

Stochastische Prozesse

Vorlesungsnotizen

Stilianos Louca

April 2009

Inhaltsverzeichnis

1	Stochastische Prozesse	5
1.1	Existenz & Eigenschaften	5
1.1.1	Definition: Zufällige Größe	5
1.1.2	Definition: Kovarianz	5
1.1.3	Definition: Stochastischer Prozess	5
1.1.4	Definition: Pfad eines Prozesses	6
1.2	Endlich-dimensionale Verteilungen	7
1.2.1	Definition: Endlich dimensionale Verteilung zufälliger Prozesse	7
1.2.2	Definition: Äquivalente stochastische Prozesse	7
1.3	Kanonische Darstellung	8
1.3.1	Definition: Zylindermenge	8
1.3.2	Lemma über $\mathcal{C}_0(S^T)$	9
1.3.3	Definition: $\mathcal{C}(S^T)$	10
1.3.4	Lemma: Messbarkeit von Pfaden	10
1.3.5	Definition: Verteilungsgesetz der Pfade (\mathcal{P}_X)	10
1.3.6	Satz: Charakterisierung von Äquivalenz stochastischer Prozesse	11
1.3.7	Kanonische Form: Vorbetrachtung	11
1.3.8	Definition: Kanonisch Form stochastischer Prozesse	11
1.3.9	Satz: Stochastische Prozesse in kanonischer Form	12
1.3.10	Definition: Unabhängige stochastische Prozesse	12
1.3.11	Lemma: Charakterisierung der Unabhängigkeit von Prozessen	12
1.3.12	Satz: Existenz unabhängiger Prozesse	13
1.4	Pfideigenschaften	13
1.4.1	Definition: Fast sichere Pfade	13
1.4.2	Maßtheoretische Vorbetrachtung	14
1.4.3	Satz über fast-Sicherheit von Pfaden	14
1.5	Existenz bei vorgegebenen Verteilungen	15
1.5.1	Notwendige Bedingungen zur Existenz eines stochastischen Prozesses	16
1.5.2	Definition: Projektives System	16
1.5.3	Satz von Kolmogorov	16
1.5.4	Theorem: Existenz stochastischer Prozesse	19
1.5.5	Spezialfall: Weißes Rauschen	20
1.5.6	Korollar: Existenz unabhängiger Prozesse	21
1.6	Zuwächse von Prozessen (increments)	21
1.6.1	Definition: Zuwachs	21
1.6.2	Satz über Zuwächse: Existenz stochastischer Prozesse	22
1.7	Prozesse mit unabhängigen Zuwächsen	23
1.7.1	Definition: Prozess mit unabhängigen Zuwächsen	23
1.7.2	Lemma: Verschiebungsinvarianz von Zuwächsen	23
1.7.3	Satz: Existenz von N_0 -zeitigen Prozessen mit unabhängigen Zuwächsen	23
1.7.4	Definition: Faltung von Maßen	24
1.7.5	Lemma: Summe unabhängiger Zufallsvariablen	25

1.7.6	Notwendige Bedingung für die Unabhängigkeit der Zuwächse	25
1.7.7	Theorem: Existenz von Prozessen mit unabhängigen Zuwächsen	26
1.7.8	Poissonprozesse	28
1.7.9	Definition: Stationäre Zuwächse	29
1.7.10	Definition: Lévy-Prozess	29
1.7.11	Definition: Faltungsgruppe	29
1.8	Markov-Prozesse	30
1.8.1	Definition: Markovkern	30
1.8.2	Definition: Markov-Prozess	31
1.8.3	Die Chapman-Kolmogorov Bedingung für Markov-Prozesse	31
1.8.4	Hinreichende Bedingung für Markov-Prozesse	32
1.8.5	Definition: Zeitlich homogene Übergangswahrscheinlichkeiten	32
1.8.6	Definition: Erzeuger zeitlich homogener Übergangswahrscheinlichkeiten	32
1.8.7	Beispiel: Prozesse mit unabhängigen Zuwächsen	33
1.8.8	Beispiel: Reflektierte Brownsche Bewegung	34
1.8.9	Spezialfälle für Markov-Prozesse	35
2	Konvergenzbegriffe	37
2.1	Konvergenz von Maßen	37
2.1.1	Definition: Schwache Konvergenz von Maßen	37
2.1.2	Satz: Charakterisierung von schwacher Konvergenz	37
2.1.3	Satz: Schwache Konvergenz unter stetigen Transformationen	37
2.1.4	Hilfslemma über die Stetigkeit der Metrik	37
2.1.5	Satz: Eindeutigkeit des schwachen Grenzwertes	38
2.1.6	Lemma zur schwachen Konvergenz von Produktmaßen	39
2.2	Konvergenz von Zufallsvariablen	39
2.2.1	Definition: Konvergenz von Zufallsvariablen	39
2.2.2	Lemma: Charakterisierung von Konvergenz in Verteilung	40
2.2.3	Satz: Übertragung der Konvergenz in Verteilung	40
2.2.4	Der zentrale Grenzwertsatz (CLT)	41
2.2.5	Der mehrdimensionale zentrale Grenzwertsatz (CLT)	41
2.3	Konvergenz stochastischer Prozesse	41
2.3.1	Definition: Konvergenz stochastischer Prozesse	41
2.3.2	Theorem von Kratoski	41
2.3.3	Bemerkung über Borel- σ -Algebren in \mathcal{C}	42
2.3.4	Satz über Messbarkeit von Pfaden	42
2.3.5	Beispiel: Schwache Konvergenz von Pfad-Suprema	42
2.3.6	Definition: Radon-Maß	43
2.3.7	Definition: Gleichmäßig Radonsch	43
2.3.8	Definition: Relative Kompaktheit	44
2.3.9	Theorem von Prokhorov (1956)	44
2.3.10	Satz: Charakterisierung der Konvergenz stochastischer Prozesse	44
3	Gaußprozesse	46
3.1	Charakteristische Funktion	46
3.1.1	Konventionen	46
3.1.2	Definition: Charakteristische Funktion	46
3.1.3	Satz: Eigenschaften der charakteristischen Funktion	46
3.1.4	Theorem von Bochner	46
3.1.5	Satz: Verknüpfungen charakteristischer Funktionen	47
3.1.6	Satz von Cramér-Wold	48
3.1.7	Satz: Charakterisierung projektiver Systeme	48
3.1.8	Theorem: Existenz stochastischer Prozesse	49
3.2	Gaußsche Vektoren	50
3.2.1	Definition: Normalverteilung	50
3.2.2	Definition: Gaußvektor	50
3.2.3	Definition: Zentrierter Gaußvektor	51
3.2.4	Definition: Standardnormalverteilter Vektor	51

3.2.5	Lemma: Existenz standardnormalverteilter Vektoren	51
3.2.6	Satz: Charakterisierung zentrierter Gauß-Vektoren	52
3.2.7	Korollar für allgemeine Gaußvektoren	54
3.2.8	Lemma: Rotationsinvarianz Gaußscher Vektoren	55
3.2.9	Lemma: Unitäre Transformationen von Gaußvektoren	56
3.3	Gaußprozesse	56
3.3.1	Definition: Gaußprozess	56
3.3.2	Satz: Charakterisierung von Gaußprozessen durch Kovarianzfunktion	57
3.3.3	Satz über die Kovarianzfunktion	58
3.3.4	Theorem: Existenz von Gaußprozessen	58
3.3.5	Der Wiener-Prozess als Gauß-Prozess	59
3.3.6	Die Brownsche Brücke	62
3.3.7	Zufällige Gaußmaße	62
3.3.8	Gaußprozesse aus Integral-Kernen	64
3.3.9	Definition: Riemann-Liouville Operator	65
3.3.10	Der Riemann-Liouville-Prozess	65
3.4	Integrabilität von Gaußprozessen	66
3.4.1	Vorbetrachtung	66
3.4.2	Definition: Zusammensetzung von Prozessen	66
3.4.3	Lemma: Rotationsinvarianz Gaußscher Prozesse	66
3.4.4	Definition: Halbnorm	67
3.4.5	Lemma: Messbarkeit des Pfad-Supremums	67
3.4.6	Korollar über die Messbarkeit des Pfad-Supremums	67
3.4.7	Hilfslemma	68
3.4.8	Theorem von Ferniques (1970)	68
3.4.9	Hilfslemma über Erwartungswerte	70
3.4.10	Theorem über Gaußprozesse	70
3.4.11	Lemma über die p -Potenz von Pfadnormen	71
3.4.12	Theorem: Monotonie von Pfadnorm-Potenzen	72
3.4.13	Gaußprozesse auf Banachräumen	73
3.5	Reihendarstellung Gaußscher Prozesse	74
3.5.1	Einführung	74
3.5.2	Lemma: Reihenkonvergenz mit zufälligen Koeffizienten	75
3.5.3	Satz: Gaußsche Prozesse als Reihen	76
3.5.4	Hilfslemma über Gaußsche Integrale	77
3.5.5	Hilfslemma	77
3.5.6	Lemma: Supremum normalverteilter Zufallsvariablen	78
3.5.7	Der Wiener-Prozess als zufällige Reihe	78
3.5.8	Vorbemerkungen zu separablen Banachräumen	80
3.5.9	Theorem: Reihenkonvergenz beim Wiener-Prozess	81
4	Empirische Prozesse	82
4.0.10	Aufgabenstellung und Definition	82
4.0.11	Satz: Verteilungsgesetz der empirischen Verteilungsfunktion	83
4.0.12	Definition: Empirischer Prozess	84
4.0.13	Endlich dimensionale Verteilungen empirischer Prozesse	85
4.0.14	Definition: Kanonische Darstellung	85
4.0.15	Lemma über F^-	85
4.0.16	Satz: Erzeugung empirischer Prozesse durch Gleichverteilungen	86
4.1	Konvergenz der endlich dimensional Verteilungen	86
4.1.1	Satz: Momente von empirischen Prozessen	86
4.1.2	Satz über die Brownsche Brücke	87
4.2	Das Theorem von Glivenko und Cantelli (1933)	88
4.2.1	Vorbetrachtung	88
4.2.2	Theorem von Glivenko-Cantelli	88
4.3	Konvergenz der empirischen Prozesse	91
4.3.1	Vorbetrachtung	91
4.3.2	Stetiger Ersatz zu U_n^*	92

4.3.3	Satz: Konvergenz des approximierten empirischen Prozesses der Gleichverteilung	93
4.3.4	Definition: Stetigkeitsmodul	94
4.3.5	Eigenschaften des Stetigkeitsmoduls	94
4.3.6	Satz von Arzela-Ascoli	94
4.3.7	Lemma über Wegbeschränktheit	95
4.3.8	Satz: Charakterisierung relativer Kompaktheit	95
4.3.9	Charakterisierung relativer Kompaktheit von Maß-Mengen	96
4.3.10	Charakterisierung relativer Kompaktheit für stochastische Prozesse	98
4.3.11	Definition: Überdeckungszahl	98
4.3.12	Satz über Überdeckungszahlen und Stetigkeitsmodule	98
4.3.13	Satz zur relativen Kompaktheit von Maß-Mengen auf $\mathcal{C}[0, 1]$	99
4.3.14	Hilfslemma über die Konvergenz der geglätteten Gleichverteilung	100
4.3.15	Theorem von Donsker (1952)	101
4.4	Anwendungen des Konvergenzsatzes von Donster	104
4.4.1	Satz über den empirischen Prozess der Gleichverteilung und die Brownsche Brücke	104
4.4.2	Korollar: Konvergenz der Betragssuprema empirischer Prozesse	105
4.4.3	Korollar: Konvergenz der Suprema empirischer Prozesse	105
4.4.4	Lemma: Konvergenz der stetigen Approximation der Gleichverteilung	106
4.4.5	Theorem von Smirnov	106
4.4.6	Satz über die Supremum-Verteilung der Brownschen Brücke	108
4.4.7	Der Kolmogorov-Test als statistische Anwendung	108
4.4.8	Bemerkung: Interpretation des Kolmogorov-Tests	109
4.4.9	Hilfssaussage über punktweise konvergente Zufallsvariablen	110
4.4.10	Satz über die Konsistenz des Kolmogorov Tests	110
4.4.11	Verhältnis zwischen den beiden Kolmogorov-Test-Varianten	111
4.4.12	Kolmogorov-Smirnov-Test	112
4.4.13	Cramér von Mises Statistik	113
A	Anhang	114
A.0.14	Hilfslemma über die Gleichverteilung	114
A.0.15	Hilfslemma zur Verschiebungsinvarianz der empirischen Gleichverteilung	114
A.0.16	Hilfslemma: Eindeutigkeit der Pfad-Supremums-Verteilung	115
A.0.17	Maximalungleichung für unabhängige, symmetrische Zufallsgrößen	116
A.0.18	Maximalungleichungen für unabhängige, zentrierte Zufallsgrößen	116
A.0.19	Lemma über 4. Momente und Zufallsgrößen	116
A.0.20	Satz über 4. Momente der empirischen Gleichverteilung	117

1 Stochastische Prozesse

1.1 Existenz & Eigenschaften

1.1.1 Definition: Zufällige Größe

Es sei $(\Omega, \mathfrak{C}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, (S, \mathcal{S}) ein messbarer Raum und $X : \Omega \rightarrow S$ messbar. Dann heißt X *zufällige, S -wertige Variable*. Es bezeichne ferner \mathcal{P}_X das durch X induzierte Bildmaß auf (S, \mathcal{S}) , definiert durch

$$\mathcal{P}_X(B) := \mathcal{P}(X^{-1}(B)) \quad , \quad B \in \mathcal{S}$$

Dann heißt \mathcal{P}_X *Verteilungsgesetz* von X . Zufallsvariablen $X_1, X_2, \dots : \Omega \rightarrow S$ heißen *unabhängig*, falls

$$\mathcal{P}_{(X_1, X_2, \dots)} = \mathcal{P}_{X_1} \otimes \mathcal{P}_{X_2} \otimes \dots$$

Zwei Zufallsvariablen $X : \Omega \rightarrow S, Y : \Omega' \rightarrow S$ heißen *identisch verteilt*, falls $\mathcal{P}_X = \mathcal{P}_Y$.¹ Gegebenfalls schreibt man $X \stackrel{d}{=} Y$. Unabhängig, identisch verteilte X_1, X_2, \dots bezeichnet man auch als **I**dentically, **I**ndependently **D**istributed (iid).

Bemerkung: Für $S = \mathbb{R}$ nimmt man in der Regel an dass $\mathcal{S} = \mathcal{B}(\mathbb{R})$ die Borel- σ -Algebra über \mathbb{R} ist. In diesem Fall spricht man von einer *reellwertigen* Zufallsvariable. Bemerke dass dann Messbarkeit von X äquivalent ist zu

$$X^{-1}((-\infty, t]) \in \mathfrak{C} \quad \forall t \in \mathbb{R}$$

1.1.2 Definition: Kovarianz

Seien $X, Y : \Omega \rightarrow \mathbb{R}$ zufällige Größen. Dann heißt

$$\text{Cov}(X, Y) := \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y))$$

Kovarianz von X und Y . Für zufälligen Vektor $X : \Omega \rightarrow \mathbb{R}^n$ heißt die Matrix $R \in \mathbb{R}^{n \times n}$ definiert durch

$$R = (\text{Cov}(X^i, X^j))_{i,j=1}^n$$

Kovarianzmatrix von X (falls existent).

Bemerkung: Sind X, Y unabhängig, so ist $\text{Cov}(X, Y) = 0$.

1.1.3 Definition: Stochastischer Prozess

Es sei $T \neq \emptyset$ eine beliebige Indexmenge, $(\Omega, \mathfrak{C}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und (S, \mathcal{S}) ein messbarer Raum. Eine Familie $(X_t)_{t \in T}$ von S -wertigen Zufallsvariablen $X_t : \Omega \rightarrow S$ heißt *S -wertiger, stochastischer Prozess auf $(\Omega, \mathfrak{C}, \mathcal{P})$, parametrisiert durch T* . Man schreibt auch einfach $X = (X_t)_{t \in T}$ und nennt S den *Zustandsraum* des Prozesses.

Bemerkungen:

- Man kann sich alternativ einen stochastischen Prozess als Abbildung $X : T \times \Omega \rightarrow S$ vorstellen, mit $X(t, \cdot) : \Omega \rightarrow S$ messbar.
- Ist $T = [0, \infty)$, so beschreibt X einen Prozess, dessen Wert zu jedem Zeitpunkt $t \geq 0$ zufällig verteilt ist, beschrieben durch X_t . Der Wert $\mathcal{P}(X_t^{-1}(B))$ entspricht dann der Wahrscheinlichkeit zum Zeitpunkt t einen Wert aus $B \in \mathcal{S}$ zu beobachten.

¹Unabhängige, identisch verteilte Zufallsvariablen X_1, \dots, X_n nennt man manchmal *unabhängige Kopien*.

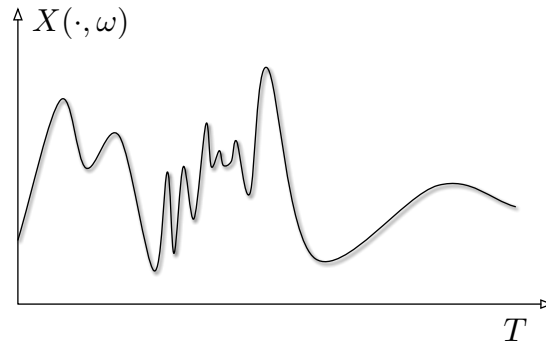


Abbildung 1.1: Beispielablauf eines stochastischen Prozesses für ein festes $\omega \in \Omega$. Jeder Wert $X_t(\omega)$ ist eigentlich zufällig verteilt.

- Die Abhängigkeit von X_t von den anderen Werten zu $s \neq t$ ist ein allgemeines mathematisches Problem.

1.1.4 Definition: Pfad eines Prozesses

Sei (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$ ein S -wertiger zufälliger Prozess auf $(\Omega, \mathfrak{C}, \mathcal{P})$. Es bezeichne

$$S^T := \{f \mid f : T \rightarrow S\}$$

Dann heißt die Abbildung $X : \Omega \rightarrow S^T$, definiert durch

$$X(\omega)(t) := X_t(\omega)$$

Pfad des stochastischen Prozesses X (beachte gleiches Symbol).²

Interpretation: Wobei der klassische Ansatz darin besteht, einen zufälligen Prozess als eine *sequenz* von zufälligen Variablen betrachtet, die zu jedem Zeitpunkt einen zufälligen Wert liefert, entspricht der Pfad des Prozesses (analytischer Ansatz) einem einmaligen *Durchlauf* des Experiments und anschließender *Auswertung* der Werte X_t .

Beispiele:

- Es sei zu jedem Zeitpunkt $t \geq 0$ der (zufällige) Luftdruck X_t an einem bestimmten Ort beobachtet. Dann wird durch den Pfad von X jeder *Weltgeschichte* ω der *Verlauf* $X_t(\omega)$ des Luftdrucks zugeordnet.

