

Stochastische Prozesse

[MAT.315UF]

gelesen von: Wolfgang Müller, Ao.Univ.-Prof. Dipl.-Ing. Dr.techn.
am Institut für Statistik
Technische Universität Universität Graz

verfasst von: Laura Philomena Mossböck
11820925

Wintersemester 2023/24



Inhaltsverzeichnis

1	Simulation stochastischer Prozesse	1
1.1	Irrfahrten	1
1.1.1	Irrfahrten auf \mathbb{Z}^d	1
1.1.2	Gambler's Ruin	2
1.2	Markov Ketten	3
1.3	Poisson-Prozesse	4
1.4	Zusammengesetzte Poisson-Prozesse	6
1.5	Zeitstetige Markov-Prozesse mit abzählbarem Zustandsraum	6
1.6	Brownsche Bewegungen	9
2	Anhang	11
2.1	Ergebnisse aus der Wahrscheinlichkeitstheorie	11

1 Simulation stochastischer Prozesse

Definition 1.1 (Stochastischer Prozess). *Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, (E, \mathcal{E}) ein Messraum¹ und I eine beliebige Indexfamilie. Ein stochastischer Prozess ist eine Familie von Zufallsvariablen $(X_i)_{i \in I}$ mit $X_i: \Omega \rightarrow E$. Die Menge E nennen wir den Zustandsraum des Prozesses. Für ein festes $\omega \in \Omega$ nennen wir die durch*

$$i \mapsto X_i(\omega)$$

festgelegte Funktion den Pfad des Prozesses. Wir nennen einen stochastischen Prozess $(X_i)_{i \in I}$ transient, wenn jeder feste Punkt $e \in E$ mit Wahrscheinlichkeit 1 höchstens endlich oft besucht wird. Wird hingegen jedes $e \in E$ mit Wahrscheinlichkeit 1 unendlich oft erreicht, so nennen wir den Prozess rekurrent.

Bemerkung 1.2. Wir werden im Folgenden meist $I = \mathbb{N}_0$ oder $I = \mathbb{R}_0^+$ annehmen. Im Fall $I = \mathbb{N}$ nennen wir den Prozess $(X_n)_{n \geq 0}$ einen Prozess in diskreter Zeit. Analog nennen wir für $I = \mathbb{R}_0^+$ $(X_t)_{t \geq 0}$ einen stochastischen Prozess in stetiger Zeit.

Viele der Standard-Prozesse können aus Folgen unabhängiger Zufallsvariablen konstruiert werden. Solche expliziten Konstruktionen erlauben die Simulation des Prozesses am Computer und ermöglichen experimentelle Untersuchung seiner Eigenschaften.

1.1 Irrfahrten

1.1.1 Irrfahrten auf \mathbb{Z}^d

Wir beschreiben zuerst die zufällige Irrfahrt eines Teilchens auf \mathbb{Z} . Von jeder ganzen Zahl n springt das Teilchen mit Wahrscheinlichkeit p zu $n + 1$ und mit Wahrscheinlichkeit $1 - p$ zu $n - 1$.

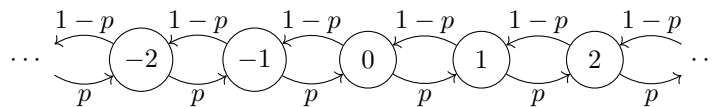


Abbildung 1: Visualisierung der möglichen Zustände einer Irrfahrt auf \mathbb{Z}

Sei $(Z_j)_{j \geq 1}$ eine Folge unabhängiger Zufallsvariablen mit

$$\mathbb{P}(Z_j = 1) = p \quad \mathbb{P}(Z_j = -1) = 1 - p.$$

Der in 0 startenden Irrfahrt entspricht der Prozess $(X_n)_{n \geq 0}$ mit $X_0 = 0$ und

$$X_n = \sum_{j=1}^n Z_j.$$

Nach dem starken Gesetz der Großen Zahlen gilt $\frac{1}{2}X_n \xrightarrow{\text{f.s.}} \mu$, wobei $\mu = \mathbb{E}(Z_1) = 2p - 1$. Im Fall $\mu \neq 0$ verhält sich X_n daher fast sicher asymptotisch wie μn und die Irrfahrt ist transient. Im Gegensatz dazu sind die symmetrischen Irrfahrten rekurrent.

Schwieriger und interessanter ist die Irrfahrt auf dem d -dimensionalen Gitter \mathbb{Z}^d . Dabei wird immer nur auf benachbarte Punkte des Gitters gesprungen. Wir bezeichnen mit e_1, \dots, e_d die kanonische Basis von \mathbb{Z}^d . Sei $(Z_j)_{j \geq 1}$ eine Folge unabhängiger d -dimensionaler Zufallsvektoren welche die Werte $\pm e_1, \dots, \pm e_d$ mit den Wahrscheinlichkeiten

$$\mathbb{P}(Z_j = \pm e_i) = p_i^\mp \quad \text{mit} \quad \sum_{i=1}^d (p_j^+ + p_j^-) = 1$$

¹Ist E höchstens abzählbar wählen wir als σ -Algebra stets die Potenzmenge. Für $E = \mathbb{R}$ oder $E = \mathbb{C}$ wählen wir die entsprechenden Borel- σ -Algebren.

annehmen. Dann beschreibt der Prozess $(X_n)_{n \geq 0}$ mit $X_n = \sum_{j=1}^n Z_j$ eine im Ursprung startende Irrfahrt auf \mathbb{Z}^d . Wie im eindimensionalen Fall folgt aus dem starken Gesetz der großen Zahlen, dass die Irrfahrt auf \mathbb{Z}^d transient ist, wenn $\mu = \mathbb{E}(Z_1) \neq 0$. Die Irrfahrt heißt symmetrisch, wenn alle p_i^\pm gleich sind, sprich

$$p_i^+ = p_i^- = \frac{1}{2d}.$$

Nach dem **Satz von Pólya** ist eine symmetrische Irrfahrt im Fall $d \leq 2$ rekurrent, aber für $d \geq 3$ transient. Wir werden das für beliebige driftlose² Irrfahrten zeigen.

