorem 3.16, dass für jedes $\varepsilon > 0$

$$\lim_{t\to 0} P(T_t > \varepsilon) = \lim_{t\to 0} P\big(\max_{s\in[0,\varepsilon]} X_s < t/\sqrt{2}\big) = 0\,,$$

    d.h., $\{T_t\}$ ist stochastisch stetig.
- Es ist nun noch zu zeigen, dass (14) gilt mit $\alpha = 1/2$.
  - Dabei können wir ähnlich wie im Beweis von Theorem 3.17 vorgehen.
  - Für beliebige $t, u \geq 0$ und $n \geq 1$ betrachten wir das Martingal $\{\mathrm{e}^{uX_s - su^2/2},\, s \geq 0\}$ und die Stoppzeit $T_t \wedge n$ bezüglich der natürlichen Filtration $\{\mathcal{F}_s^X,\, s \geq 0\}$ von $\{X_s\}$.
  - Aus Korollar 3.4 ergibt sich dann, dass für beliebige $t, u \geq 0$ und $n \geq 1$

$$\mathbb{E}\,\exp\big(uX_{T_t\wedge n} - (T_t \wedge n)u^2/2\big) = \mathbb{E}\,\mathrm{e}^{uX_0} = 1$$

    bzw.

$$\mathbb{E}\,\big[\exp\big(ut/\sqrt{2} - T_t u^2/2\big);\, T_t < n\big] + \mathbb{E}\,\big[\exp\big(uX_n - nu^2/2\big);\, T_t \geq n\big] = 1\,. \tag{16}$$

  - Weil $P(T_t < \infty) = 1$ und weil $X_n \leq t/\sqrt{2}$, wenn $T_t \geq n$, ergibt sich genauso wie im Beweis von Theorem 3.17, dass

$$\lim_{n\to\infty} \mathbb{E}\,\big[\exp\big(uX_n - nu^2/2\big);\, T_t \geq n\big] = 0 \qquad \forall\, t, u \geq 0\,.$$

  - Hieraus und aus (16) folgt, dass

$$\mathbb{E}\,\exp\big(ut/\sqrt{2} - T_t u^2/2\big) = 1 \qquad \forall\, t, u \geq 0$$

    bzw. mit der Substitution $u' = u^2/2$

$$\mathbb{E}\,\exp\big(t\sqrt{u} - T_t u\big) = 1 \qquad \forall\, t, u \geq 0\,.$$

□

Wir verallgemeinern nun das Modell von Theorem 3.18 und betrachten Prozesse von Ersterreichungszeiten für Wiener–Prozesse mit Drift.

Theorem 3.19 *Sei $\{X_t,\, t \geq 0\}$ ein Wiener–Prozess. Für beliebige $\mu \in \mathbb{R}$ und $\delta > 0$ ist dann der Prozess der Ersterreichungszeiten $\{T_t,\, t \geq 0\}$ mit*

$$T_t = \min\{s \geq 0 :\ X_s + \mu s = \delta t\} \tag{17}$$

ein Lévy–Prozess, wobei

$$\mathbb{E}\, \mathrm{e}^{-uT_t} = \exp\big(-t\delta(\sqrt{2u+\mu^2} - \mu\big) \qquad \forall\, t, u \geq 0\,. \tag{18}$$

Der *Beweis* von Theorem 3.19 verläuft ähnlich wie der Beweis von Theorem 3.18. Er wird deshalb weggelassen.

Beachte

- Für die in Theorem 3.19 betrachteten Ersterreichungszeiten T_t mit der in (18) gegebenen Laplace–Transformierten kann man zeigen, dass für jedes $x \geq 0$

$$P(T_t \leq x) = \frac{\delta t}{\sqrt{2\pi}}\, \mathrm{e}^{\delta t \mu} \int_0^x y^{-3/2} \exp\big(-\frac{1}{2}\,(t^2\delta^2 y^{-1} + \mu^2 y)\big)\, \mathrm{d}y\,. \tag{19}$$

- Man sagt, dass Zufallsvariablen mit der in (19) gegebenen Verteilungsfunktion eine *inverse Gauß–Verteilung* besitzen.

Die in den Theoremen 3.18 bzw. 3.19 betrachteten Subordinatoren $\{T_t,\, t \geq 0\}$ spielen eine wichtige Rolle bei der *Zeitransformation* von Lévy–Prozessen.

Es gilt nämlich die folgende Invarianz–Eigenschaft von Lévy–Prozessen, die wir hier ohne Beweis angeben.

Theorem 3.20

- *Sei $\{X_t,\, t \geq 0\}$ ein Lévy–Prozess und sei $\{Y_t,\, t \geq 0\}$ ein Subordinator, der über dem gleichen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ wie $\{X_t\}$ gegeben ist.*
- *Wenn $\{X_t\}$ und $\{Y_t\}$ unabhängig sind, dann ist der Prozess $\{Z_t,\, t \geq 0\}$ mit*

$$Z_t(\omega) = X_{Y_t(\omega)}(\omega) \qquad \forall\, \omega \in \Omega$$

  ebenfalls ein Lévy–Prozess.

Ein *Beweis* von Theorem 3.20 kann zum Beispiel in Applebaum (2004), S. 53–55 nachgelesen werden.

4 Stochastische Prozesse und Felder mit allgemeineren Indexmengen

In diesem Kapitel betrachten wir Beispiele von stochastischen Prozessen bzw. zufälligen Feldern $\{X_t,\, t \in I\}$, für die die Indexmenge I die gesamte reelle Achse $\mathbb{R}$ oder der d–dimensionale euklidische Raum $\mathbb{R}^d$ für $d \geq 2$ ist.

Darüber hinaus werden wir auch den Fall diskutieren, dass $\{X_t,\, t \in I\}$ ein zufälliges Maß ist, wobei I dann eine σ–Algebra ist. Dabei betrachten wir insbesondere stochastische Prozesse mit stationären Zuwächsen bzw. stationäre zufällige Maße.

4.1 Zählprozesse im $\mathbb{R}^1$

Wir modifizieren das in Abschnitt 2.1 eingeführte Konzept von Zählprozessen $\{N_t, t \geq 0\}$ dahingehend, dass wir Zählprozesse „mit Vergangenheit" betrachten, d.h., die in Abschnitt 2.1 betrachteten Ereigniszeitpunkte S_n liegen jetzt nicht nur in $[0, \infty)$, sondern sie sind auf der gesamten reellen Achse $\mathbb{R}$ verteilt.

Definition

- Sei $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum, und sei $\{S_n,\, n = 0, \pm 1, \pm 2, \ldots\}$ eine (monotone) Folge von reellwertigen Zufallsvariablen über $(\Omega, \mathcal{F}, P)$ mit $\ldots S_{-1} < S_0 \leq 0 < S_1 < S_2 < \ldots$, so dass mit Wahrscheinlichkeit 1

$$\lim_{n\to\infty} S_n = \infty \qquad \text{und} \qquad \lim_{n\to-\infty} S_n = -\infty\,.$$

- Der stochastische Prozess $\{N_t, t \in \mathbb{R}\}$ mit

$$N_t = \begin{cases} \sum_{k=1}^{\infty} \mathbb{I}(S_k \leq t) & \text{für } t \geq 0, \\ -\sum_{k=-\infty}^{0} \mathbb{I}(S_k > t) & \text{für } t < 0, \end{cases} \tag{1}$$

  wird *Zählprozess* genannt. Die Zufallsvariablen S_n heißen *Ereigniszeitpunkte*.

Beachte

- Der in (1) gegebene Zählprozess $\{N_t, t \in \mathbb{R}\}$ ist càdlàg und hat stückweise konstante Trajektorien, deren Unstetigkeitsstellen die Ereigniszeitpunkte S_n sind. Dabei gilt $N_t \geq 0$ für $t \geq 0$ und $N_t \leq 0$ für $t < 0$.
- Durch den Ansatz

$$N_B = \#\{n : S_n \in B\} \qquad \forall\, B \in \mathcal{B}(\mathbb{R}) \tag{2}$$

  ist dann ein zufälliges Zählmaß $\{N_B,\, B \in \mathcal{B}(\mathbb{R})\}$ über $(\Omega, \mathcal{F}, P)$ gegeben.

Definition

- Man sagt, dass der Zählprozess $\{N_t, t \in \mathbb{R}\}$ *stationäre Zuwächse* hat, wenn für beliebige $t_0, \ldots, t_n \in \mathbb{R}$ mit $t_0 \leq \ldots \leq t_n$ die Verteilungen der Zufallsvektoren $(N_{t_1+h} - N_{t_0+h}, \ldots, N_{t_n+h} - N_{t_{n-1}+h})$ nicht von $h > 0$ abhängen.
- Außerdem sagt man, dass $\{N_t, t \in \mathbb{R}\}$ ein *Prozess mit unabhängigen Zuwächsen* ist, wenn die Zufallsvariablen $(N_{t_1} - N_{t_0}, \ldots, N_{t_n} - N_{t_{n-1}})$ für beliebige $t_0, \ldots, t_n \in \mathbb{R}$ mit $t_0 \leq \ldots \leq t_n$ unabhängig sind.
- Das zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}(\mathbb{R})\}$ heißt *stationär*, wenn für beliebige $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ und $h \in \mathbb{R}$ die Verteilungen der Zufallsvektoren $(N_{B_1+h}, \ldots, N_{B_n+h})$ nicht von h abhängen.

Theorem 4.1 *Sei $\{N_t, t \in \mathbb{R}\}$ ein Zählprozess über $(\Omega, \mathcal{F}, P)$, und das Zählmaß $\{N_B,\, B \in \mathcal{B}(\mathbb{R})\}$ sei gegeben durch (2). Der Prozess $\{N_t, t \in \mathbb{R}\}$ hat genau dann stationäre Zuwächse, wenn $\{N_B,\, B \in \mathcal{B}(\mathbb{R})\}$ stationär ist.*

Beweis

- Wenn das zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}(\mathbb{R})\}$ stationär ist, dann gilt für beliebige $h \in \mathbb{R}$ und $t_0, \ldots, t_n \in \mathbb{R}$ mit $t_0 \leq \ldots \leq t_n$

$$\begin{aligned}(N_{t_1+h} - N_{t_0+h}, \ldots, N_{t_n+h} - N_{t_{n-1}+h}) &= (N_{(t_1,t_2]+h}, \ldots, N_{(t_{n-1},t_n]+h}) \\ \stackrel{\mathrm{d}}{=} \quad (N_{(t_1,t_2]}, \ldots, N_{(t_{n-1},t_n]}) &= (N_{t_1} - N_{t_0}, \ldots, N_{t_n} - N_{t_{n-1}})\,,\end{aligned}$$

  d.h., der Zählprozess $\{N_t, t \in \mathbb{R}\}$ hat stationäre Zuwächse.
- Sei nun umgekehrt $\{N_t, t \in \mathbb{R}\}$ ein Zählprozess mit stationären Zuwächsen.
  - Dann ist das Mengensystem $\mathcal{G} \subset \mathcal{B}(\mathbb{R})$ derjenigen Borel–Mengen $B \in \mathcal{B}(\mathbb{R})$, für die

$$N_{B+h} \stackrel{\mathrm{d}}{=} N_B$$

    für jedes $h \in \mathbb{R}$ gilt, ein sogenanntes d–System, dass die Familie aller Intervalle der Form $(t_2, t_1]$ enthält, vgl. Abschnitt WR-3.2.3.
  - Aus dem Satz über monotone Klassen (vgl. Theorem WR-3.2) ergibt sich nun, dass $\mathcal{G} = \mathcal{B}(\mathbb{R})$.
  - Auf ähnliche Weise kann man zeigen, dass

$$(N_{B_1+h}, \ldots, N_{B_n+h}) \stackrel{\mathrm{d}}{=} (N_{B_1}, \ldots, N_{B_n})$$

    für jedes $n \geq 1$, für beliebige Borel–Mengen $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ und für jedes $h \in \mathbb{R}$ gilt. □

4.1.1 Homogener Poisson–Prozess

In diesem Abschnitt betrachten wir zunächst den Fall, dass die „Zwischenankunftszeiten" $T_n = S_n - S_{n-1}$ zwischen den aufeinanderfolgenden Ereigniszeitpunkten $S_1, S_2, \ldots$

- eine Folge $\{T_n,\, n \geq 2\}$ von unabhängigen und identisch verteilten Zufalllsvariablen bilden mit $T_n \sim \mathrm{Exp}(\lambda)$ für ein $\lambda > 0$.
- Außerdem sei S_1 unabhängig von $\{T_n,\, n \geq 2\}$, und es gelte $S_1 \sim \mathrm{Exp}(\lambda)$.
- Dann ist der Zählprozess $\{N_t, t \geq 0\}$ mit $N_t = \sum_{k=1}^{\infty} \mathbb{1}(S_k \leq t)$ ein (homogener) Poisson–Prozess in $[0, \infty)$ mit der Intensität λ; vgl. Abschnitt 2.2.1.