- Es sei 3 mal gewürfelt und:

X_1 : Wert des 1. Wurfes.

X_2 : Summe des 1. und 2. Wurfes.

X_3 : Summe des 1., 2. und 3. Wurfes.

In diesem Kontext entspräche der analytische Ansatz der Wahl einer (erlaubten) Funktion $f : \{1, 2, 3\} \rightarrow \mathbb{N}$, z.B.

$$f(1) = 3, \quad f(2) = 5, \quad f(3) = 9$$

Diese entspricht dem elementaren Ereignis $\omega = (3, 2, 4)$, mit $f(\cdot) = X_{(\cdot)}(\omega)$.

²Bilder der Pfad-Abbildung werden wir auch als *Pfade* bezeichnen.

1.2 Endlich-dimensionale Verteilungen

1.2.1 Definition: Endlich dimensionale Verteilung zufälliger Prozesse

Gegeben sei ein (S, \mathcal{S}) -wertiger stochastischer Prozess $(X_t)_{t \in T}$ auf $(\Omega, \mathfrak{G}, \mathcal{P})$. Dann heißt für $t_1, \dots, t_n \in T$ das Wahrscheinlichkeitsmaß $\mathcal{P}_{t_1, \dots, t_n}^X$ definiert durch

$$\mathcal{P}_{t_1, \dots, t_n}^X(B) := \mathcal{P}[\{\omega \in \Omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\}] \quad , \quad B \in \mathcal{S}^n$$

endlich dimensionale Verteilung des Prozesses. Ferner definiert man

$$\{\mathcal{P}_{t_1, \dots, t_n}^X\}_{\substack{n \in \mathbb{N} \\ t_i \in T}}$$

als die Menge der endlich dimensionalen Verteilungen des Prozesses.

Bemerkungen:

- Das Wahrscheinlichkeitsmaß $\mathcal{P}_{t_1, \dots, t_n}^X$ ist das durch den zufälligen Vektor

$$(X_{t_1}, \dots, X_{t_n}) : \Omega \rightarrow S^n$$

induzierte Bildmaß auf (S^n, \mathcal{S}^n) .

- $\mathcal{P}_{t_1, \dots, t_n}^X$ ist eindeutig durch die Werte auf den Würfeln

$$\{B_1 \times \dots \times B_n : B_i \in \mathcal{S}\}$$

bestimmt. $\mathcal{P}_{t_1, \dots, t_n}^X(\times_i B_i)$ entspricht dabei der Wahrscheinlichkeit, dass der Prozess zu jedem Zeitpunkt t_i einen Wert in B_i annimmt.

Beispiel: Betrachten einen sog. 1-dimensionalen *random walker* der zu jedem Zeitpunkt $n \in \mathbb{N}$ einen zufälligen Einheitsschritt nach links oder rechts, mit jeweils der Wahrscheinlichkeit $(1 - p)$ und p macht.

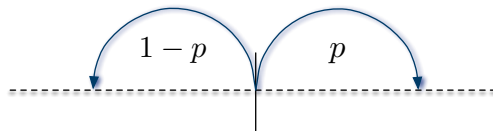


Abbildung 1.2: 1-dimensionaler random-walker.

Mit $T := \mathbb{N}_0$, $X_0 := 0$ und X_n als der Position nach n Schritten, ergibt sich beispielsweise

$$\mathcal{P}_{0,1}(\{(0, 1)\}) = p, \quad \mathcal{P}_{1,2}(\{(1, 2)\}) = p^2, \quad \mathcal{P}_{1,2}(\{(1, 0)\}) = p(1 - p)$$

1.2.2 Definition: Äquivalente stochastische Prozesse

Zwei T -parametrisierte S -wertige stochastische Prozesse $(X_t)_{t \in T}$, $(\tilde{X}_t)_{t \in T}$ jeweils auf $(\Omega, \mathfrak{G}, \mathcal{P})$ und $(\tilde{\Omega}, \tilde{\mathfrak{G}}, \tilde{\mathcal{P}})$ heißen *äquivalent (identisch verteilt)* $:\Leftrightarrow$

$$\forall t_1, \dots, t_n \in T : \mathcal{P}_{t_1, \dots, t_n}^X = \tilde{\mathcal{P}}_{t_1, \dots, t_n}^{\tilde{X}}$$

Man sagt \tilde{X} ist eine *Version* von X und schreibt $X \cong \tilde{X}$.

Bemerke: Äquivalenz stochastischer Prozesse ist eine Äquivalenzrelation.

Beispiel: Es sei $\mathcal{P} = \lambda$ das Lebesgue-Maß auf $\Omega := [0, 1]$ und

$$X_t(\omega) := \begin{cases} 1 & : \omega = t \\ 0 & : \text{sonst} \end{cases}, \quad t, \omega \in [0, 1] \quad (1.2.2.1)$$

Dann ist X äquivalent zu $\tilde{X} \equiv 0$, denn

$$\begin{aligned} \mathcal{P}(\{X_{t_i} \in B_i, i = 1, \dots, n\}) &= \mathcal{P}\left(\bigcap_i \{X_{t_i} \in B_i\}\right) = \begin{cases} 1 & : 0 \in \bigcap_i B_i \\ 0 & : \text{sonst} \end{cases} \\ &= \mathcal{P}(\{\tilde{X}_{t_i} \in B_i, i = 1, \dots, n\}) \end{aligned}$$

1.3 Kanonische Darstellung

1.3.1 Definition: Zylindermenge

Es sei $T \neq \emptyset$ eine beliebige Indexmenge und (S, \mathcal{S}) ein messbarer Raum. Dann heißt $\mathcal{Z} \subset S^T$ *Zylindermenge* $:\Leftrightarrow$

$$\mathcal{Z} = \{f \in S^T \mid (f(t_1), \dots, f(t_n)) \in B\}$$

für irgendwelche $t_1, \dots, t_n \in T$ und $B \in \mathcal{S}^n := \bigotimes_{i=1}^n \mathcal{S}$. Man nennt B die *Basis* von \mathcal{Z} und schreibt auch

$$\mathcal{Z} = \{(x_t)_{t \in T} : (x_{t_1}, \dots, x_{t_n}) \in B\}$$

Ferner setzt man

$$\mathcal{C}_0 := \mathcal{C}_0(S^T) := \{\mathcal{Z} \subset S^T : \mathcal{Z} \text{ Zylindermenge}\}$$

Beispiele:

- Für $T = \{t_1, t_2\}$, $S = \mathbb{R}$ und

$$\mathcal{Z} = \{f : T \rightarrow \mathbb{R} \mid f(t_1) \in B_1\}$$

ist

$$\mathcal{Z} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in B_1\}$$

d.h. identifizierbar mit einem Zylinder in \mathbb{R}^2 .

- Für $T := \mathbb{R}$, $S = \mathbb{R}$, $B = \underbrace{\{0\} \times \dots \times \{0\}}_{\times n}$ entspricht

$$\mathcal{Z} := \{f : \mathbb{R} \rightarrow \mathbb{R} : (f(t_1), \dots, f(t_n)) \in B\} = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f(t_i) = 0, i = 1, \dots, n\}$$

der Familie aller reellen Funktionen auf \mathbb{R} die zu den Zeitpunkten t_i den Wert 0 annehmen (siehe Abb. 1.3).

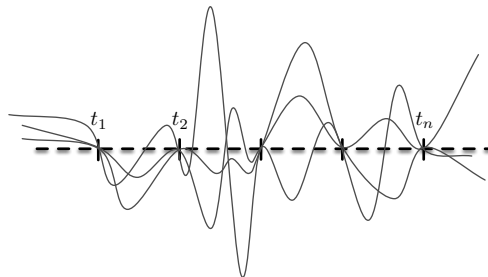


Abbildung 1.3: Zur Definition von Zylindermengen in $\mathbb{R}^{\mathbb{R}}$

Bemerkungen:

1. Für $t_1, \dots, t_n \in T$ und $f \in S^T$ setzen $\zeta_{t_1, \dots, t_n}(f) := (f(t_1), \dots, f(t_n)) \in S^n$. Somit ist

$$\zeta_{t_1, \dots, t_n} : S^T \rightarrow S^n$$

und

$$\mathcal{Z} = \zeta_{t_1, \dots, t_n}^{-1}(B)$$

eine Zylindermenge mit Basis B .

2. Für Permutation $\pi \in \text{Sym}(n)$ setzt man

$$T_\pi((x_1, \dots, x_n)) := (x_{\pi(1)}, \dots, x_{\pi(n)}) \quad (1.3.1.1)$$

Dann ist

$$T_\pi : S^n \rightarrow S^n$$

bijektiv und $T_\pi^{-1} = T_{\pi^{-1}}$.

3. Für messbare Menge $B \in \mathcal{S}^n$ und Permutation π ist offensichtlich auch $T_\pi(B) \in \mathcal{S}^n$ messbar. Mit

$$\mathcal{Z} = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\} = \{f \in S^T : (f(t_{\pi(1)}), \dots, f(t_{\pi(n)})) \in T_\pi(B)\}$$

wird ersichtlich: Durch Veränderung der Basismenge B einer Zylindermenge \mathcal{Z} kann man die entsprechenden t_1, \dots, t_n umordnen.

4. Ist $\{t_1, \dots, t_n\} \subset \{s_1, \dots, s_m\}$, $m \geq n$ so besitzt

$$\mathcal{Z} := \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\}$$

($B \in \mathcal{S}^n$) auch eine Darstellung mit s_1, \dots, s_m .

Beweis: Wegen 3. kann man o.B.d.A annehmen $t_1 = s_1, \dots, t_n = s_n$. Dann ist

$$\mathcal{Z} = \{f \in S^T : (f(s_1), \dots, f(s_m)) \in \underbrace{B \times S^{m-n}}_{\in \mathcal{S}^m}\}$$

5. Aus 4. folgt: $\mathcal{Z} \subset S^{\mathbb{N}}$ ist eine Zylindermenge \Leftrightarrow

$$\exists n \in \mathbb{N}, B \in \mathcal{S}^n : \mathcal{Z} = \{(x_i)_{i=1}^\infty : (x_1, \dots, x_n) \in B\}$$

1.3.2 Lemma über $\mathcal{C}_0(S^T)$

Es sei $T \neq \emptyset$ und (S, \mathcal{S}) ein messbarer Raum. Die Menge \mathcal{C}_0 aller Zylindermengen in S^T ist eine Algebra.

Bemerke: \mathcal{C}_0 ist im allgemeinen keine σ -Algebra.

Beweis:

- $\emptyset = \{f \in S^T : f(t_0) \in \emptyset\}$
- Sei $\mathcal{Z} = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\} \in \mathcal{C}_0$ für irgendein $B \in \mathcal{S}^n$ und $t_1, \dots, t_n \in T$. Dann ist auch

$$\mathcal{Z}^c = S^T \setminus \mathcal{Z} = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B^c\} \in \mathcal{C}_0$$

- Es seien

$$\mathcal{Z}_1 := \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B_1\}, \quad \mathcal{Z}_2 := \{f \in S^T : (f(s_1), \dots, f(s_m)) \in B_2\}$$

zwei Zylindermengen. Setzen $\{u_1, \dots, u_k\} := \{t_1, \dots, t_n\} \cup \{s_1, \dots, s_m\}$. Dann kann man nach obigen Bemerkungen schreiben

$$\mathcal{Z}_i = \{f \in S^T : (f(u_1), \dots, f(u_k)) \in \tilde{B}_i\}, \quad i = 1, 2$$

für geeignete $\tilde{B}_i \in \mathcal{S}^k$. Demnach

$$\mathcal{Z}_1 \cup \mathcal{Z}_2 = \{f \in S^T : (f(u_1), \dots, f(u_k)) \in \tilde{B}_1 \cup \tilde{B}_2\} \in \mathcal{C}_0$$

- Jedoch ist \mathcal{C}_0 allgemein keine σ -Algebra.

Gegenbeispiel: Setzen $\mathcal{Z}_n \subset \mathbb{R}^{\mathbb{N}}$ mit

$$\mathcal{Z}_n := \{(x_j)_{j=1}^{\infty} \subset \mathbb{R} : x_n = 0\} \in \mathcal{C}_0$$

und

$$\mathcal{Z} := \bigcap_{n \in \mathbb{N}} \mathcal{Z}_n = \{(x_j)_{j \geq 1} : x_j = 0 \forall j\}$$

Dann ist offensichtlich $\mathcal{Z} \notin \mathcal{C}_0$.

□

1.3.3 Definition: $\mathcal{C}(S^T)$

Sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und \mathcal{C}_0 die Algebra der Zylindermengen in S^T . Dann nennt man

$$\mathcal{C}(S^T) := \mathcal{C} := \sigma(\mathcal{C}_0(S^T))$$

(kleinste σ -Algebra die \mathcal{C}_0 enthält) die *Zylinder- σ -Algebra* in S^T .

Bemerkung: Jedes Element $A \in \mathcal{C}(S^T)$ besitzt die Darstellung

$$A = \{f \in S^T : (f(t_1), f(t_2), \dots) \in B\} \quad , \quad (t_i)_{i \in \mathbb{N}} \in T, B \in \mathcal{S}^{\mathbb{N}}$$

das heißt, hängt im wesentlichen von abzählbar vielen Parametern ab.
(Ohne Beweis)

1.3.4 Lemma: Messbarkeit von Pfaden

Es sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$ ein S -wertiger zufälliger Prozess auf $(\Omega, \mathfrak{G}, \mathcal{P})$. Dann ist der Pfad $X : \Omega \rightarrow S^T$ messbar von \mathfrak{G} nach $\mathcal{C}(S^T)$.

Beweis: Es genügt zu zeigen

$$X^{-1}(\mathcal{Z}) \in \mathfrak{G} \quad \forall \mathcal{Z} \in \mathcal{C}_0$$

Sei also

$$\mathcal{Z} := \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\} \in \mathcal{C}_0 \quad , \quad B \in \mathcal{S}^n$$

Dann ist

$$X^{-1}(\mathcal{Z}) = \{\omega \in \Omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\}$$

Da $(X_{t_1}, \dots, X_{t_n}) : \Omega \rightarrow S^n$ messbar ist, ist $X^{-1}(\mathcal{Z}) \in \mathfrak{G}$.

1.3.5 Definition: Verteilungsgesetz der Pfade (\mathcal{P}_X)

Sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$ ein S -wertiger stochastischer Prozess auf $(\Omega, \mathfrak{G}, \mathcal{P})$. Dann nennt man das Bildmaß

$$\mathcal{P}_X : \mathcal{C}(S^T) \rightarrow [0, \infty) \quad , \quad \mathcal{P}_X(Z) := \mathcal{P}(\{\omega \in \Omega : X(\omega) \in Z\}) \quad , \quad Z \in \mathcal{C}$$

das *Verteilungsgesetz der Pfade* von X .

Beispiele:

- Betrachten den reellwertigen stochastischen Prozess $(X_i)_{i \in \mathbb{N}}$ auf $(\Omega, \mathfrak{C}, \mathcal{P})$, dazu

$$A := \{(x_j)_{j \in \mathbb{N}} \in \mathbb{R} : \sup_{j \in \mathbb{N}} |x_j| \leq 1\} \in \mathcal{C}(\mathbb{R}^{\mathbb{N}})$$

Dann ist

$$\mathcal{P}_X(A) = \mathcal{P}\{\omega \in \Omega : \sup_{j \in \mathbb{N}} |X_j(\omega)| \leq 1\}$$

- Betrachten den reellwertigen Prozess $(X_t)_{0 \leq t \leq 1}$ und setzen

$$A := \{f \in \mathbb{R}^{[0,1]} : \lim_{n \rightarrow \infty} f(\frac{1}{n}) \text{ existiert}\} \in \mathcal{C}(\mathbb{R}^{[0,1]})$$

Dann ist

$$\mathcal{P}_X(A) = \mathcal{P}\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_{\frac{1}{n}}(\omega) \text{ existiert}\}$$

1.3.6 Satz: Charakterisierung von Äquivalenz stochastischer Prozesse

Es sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$, $(\tilde{X}_t)_{t \in T}$ zwei S -wertige stochastische Prozesse jeweils auf $(\Omega, \mathfrak{C}, \mathcal{P})$ und $(\tilde{\Omega}, \tilde{\mathfrak{C}}, \tilde{\mathcal{P}})$. Dann ist

$$X \cong \tilde{X} \Leftrightarrow \mathcal{P}_X = \tilde{\mathcal{P}}_{\tilde{X}}$$

Beweis: Zum einen gilt

$$X \cong \tilde{X} \Leftrightarrow \mathcal{P}\{[\omega \in \Omega : (X_{t_1}\omega, \dots, X_{t_n}\omega) \in B]\} = \tilde{\mathcal{P}}\{[\omega \in \tilde{\Omega} : (\tilde{X}_{t_1}\omega, \dots, \tilde{X}_{t_n}\omega) \in B]\}, \quad n \in \mathbb{N}, t_i \in T, B \in \mathcal{S}^n$$

$$\Leftrightarrow \mathcal{P}(\{X \in Z\}) = \tilde{\mathcal{P}}(\{\tilde{X} \in Z\}), \quad Z := \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\}$$

$$\Leftrightarrow \mathcal{P}_X|_{\mathcal{C}_0} = \tilde{\mathcal{P}}_{\tilde{X}}|_{\mathcal{C}_0}$$

Andererseits ist $S^T \in \mathcal{C}_0$, \mathcal{C}_0 ist \cap -Stabil und $\sigma(\mathcal{C}_0) = \mathcal{C}$. Nach dem Eindeutigkeitsatz für endliche Maße ist somit $X \cong \tilde{X}$ äquivalent zu $\mathcal{P}_X = \tilde{\mathcal{P}}_{\tilde{X}}$.

□

1.3.7 Kanonische Form: Vorbetrachtung

Sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und $(S^T, \mathcal{C}(S^T), \mu)$ ein Wahrscheinlichkeitsraum. Dann ist $(X_t)_{t \in T}$, definiert durch

$$X_t(f) := f(t), \quad f \in S^T \tag{1.3.7.1}$$

ein S -wertiger stochastischer Prozess auf $(S^T, \mathcal{C}(S^T), \mu)$.

Beweis: Zu zeigen wäre, dass $\forall t \in T$ die Abbildung $X_t : S^T \rightarrow S$ messbar ist. Tatsächlich ist für $B \in \mathcal{S}$:

$$X_t^{-1}(B) = \{f \in S^T : \underbrace{X_t(f)}_{f(t)} \in B\} \in \mathcal{C}_0(S^T) \subset \mathcal{C}(S^T)$$

Bemerke dass der Pfad $X : f \mapsto \underbrace{(t \mapsto f(t))}_f$ von X die Identität auf S^T ist.