1.1.2 Gambler's Ruin

Bei einem Glücksspiel (z.B. Roulette) gewinnt ein Spieler mit Wahrscheinlichkeit $0 < p < 1$. Pro Spiel setzt er einen Euro. Gewinnt er, erhält er den Einsatz und einen zusätzlichen Euro, andernfalls verliert er den eingesetzten Euro. Er startet das Spiel mit einem Anfangskapital von $a > 0$ Euro und möchte eine vorgegebene Anzahl $b > 0$ Euro gewinnen. Dazu spielt er so lange, bis er entweder b Euro gewonnen hat, oder sein Anfangskapital aufgebraucht ist. Man bestimme die Ruinwahrscheinlichkeit mit er sein ganzes Anfangskapital verliert. Der Gewinn bei den einzelnen Spielen wird durch eine Folge $(Z_j)_{j \geq 1}$ unabhängiger Zufallsvariablen mit

$$\mathbb{P}(Z_j = 1) = p \quad \text{und} \quad \mathbb{P}(Z_j = -1) = 1 - p$$

modelliert. Der Gesamtgewinn nach n Spielen ist daher $X_n = \sum_{j=1}^n Z_j$. Der Gewinnprozess $(X_n)_{n \geq 0}$ eine Irrfahrt auf \mathbb{Z} . Der Spieler spielt bis zum zufälligen Zeitpunkt

$$T_{a,b} := \inf \{n > 0 : X_n = -a \vee X_n = b\}.$$

Natürlich ist es möglich, dass $T_{a,b} = \infty$. Das geschieht aber nur mit Wahrscheinlichkeit 0. Um das einzusehen beachte man, dass man im Fall $T_{a,b} = \infty$ nicht $a + b - 1 = d$ mal hintereinander gewinnen kann. Daraus folgt mit $K \rightarrow \infty$

$$\mathbb{P}(T_{a,b} = \infty) \leq \mathbb{P}\left(\bigcap_{k=1}^K \{Z_{(k-1)d+1} = 1, \dots, Z_{kd} = 1\}^c\right) \leq (1 - p^d)^K \rightarrow 0.$$

Wegen $\mathbb{P}(T_{a,b} < \infty) = 1$ ist die Zufallsvariable

$$X_{T_{a,b}}(\omega) := X_{T_{a,b}(\omega)}(\omega)$$

fast sicher³ eindeutig definiert. Die gesuchte Ruinwahrscheinlichkeit ist $\mathbb{P}(X_{T_{a,b}} = -a)$. Sie kann mit Hilfe der Theorie der Markov-Ketten mit endlichem Zustandsraum berechnet werden, wenn man den Prozess X so modifiziert, dass er bei Erreichen des Zustands $-a$ oder b in diesen Zuständen verweilt. Dazu setzt man

$$\tilde{X}_n := X_{\min(n, T_{a,b})}.$$

Der Prozess $(\tilde{X}_n)_{n \geq 0}$ ist dann eine Markov Kette mit Zustandsraum $E = \{-a, \dots, b\}$ im Sinne des nächsten Abschnitts. Sein Verhalten wird durch den folgenden „Übergangsgraphen“ beschrieben.

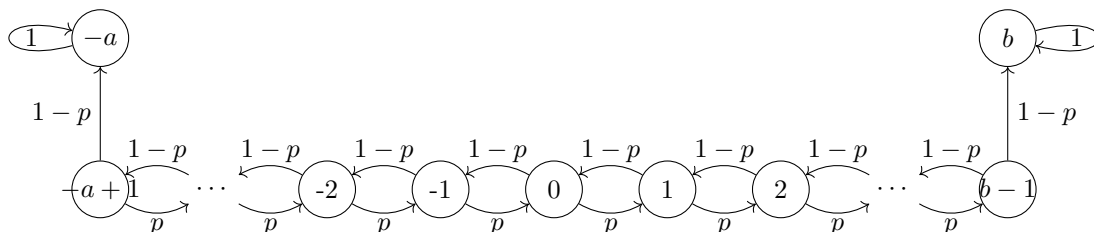


Abbildung 2: Übergangsgraph für Gambler's Ruin

²d.h. $\mu = 0$

³Auf der Nullmenge $\{T_{a,b} = \infty\}$ kann $X_{T_{a,b}}(\omega) = 0$ gesetzt werden. Dann ist $X_{T_{a,b}}$ eine Zufallsvariable, welche nur die Werte $-a$, b oder 0 annimmt, den Wert 0 nur mit Wahrscheinlichkeit 0.

1.2 Markov Ketten

Die österreichische Bonus-Malus Kfz-Haftpflichtversicherung hat 18 Stufen⁴. Jeder Versicherte starte in der Stufe 9. Die Jahresprämie ist abhängig von der Stufe, z.B: ist die Prämie in Stufe 0 nur 50% der Prämie in Stufe 9, die Prämie in Stufe 17 aber 200% der Prämie in Stufe 9. Nach jedem Jahr in dem ein Versicherter keinen Schaden an die Versicherung gemeldet hat, wird er um eine Stufe herunter gereiht (wenn er nicht schon in Stufe 0 ist). Meldet er in einem Jahr k Schäden, wird er im nächsten Jahr um $3k$ Stufen nach oben gereiht (maximal aber in die Stufe 17).

Wir betrachten eine vereinfachte Form dieser Bonus-Malus Versicherung mit den Stufen 1,2,3,4,5 und 6. Man starte in Stufe 4. Meldet man keinen Schaden wird man um eine Stufe herunter gereiht, meldet man einen oder mehrere Schäden wird man um 2 Stufen hinauf gereiht. Die Wahrscheinlichkeit mit der ein Versicherter einen oder mehrere Schäden meldet sei p (p ist abhängig von der Risikoklasse des Versicherten). Dieser Zufallsmechanismus wird durch den folgenden Übergangsgraphen beschrieben.

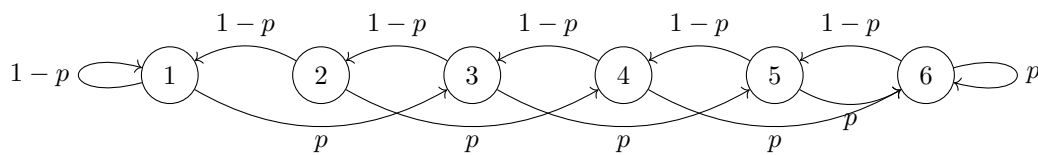


Abbildung 3: Übergangsgraph für unser vereinfachtes Bonus-Malus System

Die Gewichtung der Kanten des Graphen geben die Wahrscheinlichkeit an, von einem Knoten aus einen anderen Knoten zu erreichen⁵. Dieser Vorgang definiert einen Markov-Prozess. Das bedeutet, dass der Zufallsmechanismus der zum nächsten Zustand führt nur vom aktuellen Zustand abhängt. Die Bestimmungsstücke einer (homogenen) Markovkette mit abzählbaren Zustandsraum E sind

- Ein beliebiger Startverteilungsvektor $\nu = (p_i)_{i \in E}$ auf E , der angibt mit welcher Wahrscheinlichkeit p_i man sich zum Zeitpunkt 0 im Zustand i befindet.
- Eine Matrix von Übergangswahrscheinlichkeiten $P = (p_{ij})_{i,j \in E}$ die angibt, mit welcher Wahrscheinlichkeit p_{ij} man vom Zustand i in den Zustand j gelangt⁶. Die Matrix P ist eine stochastische Matrix, sprich ihre Einträge sind nicht negativ und alle Zeilensummen sind 1.