Wir zeigen, wie der Zählprozess $\{N_t, t \geq 0\}$ auf einfache Weise zu einem Zählprozess $\{N_t, t \in \mathbb{R}\}$ auf der gesamten reellen Achse $\mathbb{R}$ fortgesetzt werden kann, so dass $\{N_t, t \in \mathbb{R}\}$ stationäre und unabhängige (poissonverteilte) Zuwächse hat.

Theorem 4.2 *Sei $\{T_n,\, n = 0, \pm 1, \pm 2, \ldots\}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit $T_n \sim \mathrm{Exp}(\lambda)$ für ein $\lambda > 0$. Außerdem sei*

$$S_n = \begin{cases} \sum_{k=1}^{n} T_k & \text{für } n \geq 1, \\ -\sum_{k=n}^{0} T_k & \text{für } n \leq 0. \end{cases} \tag{3}$$

Dann hat der in (1) gegebene Zählprozess $\{N_t, t \in \mathbb{R}\}$ stationäre und unabhängige Zuwächse, wobei $|N_t| \sim \mathrm{Poi}(|t|\lambda)$ für jedes $t \in \mathbb{R}$.

Beweis

- Für jedes $t \geq 0$ sei $N'_t = -N_{-t}$. Mit Hilfe von Theorem 2.9 ergibt sich dann aus (1) und (3), dass
  - die stochastischen Prozesse $\{N_t, t \geq 0\}$ und $\{N'_t, t \geq 0\}$ stationäre und unabhängige Zuwächse besitzen,
  - wobei $N_t \sim \mathrm{Exp}(t\lambda)$ und $N'_t \sim \mathrm{Exp}(t\lambda)$ für jedes $t \geq 0$.
- Außerdem ergibt sich aus (1) und (3), dass
  - die beiden Prozesse $\{N_t, t \geq 0\}$ und $\{N'_t, t \geq 0\}$ unabhängig sind.
  - Hieraus und aus der Faltungsstabilität der Poisson–Verteilung ergibt sich insbesondere, dass

$$N_{t_2} - N_{t_1} = N_{t_2} + N'_{-t_1} \sim \mathrm{Poi}\big((t_2 - t_1)\lambda\big) \qquad \forall\, t_1 < 0,\, t_2 > 0\,.$$

- Damit ist klar, dass für beliebige $t_0, \ldots, t_n \in \mathbb{R}$ mit $t_0 \leq \ldots \leq t_n$ und für $h > 0$ die Komponenten der Zufallsvektoren $(N_{t_1+h} - N_{t_0+h}, \ldots, N_{t_n+h} - N_{t_{n-1}+h})$ unabhängige Zufallsvariablen sind und dass ihre Verteilungen nicht von h abhängen. □

Beachte

- Manchmal wird die in (3) eingeführte Folge von Ereigniszeitpunkten $\{S_n\}$ dahingehend modifiziert, dass der „Ereignispunkt" $S_0 = 0$ zusätzlich hinzufügt wird und die Zufallsvariablen S_n für $n \neq 0$ definiert werden durch

$$S_n = \begin{cases} \sum_{k=1}^{n} T_k & \text{für } n \geq 1, \\ -\sum_{k=n+1}^{0} T_k & \text{für } n < 0. \end{cases} \tag{4}$$

- Dann bilden die Zwischenankunftszeiten $T_n = S_n - S_{n-1}$ eine stationäre Folge $\{T_n\}$ von unabhängigen und identisch (exponentiell) verteilten Zufallsvariablen.
- Andererseits sieht man leicht, dass der in (1) gegebene Zählprozess $\{N_t,\, t \in \mathbb{R}\}$ in diesem Fall keine stationären Zuwächse hat, denn wegen $S_0 = 0$ gilt

$$\lim_{t\uparrow 0}(N_0 - N_t) = -\lim_{t\uparrow 0} N_t = 1 \qquad \text{und} \qquad \lim_{t\downarrow 0}(N_t - N_0) = \lim_{t\downarrow 0} N_t = 0\,.$$

Das folgende Resultat zeigt,

- wie sich das in (4) betrachtete Modell eines (inhomogenen) Zählprozesses vom Poisson–Typ durch Grenzübergang aus einer bedingten Version des in (3) eingeführten Modells des homogenen Poisson–Prozesses ergibt.
- Dabei wird das letztere Modell unter der Bedingung betrachtet, dass in einer (kleinen) ε–Umgebung des Nullpunktes ein Ereigniszeitpunkt liegt.

Hierfür benötigen wir die folgenden Bezeichnungen. Sei $T_{\min} = \min\{T_0, T_1\}$, und für beliebige $t_1, t_2 \in \mathbb{R}$ sei

$$N_{t_1,t_2} = \begin{cases} N_{t_2} - N_{t_1}\,, & \text{falls } t_1 \leq t_2, \\ N_{t_1} - N_{t_2}\,, & \text{falls } t_1 > t_2. \end{cases}$$

Theorem 4.3

- *Sei $\{N_t, t \in \mathbb{R}\}$ der in Theorem 4.2 betrachtete Poisson–Prozess mit stationären und unabhängigen Zuwächsen.*
- *Außerdem sei $\{N'_t, t \in \mathbb{R}\}$ ein Zählprozess mit $S'_0 = 0$, dessen Zwischenankunftszeiten $T'_n = S'_n - S'_{n-1}$ eine stationäre Folge von unabhängigen und identisch verteilten Zufallsvariablen bilden mit $T'_n \sim \mathrm{Exp}(\lambda)$.*

- *Dann gilt für jedes $n = 1, 2, \ldots$, für beliebige $t_1, \ldots, t_n \in \mathbb{R}$ mit $t_1 \leq \ldots \leq t_n$ und für beliebige $k_1, \ldots, k_n \geq 0$*

$$\begin{aligned} P(N'_{t_1} \leq k_1, \ldots, N'_{t_n} \leq k_n) &= \lim_{\varepsilon \downarrow 0} \Big(P\big(N_{t_1 - T_0, -T_0} \leq k_1, \ldots, N_{t_n - T_0, -T_0} \leq k_n, T_0 < T_1 \mid T_{\min} \leq \varepsilon\big) \\ &\quad + P\big(N_{t_1 + T_1, T_1} \leq k_1, \ldots, N_{t_n + T_1, T_1} \leq k_n, T_0 > T_1 \mid T_{\min} \leq \varepsilon\big)\Big) . \end{aligned} \tag{5}$$

Beweis

- Wir zeigen die Gültigkeit von (5) nur für $n = 1$.
  - Für beliebige $t \geq 0$ und $k \geq 0$ gilt dann

$$\begin{aligned} &\lim_{\varepsilon \downarrow 0} \Big(P\big(N_{t - T_0, -T_0} \leq k, T_0 < T_1 \mid T_{\min} \leq \varepsilon\big) + P\big(N_{t + T_1, T_1} \leq k, T_0 > T_1 \mid T_{\min} \leq \varepsilon\big)\Big) \\ &= \lim_{\varepsilon \downarrow 0} \Big(P\big(T_0 + \ldots + T_{k+1} > t, T_0 < T_1 \mid T_{\min} \leq \varepsilon\big) \\ &\qquad + P\big(T_2 + \ldots + T_{k+2} > t, T_0 > T_1 \mid T_{\min} \leq \varepsilon\big)\Big) \\ &= \lim_{\varepsilon \downarrow 0} \Big(\frac{P(T_0 \leq \varepsilon)}{P(T_{\min} \leq \varepsilon)} \, P\big(T_0 + \ldots + T_{k+1} > t, T_0 < T_1 \mid T_0 \leq \varepsilon\big) \\ &\qquad + \lim_{\varepsilon \downarrow 0} \Big(\frac{P(T_1 \leq \varepsilon)}{P(T_{\min} \leq \varepsilon)} \, P\big(T_2 + \ldots + T_{k+2} > t, T_0 > T_1 \mid T_1 \leq \varepsilon\big)\Big) \\ &= \frac{1}{2} \, P\big(T_1 + \ldots + T_{k+1} > t\big) + \frac{1}{2} \, P\big(T_2 + \ldots + T_{k+2} > t\big) \\ &= P\big(T_1 + \ldots + T_{k+1} > t\big) = P(N'_t \leq k) . \end{aligned}$$

  - Auf die gleiche Weise ergibt sich, dass für $t < 0$ und $k \geq 0$

$$\lim_{\varepsilon \downarrow 0} \Big(P\big(N_{t - T_0, -T_0} \leq k, T_0 < T_1 \mid T_{\min} \leq \varepsilon\big) + P\big(N_{t + T_1, T_1} \leq k, T_0 > T_1 \mid T_{\min} \leq \varepsilon\big)\Big) = P(N'_t \leq k) .$$

- Die Gültigkeit von (5) für $n \geq 2$ ergibt sich durch ähnliche Überlegungen. □

4.1.2 Allgemeines Konstruktionsprinzip für Zählprozesse mit stationären Zuwächsen

- Der am Ende von Abschnitt 4.1.1 erwähnte Zuhammenhang
  - zwischen (Poissonschen) Zählprozessen mit stationären Zuwächsen und stationären Folgen von (unabhängigen und exponentiell verteilten) Zwischenankunftszeiten
  - kann als Spezialfall eines wesentlich allgemeineren Szenarios aufgefasst werden.
- Dabei gehen wir nun umgekehrt vor und setzen voraus,
  - dass $S_0 = 0$ gilt und
  - dass die Zwischenankunftszeiten $T_n = S_n - S_{n-1}$ zwischen den aufeinanderfolgenden Ereigniszeitpunkten S_{n-1} und S_n eine beliebige stationäre Folge $\{T_n\}$ von (nichtnotwendig unabhängigen) positiven Zufallsvariablen über einem gewissen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ bilden.

Es gelte also

$$\{T_n\} \stackrel{\mathrm{d}}{=} \{T_{n+1}\} \tag{6}$$

und damit auch $\{T_n\} \stackrel{\mathrm{d}}{=} \{T_{n+n_0}\}$ für jede natürliche Zahl $n_0 \geq 1$. Außerdem setzen wir in diesem Abschnitt lediglich voraus, dass

$$0 < \mu = \mathbb{E}\, T_n < \infty . \tag{7}$$

Beachte

- Es ist klar, dass die Invarianzeigenschaft (6) insbesondere dann gilt, wenn $\{T_n\}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen ist.
- Weitere Beispiele von stationären Folgen $\{T_n\}$, die *nicht* aus unabhängigen Zufallsvariablen bestehen, lassen sich mit Hilfe der Sprungzeitpunkte von reversiblen Markow–Prozessen konstruieren, die ihre Werte in dem endlichen Zustandsraum $E = \{1, \ldots, \ell\}$ annehmen, wobei $\ell > 1$ eine beliebige, jedoch vorgegebene natürliche Zahl ist; vgl. Abschnitt 4.1.4.