1.3.8 Definition: Kanonisch Form stochastischer Prozesse

Sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$ ein stochastischer Prozess auf dem Wahrscheinlichkeitsraum $(S^T, \mathcal{C}(S^T), \mu)$ gegeben in der Form 1.3.7.1. Dann heißt X in *kanonischer Form*.

1.3.9 Satz: Stochastische Prozesse in kanonischer Form

Sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$ ein S -wertiger zufälliger Prozess auf $(\Omega, \mathfrak{C}, \mathcal{P})$. Dann existiert ein zu X äquivalenter Prozess \tilde{X} in kanonischer Form.

Beweis: Setzen $\mu := \mathcal{P}_X$ auf $(S^T, \mathcal{C}(S^T))$ und $\tilde{X}_t(f) := f(t)$. Zu zeigen wäre $X \cong \tilde{X}$. Tatsächlich ist für $t_1, \dots, t_n \in T$ und $B \in \mathcal{S}^n$:

$$\begin{aligned} \mu_{\tilde{X}_{t_1, \dots, t_n}}(B) &= \underbrace{\mu}_{\mathcal{P}_X}(\{f \in S^T : \underbrace{(Y_{t_1}(f), \dots, Y_{t_n}(f))}_{(f(t_1), \dots, f(t_n))} \in B\}) \\ &= \mathcal{P}\{\omega \in \Omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\} = \mathcal{P}_{t_1, \dots, t_n}^X(B) \end{aligned}$$

□

1.3.10 Definition: Unabhängige stochastische Prozesse

S -wertige stochastische Prozesse $(X_t^1)_{t \in T}, \dots, (X_t^n)_{t \in T}$ heißen *unabhängig*, wenn für alle $t_1, \dots, t_m \in T$ die Vektoren $(X_{t_1}^1, \dots, X_{t_m}^1), \dots, (X_{t_1}^n, \dots, X_{t_m}^n)$ unabhängig sind.

1.3.11 Lemma: Charakterisierung der Unabhängigkeit von Prozessen

Seien $(X_t^1)_{t \in T}, \dots, (X_t^n)_{t \in T}$ n (S, \mathcal{S}) -wertige stochastische Prozesse auf $(\Omega, \mathfrak{C}, \mathcal{P})$. Dann sind äquivalent:

1. $(X_t^1), \dots, (X_t^n)$ sind unabhängig
2. $\forall t_1^1, \dots, t_{m_1}^1, \dots, t_1^n, \dots, t_{m_n}^n \in T$ sind die Vektoren $(X_{t_1^1}^1, \dots, X_{t_{m_1}^1}^1), \dots, (X_{t_1^n}^n, \dots, X_{t_{m_n}^n}^n)$ unabhängig.
3. Für $Z_1, \dots, Z_n \in \mathcal{C}(S^T)$ gilt

$$\mathcal{P}(\{X^1 \in Z_1\} \cap \dots \cap \{X^n \in Z_n\}) = \mathcal{P}(\{X^1 \in Z_1\}) \cdot \dots \cdot \mathcal{P}(\{X^n \in Z_n\})$$

das heißt deren Pfade sind unabhängig verteilt.

Beweis: O.B.d.A. betrachten nur zwei Prozesse $(X_t)_{t \in T}, (Y_t)_{t \in T}$. Der allgemeinere Fall ist fast identisch.

1 \rightarrow 2: Seien $t_1, \dots, t_n, s_1, \dots, s_m \in T$ beliebig, dann sind

$$(X_{t_1}, \dots, X_{t_n}), (Y_{s_1}, \dots, Y_{s_m})$$

unabhängig, denn

$$\mathcal{P}\left[\underbrace{\{(X_{t_1}, \dots, X_{t_n}) \in A\}}_{\{(X_{t_1}, \dots, X_{t_n}) \in A \times S^m\}} \cap \underbrace{\{(Y_{s_1}, \dots, Y_{s_m}) \in B\}}_{\{(Y_{t_1}, \dots, Y_{s_m}) \in S^n \times B\}}\right] = \mathcal{P}\left[\underbrace{\{(X_{t_1}, \dots, X_{s_m}) \in A \times S^m\}}_{\{(X_{t_1}, \dots, X_{t_n}) \in A\}}\right] \cdot \mathcal{P}\left[\underbrace{\{(Y_{t_1}, \dots, Y_{s_m}) \in S^n \times B\}}_{\{(Y_{s_1}, \dots, Y_{s_m}) \in B\}}\right]$$

2 \rightarrow 1: Trivial.

2 \rightarrow 3: Nach (2) gilt die Aussage für $Z_1, Z_2 \in \mathcal{C}_0(S^T)$, das heißt

$$\mathcal{P}_{(X,Y)}|_{\mathcal{C}_0 \times \mathcal{C}_0} = \mathcal{P}_X \otimes \mathcal{P}_Y|_{\mathcal{C}_0 \times \mathcal{C}_0}$$

Da $\mathcal{C}_0 \times \mathcal{C}_0 \cap$ -stabil ist, $S^T \times S^T \in \mathcal{C}_0 \times \mathcal{C}_0$ und $\sigma(\mathcal{C}_0 \times \mathcal{C}_0) = \sigma(\mathcal{C}_0) \otimes \sigma(\mathcal{C}_0)$, gilt die Aussage auf ganz $\mathcal{C} \otimes \mathcal{C}$.

3 \rightarrow 2: Trivial.

□

1.3.12 Satz: Existenz unabhängiger Prozesse

Sei $(X_t)_{t \in T}$ ein (S, \mathcal{S}) -wertiger stochastischer Prozess und $n \in \mathbb{N}$. Dann existieren unabhängige, identisch zu X verteilte, (S, \mathcal{S}) -wertige stochastische Prozesse $(X_t^1)_{t \in T}, \dots, (X_t^n)_{t \in T}$.

Beweis: Es sei $X : T \times (\Omega, \mathfrak{C}, \mathcal{P}) \rightarrow S$ der *Prototyp-Prozess*. Definieren

$$\Omega' := \bigotimes_{j=1}^n \Omega, \quad \mathfrak{C}' := \bigotimes_{j=1}^n \mathfrak{C}, \quad \mathcal{P}' = \bigotimes_{j=1}^n \mathcal{P}$$

dazu die Projektionen

$$\pi_j : \Omega' \rightarrow \Omega, \quad \pi_j(\omega_1, \dots, \omega_n) := \omega_j, \quad j = 1, \dots, n$$

Setzt man

$$X_t^j := X_t \circ \pi_j, \quad t \in T, \quad j = 1, \dots, n$$

so sind $X^j : T \times \Omega' \rightarrow S$, $j = 1, \dots, n$ identisch zu X verteilte stochastische Prozesse:

$$\mathcal{P}'_{(X_{t_1}^j, \dots, X_{t_n}^j)} = \mathcal{P}' \circ (X_{t_1}^j, \dots, X_{t_n}^j)^{-1} = \underbrace{\mathcal{P}' \circ \pi_j^{-1}}_{\mathcal{P}} \circ (X_{t_1}, \dots, X_{t_n})^{-1} = \mathcal{P}_{(X_{t_1}, \dots, X_{t_n})}$$

Per Konstruktion sind sie außerdem unabhängig:

$$\begin{aligned} & \mathcal{P}'(\{\omega \in \Omega' : (X_{t_1}^1, \dots, X_{t_n}^1)(\omega) \in B_1\} \cap \dots \cap \{\omega \in \Omega' : (X_{t_1}^n, \dots, X_{t_n}^n)(\omega) \in B_n\}) \\ &= \mathcal{P}'(\{\omega \in \Omega' : (X_{t_1}, \dots, X_{t_n})(\omega_1) \in B_1\} \cap \dots \cap \{\omega \in \Omega' : (X_{t_1}, \dots, X_{t_n})(\omega_n) \in B_n\}) \\ &= \mathcal{P}'(\{\omega_1 \in \Omega : (X_{t_1}, \dots, X_{t_n})(\omega_1) \in B_1\} \times \dots \times \{\omega_n \in \Omega : (X_{t_1}, \dots, X_{t_n})(\omega_n) \in B_n\}) \\ &= \underbrace{\mathcal{P}'(\{\omega_1 \in \Omega : (X_{t_1}, \dots, X_{t_n})(\omega_1) \in B_1\})}_{\mathcal{P}'(\{\omega \in \Omega' : (X_{t_1}, \dots, X_{t_n})(\omega_1) \in B_1\})} \cdot \dots \cdot \underbrace{\mathcal{P}'(\{\omega_n \in \Omega : (X_{t_1}, \dots, X_{t_n})(\omega_n) \in B_n\})}_{\mathcal{P}'(\{\omega \in \Omega' : (X_{t_1}, \dots, X_{t_n})(\omega_n) \in B_n\})} \\ &= \mathcal{P}'(\{\omega \in \Omega' : (X_{t_1}^1, \dots, X_{t_n}^1)(\omega) \in B_1\}) \cdot \dots \cdot \mathcal{P}'(\{\omega \in \Omega' : (X_{t_1}^n, \dots, X_{t_n}^n)(\omega) \in B_n\}) \end{aligned}$$

□

1.4 Pfadigenschaften

Betrachtet sei ein von einem Parameter $t \in T$ abhängiger, zufälliger Vorgang, beschrieben durch den stochastischen Prozess $(X_t)_{t \in T}$. Dabei beschreiben per Konstruktion zu X äquivalente stochastische Prozesse den gleichen Vorgang. Ziel ist es oft, ein geeignetes X mit besonderen Eigenschaften zu finden.

1.4.1 Definition: Fast sichere Pfade

Sei $T \neq \emptyset$, (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$ ein S -wertiger stochastischer Prozess auf $(\Omega, \mathfrak{C}, \mathcal{P})$. Zu beliebiger Untermenge $E \subset S^T$ sagt man X hat *fast sichere Pfade in E* $:\Leftrightarrow$

$$\exists N \in \mathfrak{C} : \mathcal{P}(N) = 0 \wedge X(\omega) \in E \quad \forall \omega \notin N$$

Bemerkungen:

- E muss nicht unbedingt in \mathcal{S} liegen. Ist dies jedoch der Fall, so ist $\mathcal{P}_X(E) = 1$ äquivalent zur fast-Sicherheit in E . Hieraus folgt insbesondere, dass X fast-sicher Pfade in E hat, genau dann wenn alle Versionen von X fast-sicher Pfade in E haben (vgl. Satz 1.3.6)!

Beweis:

$$1 = \mathcal{P}_X(E) = \underbrace{\mathcal{P}(X^{-1}(E))}_{=: N^c \in \mathfrak{C}} \Rightarrow \mathcal{P}(N) = 0 \wedge X(N^c) \subset E$$

$$X(N^c) \subset E \wedge \mathcal{P}(N) = 0 \Rightarrow \mathcal{P}_X(E) = \underbrace{\mathcal{P}(X^{-1}(E))}_{\supset N^c} \geq \mathcal{P}(N^c) = 1$$

- Fast-Sicherheit irgendeiner Pfad Eigenschaft ist jedoch im allgemeinen **nicht** Versions-invariant. Als Beispiel sei der Prozess im Beispiel (Gl. 1.2.2.1) in Def. 1.2.2 betrachtet. Zwar sind $X \cong \tilde{X}$, doch besitzt X fast sicher keine stetigen Pfade, \tilde{X} aber schon! Somit ist insbesondere $\mathcal{C}[0, 1] \notin \mathcal{C}(\mathbb{R}^{[0,1]})$.

Beispiel: Für topologische Räume $(T, \mathcal{O}(T))$, $(S, \mathcal{O}(S))$ und S -wertigen $(S$ ausgestattet mit der Borel- σ -Algebra $\mathcal{B}(S))$ zufälligen Prozess $(X_t)_{t \in T}$ auf $(\Omega, \mathfrak{C}, \mathcal{P})$, hat X fast sicher stetige Pfade, falls für irgendein $N \in \mathfrak{C}$ mit $\mathcal{P}(N) = 0$ und alle $\omega \notin N$ gilt:

$$X_t(\omega) \xrightarrow{t \rightarrow t_0} X_{t_0}(\omega), \quad t_0 \in T$$

X hat fast sicher beschränkte Pfade, falls für $\omega \notin N$ stets gilt:

$$\sup_{t \in T} |X_t(\omega)| < \infty$$

1.4.2 Maßtheoretische Vorbetrachtung

Betrachten den endlichen Maßraum $(\Omega, \mathfrak{C}, \mu)$ und beliebige Menge $E \subset \Omega$ (nicht unbedingt messbar). Dann ist

$$\mathfrak{C}_E := \{A \cap E : A \in \mathfrak{C}\}$$

eine σ -Algebra in E (*Spur- σ -Algebra*). Ziel ist es nun μ auf \mathfrak{C}_E entsprechend *einzu*schränken (bzw. Fortzusetzen). Betrachten dazu das so genannte *äußere Maß*

$$\mu^*(B) := \inf \{\mu(A) : A \supset B, A \in \mathfrak{C}\}, \quad B \subset \Omega$$

Es ist sogar der Fall, das zu $B \subset \Omega$ ein $\tilde{B} \in \mathfrak{C}$ existiert, mit $B \subset \tilde{B}$ und $\mu(\tilde{B}) = \mu^*(B)$ (Stetigkeit von oben). Man nennt \tilde{B} *messbare Hülle* von B . Man definiert nun das Maß μ_E auf \mathfrak{C}_E :

$$\mu_E(A \cap E) := \mu(A \cap \tilde{E}), \quad A \in \mathfrak{C}$$

Tatsächlich kann gezeigt werden, dass μ_E wohldefiniert (d.h. unabhängig von der Wahl von \tilde{E}) und ein Maß auf \mathfrak{C}_E ist. Man nennt μ_E das *Spurmaß* auf E .

1.4.3 Satz über fast-Sicherheit von Pfaden

Ein (S, \mathcal{S}) -wertiger stochastischer Prozess $(X_t)_{t \in T}$ auf $(\Omega, \mathfrak{C}, \mathcal{P})$ hat genau dann eine Version mit fast-sicheren Pfaden in $E \subset S^T$, wenn $\mathcal{P}_X^*(E) = 1$ gilt.

Beweis: Richtung "⇒". Sei $Y \cong X$ ein stochastischer Prozess auf $(\tilde{\Omega}, \tilde{\mathfrak{C}}, \tilde{\mathcal{P}})$ mit Pfaden f.s. in E . Nach Satz 1.3.6 gilt $\mu := \mathcal{P}_X = \tilde{\mathcal{P}}_Y$ und

$$\exists N \in \tilde{\mathfrak{C}} : \tilde{\mathcal{P}}(N) = 0 \wedge Y(N^c) \subset E$$

Sei nun $C \in \mathcal{C}(S^T)$ mit $C \supset E$ beliebig. Wegen $Y(N^c) \subset E \subset C$ folgt

$$N^c \subset \{\omega \in \tilde{\Omega} : Y(\omega) \in C\} \Rightarrow \underbrace{\tilde{\mathcal{P}}(N^c)}_{=1} \leq \tilde{\mathcal{P}}(\{\omega : Y(\omega) \in C\})$$

das heißt $\mu(C) = 1$. Somit ist auch

$$\mu^*(E) = \inf \{ \mu(C) : C \supset E, C \in \mathcal{C} \} = 1$$

Richtung "⇐": Sei $\mu^*(E) = 1$, dazu der oben (1.4.2) eingeführte Wahrscheinlichkeitsraum

$$(E, \mathcal{C}_E, \mu_E)$$

Dann gilt

$$\mu_E(C \cap E) = \mu(C) \quad \forall C \in \mathcal{C} \quad (1.4.3.1)$$

Setzen nun $Y_t(f) := f(t)$, $f \in E$. Dann ist $(Y_t)_{t \in T}$ ein S -wertiger stochastischer Prozess auf $(E, \mathcal{C}_E, \mu_E)$, denn für $t \in T$ und $B \in \mathcal{S}$ ist

$$\{f \in E : Y_t(f) \in B\} = \{f \in E : f(t) \in B\} = E \cap \underbrace{\{f \in S^T : f(t) \in B\}}_{\in \mathcal{C}_0(S^T)} \in \mathcal{C}_E$$

das heißt $Y_t : E \rightarrow S$ ist messbar. Ferner ist $X \cong Y$, denn für $t_1, \dots, t_n \in T$ und $B \in S^T$:

$$\begin{aligned} \mu_E[f : \{(Y_{t_1}f, \dots, Y_{t_n}f) \in B\}] &= \mu_E \{f \in E : (f(t_1), \dots, f(t_n)) \in B\} = \mu_E \left[E \cap \underbrace{\{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\}}_{\in \mathcal{C}} \right] \\ &\stackrel{(1.4.3.1)}{=} \underbrace{\mu}_{\mathcal{P}_X} \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\} = \mathcal{P}(\omega \in \Omega : (X_{t_1}\omega, \dots, X_{t_n}\omega) \in B) \end{aligned}$$

Per Konstruktion hat Y fast-sicher Pfade in E .

□

Beispiel: Für reellwertigen stochastischen Prozess $(X_t)_{t \in T}$ auf $(\Omega, \mathfrak{G}, \mathcal{P})$ setze

$$\mathfrak{B}(T) := \{f \in \mathbb{R}^T : \sup_{t \in T} |f(t)| < \infty\}$$

Dann hat X eine Version mit Pfaden fast-sicher in $\mathfrak{B}(T)$ wenn $\forall C \in \mathcal{C}(\mathbb{R}^T)$ mit $\mathfrak{B}(T) \subseteq C$ genau dann wenn stets $\mathcal{P}_X(C) = 1$ gilt. Für beliebige abzählbare Teilmenge $T_0 = \{t_1, t_2, \dots\} \subset T$ setzen nun

$$\mathfrak{B}(T_0) := \{f \in \mathbb{R}^T : \sup_{t_j \in T_0} |f(t_j)| < \infty\} \in \mathcal{C}(\mathbb{R}^T)$$

Wegen $\mathfrak{B}(T) \subseteq \mathfrak{B}(T_0)$ muss dann gelten

$$\mathcal{P}_X(\mathfrak{B}(T_0)) = \mathcal{P}\{\omega : \sup_{t_j \in T_0} |X_{t_j}(\omega)| < \infty\} = 1$$

Man kann sogar zeigen, dass falls $\mathcal{P}_X(\mathfrak{B}(T_0)) = 1$ für alle solche T_0 gilt, so besitzt X eine Version mit fast-sicher Pfaden in $\mathfrak{B}(T)$.

1.5 Existenz von Prozessen mit vorgegebenen endlich-dim. Verteilungen

Betrachten (S, \mathcal{S}) -wertigen stochastischen Prozess $(X_t)_{t \in T}$ auf $(\Omega, \mathfrak{G}, \mathcal{P})$. Im allgemeinen existiert eine starke Abhängigkeit zwischen den einzelnen Zufallsvariablen X_t , die nicht allein durch die Angabe der einzelnen \mathcal{P}_{X_t} beschrieben werden kann. Benötigt wird die Angabe der endlich-dimensionalen Verteilungen $\mathcal{P}_{t_1, \dots, t_n}^X$ auf S^n .