Für unser vereinfachtes Bonus-Malus System folgt also

$$\nu = [0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0] \quad P = \begin{bmatrix} 1-p & 0 & p & 0 & 0 & 0 \\ 1-p & 0 & 0 & p & 0 & 0 \\ 0 & 1-p & 0 & 0 & p & 0 \\ 0 & 0 & 1-p & 0 & 0 & p \\ 0 & 0 & 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 0 & 1-p & p \end{bmatrix}.$$

Weitere Beispiele von Markov-Ketten sind die zuvor betrachteten Irrfahrten. Der beschriebene Zufallsmechanismus kann mit Folgen unabhängiger Zufallsvariablen nachgebildet werden. Sei dazu $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum auf dem unabhängige Zufallsvariablen

$$X_0 \quad \text{und} \quad (Z_{n,i})_{n \geq 1, i \in E}$$

existieren. Die Verteilung der Zufallsvariable X_0 sei durch ν festgelegt und die Verteilung der $Z_{n,i}$ durch die i -te Zeile von P , sprich

$$\mathbb{P}(X_0 = j) = p_j \quad \mathbb{P}(Z_{n,i} = j) = p_{ij}.$$

Durch die Rekursion

$$X_n(\omega) = Z_{n, X_{n-1}(\omega)}(\omega)$$

⁴von Stufe 0 bis 17

⁵Um Übersichtlichkeit zu bewahren werden Übergänge mit Wahrscheinlichkeit 0 nicht eingezeichnet.

⁶Im Fall einer inhomogenen Markovkette dürfen diese Übergangswahrscheinlichkeiten vom Zeitpunkt n des Übergangs abhängen.

für $n \geq 1$ eine Markov-Kette $(X_n)_{n \geq 0}$ mit Zustandsraum E , Startverteilung ν und Übergangsmatrix P festgelegt. Man beachte, dass X_n und $Z_{n+1,i}$ unabhängig sind. Daraus folgt

$$P(X_{n+1} = j | X_n = i) = \frac{\mathbb{P}(Z_{n+1,i} = j, X_n = i)}{\mathbb{P}(X_n = i)} = \mathbb{P}(Z_{n+1,i} = j) = p_{ij},$$

wenn $P(X_n = i) > 0$. Die obige Beschreibung einer Markov Kette erlaubt ihre Simulation. Man generiert zuerst eine Realisierung von X_0 gemäß der Startverteilung und wechselt dann, in Abhängigkeit vom Startzustand $i_0 = X_0(\omega)$ gemäß den Übergangswahrscheinlichkeiten $(p_{i_0 j})_{j \in E}$ in den neuen Zustand $i_1 = X_1(\omega)$ und so weiter.

Lemma 1.3 (Verteilung von X_n). Sei $\nu_n = (\mathbb{P}(X_n = i))_{i \in E}$ der Verteilungsvektor von X_n und $\nu = (p_i)_{i \in E}$ der Startverteilungsvektor, dann gilt

$$\nu_n = \nu P^n.$$

Beweis. Es genügt $\nu_n = \nu_{n-1}P$ für $n \geq 1$ zu zeigen. Dann ist $\nu_n = \nu_0 P^n = \nu P^n$.

$$\mathbb{P}(X_n = j) = \sum_{i \in E} \mathbb{P}(X_n = j | X_{n-1} = i) \mathbb{P}(X_{n-1} = i) = \sum_{i \in E} \nu_{n-1,i} p_{ij} = (\nu_{n-1}P)_j.$$

Dabei ist in der zweiten Summe die Summation auf die i mit $\mathbb{P}(X_{n-1} = i) > 0$ einzuschränken, oder man vereinbart, dass eine undefinierte Größe mit 0 multipliziert den Wert 0 ergibt. \square

Bemerkung 1.4. Wie im endlichdimensionalen ist Matrixprodukt durch $(AB)_{ij} := \sum_{l \in E} a_{il} b_{lj}$ definiert. Die auftretenden Reihen sind für stochastische Matrizen absolut konvergent und können beliebig umgeordnet werden.

Beispiel 1.5 (Bonus-Malus Versicherung). Im vereinfachten Bonus-Malus System sei die Wahrscheinlichkeit p , dass in einem Jahr ein oder mehr Schäden gemeldet werden gleich 0.2. Dann ist die Anzahl $N \geq 1$ der Jahre bis zur ersten Schadensmeldung geometrisch verteilt mit

$$\mathbb{P}(N = k) = p(1-p)^{k-1} \quad \mathbb{E}(N) = \frac{1}{p},$$

sprich im Mittel meldet man jedes 5. Jahr einen Schaden. Die 10-Jahres Übergangswahrscheinlichkeiten sind

$$P^{10} = \begin{bmatrix} 0.544 & 0.128 & 0.168 & 0.071 & 0.056 & 0.032 \\ 0.544 & 0.128 & 0.157 & 0.083 & 0.056 & 0.032 \\ 0.514 & 0.159 & 0.157 & 0.068 & 0.071 & 0.032 \\ 0.514 & 0.113 & 0.202 & 0.068 & 0.060 & 0.043 \\ 0.514 & 0.113 & 0.157 & 0.113 & 0.060 & 0.043 \\ 0.454 & 0.174 & 0.157 & 0.083 & 0.090 & 0.043 \end{bmatrix}.$$

Zehn Jahr nach dem Anfangsjahr ist die Verteilung der Versicherungsnehmer auf die 6 Stufen

$$\nu_{10} = \nu_0 P^{10} = [0.513 \quad 0.113 \quad 0.202 \quad 0.067 \quad 0.059 \quad 0.043].$$

Nur noch 6.7% der Versicherten sind in Stufe 4, aber 51% sind in Stufe 1. Lässt man die Markov-Kette weiter laufen, so ändert sich diese Verteilung nur noch wenig, für $n \geq 25$ ist (auf drei Stellen genau)

$$\nu_n = [0.513 \quad 0.132 \quad 0.166 \quad 0.074 \quad 0.060 \quad 0.033].$$

Zusätzlich beobachtet man, dass für $n \geq 25$ alle Zeilen von P^n bis auf drei Stellen genau gleich diesem Vektor sind. Sprich, nach 25 Jahren die Verteilung auf die Stufen unabhängig von der Startstufe ist.