Wir zeigen nun, wie man ausgehend von einer stationären Folge $\{T_n\}$ von positiven Zufallsvariablen über $(\Omega, \mathcal{F}, P)$ einen Wahrscheinlichkeitsraum $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ und einen Zählprozess $\{\widetilde{N}_t,\ t \in \mathbb{R}\}$ über $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ konstruieren kann, so dass $\{\widetilde{N}_t\}$ stationäre Zuwächse hat.

Hierfür führen wir zunächst einige Bezeichnungen ein.

- Mit $\mathbb{K} = \prod_{n=-\infty}^{\infty}(0, \infty)$ bezeichnen wir den Produktraum, in dem die zufällige Folge $\{T_n\}$ ihre Werte annimmt,
  - und mit $\mathcal{B}(\mathbb{K}) = \bigotimes_{n=-\infty}^{\infty} \mathcal{B}(0, \infty)$ die σ–Algebra der Borel–Mengen in $\mathbb{K}$.
  - Die Verteilung von $\{T_n\}$ über $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$ bezeichnen wir mit Q.
  - Außerdem betrachten wir den *Verschiebungsoperator* $\mathbf{U} : \mathbb{K} \to \mathbb{K}$ mit $\mathbf{U}(\{t_n\}) = \{t_{n-1}\}$.
- Weil $\{T_n\}$ stationär ist, ergibt sich aus (6), dass die Verteilung Q von $\{T_n\}$ invariant bezüglich $\mathbf{U}$ ist, d.h., es gilt
$$Q(\mathbf{U}(B)) = Q(B) \qquad \forall\, B \in \mathcal{B}(\mathbb{K})\,, \tag{8}$$
wobei $\mathbf{U}(B) = \big\{\mathbf{U}(\mathbf{t}) :\ \mathbf{t} = \{t_n\} \in B\big\}$.
- Ausgehend von der „folgenstationären" Verteilung Q über $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$, die die Invarianzeigenschaft (8) besitzt, kann man nun einen Zählprozess $\{\widetilde{N}_t,\ t \in \mathbb{R}\}$ mit stationären Zuwächsen konstruieren.
  - Hierfür sei $\mathbb{L}$ die Menge aller Folgen $\mathbf{s} = \{s_n,\ n = 0, \pm 1, \pm 2, \ldots\}$ reeller Zahlen, so dass
$$\ldots s_{-1} < s_0 \le 0 < s_1 < s_2 < \ldots \qquad \text{und} \qquad \lim_{|n| \to \infty} |s_n| = \infty\,,$$
  - Mit $\mathcal{B}(\mathbb{L})$ bezeichnen wir die σ–Algebra der Borel–Mengen in $\mathbb{L}$.
- Außerdem betrachten wir für jedes $h \in \mathbb{R}$ den (zeitkontinuierlichen) *Verschiebungsoperator* $\mathbf{T}_h : \mathbb{L} \to \mathbb{L}$
$$\mathbf{T}_h(\{s_n\}) = \{s_n - h\} \qquad \forall\, \{s_n\} \in \mathbb{L}\,, \tag{9}$$
der die „Punkte" s_n von $\mathbf{s} = \{s_n\} \in \mathbb{L}$ um h Längeneinheiten nach links bzw. den Ursprung um h Längeneinheiten nach rechts verschiebt, falls $h > 0$.
- Schließlich sei $\mathbb{L}_0 = \big\{\mathbf{s} = \{s_n\} \in \mathbb{L} :\ s_0 = 0\big\}$ und für jedes $\mathbf{t} = \{t_n\} \in \mathbb{K}$ sei die Abbildung
$$\mathbf{t} \mapsto \mathbf{s}^{(\mathbf{t})} = \{s_n^{(\mathbf{t})}\} \in \mathbb{L}_0$$
gegeben durch
$$s_n^{(\mathbf{t})} = \begin{cases} \sum_{k=1}^{n} t_k & \text{für } n \ge 1, \\ 0 & \text{für } n = 0, \\ -\sum_{k=n+1}^{0} t_k & \text{für } n < 0. \end{cases}$$

Definition Ein Wahrscheinlichkeitsmaß P über $(\mathbb{L}, \mathcal{B}(\mathbb{L}))$ heißt $\mathbf{T}$–invariant, wenn

$$P(\mathbf{T}_h(A)) = P(A) \qquad \forall\, h \in \mathbb{R},\ A \in \mathcal{B}(\mathbb{L})\,, \tag{10}$$

wobei $\mathbf{T}_h(A) = \{\mathbf{T}_h(\mathbf{s}) :\ \mathbf{s} \in A\}$.

Theorem 4.4 *Sei Q ein $\mathbf{U}$-invariantes Wahrscheinlichkeitsmaß über $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$ mit $0 < \mu = \int_{\mathbb{K}} t_1\, Q(\mathrm{d}t) < \infty$. Dann ist durch den Ansatz*

$$P_Q(A) = \frac{1}{\mu} \int_{\mathbb{K}} \int_0^{t_1} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t}) \qquad \forall\, A \in \mathcal{B}(\mathbb{L}) \tag{11}$$

ein $\mathbf{T}$–invariantes Wahrscheinlichkeitsmaß P_Q über $(\mathbb{L}, \mathcal{B}(\mathbb{L}))$ gegeben.

Beweis

- Für beliebige $\mathbf{t} = \{t_n\} \in \mathbb{K}$, $A \in \mathcal{B}(\mathbb{L})$ und $j \geq 0$ gilt offenbar

$$\int_0^{\mathbf{U}^j(\mathbf{t})_1} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{U}^j(\mathbf{t}))})\big)\, \mathrm{d}h = \int_{\sum_{k=1}^{j} t_k}^{\sum_{k=1}^{j+1} t_k} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\,. \tag{12}$$

- Weil Q ein $\mathbf{U}$-invariantes Wahrscheinlichkeitsmaß ist, ergibt sich aus (11) und (12), dass für jedes $n \geq 0$

$$\begin{aligned}
P_Q(A) &= \int_0^{t_1} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t}) \\
&= \frac{1}{(n+1)\mu} \sum_{j=0}^{n} \int_{\mathbb{K}} \int_0^{\mathbf{U}^j(\mathbf{t})_1} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{U}^j(\mathbf{t}))})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t}) \\
&= \frac{1}{(n+1)\mu} \sum_{j=0}^{n} \int_{\mathbb{K}} \int_{\sum_{k=1}^{j} t_k}^{\sum_{k=1}^{j+1} t_k} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t}) \\
&= \frac{1}{(n+1)\mu} \int_{\mathbb{K}} \int_0^{\sum_{k=1}^{n+1} t_k} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t})\,.
\end{aligned}$$

- Hieraus folgt, dass für jedes $t \in \mathbb{R}$

$$\begin{aligned}
P_Q(\mathbf{T}_t(A)) &= \frac{1}{(n+1)\mu} \int_{\mathbb{K}} \int_0^{\sum_{k=1}^{n+1} t_k} \mathbb{I}_{\mathbf{T}_t(A)}\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t}) \\
&= \frac{1}{(n+1)\mu} \int_{\mathbb{K}} \int_0^{\sum_{k=1}^{n+1} t_k} \mathbb{I}_A\big(\mathbf{T}_{h-t}(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t}) \\
&= \frac{1}{(n+1)\mu} \int_{\mathbb{K}} \int_t^{\sum_{k=1}^{n+1} t_k + t} \mathbb{I}_A\big(\mathbf{T}_h(\mathbf{s}^{(\mathbf{t})})\big)\, \mathrm{d}h\, Q(\mathrm{d}\mathbf{t})\,.
\end{aligned}$$

- Insgesamt ergibt sich also die Abschätzung

$$|P_Q(A) - P_Q(\mathbf{T}_t(A))| \leq \frac{2|t|}{(n+1)\mu} \qquad \forall\, n \geq 0\,.$$

- Weil die linke Seite des letzten Ausdruckes nicht von n abhängt und n beliebig groß gewählt werden kann, gilt also $P_Q(A) = P_Q(\mathbf{T}_t(A))$ für beliebige $A \in \mathcal{B}(\mathbb{L})$ und $t \in \mathbb{R}$. □

Beachte Durch Vertauschung der Integrationsreihenfolge ergibt sich, dass (11) äquivalent ist mit

$$P_Q(A) = \frac{1}{\mu} \int_0^\infty Q\big(\mathbf{t} : t_1 > h, \mathbf{T}_h(\mathbf{s}^{(\mathbf{t})}) \in A\big)\, \mathrm{d}h \qquad \forall\, A \in \mathcal{B}(\mathbb{L})\,. \tag{13}$$

Korollar 4.1

- *Die Folge von Zufallsvariablen $\{\widetilde{S}_n\}$ sei über dem kanonischen Wahrscheinlichkeitsraum $(\mathbb{L}, \mathcal{B}(\mathbb{L}), P_Q)$ definiert, wobei $\widetilde{S}_k(\{s_n\}) = s_k$ für jedes $k \in \{0, \pm 1, \pm 2, \ldots\}$ und das Wahrscheinlichkeitsmaß P_Q durch* (11) *gegeben ist.*
- *Durch den Ansatz*

$$\widetilde{N}_t = \begin{cases} \sum_{k=1}^{\infty} \mathbb{I}(\widetilde{S}_k \le t) & \text{für } t \ge 0, \\ -\sum_{k=-\infty}^{0} \mathbb{I}(\widetilde{S}_k > t) & \text{für } t < 0 \end{cases} \tag{14}$$

ist dann ein Zählprozess $\{\widetilde{N}_t, t \in \mathbb{R}\}$ über $(\mathbb{L}, \mathcal{B}(\mathbb{L}), P_Q)$ gegeben, der stationäre Zuwächse hat.

Beweis Die Behauptung ergibt sich unmittelbar aus der in Theorem 4.4 gezeigten $\mathbf{T}$–Invarianz des Wahrscheinlichkeitsmaßes P_Q. □

Beachte Für den in Abschnitt 4.1.1 betrachteten Spezialfall, bei dem die Zwischenankunftszeiten $T_n = S_n - S_{n-1}$ eine Folge von unabhängigen und identisch (exponentiell) verteilten Zufallsvariablen bilden, kann man sich leicht überlegen, dass dann der in (14) gegebene Zählprozess $\{\widetilde{N}_t, t \in \mathbb{R}\}$ ein homogener Poisson–Prozess ist.

4.1.3 Erneuerungsprozesse mit stationären Zuwächsen

Wir setzen nun voraus, dass $S_0 = 0$ und dass die Zwischenankunftszeiten $T_n = S_n - S_{n-1}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit einer *beliebigen* Verteilungsfunktion F bilden.

- Sei $\{\widetilde{S}_n\}$ die in Korollar 4.1 betrachtete Folge von zufälligen Zeitpunkten mit

$$\ldots \widetilde{S}_{-1} < \widetilde{S}_0 < 0 < \widetilde{S}_1 < \widetilde{S}_2 < \ldots .$$

- Wir zeigen, dass dann
  - die neuen „Zwischenankunftszeiten" $\widetilde{T}_n = \widetilde{S}_n - \widetilde{S}_{n-1}$ für $n \neq 1$ ebenfalls eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit der gleichen Verteilungsfunktion F wie T_n bilden,
  - wobei jedoch die Abstände $\widetilde{S}_1$ und $-\widetilde{S}_0$ vom Nullpunkt zum kleinsten positiven Zeitpunkt $\widetilde{S}_1$ bzw. zum größten negativen Zeitpunkt $\widetilde{S}_0$ eine besondere Rolle spielen,
  - denn im allgemeinen müssen die Zufallsvariablen $-\widetilde{S}_0$ und $\widetilde{S}_1$ weder unabhängig sein noch die Verteilungsfunktion F der Abstände $\widetilde{T}_n$ mit $n \notin \{0, 1\}$ besitzen.