Es stellt sich die Frage, inwiefern durch Vorgabe einer Familie

$$(Q_{t_1, \dots, t_n})_{\substack{n \in \mathbb{N} \\ t_i \in T}}$$

von Wahrscheinlichkeitsmaßen auf (S^n, \mathcal{S}^n) , diese Assoziiert werden kann mit den endlich-dimensionalen Verteilungen $\mathcal{P}_{t_1, \dots, t_n}^X$ irgendeines stochastischen Prozesses.

1.5.1 Notwendige Bedingungen zur Existenz eines stochastischen Prozesses

Sei (S, \mathcal{S}) ein messbarer Raum und $T \neq \emptyset$. Zu Permutation $\pi \in \text{Sym}(n)$ setze

$$T_\pi : S^n \rightarrow S^n, \quad T_\pi(x_1, \dots, x_n) := (x_{\pi(1)}, \dots, x_{\pi(n)})$$

Zu Familie von Wahrscheinlichkeitsmaßen

$$(Q_{t_1, \dots, t_n})_{\substack{n \in \mathbb{N} \\ t_i \in T \\ t_i \neq t_j \neq i}}$$

auf (S^n, \mathcal{S}^n) existiere ein S -wertiger stochastischer Prozess $(X_t)_{t \in T}$ so dass $Q_{t_1, \dots, t_n} = \mathcal{P}_{t_1, \dots, t_n}^X$. Dann muss gelten:

1. Für Permutation $\pi \in \text{Sym}(n)$:

$$Q_{t_1, \dots, t_n} = Q_{t_{\pi(1)}, \dots, t_{\pi(n)}} \circ T_\pi \quad (1.5.1.1)$$

2. Zu $B \in \mathcal{S}^{n-1}$ muss gelten

$$Q_{t_1, \dots, t_n}(B \times S) = Q_{t_1, \dots, t_{n-1}}(B) \quad (1.5.1.2)$$

Beweis:

1. Analog zu 1.3.1 ist

$$\mathcal{P}_{t_1, \dots, t_n}^X(B) = \mathcal{P}[\{(X_{t_1}, \dots, X_{t_n}) \in B\}] = \mathcal{P}[\{(X_{t_{\pi(1)}}, \dots, X_{t_{\pi(n)}}) \in T_\pi(B)\}] = \mathcal{P}_{t_{\pi(1)}, \dots, t_{\pi(n)}}^X(T_\pi(B))$$

2. Für $B \in \mathcal{S}^{n-1}$ ist

$$\mathcal{P}_{t_1, \dots, t_n}^X(B \times S) = \mathcal{P}[\{(X_{t_1}, \dots, X_{t_{n-1}}) \in B, X_{t_n} \in S\}] = \mathcal{P}[\{(X_{t_1}, \dots, X_{t_n}) \in B\}] = \mathcal{P}_{t_1, \dots, t_{n-1}}^X(B)$$

□

1.5.2 Definition: Projektives System

Es sei $T \neq \emptyset$ und (S, \mathcal{S}) ein messbarer Raum. Eine Familie von Wahrscheinlichkeitsmaßen

$$(Q_{t_1, \dots, t_n})_{\substack{n \in \mathbb{N}, t_i \in T \\ t_i \neq t_j \neq i}}$$

die die Bedingungen 1.5.1.1 und 1.5.1.2 erfüllt, heißt *projektives System* auf (S, \mathcal{S}) über T .

1.5.3 Satz von Kolmogorov

Sei S ein separabler, vollständiger, metrischer Raum und $\mathcal{S} := \mathcal{B}(S)$ die Borel- σ -Algebra von S . Für jedes projektive System $(Q_{t_1, \dots, t_n})_{\substack{t_i \in T \\ t_i \neq t_j \neq i}}$ auf (S, \mathcal{S}) existiert ein bestimmtes Wahrscheinlichkeitsmaß Q auf $(S^T, \mathcal{C}(S^T))$

mit

$$Q(\{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\}) = Q_{t_1, \dots, t_n}(B), \quad t_i \neq t_j \neq i, \quad B \in \mathcal{S}^n$$

Beweis:

- Ziel: Konstruktion eines endlichen Inhalts Q_0 auf $\mathcal{C}_0(S^T)$. Setzen

$$Q_0(Z) := Q_{t_1, \dots, t_n}(B) \quad (1.5.3.1)$$

falls

$$Z = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\}$$

(i) Zeigen: Q_0 ist wohlbestimmt, das heißt unabhängig von der Darstellung von Z . Sei also

$$Z = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B_1\} = \{f \in S^T : (f(s_1), \dots, f(s_m)) \in B_2\} ,$$

dazu $\{u_1, \dots, u_k\} := \{t_1, \dots, t_n\} \cup \{s_1, \dots, s_m\}$. Dann ist

$$Z = \{f \in S^T : (f(u_1), \dots, f(u_k)) \in B_3\}$$

für irgendein $B_3 \in \mathcal{S}^n$ (vgl. Bemerkungen in 1.3.1). Da es nun genügt zu zeigen

$$Q_{t_1, \dots, t_n}(B_1) = Q_{u_1, \dots, u_k}(B_3)$$

nehmen wir o.B.d.A an $\{t_1, \dots, t_n\} \subset \{s_1, \dots, s_m\}$.³

Sei $\pi \in \text{Sym}(m)$ eine Permutation, so dass $t_1 = s_{\pi(1)}, \dots, t_n = s_{\pi(n)}$. Dann ist

$$Z = \{f \in S^T : (f(s_{\pi(1)}), \dots, f(s_{\pi(m)})) \in T_\pi(B_2)\}$$

und somit

$$Q_{s_{\pi(1)}, \dots, s_{\pi(n)}}(T_\pi(B_2)) \stackrel{(1.5.1.1)}{=} Q_{s_1, \dots, s_m}(T_\pi^{-1}T_\pi B_2) = Q_{s_1, \dots, s_m}(B_2) .$$

Wir können also o.B.d.A annehmen $t_1 = s_1, \dots, t_n = s_n$, das heißt

$$Z = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B_1\} = \{f \in S^T : (f(t_1), \dots, f(t_n), f(s_{n+1}), \dots, f(s_m)) \in B_2\}$$

was impliziert $B_2 = B_1 \times S^{m-n}$. Nach Voraussetzung (1.5.1.2) gilt dann

$$Q_{s_1, \dots, s_m}(B_2) = Q_{t_1, \dots, t_n, s_{n+1}, \dots, s_m}(B_1 \times S^{m-n}) \stackrel{(1.5.1.1)}{=} Q_{t_1, \dots, t_n}(B_1) .$$

(ii) Es ist $Q_0(S^T) = Q_0(f \in S^T : f(t_0) \in S) = Q_{t_0}(S) = 1$ für irgendein $t_0 \in T$.

(iii) Analog ist auch $Q_0(\emptyset) = \{f \in S^T : f(t_0) \in \emptyset\} = Q_{t_0}(\emptyset) = 0$.

(iv) Zu zeigen bleibt die Additivität. Seien also

$$Z_1 = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B_1\}$$

$$Z_2 = \{f \in S^T : (f(s_1), \dots, f(s_m)) \in B_2\}$$

mit $Z_1 \cap Z_2 = \emptyset$. O.B.d.A können wir annehmen $m = n$ und $t_i = s_i$ (vgl. Bemerkungen in 1.3.1). Dann ist

$$Z_1 \cup Z_2 = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B_1 \cup B_2\}$$

Da $Z_i = \zeta_{t_1, \dots, t_n}^{-1}(B_i)$, $Z_1 \cap Z_2 = \emptyset$ und ζ_{t_1, \dots, t_n} surjektiv ist, ist auch $B_1 \cap B_2 = \emptyset$, demnach

$$Q_0(Z_1 \cup Z_2) = Q_{t_1, \dots, t_n}(B_1 \cup B_2) = Q_{t_1, \dots, t_n}(B_1) + Q_{t_1, \dots, t_n}(B_2) = Q_0(Z_1) + Q_0(Z_2) .$$

• Zu zeigen: Q_0 ist σ -additiv, das heißt für paarweise disjunkte $Z_1, Z_2, \dots \in \mathcal{C}_0$ mit $\biguplus_{i=1}^\infty Z_i \in \mathcal{C}_0$ gilt

$$Q_0\left(\bigcup_{i \in \mathbb{N}} Z_i\right) = \sum_{i=1}^\infty Q_0(Z_i) .$$

Dies ist äquivalent zur Stetigkeit in der leeren Menge

$$Z_1 \supset Z_2 \supset \dots \in \mathcal{C}_0, \bigcap_{i \in \mathbb{N}} Z_i = \emptyset \Rightarrow Q_0(Z_i) \xrightarrow{i \rightarrow \infty} 0$$

bzw.

$$Z_1 \supset Z_2 \supset \dots \in \mathcal{C}_0, \inf_{i \in \mathbb{N}} Q_0(Z_i) > 0 \Rightarrow \bigcap_{i \in \mathbb{N}} Z_i \neq \emptyset .$$

³Andernfalls würde die Behauptung aus $Q_{t_1, \dots, t_n}(B_1) \stackrel{\{t_i\} \subseteq \{u_i\}}{=} Q_{u_1, \dots, u_k}(B_3) \stackrel{\{s_i\} \subseteq \{u_i\}}{=} Q_{s_1, \dots, s_m}(B_2)$ folgen.

Zeigen also für $Z_1 \supset Z_2 \supset \dots \in \mathcal{C}_0$ mit $Q_0(Z_i) \geq \delta > 0 \quad \forall i$:

$$\bigcap_{i \in \mathbb{N}} Z_i \neq \emptyset .$$

Dabei wählen o.B.d.A Folgen $t_1, t_2, \dots \in T$ und $n_1 \leq n_2 \leq \dots \in \mathbb{N}$ mit

$$Z_i = \{f \in S^T : (f(t_1), \dots, f(t_{n_i})) \in B_i\}$$

Dies ist immer möglich, da für jede Zylindermenge die Indexmenge t_1, \dots, t_n vergrößert werden kann ohne Z zu ändern. Somit ist

$$Q_0(Z_i) = Q_{t_1, \dots, t_{n_i}}(B_i) \geq \delta \quad \forall i$$

Nach Voraussetzung an (S, \mathcal{S}) sind die Wahrscheinlichkeitsmaße Q_{t_1, \dots, t_n} inner-regulär, das heißt

$$Q_{t_1, \dots, t_{n_i}}(B_i) = \sup \{Q_{t_1, \dots, t_{n_i}}(K) : K \subset B_i \text{ kompakt}\}$$

so dass es kompakte Mengen $(K_i)_{i \in \mathbb{N}}$ gibt mit

$$K_i \subset B_i \quad \wedge \quad Q_{t_1, \dots, t_{n_i}}(B_i \setminus K_i) \leq \frac{1}{2} \frac{\delta}{2^i}$$

Setzen

$$Z'_i := \{f \in S^T : (f(t_1), \dots, f(t_{n_i})) \in K_i\}$$

Dann $Z'_i \subset Z_i$ und $Q_0(Z_i \setminus Z'_i) \leq \frac{1}{2} \frac{\delta}{2^i}$. Setzen ferner

$$V_i := Z'_1 \cap \dots \cap Z'_i \in \mathcal{C}_0$$

Dann ist $V_1 \supset V_2 \supset \dots$ und $V_i \subset Z'_i \subset Z_i$. Betrachten

$$\begin{aligned} Q_0(Z_i) - Q_0(V_i) &= Q_0(Z_i \setminus V_i) = Q_0\left[Z_i \setminus \left(\bigcap_{j=1}^i Z'_j\right)\right] \\ &= Q_0\left[Z_i \cap \bigcup_{j=1}^i (Z'_j)^c\right] = Q_0\left[\bigcup_{j=1}^i \underbrace{(Z_i \cap (Z'_j)^c)}_{\subset Z_j}\right] \\ &\leq Q_0\left[\bigcup_{j=1}^i (Z_j \setminus Z'_j)\right] \leq \sum_{j=1}^i Q_0(Z_j \setminus Z'_j) \leq \frac{1}{2} \sum_{j=1}^i \frac{\delta}{2^j} \leq \frac{\delta}{2} \end{aligned}$$

und sehen

$$Q_0(V_i) \geq \underbrace{Q_0(Z_i)}_{\geq \delta} - \frac{\delta}{2} > 0$$

das heißt $V_i \neq \emptyset \quad \forall i$. Wählen nun $f_i \in V_i$, dann

$$f_k \in V_k \subset V_i \subset Z'_i \quad \forall k \geq i$$

das heißt

$$(f_k(t_1), \dots, f_k(t_{n_i})) \in K_i \quad \forall k \geq i$$

Da K_1 kompakt ist, ist

$$(f_{k_l}(t_1), \dots, f_{k_l}(t_{n_1})) , \quad l = 1, 2, 3, \dots$$

konvergent in K_1 für irgendeine Teilfolge $(k_l)_{l \in \mathbb{N}}$. Analog ist auch

$$(f_{k_r}(t_1), \dots, f_{k_r}(t_{n_2}))$$

konvergent in K_2 für irgendeine Teilfolge $(k_r) \subset (k_l)$ u.s.w. Das Cantorsche Diagonalisierungsverfahren liefert eine Teilfolge (*Diagonalfolge*) $(k_j)_j \subset \mathbb{N}$ so dass

$$\lim_{j \rightarrow \infty} (f_{k_j}(t_1), \dots, f_{k_j}(t_{n_i})) =: z^i \in K_i$$

für alle $i = 1, 2, \dots$ existiert, wobei

$$z^{i+1} = \lim_{l \rightarrow \infty} \left(\underbrace{f_{k_l}(t_1), \dots, f_{k_l}(t_{n_i}), \dots, f_{k_l}(t_{n_{i+1}})}_{l \rightarrow \infty z^i} \right)$$

Erhalten somit eine Folge $(a_1, a_2, \dots) \subset S$ mit

$$\underbrace{(a_1, a_2, \dots, a_{n_i})}_{z^i} \in K_i \quad , \quad i = 1, 2, \dots$$

ist. Definieren nun die Funktion

$$f_0 : T \rightarrow S \quad , \quad f_0(t) := \begin{cases} a_j & : t = t_j \\ 0 & : \text{sonst} \end{cases}$$

das heißt

$$(f_0(t_1), \dots, f_0(t_{n_i})) = (a_1, \dots, a_{n_i}) = z^i \in K_i \subset B_i$$

und somit $f_0 \in Z_i$, $i = 1, 2, \dots$ bzw.

$$f_0 \in \bigcap_{i=1}^{\infty} Z_i \quad \Rightarrow \quad \bigcap_{i=1}^{\infty} Z_i \neq \emptyset$$

- Da $\mathcal{C}_0(S^T)$ eine Algebra ist und Q_0 auf \mathcal{C}_0 σ -additiv ist mit $Q_0(\emptyset) = 0$, $Q_0(S^T) = 1$, kann Q_0 eindeutig zu einem W-Maß Q auf $\mathcal{C} = \sigma(\mathcal{C}_0)$ fortgesetzt werden. Insbesondere gilt dann für Zylindermenge $Z = \{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\}$:

$$Q(Z) = Q_0(Z) = Q_{t_1, \dots, t_n}(B) \quad , \quad t_i \neq t_{j \neq i} \quad (1.5.3.2)$$

□

1.5.4 Theorem: Existenz stochastischer Prozesse

Sei S ein vollständiger, separabler metrischer Raum und $\mathcal{S} := \mathcal{B}(S)$. Zu einer Familie $(Q_{t_1, \dots, t_n})_{\substack{t_i \in T \\ t_i \neq t_{j \neq i}}}$ von Wahrscheinlichkeitsmaßen auf (S, \mathcal{S}) existiert genau dann ein S -wertiger stochastischer Prozess $(X_t)_{t \in T}$ mit

$$\mathcal{P}_{t_1, \dots, t_n}^X(B) = Q_{t_1, \dots, t_n}(B) \quad , \quad t_i \neq t_{j \neq i}$$

wenn (Q_{t_1, \dots, t_n}) ein projektives System ist, das heißt die Bedingungen 1.5.1.1 und 1.5.1.2 erfüllt.

Beweis: Die Notwendigkeit der Bedingung wurde in Abschnitt 1.5.1 behandelt. Seien nun 1.5.1.1 und 1.5.1.2 erfüllt, dazu der Wahrscheinlichkeitsraum

$$(S^T, \mathcal{C}(S^T), Q)$$

wie im Satz von Kolmogorov (1.5.3) definiert. Definieren darauf den kanonischen Prozess

$$X : T \times S^T \rightarrow S \quad , \quad X_t(f) := f(t) \quad , \quad t \in T, f \in S^T$$

Dann gilt:

$$\mathcal{P}_{t_1, \dots, t_n}^X(B) = Q \{f \in S^T : (X_{t_1}(f), \dots, X_{t_n}(f)) \in B\} = Q \underbrace{\{f \in S^T : (f(t_1), \dots, f(t_n)) \in B\}}_{\text{Zylindermenge}} = Q_{t_1, \dots, t_n}(B) \quad , \quad t_i \neq t_{j \neq i}$$

□

Spezialfälle:

a) Es sei $\emptyset \neq T \subset \mathbb{R}$ und gegeben die Wahrscheinlichkeitsmaße

$$Q_{t_1, \dots, t_n} \quad , \quad t_1 < \dots < t_n$$

auf (S^n, \mathcal{S}^n) , $n \in \mathbb{N}$. Hinsichtlich der Bedingung 1.5.1.1 definieren wir

$$Q_{t_{\pi(1)}, \dots, t_{\pi(n)}} := Q_{t_1, \dots, t_n} \circ T_{\pi}^{-1}$$

für jede Permutation $\pi \in \text{Sym}(n)$, so dass diese automatisch erfüllt ist. Gemäß der Bedingung 1.5.1.2 muss jedoch zusätzlich gefordert werden

$$Q_{t_1, \dots, t_n}(B_1 \times \dots \times B_{k-1} \times S \times \dots \times B_n) = Q_{t_1, \dots, t_{k-1}, t_{k+1}, \dots, t_n}(B_1 \times \dots \times B_{k-1} \times B_{k+1} \times \dots \times B_n) \quad (1.5.4.1)$$

für alle $1 \leq k \leq n$. In dem Fall erzeugt dann (Q_{t_1, \dots, t_n}^X) einen stochastischen Prozess $(X_t)_{t \in T}$, wie in Theorem 1.5.4 beschrieben.

b) Sei $T := \mathbb{N} = \{1, 2, \dots\}$, dazu die Familie $(Q_{1, \dots, n})_{n \in \mathbb{N}}$ von Wahrscheinlichkeitsmaßen jeweils auf (S^n, \mathcal{S}^n) . Die endlich-dimensionalen Verteilungen $\mathcal{P}_{k_1, \dots, k_n}^X$ eines von $(Q_{1, \dots, n})$ induzierten stochastischen Prozesses $(X_n)_{n \in \mathbb{N}}$ wären dann schon festgelegt, denn für $k_1 < k_2 < \dots < k_m =: n$ (keine B.d.A, vgl. Bemerkung 1.3.1 (3)) ist

$$\underbrace{\mathcal{P}_{k_1, \dots, k_m}^X}_{=: Q_{k_1, \dots, k_m}^X}(B_1 \times \dots \times B_m) = \underbrace{\mathcal{P}_{1, \dots, n}^X}_{Q_{1, \dots, n}^X}(S \times \dots \times S \times \underbrace{B_{k_1}}_{\uparrow} \times S \times \dots \times S \times \underbrace{B_{k_m}}_{\uparrow})$$

Die Existenz von X ist dabei durch die zusätzliche Forderung

$$Q_{1, \dots, n}(B \times S) = Q_{1, \dots, n-1}(B) \quad , \quad B \in \mathcal{S}^{n-1} \quad (1.5.4.2)$$

gesichert.