1.3 Poisson-Prozesse

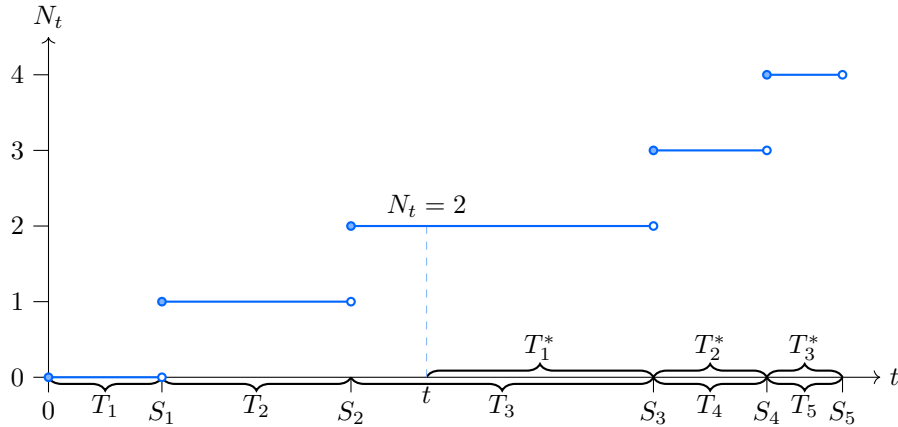
Lebensdauern von Geräten die keinem Alterungsprozess unterliegen werden durch exponentialverteilte⁷ Zufallsvariablen modelliert. Das liegt an der Gedächtnislosigkeit der Exponentialverteilung. Ähnlich werden Wartezeiten, d.h. die Zeitspannen zwischen dem Eintreten zweier Ereignisse, im einfachsten Fall durch exponentialverteilte Zufallsvariablen modelliert.

⁷Eine exponentialverteilte Zufallsvariable T hat die Dichte $f(x) = \lambda e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x)$ und Erwartung $\mathbb{E}(X) = \frac{1}{\lambda}$. Die Gedächtnislosigkeit bedeutet $\mathbb{P}(X \geq x+y | X > y) = \mathbb{P}(T > x)$ für $x, y \geq 0$.

Wir zählen wie oft Ereignisse eintreten, wenn die Zeitspannen zwischen zwei Ereignissen durch unabhängige $\exp(\lambda)$ -verteilte Zufallsvariablen $(T_n)_{n \geq 1}$ beschrieben werden. Dann ist $S_n = \sum_{j=1}^n T_j$ der Zeitpunkt des Eintretens des n -ten Ereignisses. Die Anzahl N_t der bis zum Zeitpunkt t eingetretenen Ereignisse ist

$$N_t = \sum_{n \geq 1} \mathbb{1}_{\{S_n \leq t\}}.$$

Der Prozess $N = (N_t)_{t \geq 0}$ heißt Poisson-Prozess mit Intensität $\lambda > 0$. Nach dem starken Gesetz der großen Zahlen gilt $n^{-1}S_n \xrightarrow{\text{f.s.}} \lambda^{-1}$. Daraus folgt $S_n \xrightarrow{\text{f.s.}} \infty$. Der Prozess bleibt daher fast sicher endlich. Einen typischen Pfad zeigt die folgende Abbildung.



Wir wollen die Verteilung von N_t bestimmen. Als Summe unabhängiger, exponentialverteilter Zufallsvariablen ist S_n gammaverteilt, d.h. $f_{S_n}(s) = \frac{\lambda^n}{\Gamma(n)} s^{n-1} e^{-\lambda s} \mathbb{1}_{(0, \infty)}(s)$. Es folgt

$$\begin{aligned} \mathbb{P}(N_t = n) &= \mathbb{P}(S_n \leq t < S_n + T_{n+1}) = \mathbb{P}((S_n, T_{n+1}) \in \{(s, u) : 0 \leq s \leq t, u \geq t - s\}) \\ &= \int_0^t \int_{t-s}^{\infty} f_{T_{n+1}}(u) du f_{S_n}(s) ds = \int_0^t e^{-\lambda(t-s)} \frac{1}{(n-1)!} s^{n-1} e^{-\lambda s} ds = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \end{aligned}$$

d.h. N_t ist poissonverteilt mit Parameter λt und $\mathbb{E}(N_t) = \lambda t$. Das verdeutlicht die Bedeutung der Intensität λ . Pro Zeiteinheit treten durchschnittlich λ Ereignisse ein. Die Mittlere Wartezeit ist $\frac{1}{\lambda}$.

Der Poisson-Prozess N ist eines der einfachsten Beispiele eines zeitstetigen Markov-Prozesses mit abzählbarem Zustandsraum. Die Markov-Eigenschaft bedeutet dabei: Hat man den Prozess bis zum Zeitpunkt t simuliert, muss man für die Konstruktion eines Pfades über den Zeitpunkt t hinaus nur noch den momentanen Wert $N_t(\omega)$ des Prozesses kennen, und nicht die ganze Vergangenheit $(N_s(\omega))_{0 \leq s \leq t}$. Wir zeigen, dass das im Fall eines Poisson-Prozesses aufgrund der Gedächtnislosigkeit der Wartezeiten T_j richtig ist. Dazu bedingen wir auf das Ereignis $\{N_t = n\}$.