Beachte

- Wenn die Zwischenankunftszeiten $T_n = S_n - S_{n-1}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen bilden, dann sagt man, dass der in (1) gegebene stochastische Prozess $\{N_t, t \in \mathbb{R}\}$ ein *gewöhnlicher Erneuerungszählprozess* bzw. kurz ein *Erneuerungsprozess* ist, wobei S_n der n-te *Erneuerungszeitpunkt* heißt.
- Außerdem sagt man, dass der in (14) gegebene Zählprozess $\{\widetilde{N}_t, t \in \mathbb{R}\}$ mit stationären Zuwächsen ein *verzögerter Erneuerungsprozess* ist.

Theorem 4.5

- *Die Verteilung Q von $\{T_n\}$ über $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$ sei das Produktmaß, das durch die (eindimensionale Rand–) Verteilungsfunktion F erzeugt wird, so dass $0 < \mu = \int_0^\infty (1 - F(y))\,\mathrm{d}y < \infty$.*
- *Für die in Korollar 4.1 betrachtete Folge $\{\widetilde{S}_n\}$ von zufälligen Zeitpunkten gilt dann:*
  1. *Für $n \neq 1$ sind die Zufallsvariablen $\widetilde{T}_n = \widetilde{S}_n - \widetilde{S}_{n-1}$ unabhängig und identisch verteilt mit der Verteilungsfunktion F.*
  2. *Die zufällige Folge $\{\widetilde{T}_n,\, n \neq 1\}$ und der Zufallsvektor $(-\widetilde{S}_0, \widetilde{S}_1)$ sind unabhängig.*
  3. *Die gemeinsame Verteilung von $(-\widetilde{S}_0, \widetilde{S}_1)$ ist gegeben durch*

$$P_Q(-\widetilde{S}_0 > u,\, \widetilde{S}_1 > v) = \frac{1}{\mu} \int_{u+v}^\infty (1 - F(y)\,\mathrm{d}y \qquad \forall\, u, v \geq 0\,. \tag{15}$$

Beweis

- Für $k \geq 1$, für beliebige $u, v, u_1, \ldots, u_k \geq 0$ und für beliebige $n_1, \ldots, n_k \neq 1$ mit $n_1 < \ldots < n_k$ sei

$$A = \{-\widetilde{S}_0 > u,\, \widetilde{S}_1 > v,\, \widetilde{S}_{n_1} - \widetilde{S}_{n_1-1} > u_1,\, \ldots,\, \widetilde{S}_{n_k} - \widetilde{S}_{n_k-1} > u_k\}\,.$$

- Dann ergibt sich aus (13), dass

$$\begin{aligned}
&P_Q\big(-\widetilde{S}_0 > u,\, \widetilde{S}_1 > v,\, \widetilde{S}_{n_1} - \widetilde{S}_{n_1-1} > u_1,\, \ldots,\, \widetilde{S}_{n_k} - \widetilde{S}_{n_k-1} > u_k\big)\\
&= \frac{1}{\mu} \int_0^\infty Q\big(\mathbf{t} : t_1 > h,\, \mathbf{T}_h(\mathbf{s}^{(\mathbf{t})}) \in A\big)\,\mathrm{d}h\\
&= \frac{1}{\mu} \int_0^\infty Q\big(\mathbf{t} : t_1 > h,\, h > u,\, t_1 - h > v,\, t_{n_1} > u_1,\, \ldots,\, t_{n_k} > u_k\big)\,\mathrm{d}h\\
&= \frac{1}{\mu} \int_0^\infty Q\big(\mathbf{t} : t_1 > h,\, h > u,\, t_1 - h > v\big)\,\mathrm{d}h \prod_{j=1}^k (1 - F(u_j))\\
&= \frac{1}{\mu} \int_{u+v}^\infty (1 - F(h))\,\mathrm{d}h \prod_{j=1}^k (1 - F(u_j))\,,
\end{aligned}$$

  wobei sich die vorletzte Gleichheit aus der Annahme ergibt, dass Q ein Produktmaß über $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$ ist. □

4.1.4 Semi-Markowsche Zählprozesse; Simulationsalgorithmus

Eine weitere Klasse von Zählprozessen in $\mathbb{R}$ mit stationären Zuwächsen ergibt sich, wenn das in Theorem 4.4 betrachtete $\mathbf{U}$-invariante Wahrscheinlichkeitsmaß Q über $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$ die folgende (verallgemeinerte) Produktdarstellung besitzt. Dabei verwenden wir einige grundlegende Begriffe aus der Theorie der zeitdiskreten Markow–Ketten, vgl. das Skript zur Vorlesung „Markow–Ketten und Monte–Carlo–Simulation" im SS 2003.

Theorem 4.6

- *Sei $\ell \geq 1$ eine beliebige natürliche Zahl und sei $E = \{1, \ldots, \ell\}$ ein endlicher Zustandsraum.*
  - *Außerdem sei $\mathbf{P} = (p_{ij})_{i,j=1,\ldots,\ell}$ eine irreduzible und aperiodische $\ell \times \ell$ Matrix mit*

$$p_{ij} \geq 0\,, \qquad \sum_{j=1}^\ell p_{ij} = 1\,, \tag{16}$$

– *sei* $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_\ell)^\top$ *die (eindeutig bestimmte) positive Lösung der Gleichung*

$$\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top \mathbf{P}\,, \tag{17}$$

– *und für beliebige* $i, j \in E$ *sei* $F_{ij} : \mathbb{R} \to [0,1]$ *eine Verteilungsfunktion mit* $F_{ij}(0) = 0$.

- *Die Verteilung* Q *der stationären Folge* $\{T_n\}$ *über* $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$ *sei gegeben durch*

$$Q\big(\mathbf{t} \in \mathbb{K} : t_n \le u_0,\, t_{n+1} \le u_1 \ldots,\, t_{n+m} \le u_m\big) = \sum_{i_0,\ldots,i_m \in E} \pi_{i_0} \prod_{k=0}^{m-1} p_{i_k i_{k+1}}\, F_{i_k i_{k+1}}(u_k) \tag{18}$$

für beliebige $n \in \{0, \pm 1, \pm 2, \ldots\}$, $m \ge 0$ *und* $u_0, \ldots, u_m \ge 0$, *so dass*

$$\mu\Big(= \mathbb{E}\, T_1\Big) = \sum_{i,j \in E} \pi_{i_0} p_{ij} \int_0^\infty (1 - F_{ij}(y))\, \mathrm{d}y < \infty\,.$$

- *Dann gilt für die in Korollar* 4.1 *betrachtete Folge* $\{\widetilde{S}_n\}$ *von zufälligen Zeitpunkten, dass für beliebige* $n < 0$, $m > 0$ *und* $v, u_n, \ldots, u_m \ge 0$

$$\begin{aligned} &P_Q\big(\{-\widetilde{S}_0 > u_0,\, \widetilde{S}_1 > v\} \cap \bigcap_{k \in \{n,\ldots,m\} \setminus \{0\}} \{\widetilde{S}_k - \widetilde{S}_{k-1} > u_k\}\big) \\ &= \frac{1}{\mu} \sum_{i_n,\ldots,i_m \in E} p_{i_0 i_1} \int_{u_0+v}^\infty \big(1 - F_{i_0 i_1}(y)\big)\, \mathrm{d}y \prod_{k \in \{n,\ldots,m\} \setminus \{0\}} \pi_{i_n} p_{i_k i_{k+1}} \big(1 - F_{i_k i_{k+1}}(u_k)\big)\,. \end{aligned} \tag{19}$$

Der *Beweis* von Theorem 4.6 verläuft ähnlich wie der Beweis von Theorem 4.4. Er wird deshalb weggelassen.

Definition

- Wenn die Verteilung Q der Zwischenankunftszeiten $T_n = S_n - S_{n-1}$ durch (18) gegeben ist, dann sagt man,
  - dass der in (1) gegebene stochastische Prozess $\{N_t, t \in \mathbb{R}\}$ ein *semi–Markowscher Zählprozess* ist.
  - In diesem Fall wird der in (14) gegebene Zählprozess $\{\widetilde{N}_t, t \in \mathbb{R}\}$ ein *verzögerter semi–Markowscher Zählprozess* (mit stationären Zuwächsen) genannt.
- Für $\ell = 1$ ergeben sich die in Abschnitt 4.1.3 betrachteten Erneuerungprozesse als Spezialfall.

Beachte

- Die in (16) betrachtete stochastische Matrix $\mathbf{P}$ kann als die Übergangsmatrix einer Markow–Kette aufgefasst werden, wobei $\boldsymbol{\pi}$ die zugehörige stationäre Anfangsverteilung ist, vgl. Theorem MK-2.5.
- Wenn die Verteilungsfunktionen F_{ij} nicht von j abhängen und gegeben sind durch

$$1 - F_{ij}(x) = \mathrm{e}^{-q(i)x} \qquad \forall\, x \ge 0\,, \tag{20}$$

  - wobei $q(1), \ldots, q(\ell)$ eine beliebige Folge positiver Zahlen ist,
  - dann kann man zeigen, dass $\{\widetilde{S}_n\}$ die Folge von Sprungzeitpunkten eines stationären Markow–Prozesses $\{X_t,\, t \in \mathbb{R}\}$ ist, vgl. auch Theorem 2.18.
- Wenn der Markow–Prozess $\{X_t,\, t \in \mathbb{R}\}$ zusätzlich *reversibel* ist, d.h.,
  - wenn die Eintragungen q_{ij} seiner Intensitätsmatrix $\mathbf{Q} = (q_{ij})$ dem Gleichungssystem

$$\widetilde{\pi}_i q_{ij} = \widetilde{\pi}_j q_{ji} \qquad \forall\, i, j \in E$$

    genügen, wobei $\widetilde{\boldsymbol{\pi}} = (\widetilde{\pi}_1, \ldots, \widetilde{\pi}_\ell)^\top$ die (zeit–) stationäre Anfangsverteilung des Markow–Prozesses $\{X_t\}$ ist mit $\boldsymbol{\pi}^\top \mathbf{Q} = \mathbf{0}$,
  - dann ist durch Theorem 4.6 ein Algorithmus zur Simulation des stationären Markow–Prozesses $\{X_t,\, t \in \mathbb{R}\}$ gegeben, vgl. auch Übungsaufgabe 12.3.

4.2 Poissonsche Zählmaße im $\mathbb{R}^d$

- Zur Erinnerung: Sei $d \geq 2$ eine beliebige natürliche Zahl. Die σ–Algebra $\mathcal{B}(\mathbb{R}^d)$ der Borel–Mengen im d–dimensionalen euklidischen Raum $\mathbb{R}^d$ ist die kleinste Familie $\mathcal{G}$ von Teilmengen des $\mathbb{R}^d$, die
  - alle Quader der Form $B = (a_1, b_1] \times \ldots \times (a_d, b_d]$ für beliebige $a_i, b_i \in \mathbb{R}$ mit $a_i \leq b_i$ für $i = 1, \ldots, d$ enthält und
  - die abgeschlossen ist bezüglich der Bildung des Komplementes sowie der Vereinigung von abzählbar vielen Mengen aus $\mathcal{G}$, d.h., für beliebige $A, A_1, A_2, \ldots \in \mathcal{G}$ gilt $\mathbb{R}^d \setminus A \in \mathcal{G}$ und $\bigcup_{n \geq 1} A_i \in \mathcal{G}$.
- Das d–dimensionale Lebesgue–Maß $\nu_d : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$ wird eindeutig bestimmt durch seine Werte

$$\nu_d(B) = \prod_{i=1}^{d} (b_i - a_i)$$

für alle Borel–Mengen der Form $B = (a_1, b_1] \times \ldots \times (a_d, b_d]$ mit $a_i, b_i \in \mathbb{R}$ und $a_i \leq b_i$ für jedes $i = 1, \ldots, d$.