1.5.5 Spezialfall: Weißes Rauschen

Sei S ein vollständiger, separabler, metrischer Raum. Gegeben sei eine Familie von Wahrscheinlichkeitsmaßen $(\nu_t)_{t \in T}$ auf $(S^T, \mathcal{B}(S))$, dazu

$$Q_{t_1, \dots, t_n} := \nu_{t_1} \otimes \dots \otimes \nu_{t_n} \quad , \quad t_i \neq t_j \neq i$$

das heißt insbesondere

$$Q_{t_1, \dots, t_n}(B_1 \times \dots \times B_n) = \prod_{i=1}^n \nu_{t_i}(B_i) \quad , \quad B_i \in \mathcal{B}(S)$$

Dann ist Bedingung 1.5.1.1 automatisch erfüllt. Wegen $\nu_t(S) = 1$ ist auch 1.5.1.2 erfüllt, so dass ein W-Maß Q auf $(S^T, \mathcal{C}(S^T))$ existiert mit

$$Q\{(x_t)_{t \in T} : x_{t_i} \in B_i, \quad i = 1, \dots, n\} = Q_{t_1, \dots, t_n}(B_1 \times \dots \times B_n) = \prod_{i=1}^n \nu_{t_i}(B_i)$$

Man nennt dann Q das (*unendliche*) *Produktmaß* von (ν_t) und schreibt

$$Q = \bigotimes_{t \in T} \nu_t \quad .$$

Bemerke: Diese Folgerung (Existenz des speziellen Produktmaßes) ist schwächer als der allgemeine Satz über die Existenz unendlicher Produktmaße von $(\Omega_t, \mathfrak{S}_t, \mathcal{P}_t)_{t \in T}$, da hier $\Omega_t = \Omega$ und $\mathfrak{S}_t = \mathcal{B}(\Omega)$ war.

Sei nun $(X_t)_{t \in T}$ der von obigen (Q_{t_1, \dots, t_n}) erzeugte Prozess über $(S^T, \mathcal{C}(S^T))$. Dann

$$\mathcal{P}(\{X_t \in B\}) = \mathcal{P}_t^X(B) = Q_t(B) = \nu_t(B)$$

und ferner

$$\mathcal{P}(\{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\}) = \nu_{t_1}(B_1) \dots \nu_{t_n}(B_n) = \mathcal{P}(X_{t_1} \in B_1) \dots \mathcal{P}(X_{t_n} \in B_n)$$

das heißt $(X_t)_{t \in T}$ ist eine Familie unabhängiger Zufallsvariablen.

1.5.6 Korollar: Existenz unabhängiger Prozesse

Sei S ein vollständiger, separabler, metrischer Raum und $\mathcal{S} := \mathcal{B}(S)$. Seien $(Q_{t_1, \dots, t_n})_{\substack{t_i \in T \\ t_i \neq t_j \neq i}}$, $(H_{t_1, \dots, t_n})_{\substack{t_i \in T \\ t_i \neq t_j \neq i}}$ projektive Systeme auf (S, \mathcal{S}) (vgl. Def. 1.5.2). Dann existieren unabhängige, S -wertige stochastische Prozesse $(X_t)_{t \in T}$, $(Y_t)_{t \in T}$ mit

$$\mathcal{P}_{t_1, \dots, t_n}^X(B) = Q_{t_1, \dots, t_n}(B) \quad , \quad \mathcal{P}_{t_1, \dots, t_n}^Y(B) = H_{t_1, \dots, t_n}(B) \quad , \quad t_i \neq t_j \neq i$$

(vgl. Existenzsatz 1.3.12).

Beweis: Seien Q, H jeweils die von (Q_{t_1, \dots, t_n}) , (H_{t_1, \dots, t_n}) erzeugten Wahrscheinlichkeitsmaße auf $(S^T, \mathcal{C}(S^T))$, das heißt⁴

$$Q[\{f : T \rightarrow S \mid (f(t_1), \dots, f(t_n)) \in B\}] = Q_{t_1, \dots, t_n}(B) \quad , \quad B \in \mathcal{S}^n$$

(analog auch für H). Setzen $\Omega := S^T \times S^T$, $\mathfrak{C} := \mathcal{C}(S^T) \otimes \mathcal{C}(S^T)$ und $\mathcal{P} := Q \otimes H$, dazu

$$X_t(f, g) := f(t) \quad , \quad Y_t(f, g) := g(t) \quad , \quad f, g \in S^T, \quad t \in T$$

Dann sind nach 1.3.7 die Abbildungen (X_t) , (Y_t) tatsächlich stochastische Prozesse und erfüllen sogar die Behauptungen.

□

1.6 Zuwächse von Prozessen (increments)

1.6.1 Definition: Zuwachs

Seien $G := \mathbb{R}^m$, $T \subset \mathbb{R}$ und $t, s \in T$. Dann definiert man den *Zuwachs* eines $(G, \mathcal{B}(G))$ -wertigen Prozesses $(X_t)_{t \in T}$ von t bis s als

$$Z_{t,s} := X_s - X_t$$

Bemerkung: Zu Startpunkt $t_0 \in T$ und Startwert $x_0 \in G$ mit $\mathcal{P}_{t_0}^X(\{x_0\}) = 1$ setzt man für $t_0 < t_1 < \dots < t_n \in T$:

$$Z_j := Z_{t_{j-1}, t_j} = X_{t_j} - X_{t_{j-1}} \quad , \quad j = 1, \dots, n$$

Betrachten die Abbildung $A_n^{x_0} : G^n \rightarrow G^n$ definiert durch

$$A_n^{x_0}(z_1, \dots, z_n) := (x_0 + z_1, x_0 + z_1 + z_2, \dots, x_0 + z_1 + z_2 + \dots + z_n) \quad (1.6.1.1)$$

formal dargestellt durch die *affine Transformation*

$$A_n^{x_0} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} x_0 \\ \vdots \\ x_0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ & & & \ddots & \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

mit

$$(A_n^{x_0})^{-1}(x_1, \dots, x_n) = (x_1 - x_0, x_2 - x_1, \dots, x_n - x_{n-1}) \quad .$$

Bemerke dass $A_n^{x_0}$ bzw. $(A_n^{x_0})^{-1}$ messbar sind. Dabei ist

$$A_n^{x_0}((Z_1, \dots, Z_n)) = (X_{t_1} - X_{t_0} + x_0, \dots, X_{t_n} - X_{t_0} + x_0) \stackrel{\text{fast sicher}}{=} (X_{t_1}, \dots, X_{t_n})$$

das heißt

$$\mathcal{P}_{t_1, \dots, t_n}^X(B) = \mathcal{P}\{A_n^{x_0}(Z_1, \dots, Z_n) \in B\} = \mathcal{P}\{(Z_1, \dots, Z_n) \in (A_n^{x_0})^{-1}(B)\} =: Q_{t_1, \dots, t_n}^X((A_n^{x_0})^{-1}(B))$$

⁴Alternativ kann man sich Q als das Bildmaß des Pfades $X : \Omega \rightarrow S^T$ vorstellen, wobei $X : \Omega \rightarrow S$ der zu (Q_{t_1, \dots, t_n}) gehörige Prozess ist.

mit

$$Q_{t_1, \dots, t_n}^X(B) := \mathcal{P} \{(Z_1, \dots, Z_n) \in B\}$$

bzw.

$$\mathcal{P}_{t_1, \dots, t_n}^X = Q_{t_1, \dots, t_n}^X \circ (A_n^{x_0})^{-1} \quad (1.6.1.2)$$

Die Q_{t_1, \dots, t_n}^X entsprechen also genau den Verteilungen der Prozess-Zuwächse. Beachte: Auch wenn diese explizit gegeben sind, sind die $\mathcal{P}_{t_1, \dots, t_n}^X$ oft sehr schwer auszurechnen.

1.6.2 Satz über Zuwächse: Existenz stochastischer Prozesse

Sei $G := \mathbb{R}^m$ und $(Q_{1, \dots, n})_{n \in \mathbb{N}}$ eine Familie von Wahrscheinlichkeitsmaßen auf jeweils $(G^n, \mathcal{B}(G)^n)$. Dann existiert ein G -wertiger stochastischer Prozess $(X_n)_{n \in \mathbb{N}_0}$ mit beliebigem X_0 (e.g. $X_0 \equiv x_0$ für ein $x_0 \in G$) und

$$\mathcal{P} \{(X_1 - X_0, X_2 - X_1, \dots, X_n - X_{n-1}) \in B\} = Q_{1, \dots, n}(B) \quad , \quad B \in \mathcal{V}^n, \quad n \in \mathbb{N} \quad (1.6.2.1)$$

genau dann wenn

$$Q_{1, \dots, n}(B \times G) = Q_{1, \dots, n-1}(B) \quad , \quad B \in \mathcal{B}(G)^{n-1} \quad (1.6.2.2)$$

Beweis: Richtung "⇒"

$Q_{1, \dots, n}$ erfülle 1.6.2.1. Dann:

$$\begin{aligned} Q_{1, \dots, n}(B \times G) &= \mathcal{P} \{(X_1 - X_0, \dots, X_{n-1} - X_{n-2}) \in B, (X_n - X_{n-1}) \in G\} \\ &= \mathcal{P} \{(X_1 - X_0, \dots, X_{n-1} - X_{n-2}) \in B\} = Q_{1, \dots, n-1}(B) \end{aligned}$$

Richtung "⇐"

$Q_{1, \dots, n}$ erfülle 1.6.2.2. Setzen

$$\mathcal{P}_{1, \dots, n} := Q_{1, \dots, n} \circ (A_n^0)^{-1}$$

(vgl. 1.6.1.1), dann gilt für $B \in \mathcal{B}(G)^{n-1}$:

$$\begin{aligned} \mathcal{P}_{1, \dots, n}(\{B \times S\}) &= Q_{1, \dots, n}(\{(z_1, \dots, z_n) : (z_1, \dots, z_1 + \dots + z_{n-1}) \in B, (z_1 + \dots + z_n) \in V\}) \\ &= Q_{1, \dots, n}(\{(z_1, \dots, z_n) : A_{n-1}^0(z_1, \dots, z_{n-1}) \in B, z_n \in V\}) \\ &= Q_{1, \dots, n}(\{(z_1, \dots, z_n) : (z_1, \dots, z_{n-1}) \in \underbrace{(A_{n-1}^0)^{-1}(B)}_{\tilde{B}}, z_n \in V\}) \\ &= Q_{1, \dots, n}(\tilde{B} \times S) = Q_{1, \dots, n-1}(\tilde{B}) = Q_{1, \dots, n-1}((A_{n-1}^0)^{-1}(B)) = \mathcal{P}_{1, \dots, n-1}(B) \end{aligned}$$

(vgl. Forderung 1.5.4.2) das heißt es existiert ein stochastischer Prozess $(X_n)_{n \in \mathbb{N}}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$ mit

$$\mathcal{P} \{(X_1, X_2 - X_1, \dots, X_n - X_{n-1}) \in B\} = Q_{1, \dots, n}(B)$$

(vgl. Theorem 1.5.4). Dabei setzen $X_0 := 0$. Sei nun $\tilde{X}_0 : \Omega \rightarrow G$ beliebig, dann erfüllt der Prozess

$$\tilde{X}_n := X_n + \tilde{X}_0 \quad , \quad n \in \mathbb{N}_0$$

die gewünschte Bedingung

$$\mathcal{P} \left(\left\{ (\tilde{X}_1 - \tilde{X}_0, \dots, \tilde{X}_n - \tilde{X}_{n-1}) \in B \right\} \right) = \mathcal{P} \left(\left\{ (X_1, \dots, X_n - X_{n-1}) \in B \right\} \right) = Q_{1, \dots, n}(B)$$

□

1.7 Prozesse mit unabhängigen Zuwächsen

1.7.1 Definition: Prozess mit unabhängigen Zuwächsen

Es sei $G := \mathbb{R}^m$, $T \subset \mathbb{R}$ und $(X_t)_{t \in T}$ ein $(G, \mathcal{B}(G))$ -wertiger zufälliger Prozess. Dann hat X *unabhängige Zuwächse* falls

$$\forall t_0 < t_1 < \dots < t_n \in T, n \in \mathbb{N}$$

die Zuwächse

$$Z_j := X_{t_j} - X_{t_{j-1}}, \quad j = 1, \dots, n$$

unabhängige Zufallsvariablen sind.

Spezialfall: Sei $T = \mathbb{N}_0$ und $Z_n := X_n - X_{n-1}$ der n -te Zuwachs. Dann hat X unabhängige Zuwächse genau dann wenn

$$\mathcal{P}(\{Z_1 \in B_1, \dots, Z_n \in B_n\}) = \mathcal{P}(\{Z_1 \in B_1\}) \cdot \dots \cdot \mathcal{P}(\{Z_n \in B_n\})$$

für beliebige $B_i \in \mathcal{B}(G)$, $n \in \mathbb{N}$ gilt⁵. Setzt man nun $\nu_n := \mathcal{P}_{Z_n}$ so ist

$$\mathcal{P}(\{(Z_1, \dots, Z_n) \in B\}) = (\nu_1 \otimes \dots \otimes \nu_n)(B)$$

1.7.2 Lemma: Verschiebungsinvarianz von Zuwächsen

Sei $(X_t)_{t \in T}$ ein \mathbb{R}^m -wertiger stochastischer Prozess über $(\Omega, \mathfrak{C}, \mathcal{P})$ mit

$$\mathcal{P}[\{(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}) \in B\}] = Q_{t_0, t_1, \dots, t_n}(B)$$

für $t_0 < t_1 < \dots < t_n \in T$, $B \in \mathcal{B}(\mathbb{R}^m)^n$. Sei ferner $Y : \Omega \rightarrow \mathbb{R}^m$ irgendeine Zufallsvariable. Dann erfüllt der stochastische Prozess

$$\tilde{X}_t := X_t + Y, \quad t \in T$$

ebenfalls

$$\mathcal{P}[\{(\tilde{X}_{t_1} - \tilde{X}_{t_0}, \tilde{X}_{t_2} - \tilde{X}_{t_1}, \dots, \tilde{X}_{t_n} - \tilde{X}_{t_{n-1}}) \in B\}] = Q_{t_0, t_1, \dots, t_n}(B)$$

für $t_i \in T$. Die Verteilung der Zuwächse von \tilde{X} ist sozusagen *invariant* gegenüber *Verschiebungen*.

Beweis: Per Konstruktion

$$\begin{aligned} & \mathcal{P}[\{(\tilde{X}_{t_1} - \tilde{X}_{t_0}, \tilde{X}_{t_2} - \tilde{X}_{t_1}, \dots, \tilde{X}_{t_n} - \tilde{X}_{t_{n-1}}) \in B\}] \\ &= \mathcal{P}[\{(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}) \in B\}] \\ &= Q_{t_0, t_1, \dots, t_n}(B) \end{aligned}$$

□

1.7.3 Satz: Existenz von \mathbb{N}_0 -zeitigen Prozessen mit unabhängigen Zuwächsen

Sei $G := \mathbb{R}^m$. Gegeben seien die Wahrscheinlichkeitsmaße $(\nu_n)_{n \in \mathbb{N}}$ auf $(G, \mathcal{B}(G))$. Dann existiert ein bis auf Äquivalenz eindeutig bestimmter $(G, \mathcal{B}(G))$ -wertiger Prozess $(X_n)_{n \in \mathbb{N}_0}$ mit beliebigem X_0 und unabhängigen Zuwächsen

$$Z_i = X_i - X_{i-1}, \quad i \in \mathbb{N}$$

verteilt gemäß den Wahrscheinlichkeitsmaßen ν_1, ν_2, \dots auf $(G, \mathcal{B}(G))$.

⁵Beachte: Für unabhängige Zufallsvariablen Y_1, \dots, Y_n auf $(\Omega, \mathfrak{C}, \mathcal{P})$ folgt auch die Unabhängigkeit von $Y_1 + Y_2, Y_3, \dots, Y_n$. Da $Z_{n,k} = Z_n + Z_{n+1} + \dots + Z_k$ folgt induktiv die Behauptung.

Beweis: Setzen

$$Q_{1,\dots,n} := \nu_1 \otimes \cdots \otimes \nu_n$$

Dann ist

$$Q_{1,\dots,n+1}(B \times G) = (\nu_1 \otimes \cdots \otimes \nu_n)(B) \cdot \underbrace{\nu_{n+1}(G)}_1 = Q_{1,\dots,n}(B)$$

für $B \in \mathcal{B}(G)$ und $n \in \mathbb{N}$. Somit existiert nach Satz 1.6.2 ein G -wertiger stochastischer Prozess $(X_n)_{n \in \mathbb{N}_0}$ mit

$$\mathcal{P}[\{(X_1 - X_0, \dots, X_n - X_{n-1}) \in B\}] = Q_{1,\dots,n}(B) = (\nu_1 \otimes \cdots \otimes \nu_n)(B) \quad , \quad B \in \mathcal{B}(G) \quad , \quad n \in \mathbb{N}$$

und X_0 beliebig. Insbesondere sind dann die $X_i - X_{i-1}$ unabhängig verteilt gemäß ν_i . Eindeutigkeit folgt aus Darstellung (1.6.1.2) und Verschiebungsinvarianz 1.7.2.

□

Beispiel: Der Random Walker Sei $G = \mathbb{R}^m$ und $T = \mathbb{N}_0$, dazu die Zuwächse Z_i eines stochastischen Prozesses mit

$$\mathcal{P}(\{Z_i = \mathbf{e}_k\}) = \frac{1}{2m} \quad , \quad \mathcal{P}(\{Z_i = -\mathbf{e}_k\}) = \frac{1}{2m}$$

und $X_0 \equiv 0$. Dann beschreibt der zufällige Prozess

$$X_n := X_0 + Z_1 + \cdots + Z_n \quad , \quad n \in \mathbb{N}_0$$

die *Position* eines *Teilchens* im \mathbb{R}^m nach n Schritten, das zu jeden Zeitpunkt mit gleicher Wahrscheinlichkeit in eine der 8 Achsenrichtungen die Streckenlänge 1 *hüpf*t.