Satz 1.6 (Markov-Eigenschaft eines Poisson-Prozesses). *Wir betrachten das Wahrscheinlichkeitsmaß $\mathbb{P}_t(A) = \mathbb{P}(A | N_t = n)$. Bezüglich dieses Wahrscheinlichkeitsmaßes sind die nach t verbleibenden Wartezeiten*

$$T_1^* := S_n + T_{n+1} - t \quad T_j^* = T_{n+j}$$

exp(λ)-verteilt und unabhängig. Damit ist der aus $(T_n^)_{n \geq 0}$ konstruierte Prozess $N^* = (N_{s+t} - N_t)_{s \geq 0}$ bezüglich \mathbb{P}_t wieder ein Poisson-Prozess mit Intensität λ .*

Beweis. Für den Beweis drücken wir alle Ereignisse in Termen der unabhängigen T_n aus, z.B. ist $\{N_t = n\} = \{S_n \leq t < S_n + T_{n+1}\}$. Zunächst erhält man für $j \geq 2$ und beliebige Borelmengen B_j

$$\mathbb{P}_t(T_j^* \in B_j) = \frac{\mathbb{P}(T_{n+j} \in B_j, S_n \leq t < S_n + T_{n+1})}{\mathbb{P}(N_t = n)} = \mathbb{P}(T_{n+j} \in B_j).$$

T_j^* ist also auch unter \mathbb{P}_t eine $\exp(\lambda)$ -verteilte Zufallsvariable. Die Gedächtnislosigkeit der Exponentialverteilung ist gleichbedeutend mit der für $x \geq 0$ und $t \geq s$ gültigen trivialen Identität

$$\mathbb{P}(T_{n+1} > x + t - s) = e^{-\lambda(x+t-s)} = e^{-\lambda x} \mathbb{P}(T_{n+1} > t - s).$$

Damit erhält man für $x \geq 0$

$$\begin{aligned}
& \mathbb{P}_t(T_1^* > x, T_j^* \in B_j \text{ für } 2 \leq j \leq k) \\
&= \frac{\mathbb{P}(S_n + T_{n+1} > x + t, S_n \leq t < S_n + T_{n+1}, T_{n+j} \in B_j \text{ für } 2 \leq j \leq k)}{\mathbb{P}(N_t = n)} \\
&= \frac{\mathbb{P}(S_n + T_{n+1} > x + t, S_n \leq t, T_{n+j} \in B_j \text{ für } 2 \leq j \leq k)}{\mathbb{P}(N_t = n)} \\
&= \int_0^t \mathbb{P}(T_{n+1} > x + t - s) f_{S_n}(s) \, ds \frac{1}{\mathbb{P}(N_t = n)} \prod_{j=2}^k \mathbb{P}(T_{n+j} \in B_j) \\
&= e^{-\lambda x} \int_0^t \mathbb{P}(T_{n+1} > t - s) f_{S_n}(s) \, ds \frac{1}{\mathbb{P}(N_t = n)} \prod_{j=2}^k \mathbb{P}(T_{n+j} \in B_j).
\end{aligned}$$

Der Spezialfall $x = 0$ und $B_j = \mathbb{R}$ für $2 \leq j \leq k$ liefert

$$\int_0^t \mathbb{P}(T_{n+1} > t - s) f_{S_n}(s) \, ds \frac{1}{\mathbb{P}(N_t = n)} = \mathbb{P}_t(T_1^* > 0) = 1.$$

Zusammen erhält man

$$\mathbb{P}_t(T_1^* > x, T_j^* \in B_j \text{ für } 2 \leq j \leq k) = e^{-\lambda x} \prod_{j=2}^k \mathbb{P}_t(T_j^* \in B_j). \quad (1)$$

Mit $B_j = \mathbb{R}$ für alle j folgt, dass $T_1^* \exp(\lambda)$ -verteilt bezüglich \mathbb{P}_t ist. dann liefert Gleichung (1) die Unabhängigkeit aller T_j^* bezüglich \mathbb{P}_t . \square

Die analog definierten Zählprozesse mit unabhängigen, identisch verteilten Wartezeiten $(T_n)_{n \geq 1}$ heißen Erneuerungsprozesse. Sie sind im allgemeinen keine Markov-Prozesse. Zur Fortsetzung der Simulation über den Zeitpunkt t hinaus muss neben $N_t = n$ noch die Zeit $A_t = t - S_n$ die seit dem letzten Ereignis vergangen ist bekannt sein.

Bemerkung 1.7. Der zweidimensionale Prozess $(N_t, A_t)_{t \geq 0}$ ist daher ein Markov-Prozess.

1.4 Zusammengesetzte Poisson-Prozesse

Eine Versicherung modelliert die Anzahl der einlangenden Schadensmeldungen durch einen Poisson-Prozess $(N_t)_{t \geq 0}$ mit Intensität $\lambda > 0$. Die Folge der Schadenshöhen wird durch eine Folge unabhängiger Zufallsvariablen $(U_j)_{j \geq 1}$ modelliert, die alle die selbe Verteilung ν haben und unabhängig von den Wartezeiten $(T_j)_{j \geq 1}$ des Poisson-Prozesses sind. Der durch

$$X_t = \sum_{j=1}^{N_t} U_j$$

definierte Prozess $(X_t)_{t \geq 0}$ beschreibt den zeitlichen Verlauf der Gesamtschadenshöhe. Er heißt zusammengesetzter Poisson-Prozess mit Intensität λ und Sprunghöhenverteilung ν .

1.5 Zeitstetige Markov-Prozesse mit abzählbarem Zustandsraum

Beispiel 1.8 (Call-Center). Bei einem Call-Center langen Anrufe gemäß einem Poisson-Prozess mit Intensität λ_W ein. Diese werden auf $k \geq 1$ Operatoren aufgeteilt (ist keiner der Operatoren frei wird der Anruf der Einfachheit halber abgewiesen). Jeder der Operatoren benötigt zur Behandlung des Anliegens eine $\exp(\lambda_S)$ -verteilte zufällige Servicezeit. Die Operatoren arbeiten unabhängig voneinander. Sei X_t die Anzahl der im System vorhandenen Anrufe, dann ist $(X_t)_{t \geq 0}$ Beispiel eines zeitstetigen Markov-Prozesses mit endlichem Zustandsraum $E = \{0, 1, \dots, k\}$.

Wie wir uns im Folgenden überlegen, entspricht der beschriebene Zufallsmechanismus einer Markov-Kette, bei der die Zustände nicht zu vorgegebenen Zeitpunkten $n \in \mathbb{N}$, sondern nach unabhängigen exponentialverteilten Wartezeiten geändert werden.

Bemerkung 1.9. Sind die Service- und Wartezeiten nicht exponentialverteilt, so ist der Prozess kein Markov-Prozess.

Lemma 1.10 (Exponentialverteilung).

i) Das Minimum M_k unabhängiger $\exp(\lambda_j)$ -verteilter Zufallsvariablen τ_j ist wieder exponentialverteilt, genauer

$$M_k = \min_{j=1, \dots, k} \tau_j \sim \exp\left(\sum_{j=1}^k \lambda_j\right) \quad \text{und} \quad \mathbb{P}(\tau_1 < \tau_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

ii) Bedingt auf das Ereignis $\{M_k = \tau_k\}$ sind die verbleibenden Zeiten $\tau_j - \tau_k$ für $j < k$ unabhängig und $\exp(\lambda_j)$ -verteilt.