Theorem 4.7

- *Sei $\{N_B,\ B \in \mathcal{B}(\mathbb{R}^d)\}$ ein zufälliges Zählmaß, d.h., für paarweise disjunkte $B_1, B_2, \ldots \in \mathcal{B}(\mathbb{R}^d)$ gelte*

$$N_{\bigcup_{n=1}^{\infty} B_n} = \sum_{n=1}^{\infty} N_{B_n}\,, \tag{1}$$

und sei $\mathcal{M}^d$ die Familie der halboffenen d–dimensionalen Quader in $\mathbb{R}^d$, d.h.,

$$\mathcal{M}^d = \left\{B \subset \mathbb{R}^d : B = (a_1, b_1] \times \ldots \times (a_d, b_d],\ a_i, b_i \in \mathbb{R},\ a_i \leq b_i \,\forall\, i = 1, \ldots, d\right\}. \tag{2}$$

- *Dann ist die Verteilung des zufälligen Zählmaßes $\{N_B\}$ eindeutig bestimmt durch die Familie der „endlich–dimensionalen" Wahrscheinlichkeiten $\{P_{B_1,\ldots,B_n}(k_1, \ldots, k_n) : n \geq 1,\ k_1, \ldots, k_n \geq 0,\ B_1, \ldots, B_n \in \mathcal{M}^d\}$, wobei*

$$P_{B_1,\ldots,B_n}(k_1, \ldots, k_n) = P(N_{B_1} = k_1, \ldots, N_{B_n} = k_n)\,.$$

Der *Beweis* von Theorem 4.7 kann mit Hilfe des Satzes über monotone Klassen geführt werden (vgl. Theorem WR-3.2), wobei man ähnlich wie im Beweis von Theorem 4.1 vorgehen kann.

4.2.1 Definition und elementare Eigenschaften

Wir führen nun den Begriff des Poissonschen Zählmaßes im d–dimensionalen euklidischen Raum $\mathbb{R}^d$ ein.

Definition Sei $\mathcal{B}_0(\mathbb{R}^d)$ die Familie aller beschränkten Borel–Mengen in $\mathbb{R}^d$ und sei $\mu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$ ein beliebiges lokal–endliches Maß, d.h., $\mu(B) < \infty$ gilt für jedes $B \in \mathcal{B}_0(\mathbb{R}^d)$.

- Man sagt, dass $\{N_B,\ B \in \mathcal{B}(\mathbb{R}^d)\}$ ein (inhomogenes) *Poissonsches Zählmaß* mit dem *Intensitätsmaß* μ ist, wenn
  1. die Zufallsvariablen $N_{B_1}, N_{B_2}, \ldots$ unabhängig sind für paarweise disjunkte $B_1, B_2, \ldots \in \mathcal{B}_0(\mathbb{R}^d)$ und
  2. $N_B \sim \mathrm{Poi}(\mu(B))$ für jedes $B \in \mathcal{B}_0(\mathbb{R}^d)$ gilt.

- Wenn zusätzlich vorausgesetzt wird, dass das Intensitätsmaß μ proportional zum d–dimensionalen Lebesgue–Maß ν_d ist, d.h., für eine Konstante $\lambda < \infty$ gilt

$$\mu(B) = \lambda \nu_d(B) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^d)\,,$$

dann sagt man, dass $\{N_B\}$ ein *homogenes Poissonsches Zählmaß* mit der *Intensität* λ (bzw. kurz ein *homogener Poisson–Prozess*) im $\mathbb{R}^d$ ist.

Beachte Aus Theorem 4.7 ergibt sich, dass die Bedingungen 1 und 2 in der Definition des Poisson–Prozesses durch die folgenden (scheinbar schwächeren) Bedingungen ersetzt werden können:

1*. Die Zufallsvariablen $N_{B_1}, N_{B_2}, \ldots$ sind unabhängig für paarweise disjunkte $B_1, B_2, \ldots \in \mathcal{M}^d$ und

2*. $N_B \sim \mathrm{Poi}(\mu(B))$ gilt für jedes $B \in \mathcal{M}^d$.

Wir diskutieren zunächst einige elementare Eigenschaften von Poisson–Prozessen im $\mathbb{R}^d$.

Theorem 4.8 *Sei $\{N_B,\, B \in \mathcal{B}(\mathbb{R}^d)\}$ ein Poisson–Prozess mit dem Intensitätsmaß μ.*

- *Dann gilt*

$$P(N_{B_1} = k_1, \ldots, N_{B_n} = k_n) = \frac{\mu^{k_1}(B_1) \ldots \mu^{k_n}(B_n)}{k_1! \ldots k_n!} \, \exp\Big(-\sum_{i=1}^{n} \mu(B_i)\Big) \tag{3}$$

für beliebige $n \geq 1$, $k_1, \ldots, k_n \geq 0$ und paarweise disjunkte $B_1, \ldots, B_n \in \mathcal{B}_0(\mathbb{R}^d)$.

- *Außerdem genügen die bedingten Wahrscheinlichkeiten $P(N_{B_1} = k_1, \ldots, N_{B_n} = k_n \mid N_B = k)$ für jedes $B \in \mathcal{B}_0(\mathbb{R}^d)$ mit $0 < \mu(B) < \infty$ und für paarweise disjunkte $B_1, \ldots, B_n \in \mathcal{B}_0(\mathbb{R}^d)$ mit $\bigcup_{i=1}^n B_i = B$ einer Multinomialverteilung, d.h., es gilt*

$$P(N_{B_1} = k_1, \ldots, N_{B_n} = k_n \mid N_B = k) = \frac{k!}{k_1! \ldots k_n!} \, \frac{\mu^{k_1}(B_1) \ldots \mu^{k_n}(B_n)}{\mu^k(B)} \tag{4}$$

für beliebige $k, k_1, \ldots, k_n \geq 0$ mit $k = k_1 + \ldots + k_n$.

Der *Beweis* von Theorem 4.8 ergibt sich unmittelbar aus den Bedingungen 1 und 2 in der Definition des homogenen Poisson–Prozesses.

Mit Hilfe der Theoreme 4.7 und 4.8 lässt sich eine einfache Methode zur Konstruktion von Poisson–Prozessen mit endlichem Intensitätsmaß herleiten.

Korollar 4.2 *Sei $\mu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty)$ ein beliebiges Maß mit $0 < \mu(\mathbb{R}^d) < \infty$, und $N : \Omega \to [0, \infty)$ bzw. $S_1, S_2, \ldots : \Omega \to \mathbb{R}^d$ seien unabhängige Zufallsvariablen mit*

$$N \sim \mathrm{Poi}(\mu(\mathbb{R}^d))\,, \qquad S_i \sim \frac{\mu(\cdot)}{\mu(\mathbb{R}^d)} \quad \forall\, i \geq 1\,. \tag{5}$$

Dann ist das zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}(\mathbb{R}^d)\}$ mit

$$N_B = \#\{i : 1 \leq i \leq N, S_i \in B\} \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^d) \tag{6}$$

ein Poisson–Prozess mit dem Intensitätsmaß μ.

Beweis

- Mit der Schreibweise
$$p(B) = \frac{\mu(B)}{\mu(\mathbb{R}^d)} \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^d)$$
ergibt sich aus (5) und (6), dass
$$P(N_{B_1} = k_1, \ldots, N_{B_n} = k_n \mid N = k) = \frac{k!}{k_0! k_1! \ldots k_n!}\, p^{k_0}(B_0)\, p^{k_1}(B_1) \ldots p^{k_n}(B_n)$$
für jedes $n \geq 1$, für beliebige, paarweise disjunkte $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}^d)$ und für $k_0, k_1, \ldots, k_n \geq 0$ mit $k = k_0 + k_1 + \ldots + k_n$, wobei $B_0 = \mathbb{R}^d \setminus \bigcup_{i=1}^n B_i$.
- Hieraus folgt, dass
$$\begin{aligned} P(N_{B_1} = k_1, \ldots, N_{B_n} = k_n) &= \sum_{k=k_1+\ldots+k_n}^{\infty} P(N = k)\, P(N_{B_1} = k_1, \ldots, N_{B_n} = k_n \mid N = k) \\ &= \sum_{k=k_1+\ldots+k_n}^{\infty} \frac{\mathrm{e}^{-\mu(\mathbb{R}^d)} \mu^k(\mathbb{R}^d)}{k!} \frac{k!}{(k - k_1 - \ldots - k_n)! k_1! \ldots k_n!}\, p^{k-k_1-\ldots-k_n}(B_0)\, p^{k_1}(B_1) \ldots p^{k_n}(B_n) \\ &= \sum_{k=k_1+\ldots+k_n}^{\infty} \frac{\mathrm{e}^{-\mu(B_0)} \mu^{k-k_1-\ldots-k_n}(B_0)}{(k - k_1 - \ldots - k_n)!} \frac{\mathrm{e}^{-\mu(\mathbb{R}^d \setminus B_0)}}{k_1! \ldots k_n!}\, \mu^{k_1}(B_1) \ldots \mu^{k_n}(B_n) \\ &= \frac{\mathrm{e}^{-\mu(\mathbb{R}^d \setminus B_0)}}{k_1! \ldots k_n!}\, \mu^{k_1}(B_1) \ldots \mu^{k_n}(B_n) = \prod_{i=1}^n \frac{\mathrm{e}^{-\mu(B_i)} \mu^{k_i}(B_i)}{k_i!}\,. \end{aligned}$$
- Vergleicht man dieses Ergebnis mit der Formel (3) in Theorem 4.8, dann erkennt man, dass das in (6) gegebene zufällige Zählmaß die gleichen endlich-dimensionalen Verteilungen hat wie ein Poisson–Prozess mit dem Intensitätsmaß μ.
- Hieraus und aus Theorem 4.7 ergibt sich sich nun die Behauptung. □

Beachte

- Wenn das Intensitätsmaß μ in Korollar 4.2 die Form $\mu(B) = \lambda \nu_d(B \cap C)$ hat für ein $\lambda < \infty$ und eine beschränkte Borel–Menge $C \in \mathcal{B}_0(\mathbb{R}^d)$ mit $0 < \nu_d(C) < \infty$, dann sind die in (5) betrachteten (unabhängigen) Zufallsvektoren $S_1, S_2, \ldots$ gleichverteilt in C.
- Wenn zusätzlich angenommen wird, dass die Menge C ein d–dimensionaler Quader der Form
$$C = (a_1, b_1] \times \ldots \times (a_d, b_d] \tag{7}$$
ist, dann hat der Zufallsvektor $S_i = (S_{i1}, \ldots, S_{id})$ für jedes $i = 1, \ldots, n$ unabhängige Komponenten $S_{i1}, \ldots, S_{id}$, wobei $S_{ij} \sim \mathrm{U}(a_j, b_j)$ für jedes $j = 1, \ldots, d$.
- Die Aussage von Korollar 4.2 wird deshalb manchmal die *bedingte Gleichverteilungseigenschaft* von homogenen Poisson–Prozessen in beschränkten Borel–Mengen genannt.