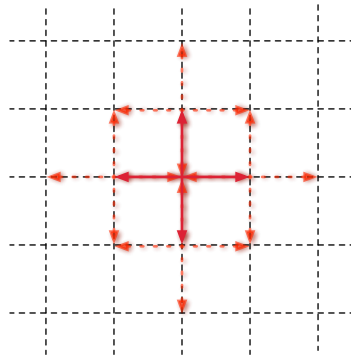


Abbildung 1.4: 2-dimensionaler Random-Walker

1.7.4 Definition: Faltung von Maßen

Es seien ν, μ zwei Maße auf der topologischen Gruppe⁶ $(G, \mathcal{B}(G))$. Dann definiert man die *Faltung* $\nu * \mu$ gemäß

$$(\nu * \mu)(B) := \int_G \nu(B - x) d\mu(x) \quad , \quad B \in \mathcal{B}(G)$$

⁶Eine Gruppe G heißt *topologische Gruppe*, wenn sie mit einer Topologie so versehen ist, dass gilt:

- (i) Die Gruppenverknüpfung $\circ : G \times G \rightarrow G$ ist stetig, wobei $G \times G$ mit der Produkttopologie versehen ist.
- (ii) Die Inversenabbildung $G \rightarrow G$ ist stetig.

Bemerkungen:

- (i) $\nu * \mu$ ist tatsächlich ein Maß auf $(G, \mathcal{B}(G))$.
(ii) Sind ν, μ Wahrscheinlichkeitsmaße auf $(G, \mathcal{B}(G))$, so ist es auch $\nu * \mu$.
(iii) Die Operation $(\nu, \mu) \mapsto \nu * \mu$ ist kommutativ und assoziativ. Somit macht insbesondere Sinn

$$\nu_1 * \nu_2 * \cdots * \nu_n$$

für Maße ν_1, \dots, ν_n .

- (iv) Für das Dirac-Maß δ_{x_0} ist

$$(\nu * \delta_{x_0})(B) = \nu(B - x_0)$$

- (v) Ist 0 das neutrale Element in G , so ist δ_0 neutral bzgl. $*$, das heißt

$$\nu * \delta_0 = \nu$$

Demnach ist die Menge aller W-Maße $\mathfrak{M}_1(G)$ auf $(G, \mathcal{B}(G))$ bzgl. der Faltung ein Monoid.

1.7.5 Lemma: Summe unabhängiger Zufallsvariablen

Seien $X_1, X_2 : \Omega \rightarrow \mathbb{R}^m$ unabhängige Zufallsvariablen von $(\Omega, \mathfrak{G}, \mathcal{P})$ nach $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$. Dann gilt:

$$\mathcal{P}_{X_1+X_2} = \mathcal{P}_{X_1} * \mathcal{P}_{X_2}$$

Beweis: Die Abbildung

$$h : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m, (x_1, x_2) \mapsto x_1 + x_2$$

ist stetig und somit messbar, insbesondere ist auch $h(X_1, X_2) = X_1 + X_2$ messbar. Für den zufälligen Vektor $X := (X_1, X_2) : \Omega \rightarrow \mathbb{R}^m \times \mathbb{R}^m$ gilt

$$X_1 + X_2 = h \circ X \Rightarrow (X_1 + X_2)^{-1}(B) = X^{-1}(h^{-1}(B)) .$$

Man kann also für $B \in \mathcal{B}(\mathbb{R}^m)$ schreiben

$$\begin{aligned} \mathcal{P}_{X_1+X_2} &= P[(X_1 + X_2)^{-1}(B)] = P_X(h^{-1}(B)) \stackrel{X_1, X_2 \text{ unabhängig}}{=} (\mathcal{P}_{X_1} \otimes \mathcal{P}_{X_2})(h^{-1}(B)) \\ &= \int_{G^2} \underbrace{1_{h^{-1}(B)}(x_1, x_2)}_{1_B(x_1+x_2)} \underbrace{d(\mathcal{P}_{X_1} \otimes \mathcal{P}_{X_2})}_{d\mathcal{P}_{X_1} d\mathcal{P}_{X_2}} = \int_G \int_G 1_{B-x_2}(x_1) d\mathcal{P}_{X_1}(x_1) d\mathcal{P}_{X_2}(x_2) \\ &= \int_G \mathcal{P}_{X_1}(B - x_2) d\mathcal{P}_{X_2}(x_2) = (\mathcal{P}_{X_1} * \mathcal{P}_{X_2})(B) \end{aligned}$$

□

1.7.6 Notwendige Bedingung für die Unabhängigkeit der Zuwächse

Sei $G := \mathbb{R}^m$, $T \subset \mathbb{R}$ und gegeben eine Familie $(\nu_{t,s})_{t < s \in T}$ von Wahrscheinlichkeitsmaßen auf $(G, \mathcal{B}(G))$. Dann ist eine notwendige Bedingung für die Existenz eines $(G, \mathcal{B}(G))$ -wertigen Prozesses $(X_t)_{t \in T}$ mit unabhängigen Zuwächsen $Z_{t,s}$ und

$$\mathcal{P}_{Z_{t,s}} = \nu_{t,s} \quad \forall t < s \in T$$

dass gilt:

$$\nu_{t,u} = \nu_{t,s} * \nu_{s,u} \quad \forall t < s < u \in T \tag{1.7.6.1}$$

Beweis: Da $Z_{t,s}$, $Z_{s,u}$ unabhängig sind, gilt nach Lemma 1.7.5:

$$\underbrace{\mathcal{P}_{Z_{t,s}+Z_{s,u}}}_{\mathcal{P}_{Z_{t,u}}} = \mathcal{P}_{Z_{t,s}} * \mathcal{P}_{Z_{s,u}} \quad .$$

□

1.7.7 Theorem: Existenz von Prozessen mit unabhängigen Zuwächsen

Es sei $G := \mathbb{R}^m$, $T \subset \mathbb{R}$, $\min(T) = t_0$ und $(\nu_{t,s})_{t < s \in T}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(G, \mathcal{B}(G))$ mit der Faltungseigenschaft 1.7.6.1. Dann existiert für beliebiges X_{t_0} ein stochastischer Prozess $X = (X_t)_{t \in T}$ mit unabhängigen Zuwächsen und

$$\mathcal{P}_{Z_{t,s}} = \nu_{t,s} \quad , \quad t < s \in T$$

Beweis: Betrachten zunächst den Spezialfall $X_{t_0} \equiv 0$. Konstruieren für $t_0 < t_1 < t_2 < \dots < t_n$ die Wahrscheinlichkeitsmaße $\mathcal{P}_{t_1, \dots, t_n}$ auf $(G^n, \mathcal{B}(G)^n)$ die Bedingung 1.5.4.1 erfüllen, das heißt

$$\mathcal{P}_{t_1, \dots, t_n}(B_1 \times \dots \times B_{k-1} \times G \times B_{k+1} \times \dots \times B_n) = \mathcal{P}_{t_1, \dots, t_{k-1}, t_{k+1}, \dots, t_n}(B_1 \times \dots \times B_{k-1} \times B_{k+1} \times \dots \times B_n)$$

Da für den Prozess X gelten müsse:

$$\mathcal{P}(\{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\}) = \mathcal{P}(\{Z_{t_0, t_1} \in B_1, Z_{t_0, t_1} + Z_{t_1, t_2} \in B_2, \dots, Z_{t_0, t_1} + \dots + Z_{t_{n-1}, t_n} \in B_n\})$$

machen wir den Ansatz

$$\begin{aligned} \mathcal{P}_{t_1, \dots, t_n}(B) &:= (\nu_1 \otimes \dots \otimes \nu_n) (\{x \in G^n : \underbrace{(x_1, x_1 + x_2, \dots, x_1 + \dots + x_n)}_{A_n^0(x)} \in B\}) \\ &= (\nu_1 \otimes \dots \otimes \nu_n) ((A_n^0)^{-1}(B)) \end{aligned}$$

wobei $\nu_j := \nu_{t_{j-1}, t_j}$, $j = 1, \dots, n$. Für beliebige messbare $f : G^n \rightarrow \mathbb{R}$ gilt dann

$$\int_{G^n} f(x) d\mathcal{P}_{t_1, \dots, t_n}(x) = \int_{G^n} f(A_n^0(x)) d\nu_1(x_1) \dots d\nu_n(x_n) \quad (1.7.7.1)$$

Bemerke dass⁷

$$\int_{G^2} g(x_k + x_{k+1}) d\nu_k(x_k) d\nu_{k+1}(x_{k+1}) = \int_G g(y) d(\nu_k * \nu_{k+1})(y) \stackrel{(1.7.6.1)}{=} \int_G g(y) d\nu_{t_{k-1}, t_{k+1}}(y) \quad (1.7.7.2)$$

Nun seien

$$B := B_1 \times \dots \times B_{k-1} \times G \times B_{k+1} \times \dots \times B_n \quad , \quad B_i \in \mathcal{B}(G)$$

$$\tilde{B} := B_1 \times \dots \times B_{k-1} \times B_{k+1} \times \dots \times B_n$$

⁷Beachte: Für W-Maße ν, μ auf $(G, \mathcal{B}(G))$ und $h : (x, y) \mapsto x + y$ ist $(\nu \otimes \mu) \circ h^{-1} = \nu * \mu$.

dann

$$\begin{aligned}
\mathcal{P}_{t_1, \dots, t_n}(B) &= \int_{G^n} 1_B(x) d\mathcal{P}_{t_1, \dots, t_n}(x) \\
&\stackrel{(1.7.7.1)}{=} \int_G \dots \int_G \underbrace{1_B(x_1, x_1 + x_2, \dots, x_1 + \dots + x_n)}_{1_{B_1}(x_1)1_{B_2}(x_1+x_2)\dots 1_{B_{k-1}}(x_1+\dots+x_{k-1})1_{B_{k+1}}(x_1+\dots+x_{k+1})\dots 1_{B_n}(x_1+\dots+x_n)} d\nu_1(x_1)\dots d\nu_n(x_n) \\
&= \int_G \dots \int_G 1_{B_1}(x_1) \dots 1_{B_{k-1}}(x_1 + \dots + x_{k-1}) d\nu_1(x_1)\dots d\nu_{k-1}(x_{k-1}) \\
&\quad \int_G \dots \int_G \underbrace{1_{B_{k+1}}(x_1 + \dots + x_k + x_{k+1}) \dots 1_{B_{k+2}}(x_1 + \dots + x_n)}_{g(x_k+x_{k+1})} d\nu_k(x_k) d\nu_{k+1}(x_{k+1})\dots d\nu_n(x_n) \\
&\stackrel{(1.7.7.2)}{=} \underbrace{\int_G \dots \int_G}_{\times(n-1)} 1_{B_1}(x_1)\dots 1_{B_{k-1}}(x_1 + \dots + x_{k-1})1_{B_{k+1}}(x_1 + \dots + x_{k-1} + z)\dots 1_{B_n}(x_1 + \dots + z + x_{k+2}\dots x_n) \\
&\quad d\nu_1(x_1)\dots d\nu_{k-1, k+1}(z)\dots d\nu_n(x_n) \\
&\stackrel{(1.7.7.1)}{=} \mathcal{P}_{t_1, \dots, t_{k-1}, t_{k+1}, \dots, t_n}(\tilde{B})
\end{aligned}$$

Nach Bemerkung (a) in Theorem 1.5.4 existiert somit ein stochastischer Prozess $(X_t)_{t \in T \setminus \{t_0\}}$ mit

$$\mathcal{P}(\{(X_{t_1}, \dots, X_{t_n}) \in C\}) = (\nu_{t_0, t_1} \otimes \dots \otimes \nu_{t_{n-1}, t_n}) ((A_n^0)^{-1}(C)) \quad , \quad C \in \mathcal{B}(G)^n, \quad t_0 < t_1 < \dots < t_n \in T$$

bzw.

$$\begin{aligned}
\mathcal{P}[\{((A_n^0)^{-1}(X_{t_1}, \dots, X_{t_n})) \in C\}] &= (\nu_{t_0, t_1} \otimes \dots \otimes \nu_{t_{n-1}, t_n})(C) \\
\mathcal{P}[\{(X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}) \in C\}] &
\end{aligned}$$

Dabei setzen wir $X_{t_0} := 0$ also

$$\mathcal{P}[\{(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}) \in C\}] = (\nu_{t_0, t_1} \otimes \dots \otimes \nu_{t_{n-1}, t_n})(C)$$

und ferner

$$\mathcal{P}_{X_{t_2} - X_{t_1}} = \nu_{t_1, t_2} \quad .$$

Offensichtlich sind die Zuwächse $X_{t_k} - X_{t_{k-1}}$, $k = 1, \dots, n$ unabhängig. Nach Verschiebungslemma 1.7.2 folgt dann der Fall mit X_{t_0} beliebig.

□

Beispiele:

(i) Mit $\nu_{t,s} := \mathcal{N}_{0,s-t}$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ für $0 \leq t < s$ ist tatsächlich

$$\nu_{t,s} * \nu_{s,u} = \mathcal{N}_{0,s-t} * \mathcal{N}_{0,u-s} = \mathcal{N}_{0,(s-t)+(u-s)} = \nu_{t,u} \quad .$$

Somit existiert ein reellwertiger stochastischer Prozess $(W_t)_{t \geq 0}$ mit unabhängigen Zuwächsen $Z_{t,s}$ verteilt gemäß

$$\mathcal{P}_{Z_{t,s}} = \mathcal{N}_{0,s-t}$$

und $W_0 = 0$ fast sicher. W heißt *Wiener-Prozess* (*Prozess der Brownschen Bewegung*).

(ii) Betrachten die Poissonverteilung mit Parameter λ :

$$\pi_\lambda(\{k\}) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad , \quad k \in \mathbb{N}_0$$

Bekanntlich $\pi_\lambda * \pi_\mu = \pi_{\lambda+\mu}$. Setzt man nun für festes $\lambda > 0$ und $0 \leq t < s < \infty$:

$$\nu_{t,s}(B) := \pi_{\lambda(s-t)}(B) = \sum_{0 \leq k \in B} \frac{\lambda^k (s-t)^k}{k!} \cdot e^{-\lambda(s-t)}$$

so gilt

$$\nu_{t,s} * \nu_{s,u} = \pi_{\lambda(s-t)} * \pi_{\lambda(u-s)} = \pi_{\lambda(u-t)} = \nu_{t,u} \quad .$$

Es existiert also ein \mathbb{N}_0 -wertiger stochastischer Prozess $(N_t)_{t \geq 0}$ (*Poissonprozess*) mit unabhängigen Zuwächsen, $N_0 = 0$ fast sicher und

$$\mathcal{P}(\{N_s - N_t = k\}) = \frac{\lambda^k (s-t)^k}{k!} \cdot e^{-\lambda(s-t)} \quad .$$

1.7.8 Poissonprozesse

Betrachten den Poissonprozess $(N_t)_{t \geq 0}$ über $(\Omega, \mathfrak{S}, \mathcal{P})$ mit Intensität λ und $N_0 \equiv 0$. Dessen Pfade sind typischerweise fast-überall konstante \mathbb{N}_0 -wertige Funktionen über $[0, \infty)$ die an *zufälligen* Stellen positive Sprünge aufweisen. Dabei existiert eine Version mit rechtsseitig stetigen Pfaden[1]. Die Wahrscheinlichkeit dass ein Pfad von t nach $s > t$ genau k Sprünge macht ist dabei gegeben durch

$$\mathcal{P}(\{N_s - N_t = k\}) = \pi_{\lambda(s-t)}(\{k\})$$

Für $h > 0$ ist

$$\mathcal{P}(N_{t+h} - N_t \geq 2) = 1 - e^{-\lambda h} - \frac{\lambda h}{1!} \cdot e^{-\lambda h} \approx \lambda h \quad \text{für } h \approx 0$$

Die Wahrscheinlichkeit in einem Zeitintervall $[t, t+h]$ 2 oder mehr Sprünge zu machen ist von Ordnung $\mathcal{O}(h)$ für $h \rightarrow 0$.

Definiert man nun die *Eintreffzeit*

$$T_1(\omega) := \inf \{t > 0 : N_t(\omega) = 1\}$$

(Konvention: $\inf(\emptyset) := \infty$) so ist T_1 eine Zufallsvariable auf Ω und es gilt:

$$\mathcal{P}(\{T_1 > t\}) = \mathcal{P}(\{N_t = 0\}) = e^{-\lambda t} \quad , \quad 0 \leq t$$

$$\mathcal{P}(\{T_1 \in [a, b]\}) = \lambda \int_a^b e^{-\lambda x} dx \quad , \quad 0 \leq a \leq b$$

Analog definiert man

$$T_n(\omega) := \inf \{t > 0 : N_t(\omega) = n\}$$

Betrachten wir nun den Prozess

$$\tilde{N}_t(\omega) := N_{t+T_1(\omega)}(\omega) - 1 \quad , \quad \omega \in \Omega, t \geq 0$$

Dann lässt sich zeigen dass \tilde{N} ebenfalls ein Poissonprozess ist, unabhängig von N_t für $t < T_1(\omega)$. Allgemein sind die Zufallsvariablen $T_1, T_2 - T_1, \dots, T_{n+1} - T_n$ für $n \in \mathbb{N}$ unabhängig Exp_λ -verteilt, das heißt

$$\mathcal{P}(T_{n+1} - T_n > t) = e^{-\lambda t}$$

Umgekehrt, sind ξ_1, \dots, ξ_n unabhängige Exp_λ -verteilte Zufallsvariablen auf $(\Omega, \mathfrak{S}, \mathcal{P})$ (z.B. Lebensdauern von Glühbirnen), so ist

$$N_t(\omega) := \sup \{k \in \mathbb{N}_0 : \xi_1(\omega) + \dots + \xi_k(\omega) < t\}$$

ein Poissonprozess mit Intensität λ (Anzahl der Abwechslungen von Glühbirnen bis zum Zeitpunkt t).

1.7.9 Definition: Stationäre Zuwächse

Es sei $T \subset \mathbb{R}$ und $(X_t)_{t \in T}$ ein \mathbb{R}^m -wertiger zufälliger Prozess. Dann besitzt X *zeitlich homogene (stationäre) Zuwächse* $:\Leftrightarrow$

Das Verteilungsgesetz jedes beliebigen Zuwachses $Z_{t,s} := X_s - X_t$, $s > t \in T$ ist nur von der Differenz $(s - t)$ abhängig, das heißt

$$\mathcal{P}(\{X_s - X_t \in B\}) = \nu_{s-t}(B) \quad , \quad B \in \mathcal{B}(\mathbb{R}^m)$$

für irgendeine Familie $(\nu_{s-t})_{s>t \in T}$ von Wahrscheinlichkeitsmaßen.