Beweis.

i) Die erste Behauptung folgt aus

$$\mathbb{P}(\min(\tau_1, \dots, \tau_k) > t) = \prod_{j=1}^k \mathbb{P}(\tau_j > t) e^{-(\lambda_1 + \dots + \lambda_k)t}$$

und die zweite aus

$$\mathbb{P}(\tau_1 < \tau_2) = \int_0^\infty \int_{t_1}^\infty f_{\tau_1}(t_1) f_{\tau_2}(t_2) dt_2 dt_1 = \int_0^\infty \lambda_1 e^{-(\lambda_1 + \lambda_2)t_1} dt_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

ii) Wegen $\mathbb{P}(\tau_i > x_i + t) = \mathbb{P}(\tau_i > x_i) \mathbb{P}(\tau_i > t)$ für $x_j \geq 0$ und $t \geq 0$ gilt

$$\begin{aligned} \mathbb{P}(\tau_j - \tau_k > x_j \text{ für } j < k, M_k = \tau_k) &= \mathbb{P}(\tau_j > x_j + t \text{ für } j < k) \\ &= \int_0^\infty \mathbb{P}(\tau_j > x_j + t \text{ für } j < k) d\mathbb{P}_{\tau_k}(t) = \left(\prod_{j < k} \mathbb{P}(\tau_j > x_j) \right) \int_0^\infty \mathbb{P}(\tau_j > t \text{ für } j < k) d\mathbb{P}_{\tau_k}(t) \\ &= \left(\prod_{j < k} \mathbb{P}(\tau_j > x_j) \right) \mathbb{P}(M_k = \tau_k). \end{aligned}$$

Division durch $\mathbb{P}(M_k = \tau_k) > 0$ liefert die Behauptung. □

Wir bezeichnen mit $(T_j)_{j \geq 1}$ die Folge der Wartezeiten zwischen den Zustandswechseln und mit $S_n = \sum_{j=1}^n T_j$ den Zeitpunkt des n -ten Zustandswechsels. Im Intervall $[S_n, S_{n+1})$ ist X_t konstant gleich dem Zustand Z_n .

- Zum Zeitpunkt 0 seien keine Anrufer im System, d.h. $Z_0 = X_0 = 0$. Der erste Kunde kommt zum Zeitpunkt $T_1 \sim \exp(\lambda_W)$ ins System. Auf $[0, S_1) = [0, T_1)$ ist $X_t = 0$.
- Ist der Prozess X bis zum Zeitpunkt S_n für $n \geq 1$ konstruiert, dann ist Z_{n-1} bekannt. Der Zustandswechsel erfolgt zur Zeit S_n , weil entweder ein neuer eingetroffen ist, oder eine Servicezeit zu Ende gegangen ist. Nach Lemma 1.3 ist beiden Fällen die Wartezeit bis zum Eintreffen des nächsten Kunden $\exp(\lambda_W)$ -verteilt und die Rest-Servicezeiten sind $\exp(\lambda_S)$ -verteilt. Alle diese Zeiten sind unabhängig. Es sei $Z_{n-1} = i$, sprich i Operatoren sind belegt.

Ist $i = 0$, dann wechselt man mit Wahrscheinlichkeit $p_{01} = 1$ in den Zustand $Z_n = 1$. Das geschieht nach einer $\exp(\lambda_W)$ -verteilten Wartezeit T_{n+1} .

Ist $i = k$, dann ist der nächste Zustand mit Wahrscheinlichkeit $p_{k,k-1} = 1$ gleich $Z_n = k - 1$. Das geschieht, wenn der erste Operator frei wird. Nach Lemma 1.3 erfolgt das nach einer $\exp(k\lambda_S)$ -verteilten Wartezeit T_{n+1} .

Ist $0 < i < k$, dann ist der nächste Zustand $Z_n = i + 1$ oder $Z_n = i - 1$. Ein neuer Kunde kommt nach einer $\exp(\lambda_W)$ -verteilten Zeit τ_W . Nach Lemma 1.3 wird von den i besetzten Operatoren der erste Operator

nach einer $\exp(i\lambda_S)$ -verteilten Servicezeit τ_S frei. Der Zustandswechsel von i nach $i + 1$ tritt ein, wenn $\tau_W < \tau_S$, andernfalls wechselt man in den Zustand $i - 1$. Die Übergangswahrscheinlichkeiten sind daher

$$p_{i,i+1} = \mathbb{P}(\tau_W < \tau_S) = \frac{\lambda_W}{\lambda_W + i\lambda_S} \quad \text{und} \quad p_{i,i-1} = \frac{i\lambda_S}{\lambda_W + i\lambda_S}.$$

Die Wartezeit $T_{n+1} = \min(\tau_W, \tau_S)$ bis zum Sprung in den nächsten Zustand ist nach Lemma 1.3 eine $\exp(\lambda_W + i\lambda_S)$ -verteilte Zufallsvariable.

Die Folge der Zustände wird durch eine Markov-Kette $(Z_n)_{n \geq 0}$ mit Übergangsmatrix

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{\lambda_S}{\lambda_W + \lambda_S} & 0 & \frac{\lambda_W}{\lambda_W + \lambda_S} & 0 & 0 & 0 \\ 0 & \frac{2\lambda_S}{\lambda_W + 2\lambda_S} & 0 & \frac{\lambda_W}{\lambda_W + 2\lambda_S} & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \frac{(k-1)\lambda_S}{\lambda_W + (k-1)\lambda_S} & 0 & \frac{\lambda_W}{\lambda_W + (k-1)\lambda_S} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (2)$$

beschrieben. Ist man im Zustand i ist die Wartezeit bis zum Sprung in den nächsten Zustand eine $\exp(\lambda_i)$ -verteilte Zufallsvariable mit

$$\lambda_0 = \lambda_W, \quad \lambda_i = \lambda_W + i\lambda_S \quad \text{und} \quad \lambda_k = k\lambda_S.$$

Das rekursive Verfahren legt den Prozess $(X_t)_{t \geq 0}$ auf dem Intervall $[0, S_\infty)$ fest, wobei

$$S_\infty = \lim_{n \rightarrow \infty} S_n$$

die Explosionszeit des Prozesses bezeichnet. Der Prozess heißt nicht explosiv, wenn fast sicher $S_\infty = \infty$. In diesem Fall ist $(X_t)_{t \geq 0}$ fast sicher auf ganz $[0, \infty)$ festgelegt.