Das folgende Resultat über die Summation von unabhängigen Poisson–Prozessen ist ein Analogon der Faltungsstabilität von Poisson–Verteilungen.

Theorem 4.9 *Sei $\{N_B^{(1)}\}, \{N_B^{(2)}\}, \ldots$ eine Folge unabhängiger Poisson–Prozesse in $\mathbb{R}^d$ mit den Intensitätsmaßen $\mu_1, \mu_2, \ldots$, so dass das Maß $\mu = \sum_{i=1}^\infty \mu_i$ lokal endlich ist. Dann ist das zufällige Zählmaß $\{N_B\}$ mit $N_B = \sum_{i=1}^\infty N_B^{(i)}$ ein Poisson–Prozess mit dem Intensitätsmaß μ.*

Beweis

- Wir zeigen zunächst, dass
$$N_B \sim \mathrm{Poi}(\mu(B)) \qquad \forall\, B \in \mathcal{B}_0(\mathbb{R}^d)\,. \tag{8}$$
  - Weil $\{N_B^{(1)}\}$, $\{N_B^{(2)}\}$, ... unabhängige Zählmaße sind, sind die Zufallsvariablen $N_B^{(1)}, N_B^{(2)}, \ldots$ unabhängig, wobei $N_B^{(i)} \sim \mathrm{Poi}(\mu_i(B))$ für jedes $i \geq 1$ und für jedes $B \in \mathcal{B}_0(\mathbb{R}^d)$.
  - Aus der Faltungsstabilität von Poisson–Verteilungen ergibt sich somit, dass für jedes $n \geq 1$ und für jedes $B \in \mathcal{B}_0(\mathbb{R}^d)$
$$\sum_{i=1}^n N_B^{(i)} \sim \mathrm{Poi}(\sum_{i=1}^n \mu_i(B))\,.$$
  - Hieraus und aus der Monotonie von Wahrscheinlicheitsmaßen ergibt sich, dass für beliebige $k \geq 0$ und $B \in \mathcal{B}_0(\mathbb{R}^d)$
$$\begin{aligned} P(N_B \leq k) &= P\big(\sum_{i=1}^\infty N_B^{(i)} \leq k\big) = \lim_{n\to\infty} P\big(\sum_{i=1}^n N_B^{(i)} \leq k\big) \\ &= \lim_{n\to\infty} \sum_{j=0}^k \frac{\exp\big(-\sum_{i=1}^n \mu_i(B)\big)\,\big(\sum_{i=1}^n \mu_i(B)\big)^j}{j!} \\ &= \sum_{j=0}^k \frac{\exp\big(-\mu(B)\big)\,\big(\mu(B)\big)^j}{j!}\,. \end{aligned}$$
  - Damit ist (8) bewiesen.
- Um den Beweis zu beenden, ist noch zu zeigen, dass die Zufallsvariablen $N_{B_1}, N_{B_2}, \ldots, N_{B_n}$ für jedes $n \geq 2$ und paarweise disjunkte $B_1, \ldots, B_n \in \mathcal{B}_0(\mathbb{R}^d)$ unabhängig sind.
  - Weil die Zufallsvariablen $\{N_{B_j}^{(i)},\, i = 1, \ldots, m,\, j = 1, \ldots, n\}$ für beliebige $m, n \geq 1$ unabhängig sind, ergibt sich aus dem Satz über die Unabhängigkeit zusammengesetzter Abbildungen (vgl. Theorem WR-3.18), dass auch die Zufallsvariablen $\{\sum_{i=1}^m N_{B_j}^{(i)},\, j = 1, \ldots, n\}$ unabhängig sind.
  - Wegen der Monotonie von Wahrscheinlicheitsmaßen kann man nun leicht zeigen, dass die Zufallsvariablen
$$\big\{N_{B_j},\, j = 1, \ldots, n\big\} = \big\{\sum_{i=1}^\infty N_{B_j}^{(i)},\, j = 1, \ldots, n\big\}$$
    ebenfalls unabhängig sind. □

Wir zeigen nun noch, dass die Einschränkung von Poisson–Prozessen auf Borelsche Teilmengen des $\mathbb{R}^d$ erneut zu Poisson–Prozessen führt.

Theorem 4.10 *Sei $\{N_B,\, B \in \mathcal{B}(\mathbb{R}^d)\}$ ein Poisson–Prozess mit dem Intensitätsmaß μ, und sei $B_0 \in \mathcal{B}(\mathbb{R}^d)$ eine beliebige Borel–Menge. Dann ist das zufällige Zählmaß $\{\widetilde{N}_B,\, B \in \mathcal{B}(\mathbb{R}^d)\}$ mit $\widetilde{N}_B = N_{B\cap B_0}$ ein Poisson–Prozess mit dem Intensitätsmaß $\widetilde{\mu}$, wobei $\widetilde{\mu}(B) = \mu(B \cap B_0)$ für jedes $B \in \mathcal{B}(\mathbb{R}^d)$.*

Beweis

- Weil $\{N_B\}$ ein Poisson–Prozess mit dem Intensitätsmaß μ ist, gilt
$$\widetilde{N}_B = N_{B\cap B_0} \sim \mathrm{Poi}\big(\mu(B \cap B_0)\big) = \mathrm{Poi}(\widetilde{\mu}(B)) \qquad \forall\, B \in \mathcal{B}_0(\mathbb{R}^d)\,.$$
- Weil für beliebige paarweise disjunkte $B_1, \ldots, B_n \in \mathcal{B}_0(\mathbb{R}^d)$ auch die Mengen $B_1 \cap B_0, \ldots, B_n \cap B_0$ paarweise disjunkt sind, sind die Zufallsvariablen $\widetilde{N}_{B_1} = N_{B_1\cap B_0}, \ldots, \widetilde{N}_{B_n} = N_{B_n\cap B_0}$ unabhängig. □

4.2.2 Messbare Indizierung der Atome

In diesem Abschnitt betrachten wir den Begriff der messbaren Indizierung der (zufälligen) Atome von Poisson–Prozessen, der einen konstruktiven Zugang zu Poisson–Prozessen im $\mathbb{R}^d$ und somit die mathematische Grundlage von Simulationsalgorithmen bildet, vgl. auch die Abschnitte 4.2.3 und 4.2.4.

Beachte Man sagt, dass die Folge $\{\widetilde{S}_i\}$ von Zufallsvektoren $\widetilde{S}_1, \widetilde{S}_2, \ldots : \Omega \to \mathbb{R}^d \cup \{\infty\}$ mit

$$\widetilde{S}_i = \begin{cases} S_i\,, & \text{falls } N \geq i, \\ \infty\,, & \text{sonst} \end{cases} \tag{9}$$

eine *messbare Indizierung* der (zufälligen) Atome des in (6) gegebenen zufälligen Zählmaßes $\{N_B\}$ ist.

Von nun an werden wir stets voraussetzen, dass das Intensitätsmaß $\mu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$ *diffus* ist, d.h., es gelte

$$\mu(\{x\}) = 0 \qquad \forall\, x \in \mathbb{R}^d\,. \tag{10}$$

Der folgende Hilfssatz wird manchmal *Disjunktheitstheorem* genannt. Wir nutzen dieses Ergebnis, um zu zeigen, dass man auch für Poissonsche Zählmaße mit einem beliebigen (diffusen und lokal endlichen) Intensitätsmaß eine messbare Indizierung der Atome konstruieren kann.

Lemma 4.1

- *Seien $\{N_B^{(1)},\, B \in \mathcal{B}(\mathbb{R}^d)\}$ und $\{N_B^{(2)},\, B \in \mathcal{B}(\mathbb{R}^d)\}$ zwei unabhängige Poisson–Prozesse mit den Intensitätsmaßen μ_1 bzw. μ_2, so dass $0 < \mu_1(\mathbb{R}^d),\, \mu_2(\mathbb{R}^d) < \infty$.*
- *Außerdem seien $\{\widetilde{S}_i^{(1)}\}$ und $\{\widetilde{S}_i^{(2)}\}$ unabhängige messbare Indizierungen der Atome von $\{N_B^{(1)}\}$ bzw. $\{N_B^{(2)}\}$, die gemäß (9) gegeben sind.*
- *Dann gilt für beliebige $i, j \geq 1$*

$$P\big(\{\widetilde{S}_i^{(1)} \neq \widetilde{S}_j^{(2)}\} \cup \{\widetilde{S}_i^{(1)} = \widetilde{S}_j^{(2)} = \infty\}\big) = 1\,. \tag{11}$$

Beweis

- Es genügt zu zeigen, dass für beliebige $i, j \geq 1$ mit $i \neq j$

$$P\big(\{\widetilde{S}_i^{(1)} = \widetilde{S}_j^{(2)}\} \cap \{\widetilde{S}_j^{(2)} \neq \infty\}\big) = 0\,.$$

- Mit der Schreibweise $p^{(i)}(B) = \mu_i(B)/\mu_i(\mathbb{R}^d)$ für $i = 1, 2$ und $B \in \mathcal{B}(\mathbb{R}^d)$ ergibt sich aus der Unabhängigkeit der Zufallsvektoren $\widetilde{S}_i^{(1)}$ und $\widetilde{S}_j^{(2)}$, dass

$$\begin{aligned} &P\big(\{\widetilde{S}_i^{(1)} = \widetilde{S}_j^{(2)}\} \cap \{\widetilde{S}_j^{(2)} \neq \infty\}\big) = P\big(\{S_i^{(1)} = S_j^{(2)}\} \cap \{N^{(2)} \geq j\}\big) \\ &= P(N^{(2)} \geq j) P(S_i^{(1)} = S_j^{(2)}) = P(N^{(2)} \geq j) \int_{\mathbb{R}^d} P\big(S_i^{(1)} = S_j^{(2)} \mid S_j^{(2)} = s\big)\, P(S_j^{(2)} \in \mathrm{d}s) \\ &= P(N^{(2)} \geq j) \int_{\mathbb{R}^d} P(S_i^{(1)} = s)\, p^{(2)}(\mathrm{d}s) = P(N^{(2)} \geq j) \int_{\mathbb{R}^d} p^{(1)}(\{s\})\, p^{(2)}(\mathrm{d}s) = 0\,, \end{aligned}$$

  wobei in der letzten Gleichheit die Tatsache genutzt wurde, dass das Intensitätsmaß μ_1 diffus ist und dass somit $p^{(1)}(\{s\}) = 0$ für jedes $s \in \mathbb{R}^d$. □

Beachte Aus Lemma 4.1 ergibt sich insbesondere, dass durch $\{\widetilde{S}_i\}$ mit

$$\widetilde{S}_i = \begin{cases} S_i^{(1)}, & \text{falls } N_{\mathbb{R}^d}^{(1)} \geq i, \\ S_{i-N_{\mathbb{R}^d}^{(1)}}^{(2)}, & \text{falls } N_{\mathbb{R}^d}^{(1)} + N_{\mathbb{R}^d}^{(2)} \geq i > N_{\mathbb{R}^d}^{(1)}, \\ \infty & \text{sonst} \end{cases} \tag{12}$$

eine messbare Indizierung der Atome des zufälligen Maßes $\{N_B\}$ mit $N_B = N_B^{(1)} + N_B^{(2)}$ gegeben ist.

Wir übertragen nun den in (12) gegebenen Ansatz auf den Fall von beliebigen (endlichen bzw. abzählbar unendlichen) Summen unabhängiger Poisson–Prozesse.