Bemerkungen:

- (i) Ist $T = [0, \infty)$ und $X_0 = 0$ fast sicher, so ist

$$\nu_t(B) = \mathcal{P}(\{X_t \in B\}) \quad , \quad B \in \mathcal{B}(\mathbb{R}^m), t > 0$$

- (ii) Besitzt der Prozess ferner unabhängige Zuwächse, so muss entsprechend (1.7.6.1) gelten

$$\nu_{t_1+t_2} = \nu_{t_1} * \nu_{t_2}$$

für $t_1, t_2 > 0$.

1.7.10 Definition: Lévy-Prozess

Sei $0 \in T \subset \mathbb{R}$. Ein \mathbb{R}^m -wertiger stochastischer Prozess $(X_t)_{t \in T}$ heißt *Lévy-Prozess* falls:

- $\mathcal{P}(X_0 = 0) = 1$
- X hat stationäre, unabhängige Zuwächse.
- Die Pfade von X sind càdlàg⁸.

1.7.11 Definition: Faltungsgruppe

Sei G eine topologische Gruppe und $T \subseteq \mathbb{R}$. Eine Familie $(\nu_t)_{t \in T}$ von Maßen auf $(G, \mathcal{B}(G))$ heißt *Faltungshalbgruppe* falls für $t_1, t_2 \in T$ gilt

$$\nu_{t_1} * \nu_{t_2} = \nu_{t_1+t_2}$$

Beispiele:

- (i) Für \mathbb{R}^m -wertigen Lévy-Prozess $(X_t)_{t \geq 0}$ bilden

$$\nu_t(B) := \mathcal{P}(X_t - X_0 \in B) \quad , \quad B \in \mathcal{B}(\mathbb{R}^m), t > 0$$

eine Faltungshalbgruppe.

- (ii) Die Dirac-Maße

$$\nu_t := \delta_t \quad , \quad t > 0$$

auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ bilden eine Faltungshalbgruppe. Ein Prozess $(X_t)_{t \geq 0}$ mit

$$\mathcal{P}(X_{s+t} - X_s \in B) = \delta_t(B) = \begin{cases} 1 & : t \in B \\ 0 & : \text{sonst} \end{cases}$$

ist dabei genau ein (fast-) deterministischer Prozess, mit $X_t = t$ fast sicher.

⁸Eine Abbildung $f : T \rightarrow S$ von $T \subset \mathbb{R}$ in einen topologischen Raum S heißt *càdlàg*, falls:

- $\lim_{s \rightarrow t^-} f(s)$ existiert
- $\lim_{s \rightarrow t^+} f(s) = f(t)$

für beliebiges $t \in T$, das heißt f ist rechtsseitig stetig mit linksseitigen Grenzwerten.

(iii) Die Verteilungen der Zuwächse des Wiener-Prozesses (Bsp. (i) in Theorem 1.7.7) bilden die Faltungshalbgruppe:

$$(\mathcal{N}_{0,t})_{t>0}$$

Analog induzieren auch die Zuwächse des Poisson-Prozesses die Faltungshalbgruppe

$$(\pi_{\lambda t})_{t>0}$$

(iv) Die Γ -Verteilungen

$$\Gamma_{a,b}(B) := \frac{1}{a^b \Gamma(b)} \int_B x^{b-1} e^{-\frac{x}{a}} dx, \quad B \in \mathcal{B}([0, \infty))$$

induzieren die Faltungshalbgruppe $(\Gamma_{a,b})_{b \geq 0}$, denn

$$\Gamma_{a,b_1} * \Gamma_{a,b_2} = \Gamma_{a,b_1+b_2}$$

Ein Γ -Prozess ist entsprechend ein reellwertiger Prozess $(X_t)_{t \geq 0}$ mit unabhängigen Zuwächsen und

$$\mathcal{P}(X_{s+t} - X_s \geq x) = \frac{1}{a^t \Gamma(t)} \int_x^\infty s^{t-1} e^{-\frac{s}{a}} ds, \quad t > 0$$

1.8 Markov-Prozesse

1.8.1 Definition: Markovkern

Sei $T \subset \mathbb{R}$ und (S, \mathcal{S}) ein messbarer Raum. Dazu sei

$$(\mathcal{P}_{t,s})_{t < s \in T}$$

eine Familie von Abbildungen

$$\mathcal{P}_{t,s} : S \times \mathcal{S} \rightarrow [0, 1]$$

mit folgender Eigenschaft:

1. $\mathcal{P}_{t,s}(\cdot, B) : x \mapsto \mathcal{P}_{t,s}(x, B)$ ist messbar für jedes $B \in \mathcal{S}$.
2. $\mathcal{P}_{t,s}(x, \cdot) : B \mapsto \mathcal{P}_{t,s}(x, B)$ ist für jedes $x \in S$ ein Wahrscheinlichkeitsmaß auf (S, \mathcal{S}) .

Dann heißen die $\mathcal{P}_{t,s}$ *Markovkerne* (auf (S, \mathcal{S}) über T).

Interpretation: Für S -wertigen stochastischen Prozess $(X_t)_{t \in T}$ entsprechen

$$\mathcal{P}_{t,s}(x, B) := \mathcal{P}(X_s \in B \mid X_t = x), \quad x \in S, B \in \mathcal{S}$$

den bedingten Wahrscheinlichkeiten, zur Zeit s in B anzukommen unter der Voraussetzung $X_t = x$. Die Markov-Kerne $(\mathcal{P}_{t,s})$ heißen in dem Falle *Übergangswahrscheinlichkeiten* von X .

Bemerke dass im Fall stationärer Zuwächse $\mathcal{P}_{t,s}$, $t < s \in T$ nur von $s - t$ abhängen.

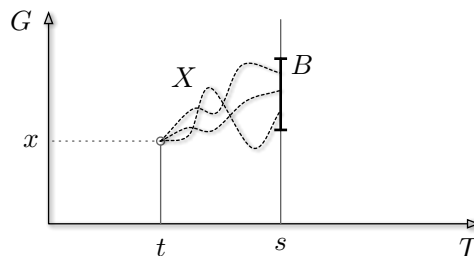


Abbildung 1.5: Zur Definition von Markov-Kernen. Abgebildet sind zufällige Pfade eines reellwertigen stochastischen Prozesses $(X_t)_{t \geq 0}$, die am Zeitpunkt t durch x gehen.

Bemerke dass die Definition von $\mathcal{P}(X_s \in B \mid X_t = x)$ nicht-trivial ist, da im allgemeinen $\mathcal{P}(X_t = x) = 0$ sein wird. Es stellt sich heraus, dass eine bis auf Nullmengen eindeutige, messbare Abbildung

$$\mathcal{P}_{t,s} : S \times \mathcal{S} \rightarrow [0, 1]$$

existiert mit

$$\mathcal{P}(\{X_s \in B, X_t \in A\}) = \int_A \mathcal{P}_{t,s}(x, B) \mathcal{P}_{X_t}(dx) \quad , \quad A, B \in \mathcal{S} \quad (1.8.1.1)$$

1.8.2 Definition: Markov-Prozess

Seien $T \subset \mathbb{R}$ und (S, \mathcal{S}) ein messbarer Raum. Ein S -wertiger stochastischer Prozess $(X_t)_{t \in T}$ mit Übergangswahrscheinlichkeiten $(\mathcal{P}_{t,s})_{t < s \in T}$ heißt *Markov-Prozess* falls gilt:

$$\mathcal{P}(\{X_{t_0} \in B_0, \dots, X_{t_n} \in B_n\}) = \int_{B_0} \dots \int_{B_n} \mathcal{P}_{t_{n-1}, t_n}(x_{n-1}, dx_n) \dots \mathcal{P}_{t_0, t_1}(x_0, dx_1) \mathcal{P}_{X_{t_0}}(dx_0) \quad (1.8.2.1)$$

für $t_0 < t_1 < \dots < t_n \in T$ und $B_1, \dots, B_n \in \mathcal{S}$.

1.8.3 Die Chapman-Kolmogorov Bedingung für Markov-Prozesse

Sei $T \subset \mathbb{R}$, (S, \mathcal{S}) ein messbarer Raum und $(X_t)_{t \in T}$ ein S -wertiger Markov-Prozess mit Übergangswahrscheinlichkeiten $(\mathcal{P}_{t,s})_{t < s \in T}$. Dann muss für \mathcal{P}_{X_t} -fast alle $x \in \mathcal{B}(S)$ gelten:

$$\boxed{\mathcal{P}_{t,u}(x, B) = \int_S \mathcal{P}_{s,u}(y, B) \mathcal{P}_{t,s}(x, dy)} \quad (1.8.3.1)$$

für $B \in \mathcal{S}$, $x \in S$, $t < s < u \in T$ (Chapman-Kolmogorov).

Interpretation: Die C.K. Gleichung erlaubt folgende Interpretation: Der Prozess, beginnend am Punkt x , muss von t zu u zwischendurch zur Zeit $s \in (t, u)$ einen Wert y annehmen. Dieser Übergang trifft mit einer gewissen Wahrscheinlichkeit zu, wobei der Übergang von y zu B mit einer vom Startpunkt x unabhängigen und nur von y abhängigen Wahrscheinlichkeitsverteilung stattfindet. Bei Übergängen spielt gewissermaßen nur die *Gegenwart* eine Rolle.

Insbesondere sind die endlich dimensionalen Verteilungen (und so auch die Zuwächse) eindeutig gegeben durch die Übergangswahrscheinlichkeiten.

Beweis: Per Definition

$$\mathcal{P}(X_u \in B, X_t \in A) = \int_A \mathcal{P}_{t,u}(x, B) \mathcal{P}_{X_t}(dx) = \mathcal{P}(X_t \in A, X_s \in S, X_u \in B)$$

$$\stackrel{\text{Markov}}{=} \int_A \int_S \mathcal{P}_{s,u}(y, B) \mathcal{P}_{t,s}(x, dy) \mathcal{P}_{X_t}(dx)$$

für beliebiges $A \in \mathcal{B}(S)$, das heißt

$$\int_S \mathcal{P}_{s,u}(y, B) \mathcal{P}_{t,s}(x, dy) = \mathcal{P}_{t,u}(x, B) \pmod{0}$$

□

1.8.4 Hinreichende Bedingung für Markov-Prozesse

Sei $T \subset \mathbb{R}$, S ein vollständiger, separabler, metrischer Raum und $(\mathcal{P}_{t,s})_{t < s \in T}$ Markov-Kerne auf S , die die Chapman-Kolmogorov Bedingung (1.8.3.1) erfüllen. Dann existiert ein $(S, \mathcal{B}(S))$ -wertiger Markov-Prozess $(X_t)_{t \in T}$ mit Übergangswahrscheinlichkeiten $(\mathcal{P}_{t,s})_{t < s \in T}$ (vgl. 1.8.2.1).

Beweis: Definieren die Wahrscheinlichkeitsmaße Q_{t_1, \dots, t_n} auf $\mathcal{B}(S)^n$ durch

$$Q_{t_1, \dots, t_n}(B_1 \times \dots \times B_n) := \int_{B_1} \dots \int_{B_n} \mathcal{P}_{t_{n-1}, t_n}(x_{n-1}, dx_n) \dots \mathcal{P}_{t_1, t_2}(x_1, dx_2) \mathcal{P}_{X_{t_1}}(dx_1)$$

für $B_i \in \mathcal{B}(S)$ und $t_1 < \dots < t_n \in T$. Dann lässt sich zeigen dass die (Q_{t_1, \dots, t_n}) die Bedingung 1.5.4.1 erfüllen[1], was die Behauptung impliziert.

□

1.8.5 Definition: Zeitlich homogene Übergangswahrscheinlichkeiten

Übergangswahrscheinlichkeiten $(\mathcal{P}_{t,s})_{t < s \in T}$ über (S, \mathcal{S}) heißen *zeitlich homogen* falls die $\mathcal{P}_{t,s}$ nur von der Differenz $(s - t)$ abhängen, das heißt

$$\mathcal{P}_{t,s}(x, B) = \mathcal{P}_{s-t}(x, B)$$

Markov-Prozesse mit zeitlich homogenen Übergangswahrscheinlichkeiten heißen *zeitlich homogen*.

Bemerkung:

- Die Chapman-Kolmogorov-Bedingung (1.8.3.1) lautet für zeitlich homogene Übergangswahrscheinlichkeiten:

$$\mathcal{P}_{t+s}(x, B) = \int_S \mathcal{P}_t(y, B) \mathcal{P}_s(x, dy) = \int_S \mathcal{P}_s(y, B) \mathcal{P}_t(x, dy)$$

- Ist ferner $\mathcal{P}_t(x, \cdot) \ll \lambda$ (Lebesgue-Maß λ auf \mathbb{R}) das heißt

$$\mathcal{P}_t(x, B) = \int_B \rho_t(x, z) dz \quad , \quad B \in \mathcal{B}(\mathbb{R})$$

so lautet diese

$$\rho_{t+s}(x, z) = \int_{\mathbb{R}} \rho_t(x, y) \rho_s(y, z) dy \tag{1.8.5.1}$$

1.8.6 Definition: Erzeuger zeitlich homogener Übergangswahrscheinlichkeiten

Es seien (S, \mathcal{S}) ein messbarer Raum und $(\mathcal{P}_t)_{t \in T}$ zeitlich homogene Markov-Kerne über (S, \mathcal{S}) . Dann definiert man für messbare, beschränkte Funktion $f \in \mathfrak{B}(S)$ und $t \in T$ den linearen Operator:

$$\mathcal{S}_t : \mathfrak{B}(S) \rightarrow \mathfrak{B}(S) \quad , \quad (\mathcal{S}_t f)(x) := \int_S f(y) \mathcal{P}_t(x, dy) \quad , \quad x \in E$$

Dabei sind alle \mathcal{S}_t beschränkt (genauer: Isometrien) und nach Chapman-Kolmogorov (1.8.3.1) gilt

$$\mathcal{S}_{t_1} \circ \mathcal{S}_{t_2} = \mathcal{S}_{t_1+t_2}$$

für $t_1, t_2, t_1 + t_2 \in T$. Ist also $T \subset \mathbb{R}$ eine Halbgruppe bzgl. Addition, dann ist $(\mathcal{S}_t)_{t \in T}$ eine Halbgruppe beschränkter, linearer Operatoren auf $\mathfrak{B}(S)$. Ist 0 ein Häufungspunkt von T so setzt man

$$\mathcal{G}f := \left. \frac{d}{dt} (\mathcal{S}_t f) \right|_{t=0} \quad , \quad (\mathcal{G}f)(x) := \lim_{t \rightarrow 0} \frac{(\mathcal{S}_t f)(x) - f(x)}{t} \quad (\text{wo möglich})$$

und nennt \mathcal{G} den *Erzeuger* von $(\mathcal{P}_t)_{t \in T}$.

Bemerkung: Die Bilder $\mathcal{G}f$ sind auch wieder Funktionen auf S , jedoch oft mit beschränktem Definitionsgebiet. Da die $\mathcal{G}f$ unbeschränkt sein können, ist auch \mathcal{G} im allgemeinen unbeschränkt.

1.8.7 Beispiel: Prozesse mit unabhängigen Zuwächsen

Sei $(X_t)_{t \geq 0}$ ein $(G, \mathcal{B}(G))$ -wertiger Prozess mit unabhängigen Zuwächsen und $X_0 = 0$ fast sicher, dazu

$$\mu_{t,s}(B) := \mathcal{P}(X_s - X_t \in B) \quad , \quad B \in \mathcal{B}(G)$$

und Übergangswahrscheinlichkeiten $\mathcal{P}_{t,s}$. Dann ist für $0 \leq t < s$:

$$\begin{aligned} \mathcal{P}_{t,s}(x, B) &:= \mathcal{P}(X_s \in B \mid X_t = x) = \mathcal{P}(X_s - x \in B - x \mid X_t = x) \\ &= \mathcal{P}(X_s - X_t \in B - x \mid \underbrace{X_t}_{X_t - X_0} = x) \stackrel{\text{unabhängige}}{\text{Zuwächse}} \mathcal{P}(X_s - X_t \in B - x) = \mu_{t,s}(B - x) \end{aligned}$$

Dabei lautet die Chapman-Kolmogorov Bedingung (1.8.3.1) an die Übergangswahrscheinlichkeiten $(\mathcal{P}_{t,s})$:

$$\int_G \mu_{s,u}(B - x - y) \mu_{t,s}(dy) \stackrel{!}{=} \mu_{t,u}(B - x)$$

bzw. äquivalent dazu

$$\underbrace{\int_G \mu_{s,u}(B - y) \mu_{t,s}(dy)}_{(\mu_{t,s} * \mu_{s,u})(B)} = \mu_{t,u}(B)$$

was aus der Unabhängigkeit der Zuwächse folgt (vgl. 1.7.6).

Im zeitlich homogenen Fall $\mu_{0,t} =: \mu_t$ ist dann

$$(\mathcal{S}_t f)(x) = \int_G f(y) \underbrace{\mathcal{P}_t(x, dy)}_{\mu_t(dy-x)} = \int_G f(x+y) \mu_t(dy) =$$

Ist z.B. $\mu_t = \mathcal{N}_{0,t}$ auf $G = \mathbb{R}$ so folgt

$$\begin{aligned} (\mathcal{S}_t f)(x) &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} f(x+y) e^{-\frac{y^2}{2t}} dy \stackrel{y=:s\sqrt{t}}{=} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x + \sqrt{t}s) e^{-\frac{s^2}{2}} ds \\ \Rightarrow (\mathcal{G}f)(x) &= \lim_{t \searrow 0} \frac{\mathcal{S}_t f - f}{t}(x) = \lim_{t \rightarrow 0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{f(x + \sqrt{t}s) - f(x)}{t} e^{-\frac{s^2}{2}} ds \end{aligned}$$

Speziell für $f \in \mathcal{C}^2(\mathbb{R})$:

$$\begin{aligned} (\mathcal{G}f)(x) &= \lim_{t \searrow 0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{\sqrt{t}s f'(x) + t \frac{s^2}{2} f''(x) + o(t^2)}{t} e^{-\frac{s^2}{2}} ds \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{s^2}{2} f''(x) e^{-\frac{s^2}{2}} ds = \frac{f''(x)}{2} \end{aligned}$$

das heißt

$$\mathcal{G}|_{\mathcal{C}^2} = \frac{1}{2} \frac{d^2}{dx^2}$$

1.8.8 Beispiel: Reflektierte Brownsche Bewegung

Sei $(W_t)_{t \geq 0}$ ein reellwertiger Wiener-Prozess (vgl. Beispiel (i) in 1.7.7) mit (homogenen) Übergangswahrscheinlichkeiten

$$\mathcal{P}_t(x, B) = \mathcal{N}_{x,t}(B) = \frac{1}{\sqrt{2\pi t}} \int_B e^{-\frac{(z-x)^2}{2t}} dz$$

(vgl. 1.8.7), dazu

$$X_t := |x_0 + W_t|, \quad t \geq 0$$

der *reflektierende* Prozess mit *Anfangsposition* $x_0 \geq 0$ (d.h. $X_0 = x_0$ fast sicher).