Lemma 1.11 (nicht explosive Markov-Prozesse). *Die Bedingung*

$$c := \sup_{i \in E} \lambda_i < \infty \quad (3)$$

ist hinreichend dafür, dass der Markov-Prozess $(X_t)_{t \geq 0}$ nicht explodiert. Prozesse mit endlichen Zustandsraum sind stets nicht explosiv.

Beweis. Für eine $\exp(\lambda)$ -verteilte Zufallsvariable T ist

$$\mathbb{E}(e^{-T}) = \int_0^\infty \lambda e^{-(1+\lambda)t} dt = (1 + \lambda^{-1})^{-1}.$$

Wegen der Unabhängigkeit von $(Z_n)_{n \geq 0}$ und $(T_n)_{n \geq 1}$ gilt

$$\begin{aligned} \mathbb{E}(e^{-S_n}) &= \mathbb{E}\left(\prod_{\ell=1}^n e^{-T_\ell}\right) = \sum_{i_0, \dots, i_{n-1} \in E} \mathbb{E}\left(\prod_{\ell=1}^n e^{-T_\ell} \mathbb{1}_{\{Z_0=i_0, \dots, Z_{n-1}=i_{n-1}\}}\right) \\ &= \sum_{i_0, \dots, i_{n-1} \in E} \prod_{\ell=1}^n (1 + \lambda_{i_{\ell-1}}^{-1})^{-1} \mathbb{P}(Z_0 = i_0, \dots, Z_{n-1} = i_{n-1}) = (1 + c^{-1})^{-n}. \end{aligned}$$

Mittels monotoner Konvergenz folgt daraus $\mathbb{E}(e^{-S_\infty}) = 0$ und damit $\mathbb{P}(S_\infty = \infty) = 1$. \square

Im allgemeinen hat ein zeitstetiger Markov-Prozess mit abzählbaren Zustandsraum E folgende Bestimmungsstücke:

- ein beliebiger Startverteilungsvektor $\nu = (p_i)_{i \in E}$, der angibt mit welcher Wahrscheinlichkeit p_i der Prozess sich zum Zeitpunkt 0 im Zustand i befindet
- eine stochastische Matrix von Übergangswahrscheinlichkeiten $P = (p_{ij})_{i,j \in E}$ mit $p_{ii} = 0$ oder $p_{ii} = 1$ für alle $i \in E$. Sie gibt an, mit welcher Wahrscheinlichkeit p_{ij} man vom Zustand i in den Zustand j wechselt

- Einen Vektor von Intensitäten $\lambda = (\lambda_i)_{i \in E}$ mit $\lambda_i > 0$, wenn $p_{ii} = 0$ und $\lambda_i = 0$, wenn $p_{ii} = 1$ für alle $i \in E$. Ist $p_{ii} = 1$, dann nennt man den Zustand i absorbierend ($\lambda_i = 0$ entspricht der unendlichen Wartezeit bis zum nächsten Zustandswechsel). Andernfalls ist die Verweildauer im Zustand i eine $\exp(\lambda_i)$ -verteilte Zufallsvariable

Analog zum einleitenden Beispiel kann der Markov-Prozess $(X_t)_{t \geq 0}$ wie folgt simuliert werden:

1. Mit Wahrscheinlichkeit p_i ist $Z_0 = X_0 = i$ und $S_0 = 0$. Die Wartezeit bis zur ersten Zustandsänderung ist $T_1 \sim \exp(\lambda_i)$. Auf $[0, S_1) = [0, T_1)$ ist $X_t = i$.
2. Ist der Prozess X bis zum Zeitpunkt S_n mit $n \geq 1$ konstruiert, dann ist Z_{n-1} bekannt. Ist $Z_{n-1} = i$ und i absorbierend, dann verbleibt X in diesem Zustand und $T_{n+1} = \infty$. Andernfalls ist die Wartezeit T_{n+1} bis zum nächsten Zustandswechsel $\exp(\lambda_i)$ -verteilt und der nächste Zustand mit Wahrscheinlichkeit p_{ij} gleich $Z_n = j$. Auf $[S_n, S_{n+1})$ ist $X_t = i$.

Das legt den Prozess auf $[0, S_\infty)$ fest. Ein nicht explosiver Prozess ist auf ganz $[0, \infty)$ definiert. Dafür ist die Bedingung 3 hinreichend.

Bemerkung 1.12. *Voll ausgeschrieben bedeutet diese Konstruktion folgendes. Man wählt einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ auf dem eine Markov-Kette $Z = (Z_n)_{n \geq 0}$ mit Startverteilung ν_0 und Übergangsmatrix P , und eine davon unabhängige Folge unabhängiger Zufallsvariablen*

$$(\tau_{n,i})_{n \geq 1, i \in E}$$

existiert. Die $\tau_{n,i}$ seien $\exp(\lambda_i)$ -verteilt. Ist S_n für $n \geq 0$ konstruiert, setzt man

$$T_{n+1} := \begin{cases} \tau_{n+1,i} & Z_n = i, \lambda_i > 0 \\ \infty & Z_n = i, \lambda_i = 0 \end{cases}$$

und $X_t := Z_n$ für $t \in [S_n, S_{n+1})$, wobei $S_{n+1} = S_n + T_{n+1}$. Die Kette $Z = (Z_n)_{n \geq 0}$ heißt die in X eingebettete Markov-Kette.

1.6 Brownsche Bewegungen

Brownsche Bewegungen sind fundamentale zeitstetige Prozesse. Sie entstehen als Grenzprozesse normalisierter Irrfahrten, ähnlich wie Normalverteilungen als Grenzverteilungen normierter Summen unabhängiger Zufallsvariablen entstehen. Wir werden sehen, dass der Grenzprozess (sofern existent) die folgenden definierenden Eigenschaften haben muss.