Theorem 4.11 *Sei $\mu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$ ein beliebiges diffuses und lokal endliches Maß. Dann gibt es eine Folge $S_1, S_2, \ldots : \Omega \to \mathbb{R}^d \cup \{\infty\}$ von Zufallsvektoren, so dass das zufällige Zählmaß $\{N_B,\, B \in \mathcal{B}(\mathbb{R}^d)\}$ mit*

$$N_B = \#\{i : S_i \in B\} \qquad \forall\, B \in \mathbb{R}^d \tag{13}$$

ein Poisson–Prozess mit dem Intensitätsmaß μ ist.

Beweis

- Man kann sich leicht überlegen, dass sich μ als (abzählbar unendliche) Summe von endlichen Maßen $\mu_1, \mu_2, \ldots : \mathcal{B}(\mathbb{R}^d) \to [0, \infty)$ darstellen lässt, so dass

$$\mu(B) = \sum_{j=1}^{\infty} \mu_j(B) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^d)\,. \tag{14}$$

- Dabei können wir o.B.d.A. annehmen, dass $\mu_j(\mathbb{R}^d) > 0$ für jedes $j \geq 1$.
- Gemäß Korollar 4.2 gibt es dann für jedes $j \geq 1$ unabhängige Zufallsvariablen $N^{(j)}, S_1^{(j)}, S_2^{(j)}, \ldots$ mit

$$N^{(j)} \sim \mathrm{Poi}(\mu_j(\mathbb{R}^d))\,, \qquad S_i^{(j)} \sim \frac{\mu_j(\cdot)}{\mu_j(\mathbb{R}^d)} \quad \forall\, i \geq 1\,,$$

  - so dass durch den Ansatz

$$N_B^{(j)} = \#\{i : 1 \leq i \leq N^{(j)}, S_i^{(j)} \in B\} \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^d)$$

    ein Poisson–Prozess $\{N_B^{(j)},\, B \in \mathcal{B}(\mathbb{R}^d)\}$ mit dem Intensitätsmaß μ_j gegeben ist.
  - Dabei können die Folgen $\{N^{(1)}, S_1^{(1)}, S_2^{(1)}, \ldots\}$, $\{N^{(2)}, S_1^{(2)}, S_2^{(2)}, \ldots\}$, ... so gewählt werden, dass sie ihrerseits unabhängig sind.
  - Damit sind auch die Poisson–Prozesse $\{N_B^{(1)}\}$, $\{N_B^{(2)}\}$, ... unabhängig.
- Aus Theorem 4.9 ergibt sich nun, dass das zufällige Zählmaß $\{N_B\}$ mit $N_B = \sum_{i=1}^{\infty} N_B^{(i)}$ ein Poisson–Prozess mit dem Intensitätsmaß μ ist.
- Darüber hinaus ergibt sich aus Lemma 4.1, dass durch die Folge $\{S_i\}$ von Zufallsvektoren mit

$$S_i = \begin{cases} S_i^{(1)}, & \text{falls } N^{(1)} \geq i, \\ S_{i-N^{(1)}}^{(2)}, & \text{falls } N^{(1)} + N^{(2)} \geq i > N^{(1)}, \\ \vdots & \\ S_{i-N^{(1)}-\ldots-N^{(j-1)}}^{(j)}, & \text{falls } N^{(1)} + \ldots + N^{(j)} \geq i > N^{(1)} + \ldots + N^{(j-1)} \text{ für ein } j > 2, \\ \infty & \text{sonst} \end{cases}$$

  eine messbare Indizierung der Atome von $\{N_B\}$ gegeben ist, so dass (13) gilt. □

4.2.3 Simulationsalgorithmus; Akzeptanz– und Verwerfungsmethode

Sei $\{N_B,\, B \in \mathcal{B}(\mathbb{R}^d)\}$ ein Poisson–Prozess mit dem (diffusen und lokal endlichen) Intensitätsmaß μ, und seien $B_1, \ldots, B_n \in \mathcal{M}^d$ paarweise disjunkte Quader mit der in (2) gegebenen Form, so dass $\mu(B_i) > 0$ für jedes $i = 1, \ldots, n$.

Um den Poisson–Prozess $\{N_B\}$ in der beschränkten Borel–Menge $C = \bigcup_{i=1}^n B_i$ zu simulieren, genügt es zu beachten,

- dass das zufällige Zählmaß $\{\widetilde{N}_B\}$ mit $\widetilde{N}_B = N_{B \cap C}$ gemäß Theorem 4.10 ein Poisson–Prozess mit dem endlichen Zählmaß $\widetilde{\mu}$ ist wobei $\widetilde{\mu}(B) = \mu(B \cap C)$,
- und dass man deshalb gemäß Theorem 4.8 bzw. Korollar 4.2 wie folgt vorgehen kann:

  Schritt 0 Generiere eine Realisierung von $N_C \sim \mathrm{Poi}(\mu(C))$.

  Schritt 1 Falls $N_C = k$, dann generiere eine Realisierung des multinomial verteilten Zufallsvektors

  $$(N_1, \ldots, N_n) \sim \mathrm{Mult}(k; p_1, \ldots, p_n)\,,$$

  wobei $p_i = \mu(B_i)/\mu(C)$ für jedes $i = 1, \ldots, n$.

  Schritt 2 Falls $(N_1, \ldots, N_n) = (k_1, \ldots, k_n)$, dann generiere

  k_1 unabhängige Zufallsvektoren $S_1^{(1)}, \ldots, S_{k_1}^{(1)} \sim \mu(\cdot \cap B_1)/\mu(B_1)$,

  $\vdots$

  k_n unabhängige Zufallsvektoren $S_1^{(n)}, \ldots, S_{k_n}^{(n)} \sim \mu(\cdot \cap B_n)/\mu(B_n)$,

  wobei die Zufallsvektoren $(S_1^{(1)}, \ldots, S_{k_1}^{(1)})$, …, $(S_1^{(n)}, \ldots, S_{k_n}^{(n)})$ ebenfalls unabhängig sind.

Beachte

- Seien $(s_1^{(1)}, \ldots, s_{k_1}^{(1)})$, …, $(s_1^{(n)}, \ldots, s_{k_n}^{(n)})$ Realisierungen von $(S_1^{(1)}, \ldots, S_{k_1}^{(1)})$, …, $(S_1^{(n)}, \ldots, S_{k_n}^{(n)})$.
- Dann ist die (nichtgeordnete) Menge $\{s_1^{(1)}, \ldots, s_{k_1}^{(1)}, \ldots, s_1^{(n)}, \ldots, s_{k_n}^{(n)}\}$ von Punkten im $\mathbb{R}^d$ eine Realisierung der Atome des Poisson–Prozesses $\{N_B\}$ in der Menge $C = \bigcup_{i=1}^n B_i$.

Wenn die Anzahl n der Quader $B_1, \ldots, B_n \in \mathcal{M}^d$ groß ist, aus denen die Menge $C = \bigcup_{i=1}^n B_i$ besteht, dann kann die praktische Umsetzung der Simulationsschritte 1 und 2 mit einem großen Rechenaufwand verbunden sein.

- In diesem Fall kann es effizienter sein, den Poisson–Prozess $\{N_B\}$ mit der folgenden *Akzeptanz– und Verwerfungsmethode* in C zu simulieren.
- Diese Methode hat darüber hinaus den Vorteil, dass das Gebiet $C \subset \mathbb{R}^d$, in dem Poisson–Prozess $\{N_B\}$ simuliert wird, eine *beliebige* beschränkte Borel–Menge sein kann.
  - Sei also $C \in \mathcal{B}_0(\mathbb{R}^d)$ eine beliebige beschränkte Borel–Menge mit $0 < \mu(C) < \infty$, und sei

    $$\widetilde{C} = (a_1, b_1] \times \ldots \times (a_d, b_d]$$

    ein (beschränkter) d–dimensionaler Quader mit $C \subset \widetilde{C}$.
  - Um den Poisson–Prozess $\{N_B\}$ in der Menge C zu simulieren, kann man nun gemäß Theorem 4.10 wie folgt vorgehen:

Schritt 0 Generiere eine Realisierung von $N_C \sim \mathrm{Poi}(\mu(C))$.

Schritt 1 Falls $N_C = k$, dann generiere so lange Realisierungen $s_1, s_2, \ldots$ der unabhängigen Zufallsvektoren $S_1, S_2, \ldots \sim \mu(\cdot \cap \widetilde{C})/\mu(\widetilde{C})$, bis k der Pseudozufallszahlen $s_1, \ldots, s_n$ in der Menge C liegen, wobei

$$n = \min_{j \geq 1} \big\{ \#\{i : s_i \in C, 1 \leq i \leq j\} \geq k \big\} .$$

Schritt 3 Dann ist die (nichtgeordnete) Menge $\{s_i : s_i \in C, 1 \leq i \leq n\}$ von Punkten im $\mathbb{R}^d$ eine Realisierung der Atome des Poisson–Prozesses $\{N_B\}$ in der Menge C.

4.2.4 Transformationssätze; radiale Simulation von homogenen Poisson–Prozessen

In diesem Abschnitt betrachten wir zwei verschiedene Arten von Transformationen von Poisson–Prozesse im $\mathbb{R}^d$. Für beliebige $d, d' \geq 1$ seien die Borel–Mengen $E \in \mathcal{B}(\mathbb{R}^d)$ und $E' \in \mathcal{B}(\mathbb{R}^{d'})$ gegeben.

- Außerdem sei $\mathbf{T} : E \to E'$ eine Borel–messbare Abbildung, d.h., es gelte

$$\mathbf{T}^{-1}(B') = \{x \in E : \mathbf{T}(x) \in B'\} \in \mathcal{B}(E) \qquad \forall\, B' \in \mathcal{B}(E') ,$$

- wobei die Urbilder von beschränkten Borel–Mengen beschränkt seien, d.h., es gelte

$$\mathbf{T}^{-1}(B') \in \mathcal{B}_0(E) \qquad \forall\, B' \in \mathcal{B}_0(E') .$$

Beachte Sei $\{N_B,\, B \in \mathcal{B}(E)\}$ ein beliebiges Zählmaß in E mit dem Intensitätsmß $\mu : \mathcal{B}(E) \to [0, \infty]$. Man kann sich leicht überlegen, dass dann durch den Ansatz

$$N'_{B'} = N_{\mathbf{T}^{-1}(B')} \qquad \forall\, B' \in \mathcal{B}(E') \tag{15}$$

ein zufälliges Zählmaß $\{N'_{B'},\, B' \in \mathcal{B}(E')\}$ in E' gegeben ist, wobei das Intensitätsmaß $\mu' : \mathcal{B}(E') \to [0, \infty]$ von $\{N'_{B'}\}$ gegeben ist durch

$$\mu'(B') = \mathbb{E}\, N'_{B'} \overset{(15)}{=} \mathbb{E}\, N_{\mathbf{T}^{-1}(B')} = \mu(\mathbf{T}^{-1}(B')) \qquad \forall\, B' \in \mathcal{B}(E') . \tag{16}$$

Theorem 4.12 *Sei $\{N_B,\, B \in \mathcal{B}(E)\}$ ein Poisson–Prozess in E mit dem (diffusen und lokal endlichem) Intensitätsmaß μ.*

- *Dann ist das in* (15) *gegebene zufällige Zählmaß $\{N'_{B'},\, B' \in \mathcal{B}(E')\}$ ein Poisson–Prozess in E', dessen Intensitätsmaß μ' durch* (16) *gegeben ist.*
- *Wenn $\{S_i\}$ eine messbare Indizierung der Atome von $\{N_B\}$ ist und wenn die Abbildung $\mathbf{T} : E \to E'$ eineindeutig ist, dann ist durch*

$$S'_i = \mathbf{T}'(S_i) \tag{17}$$

eine messbare Indizierung $\{S'_i\}$ der Atome von $\{N'_{B'}\}$ gegeben, wobei die Abbildung $\mathbf{T}' : E \cup \{\infty\} \to E' \cup \{\infty\}$ gegeben ist durch

$$\mathbf{T}'(x) = \begin{cases} \mathbf{T}(x) , & \textit{falls } x \in E, \\ \infty , & \textit{falls } x = \infty. \end{cases}$$