Definition 1.13 (Brownsche Bewegung). *Ein reellwertiger Prozess $B = (B_t)_{t \geq 0}$ heißt (eindimensionale, standard) Brownsche Bewegung, wenn er folgende Eigenschaften hat:*

- i) $B_0 = 0$
- ii) Für $0 = t_0 < t_1 < \dots < t_n$ sind die Zuwächse $B_{t_1} - B_{t_0}, \dots, B_{t_n} - B_{t_{n-1}}$ unabhängig und $B_{t_i} - B_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$
- iii) alle Pfade von B sind stetig

Bemerkung 1.14. *Häufig wird nur verlangt, dass i) und iii) nur fast sicher erfüllt sind. Ist \tilde{B} so ein Prozesse und bezeichnet N die Ausnahmemenge auf denen i) oder iii) nicht gelten, dann ist der Prozess B der auf N^c mit \tilde{B} übereinstimmt und auf N identisch 0 ist, eine Brownsche Bewegung im Sinne von Definition 1.13.*

Die Definition kann direkt dazu benutzt werden eine Brownsche Bewegung auf einer beliebigen Menge von Stützstellen $0 = t_0 < t_1 < \dots < t_n$ als Summe der unabhängigen normalverteilten Zuwächse zu simulieren

$$B_0 = 0 \quad B_{t_1} = B_{t_1} - B_{t_0} \quad B_{t_2} = (B_{t_2} - B_{t_1}) + (B_{t_1} - B_{t_0}) \quad \dots \quad B_{t_n} = \sum_{j=1}^n (B_{t_j} - B_{t_{j-1}}).$$

Im Jahr 1827 beobachtete der Biologe Robert Brown erstmals unter dem Mikroskop die irreguläre Bewegung von Pollen in einer Flüssigkeit. Er erklärte diese Bewegung durch die Kollision mit den sich zufällig bewegendenden Molekülen der Flüssigkeit. Die Position eines im Ursprung startenden Pollenkorns kann durch den zweidimensionalen Prozess

$$\sigma(B_t^1, B_t^2)_{t \geq 0}$$

beschrieben werden, wobei B^1 und B^2 zwei unabhängige eindimensionale Brownsche Bewegungen sind, und $\sigma > 0$ eine von der Flüssigkeit abhängende Konstante.

Satz 1.15. *Brownsche Bewegungen existieren.*

Die erste mathematisch rigorose Konstruktion einer Brownschen Bewegung wurde 1923 von Norbert Wiener angegeben. Ihm zu Ehren heißen Brownsche Bewegungen auch Wiener Prozesse. Dazu muss man einen Wahrscheinlichkeitsraum und einen auf diesem Raum definierten Prozess $B = (B_t)_{t \geq 0}$ konstruieren, der die geforderten Eigenschaften hat. Dass das nicht selbstverständlich ist wird durch die merkwürdigen Eigenschaften einer Brownschen Bewegung deutlich. Man kann zeigen, dass fast sicher jeder Pfad einer Brownschen Bewegung an keiner einzigen Stelle differenzierbar ist. Der typische Pfad ist also eine Funktion im Sinne der klassischen Analysis ein pathologischer Sonderfall, nämlich überall stetig und nirgends differenzierbar.

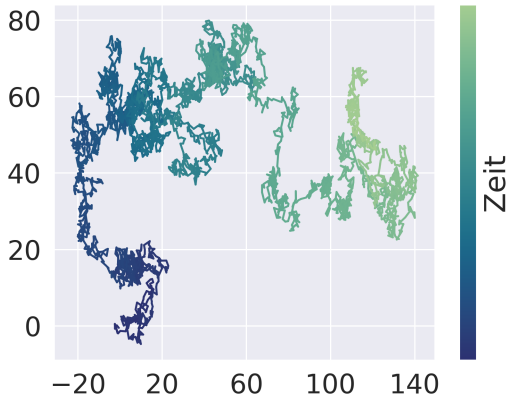


Abbildung 5: Beispiel einer zweidimensionalen Brownschen Bewegung

Zur Untersuchung ihrer Verteilungseigenschaften zerlegen wir die normalisierte Irrfahrt $S_t^{(n)} = \tilde{S}_t^{(n)} + \xi_t^{(n)}$ in die Teile

$$\tilde{S}_t^{(n)} := \frac{1}{\sqrt{\sigma^2 n}} S_{[nt]} \quad \text{und} \quad \xi_t^{(n)} = \frac{1}{\sqrt{\sigma^2 n}} (tn - [tn]) X_{[tn]+1}.$$

Dann gilt $\xi_t^{(n)} \xrightarrow{p} 0$ und $\mathbb{E}(\tilde{S}_t^{(n)}) = 0$ sowie $\text{var}(\tilde{S}_t^{(n)} - \tilde{S}_s^{(n)}) = \frac{[nt] - [ns]}{n} \sim t - s$. Aus dem zentralen Grenzwertsatz erhält man $\tilde{S}_t^{(n)} - \tilde{S}_s^{(n)} \xrightarrow{d} \mathcal{N}(0, t - s)$. Mit dem Lemma von Slutsky folgt daraus für $n \rightarrow \infty$

$$S_t^{(n)} - S_s^{(n)} \xrightarrow{d} \mathcal{N}(0, t - s).$$



Abbildung 4: Drei verschiedene eindimensionale Brownsche Bewegungen

Wir überlegen uns hier, dass die Zuwächse von normalisierten Irrfahrten gegen Zufallsvariablen konvergieren, welche die selben Verteilungseigenschaften haben wie die Zuwächse einer Brownschen Bewegung.

Sei $(X_n)_{n \geq 1}$ eine Folge unabhängiger identisch verteilter Zufallsvariablen mit $\mathbb{E}(X_n) = 0$ und Varianz $\text{var}(X_n) = \sigma^2 < \infty$. Setze $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$. Der Prozess $(S_n)_{n \geq 0}$ heißt allgemeine Irrfahrt auf \mathbb{R} . Wir betrachten den zeitstetigen Prozess $(S_u)_{u \geq 0}$ der zwischen den Werten der Irrfahrt $(S_n)_{n \geq 0}$ linear interpoliert, d.h.

$$S_u = S_n + (u - n)X_{n+1}.$$

Für jedes $n \geq 1$ ist die normalisierte Irrfahrt $S^{(n)} = (S_t^{(n)})_{t \geq 0}$ definiert durch

$$S_T^{(n)} := \frac{1}{\sqrt{\sigma^2 n}} S_{nt}.$$

Zur Untersuchung ihrer Verteilungseigenschaften zer-

2 Anhang

2.1 Ergebnisse aus der Wahrscheinlichkeitstheorie

Satz 2.1 (Lemma von Slutsky). *Seien X und $(X_n)_{n \geq 1}$ k -dimensionale Zufallsvektoren, $(Y_n)_{n \geq 1}$ ℓ -dimensionale Zufallsvektoren und $c \in \mathbb{R}^\ell$. Sei weiters $h: \mathbb{R}^k \times \mathbb{R}^\ell \rightarrow \mathbb{R}^m$ eine stetige Funktion, dann gilt*

$$X_n \xrightarrow{p} X \wedge Y_n \xrightarrow{p} \implies h(X_n, Y_n) \xrightarrow{p} h(X, c).$$