Beweis

- Offenbar gilt $N'_{B'} = N_{\mathbf{T}^{-1}(B')} \sim \mathrm{Poi}(\mu(\mathbf{T}^{-1}(B'))$.
- Wenn $B'_1, \ldots, B'_n \in \mathcal{B}_0^{d'}(E')$ paarweise disjunkte Borel–Mengen sind, dann sind auch die Urbilder $\mathbf{T}^{-1}(B'_1), \ldots, \mathbf{T}^{-1}(B'_n)$ paarweise disjunkt und die Zufallsvariablen

$$N'_{B'_1} = N_{\mathbf{T}^{-1}(B'_1)}, \ldots, N'_{B'_n} = N_{\mathbf{T}^{-1}(B'_n)}$$

  sind somit unabhängig.
- Außerdem ist die Abbildung $S'_i = \mathbf{T}'(S_i) : \Omega \to E' \cup \{\infty\}$ für jedes $i \geq 1$ messbar, weil die Abbildungen $S_i : \Omega \to E \cup \{\infty\}$ und $\mathbf{T}' : E \cup \{\infty\} \to E' \cup \{\infty\}$ messbar sind, und es gilt für jedes $B' \in \mathcal{B}(E')$

$$N'_{B'} = N_{\mathbf{T}^{-1}(B')} = \#\{i : S_i \in \mathbf{T}^{-1}(B')\} = \#\{i : \mathbf{T}'(S_i) \in B'\}.$$

- Also ist $\{\mathbf{T}'(S_i)\}$ eine messbare Indizierung der Atome von $\{N'_{B'}\}$. □

Beispiele Sei $\{N_B, B \in \mathcal{B}(E)\}$ ein homogener Poisson–Prozess in $E = (0, \infty)$ mit der Intensität $\lambda = 1$.

1. Sei $E' = E = (0, \infty)$ und $\mathbf{T}(x) = x^2$.
   - Dann ist $\{N'_{B'}\}$ ein Poisson–Prozess in $(0, \infty)$.
   - Das Intensitätsmaß μ' von $\{N'_{B'}\}$ ist absolutstetig bezüglich dem (1–dimensionalen) Lebesgue–Maß $\nu_1(\cdot \cap (0, \infty))$, wobei die Dichte gegeben ist durch

$$\frac{\mathrm{d}\mu'}{\mathrm{d}x}(x) = \frac{1}{2\sqrt{x}} \qquad \forall\, x > 0\,.$$

2. Sei $E' = E \times E$ und die Abbildung $\mathbf{T} : E \to E \times E$ sei gegeben durch $\mathbf{T}(x) = (x, x^2)$.
   - Dann ist die Folge $\{S_i\}$ der Sprungzeitpunkte des (Poissonschen) Zählprozesses $\{N_t, t > 0\}$ mit $N_t = N_{(0,t]}$ eine messbare Indizierung der Atome von $\{N_B\}$.
   - Außerdem ist $\{S'_i\} = \{(S_i, S_i^2)\}$ eine messbare Indizierung der Atome eines Poisson–Prozesses $\{N'_{B'}\}$ in $(0, \infty)^2$,
   - dessen Intensitätsmaß μ' auf dem Funktionsgraphen $\{(x, x^2) : x > 0\}$ konzentriert ist.
   - Beachte: Das heißt insbesondere, dass die Atome von $\{N'_{B'}\}$ mit Wahrscheinlichkeit 1 in einer 1–*dimensionalen* Teilmenge von $(0, \infty)^2$ liegen.

Wir betrachten nun noch eine andere Art der Transformation von Poisson–Prozessen in $\mathbb{R}^d$, mit deren Hilfe man Poisson–Prozesse $\{N'_{B'}\}$ in höherdimensionalen Räumen $E' \subset \mathbb{R}^{d'}$ mit $d' > d$ konstruieren kann, so dass der Träger des Intensitätsmaßes μ' von $\{N'_{B'}\}$ eine d'–dimensionale Menge ist.

Theorem 4.13

- *Sei $\{N_B, B \in \mathcal{B}(E)\}$ ein Poisson–Prozess in E mit dem Intensitätsmaß μ, so dass $\mu(E) > 0$, und sei $\{S_i\}$ eine messbare Indizierung der Atome von $\{N_B\}$.*
- *Außerdem sei $m \geq 1$ eine beliebige natürliche Zahl und $U_1, U_2, \ldots : \Omega \to \mathbb{R}^m$ sei eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit der Verteilung P_U, die von $\{S_i\}$ unabhängig sind.*
- *Dann ist durch*

$$N'(B \times C) = \#\{i : (S_i, U_i) \in B \times C\} \qquad \forall\, B \in \mathcal{B}(E), C \in \mathcal{B}(\mathbb{R}^m) \tag{18}$$

  ein Poisson–Prozess $\{N'_{B'}\}$ in $E' = E \times \mathbb{R}^m$ mit dem Intensitätsmaß $\mu' = \mu \times P_U$ gegeben.

Beweis

- Wir betrachten zunächst den Fall, dass μ ein endliches Maß ist.
  - Dann können wir o.B.d.A. annehmen, dass die messbare Indizierung $\{S_i\}$ der Atome von $\{N_B\}$ durch (5) und (9) gegeben ist.
  - Dann sind die Zufallsvektoren $S_1', S_2', \ldots$ mit $S_i' = (S_i, U_i)$ unabhängig, und es gilt

$$S_i' \sim \frac{\mu(\cdot)}{\mu(E)} \times P_U \qquad \forall\, i \geq 1\,.$$

  - Aus Korollar 4.2 ergibt sich nun, dass durch den Ansatz

$$N_{B'}' = \#\{i : 1 \leq i \leq N, S_i' \in B'\} \qquad \forall\, B' \in^{(} \mathcal{E} \times \mathbb{R}^{\Updownarrow})$$

    ein Poisson–Prozess $N_{B'}'$ in $E' = E \times \mathbb{R}^m$ mit dem Intensitätsmaß μ' gegeben ist, wobei

$$\mu'(B \times C) = \mu(E)\big(\frac{\mu(B)}{\mu(E)} \times P_U(C)\big) = \mu(B) \times P_U(C) \qquad \forall\, B \in \mathcal{B}(E),\, C \in \mathcal{B}(\mathbb{R}^m)\,.$$

  - Damit ist die Behauptung unter der zusätzlichen Annahme bewiesen, dass μ endlich ist.
- Wenn μ ein beliebiges (diffuses und lokal endliches) Maß ist,
  - dann können wir $\mu' = \mu \times P_U$ als (abzählbar unendliche) Summe von endlichen Maßen $\mu_1', \mu_2', \ldots$ darstellen
  - und anschließend so wie im Beweis von Theorem 4.11 vorgehen. □

Beachte

- Ähnlich wie bei den zusammengesetzten Poisson–Prozessen in $[0, \infty)$, die in Abschnitt 2.2.2 eingeführt worden sind, können die in Theorem 4.13 betrachteten Zufallsvariablen U_i als „Marken" der Atome S_i aufgefasst werden.
- Dabei sagt man, dass die Folge $\{(S_i, U_i)\}$ der markierten Atome eine messbare Indizierung eines *unabhängig markierten* Poisson–Prozesses ist.

Mit Hilfe der Theoreme 4.12 und 4.13 konstruieren wir nun einen Algorithmus zur *radialen Simulation* von homogenen Poisson–Prozessen im $\mathbb{R}^2$.

- Sei $T_1, T_2, \ldots : \Omega \to [0, \infty)$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit $T_i \sim \mathrm{Exp}(1)$ für jedes $i \geq 1$,
- Für jedes $\lambda > 0$ ergibt sich dann aus Theorem 4.12, dass durch

$$N_B = \#\Big\{i : \sqrt{\sum_{k=1}^{i} \frac{T_k}{\pi\lambda}} \in B\Big\} \qquad \forall\, B \in \mathcal{B}([0, \infty))$$

  ein Poisson–Prozess $\{N_B,\, B \in \mathcal{B}([0, \infty))\}$ in $[0, \infty)$ gegeben ist, dessen Intensitätsmaß μ absolutstetig ist mit der Dichte

$$\frac{\mathrm{d}\mu}{\mathrm{d}x}\,(x) = 2\pi\lambda x \qquad \forall\, x \geq 0\,.$$

- Dabei ist durch

$$S_i = \sqrt{\sum_{k=1}^{i} \frac{T_i}{\pi\lambda}} \qquad \forall\, i \geq 1$$

  eine messbare Indizierung $\{S_i\}$ der Atome von $\{N_B\}$ gegeben.

- Außerdem sei $U_1, U_2, \ldots : \Omega \to [0, 2\pi)$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit $U_i \sim \mathrm{U}([0, 2\pi))$, die unabhängig von $\{T_i\}$ ist.
- Dann ergibt sich aus Theorem 4.13, dass $\{(S_i, U_i)\}$ eine messbare Indizierung der Atome eines Poisson–Prozesses in $[0, \infty) \times [0, 2\pi)$ ist, dessen Intensitätsmaß durch $\mu \times \mathrm{U}(0, 2\pi)$ gegeben ist.
- Die erneute Anwendung von Theorem 4.12 auf die Abbildung $\mathbf{T} : [0, \infty) \times [0, 2\pi) \to \mathbb{R}^2$ mit

$$\mathbf{T}(s, u) = (s \cos u,\ s \sin u) \qquad \forall\, s \geq 0,\ u \in [0, 2\pi) \tag{19}$$

  ergibt schließlich, dass $\{\mathbf{T}(S_i, U_i)\}$ eine messbare Indizierung der Atome eines homogenen Poisson–Prozessen in $\mathbb{R}^2$ mit der Intensität λ ist.

Um einen homogenen Poisson–Prozess mit der Intensität λ im Kreis $B(0, r) = \{x \in \mathbb{R}^2 : |x| \leq r\}$ mit Radius $r > 0$ zu simulieren, kann man also wie folgt vorgehen:

Schritt 0 Generiere die Pseudozufallszahlen $t_1, t_2, \ldots, t_{n(r)}$ gemäß der Verteilung Exp(1), wobei

$$n(r) = \max\Big\{ i : \sqrt{\sum_{k=1}^{i} \frac{t_k}{\pi\lambda}} \leq r \Big\} .$$

Schritt 1 Generiere die Pseudozufallszahlen $u_1, u_2, \ldots, u_{n(r)}$ gemäß der Verteilung $\mathrm{U}(0, 2\pi)$.

Schritt 2 Berechne die Pseudozufallsvektoren $(s_1, u_1), \ldots, (u_{n(r)}, u_{n(r)})$, wobei

$$s_i = \sqrt{\sum_{k=1}^{i} \frac{t_k}{\pi\lambda}} \qquad \forall\, i \geq 1 .$$

Schritt 3 Transformiere die Pseudozufallsvektoren $(s_1, u_1), \ldots, (u_{n(r)}, u_{n(r)})$ mit Hilfe der in (19) gegebenen Abbildung $\mathbf{T} : [0, \infty) \times [0, 2\pi) \to \mathbb{R}^2$.

Schritt 4 Generiere so die Realisierung $\mathbf{T}(s_1, u_1), \ldots, \mathbf{T}(u_{n(r)}, u_{n(r)})$ eines homogenen Poisson–Prozesses mit der Intensität λ im Kreis $B(0, r) = \{x \in \mathbb{R}^2 : |x| \leq r\}$.