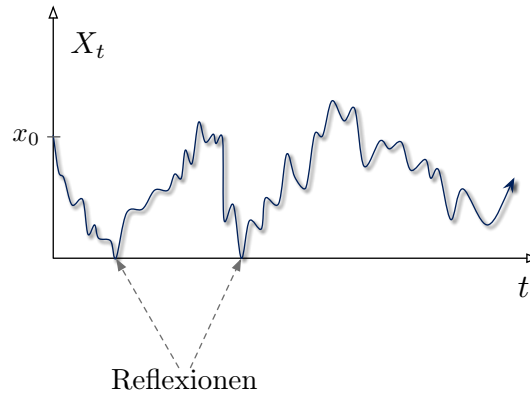


Abbildung 1.6: Typischer Pfad einer reflektierten Brownschen Bewegung mit Anfangsposition x_0 .

Zu $B \in \mathcal{B}([0, \infty))$ ist dann die Übergangswahrscheinlichkeit von X gegeben durch:

$$\begin{aligned} \tilde{\mathcal{P}}_t(x_0, B) &= \mathcal{P}(\{|x_0 + W_t| \in B\}) = \mathcal{P}(\{x_0 + W_t \in B \vee x_0 + W_t \in -B\}) \\ &= \mathcal{P}(\{x_0 + W_t \in B\}) + \mathcal{P}(\{x_0 + W_t \in -B\}) = \int_B \underbrace{\frac{1}{\sqrt{2\pi t}} [e^{-(y-x_0)^2/2t} + e^{-(y+x_0)^2/2t}]}_{\rho_t(x_0, y)} dy \end{aligned}$$

mit der Dichte $\rho_t(x_0, \cdot) := \frac{d\tilde{\mathcal{P}}_t(x_0, \cdot)}{d\lambda}$ bzgl. des Lebesgue-Maßes λ . Zu zeigen sei nun die Chapman-Kolmogorov Bedingung (vgl. 1.8.5.1). Tatsächlich ist

$$\begin{aligned} \int_0^\infty f(y) \rho_s(y, z) dy &= \frac{1}{\sqrt{2\pi s}} \int_0^\infty f(y) e^{-(y-z)^2/2s} dy + \underbrace{\frac{1}{\sqrt{2\pi s}} \int_0^\infty f(y) e^{-(y+z)^2/2s} dy}_{\frac{1}{\sqrt{2\pi s}} \int_{-\infty}^0 f(-y) e^{-(y-z)^2/2s} dy} \\ &= \frac{1}{\sqrt{2\pi s}} \int_{-\infty}^\infty f(|y|) e^{-(y-z)^2/2s} dy \end{aligned}$$

für beliebige messbare f . Ist nun speziell $f(y) := \rho_t(x_0, y)$ (x_0 fest) so folgt

$$\int_{\mathbb{R}} \rho_t(x_0, y) \rho_s(y, z) dy = \int_{\mathbb{R}} q_{x_0, t}(y) q_{z, s}(y) dy + \int_{\mathbb{R}} q_{-x_0, t}(y) q_{z, s}(y) dy = q_{x_0+z, t+s} + q_{z-x_0, t+s} = \rho_{t+s}(x_0, z)$$

mit der Dichte

$$q_{x_0, t}(y) := \frac{1}{\sqrt{2\pi t}} e^{-(y-x_0)^2/2t} = \frac{d\mathcal{N}_{x_0, t}}{d\lambda}(y)$$

Es existiert somit nach 1.8.4 ein Markov Prozess mit genau den obigen Übergangswahrscheinlichkeiten.

1.8.9 Spezialfälle für Markov-Prozesse

Sei $(X_t)_{t \in T}$ ein (S, \mathcal{S}) -wertiger Markov-Prozess mit Übergangswahrscheinlichkeiten $(\mathcal{P}_{t,s})_{t < s \in T}$.

1. Kontinuierliche Zeit $T = [0, \infty)$ und diskreter Zustandsraum $S = \mathbb{Z}$. Dann ist

$$\underbrace{\mathcal{P}_{t,s}(k, \{l\})}_{\mathcal{P}(X_s=l|X_t=k)} =: \mathcal{P}_{t,s}(k, l) \quad , \quad k, l \in \mathbb{Z}$$

wobei die Chapman-Kolmogorov Gleichungen lauten

$$\mathcal{P}_{t,s}(k, m) = \sum_{l=-\infty}^{\infty} \mathcal{P}_{t,s}(k, l) \mathcal{P}_{s,u}(l, m)$$

bzw. im homogenen Fall

$$\mathcal{P}_{t+s}(k, m) = \sum_{l=-\infty}^{\infty} \mathcal{P}_t(k, l) \mathcal{P}_s(l, m)$$

2. Zeit diskret $T = \mathbb{N}_0$, zeitlich homogener Fall:

$$\mathcal{P}_{n,n+1}(x, B) = \mathcal{P}_{0,1}(x, B) = \mathcal{P}(X_{n+1} \in B | X_n = x) = \mathcal{P}(X_1 \in B | X_0 = x) =: \mathcal{P}_1(x, B)$$

Analog definiert sei

$$\mathcal{P}_n(x, B) := \mathcal{P}(X_n \in B | X_0 = x)$$

Dann lauten die Chapman-Kolmogorov Gleichungen

$$\begin{aligned} \mathcal{P}_2(x, B) &= \int_S \mathcal{P}_1(y, B) \mathcal{P}_1(x, dy) \\ \mathcal{P}_3(x, B) &= \int_S \mathcal{P}_2(y, B) \mathcal{P}_1(x, dy) = \int_S \int_S \mathcal{P}_1(z, B) \mathcal{P}_1(y, dz) \mathcal{P}_1(x, dy) \\ &\vdots \end{aligned}$$

3. Diskrete Zeit $T = \mathbb{N}_0$ und diskreter Zustandsraum $S = \mathbb{Z}$, zeitlich homogener Fall:

$$\mathcal{P}_n(k, l) = \mathcal{P}(X_n = l | X_0 = k) \quad , \quad n \in \mathbb{N}$$

Dann kann die Familie

$$(\mathcal{P}_n(k, l))_{k, l = -\infty}^{\infty}$$

als unendliche Matrix interpretiert werden, wobei die Chapman-Kolmogorov Gleichung

$$\mathcal{P}_{n+1}(k, m) = \sum_{l=-\infty}^{\infty} \mathcal{P}_n(k, l) \mathcal{P}_1(l, m) \tag{1.8.9.1}$$

einer *Matrix-Multiplikation* entspricht.

Beispiel: Betrachten eine durch eine halbdurchlässige Membran aufgeteilte Doppelkammer mit Gasteilchen, in der zu diskreten Zeitpunkten jeweils ein Teilchen von der einen Hälfte in die andere Hälfte übergeht.

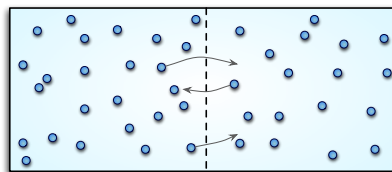


Abbildung 1.7: Verbundene Doppelkammer mit Gasteilchen.

Dabei sei $(X_n)_{n \in \mathbb{N}_0}$ der den Ablauf beschreibende Prozess, mit X_n als Teilchenzahl in der linken Kammer nach n Zeitschritten. Die (zeitlich homogenen) Übergangswahrscheinlichkeiten sind gegeben durch

$$\mathcal{P}(X_{n+1} = k + 1 \mid X_n = k) = \frac{2N - k}{2N}$$

$$\mathcal{P}(X_{n+1} = k - 1 \mid X_n = k) = \frac{k}{2N}$$

das heißt

$$\mathcal{P}_1(k, \{l\}) = \begin{cases} \frac{2N-k}{2N} & : l = k + 1 \\ \frac{k}{2N} & : l = k - 1 \\ 0 & : \text{sonst} \end{cases}$$

Über Gl. 1.8.9.1 ergeben sich dann (hinsichtlich der Konstruktion eines Markov-Prozesses) die restlichen Übergangswahrscheinlichkeiten.

2 Konvergenzbegriffe

2.1 Konvergenz von Maßen

2.1.1 Definition: Schwache Konvergenz von Maßen

Sei (E, d) ein metrischer Raum und $Q, (Q_n)_{n \in \mathbb{N}}$ Maße auf $(E, \mathcal{B}(E))$. Dann konvergieren Q_n *schwach* oder *in Verteilung* (engl. *weak*) gegen Q falls:

$$\forall f \in \mathcal{C}_b(E) : \int_E f dQ_n \xrightarrow[n \rightarrow \infty]{w} \int_E f dQ$$

Man schreibt dann: $Q_n \xrightarrow[n \rightarrow \infty]{w} Q$.

2.1.2 Satz: Charakterisierung von schwacher Konvergenz

Seien (E, d) metrisch und $Q, (Q_n)_{n \in \mathbb{N}}$ Wahrscheinlichkeitsmaße auf $(E, \mathcal{B}(E))$. Dann sind folgende Aussagen äquivalent:

1. $Q_n \xrightarrow[n \rightarrow \infty]{w} Q$

2. $\forall C \subseteq E$ abgeschlossen gilt:

$$\limsup_{n \rightarrow \infty} Q_n(C) \leq Q(C)$$

3. $\forall U \subseteq E$ offen gilt:

$$\liminf_{n \rightarrow \infty} Q_n(U) \geq Q(U)$$

4. $\forall B \in \mathcal{B}(E)$ mit $Q(\partial B) = 0$ gilt:

$$Q_n(B) \xrightarrow[n \rightarrow \infty]{w} Q(B)$$

2.1.3 Satz: Schwache Konvergenz unter stetigen Transformationen

Seien $(M, d_M), (N, d_N)$ metrisch und $(\mu_n), \mu$ Wahrscheinlichkeitsmaße auf $(M, \mathcal{B}(M))$ mit $\mu_n \xrightarrow[n \rightarrow \infty]{w} \mu$. Dann gehen auch für jede stetige $h : M \rightarrow N$ die Bildmaße

$$\mu_n \circ h^{-1} \xrightarrow[n \rightarrow \infty]{w} \mu \circ h^{-1}$$

Beweis: Für $f \in \mathcal{C}_b(N)$ ist auch $f \circ h : M \rightarrow \mathbb{R}$ stetig, beschränkt, also

$$\int_N f d(\mu_n)_h = \int_M f \circ h d\mu_n \xrightarrow[n \rightarrow \infty]{w} \int_M f \circ h d\mu = \int_N f d\mu_n$$

□

Bemerkung: Es kann sogar gezeigt werden: Die Behauptung gilt auch wenn $h : M \rightarrow N$ lediglich μ -fast überall stetig auf M ist.

2.1.4 Hilfslemma über die Stetigkeit der Metrik

Sei (E, d) ein metrischer Raum und

$$d(x, A) := \inf_{y \in A} d(x, y) \quad , \quad x \in E, A \subseteq E$$

Dann ist für jedes $A \subseteq E$ die Abbildung $d(\cdot, A) : x \mapsto d(x, A)$ stetig.

Beweis: Es gilt die Dreiecksungleichung

$$d(x, y) + d(y, A) = d(x, y) + \inf_{z \in A} d(y, z) = \inf_{z \in A} \underbrace{(d(x, y) + d(y, z))}_{\geq d(x, z)} \geq \inf_{z \in A} d(x, z) = d(x, A)$$

Dementsprechend gilt für Folge $x_n \xrightarrow{n \rightarrow \infty} x \in E$:

$$\underbrace{d(x, A) - d(x_n, x)}_{\xrightarrow{n \rightarrow \infty} d(x, A)} \leq d(x_n, A) \leq \underbrace{d(x_n, x) + d(x, A)}_{\xrightarrow{n \rightarrow \infty} d(x, A)}$$

das heißt $d(x_n, A) \xrightarrow{n \rightarrow \infty} d(x, A)$.

□

Bemerkung: Ist $A \subseteq E$ abgeschlossen, so ist $d(x, A) > 0$ für $x \notin A$, denn sonst gäbe es eine Folge $(y_n) \subset A$ mit $y_n \xrightarrow{n \rightarrow \infty} x$.

2.1.5 Satz: Eindeutigkeit des schwachen Grenzwertes

Sei (E, d) ein metrischer Raum und (μ_n) eine Folge von Maßen, die schwach gegen das endliche Maß μ konvergieren. Dann ist μ der einzige schwache Grenzwert von (μ_n) .

Beweis: Sei $A \subseteq E$ abgeschlossen. Betrachten zunächst die Familie von Funktionen

$$f_\varepsilon(x) := 1 - \min \left\{ \frac{1}{\varepsilon} d(x, A), 1 \right\}, \quad x \in E, \quad \varepsilon > 0$$

Nach Hilfslemma 2.1.4 ist jedes $f_\varepsilon : E \rightarrow \mathbb{R}$ stetig und beschränkt. Andererseits konvergieren die f_ε für $\varepsilon \searrow 0$ punktweise gegen 1_A , denn für festes $x \in E$ ist

$$\lim_{\varepsilon \searrow 0} f_\varepsilon(x) = 1 - \min \left\{ \underbrace{\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} d(x, A)}_{\infty \cdot 1_{A^c}(x)}, 1 \right\} = 1_A(x)$$

Sei nun ν ein weiterer schwacher Grenzwert von (μ_n) , dann ist auch ν endlich, denn für $1 \in \mathcal{C}_b(E)$ müsse gelten

$$\mu(E) = \int_E 1 \, d\mu = \lim_{n \rightarrow \infty} \int_E 1 \, d\mu_n = \int_E 1 \, d\nu = \nu(E)$$

Per Definition müsse außerdem gelten

$$\int_E f_\varepsilon \, d\mu = \lim_{n \rightarrow \infty} \int_E f_\varepsilon \, d\mu_n = \int_E f_\varepsilon \, d\nu \quad \forall \varepsilon > 0$$

Andererseits ist die Majorante $1 \geq |f_\varepsilon|$ bzgl. μ und ν integrierbar, das heißt nach Lebesgue

$$\mu(A) = \int_E 1_A \, d\mu = \lim_{\varepsilon \searrow 0} \int_E f_\varepsilon \, d\mu = \lim_{\varepsilon \searrow 0} \int_E f_\varepsilon \, d\nu = \int_E 1_A \, d\nu = \nu(A)$$

für jede abgeschlossene $A \subseteq E$. Da $\mathcal{C} := \{A \subseteq E : A \text{ abgeschlossen}\}$ durchschnittsstabil und ein Erzeuger von $\mathcal{B}(E)$ ist, mit $E \in \mathcal{C}$, folgt dass μ und ν sogar auf ganz $\mathcal{B}(E)$ übereinstimmen.

□

2.1.6 Lemma zur schwachen Konvergenz von Produktmaßen

Seien $\mathcal{P}_n, \mathcal{Q}_n, \mathcal{P}, \mathcal{Q}$ Borelmaße auf dem separablen, metrischen Raum (E, d) mit

$$\mathcal{P}_n \xrightarrow[\text{w}]{n \rightarrow \infty} \mathcal{P}, \quad \mathcal{Q}_n \xrightarrow[\text{w}]{n \rightarrow \infty} \mathcal{Q}$$

Dann geht

$$\mathcal{P}_n \otimes \mathcal{Q}_n \xrightarrow[\text{w}]{n \rightarrow \infty} \mathcal{P} \otimes \mathcal{Q}$$

auf $(E \times E, d \times d)$ wobei⁹

$$(d \times d)((x_1, x_2), (y_1, y_2)) := \max \{d(x_1, y_1), d(x_2, y_2)\}$$

2.2 Konvergenz von Zufallsvariablen

2.2.1 Definition: Konvergenz von Zufallsvariablen

Sei (E, d) ein separabler metrischer Raum, $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen $X_n : \Omega \rightarrow E$.

Dann:

- (X_n) konvergiert *stochastisch* oder *in Wahrscheinlichkeit* gegen die Zufallsvariable $X : \Omega \rightarrow E$ falls

$$\forall \varepsilon > 0 : \lim_{n \rightarrow \infty} \underbrace{\mathcal{P}(d(X_n, X) > \varepsilon)}_{\mathcal{P}\{\omega : d(X_n(\omega), X(\omega)) > \varepsilon\}} = 0$$

Man schreibt $X_n \xrightarrow[\text{p}]{n \rightarrow \infty} X$.

- (X_n) konvergiert *fast sicher* gegen X , falls

$$\mathcal{P}\left(\lim_{n \rightarrow \infty} X_n = X\right) = \mathcal{P}\left(\left\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

das heißt X_n konvergiert \mathcal{P} -fast-überall punktweise (in E) gegen X .

- X_n konvergiert *sicher*, oder *punktweise* gegen X , falls

$$X_n(\omega) \xrightarrow{n \rightarrow \infty} X(\omega) \quad \forall \omega \in \Omega$$

- (X_n) konvergiert *in Verteilung* oder *schwach* gegen X , wenn die Folge der induzierten Bildmaße \mathcal{P}_{X_n} schwach gegen das Bildmaß \mathcal{P}_X konvergiert, das heißt für jede beschränkte, stetige $f : E \rightarrow \mathbb{R}$ geht

$$\lim_{n \rightarrow \infty} \underbrace{\int_E f d\mathcal{P}_{X_n}}_{\mathbb{E}f(X_n)} = \underbrace{\int_E f d\mathcal{P}_X}_{\mathbb{E}f(X)}$$

Man schreibt $X_n \xrightarrow[\text{d}]{n \rightarrow \infty} X$. Ähnlich definiert man auch *Konvergenz in Verteilung* von (X_n) gegen irgendein Wahrscheinlichkeitsmaß \mathcal{P} .

- Ist $(E, \|\cdot\|_E)$ ein separabler Banachraum, so konvergiert (X_n) gegen X *im p -ten Mittel* ($0 < p < \infty$), falls

$$\lim_{n \rightarrow \infty} \mathbb{E} \|X_n - X\|_E^p = \lim_{n \rightarrow \infty} \int_{\Omega} \|X_n(\omega) - X(\omega)\|_E^p d\omega = 0$$

⁹Die Separabilität von E sichert, dass $d \times d$ genau die σ -Algebra $\mathcal{B}(E) \otimes \mathcal{B}(E)$ erzeugt.

Bemerkungen:

- (i) Fast sichere Konvergenz bzw. Konvergenz im p -ten Mittel implizieren stochastische Konvergenz.
- (ii) Stochastische Konvergenz impliziert stets Konvergenz in Verteilung.
- (iii) Konvergiert X_n in Verteilung gegen eine Konstante C , so konvergiert X_n auch stochastisch gegen C .
- (iv) Seien $a_n \xrightarrow[n \rightarrow \infty]{d} a \in E$ und $X_n \xrightarrow[n \rightarrow \infty]{d} X$. Dann gehen auch

$$a_n X_n \xrightarrow[n \rightarrow \infty]{d} aX$$

- (v) Sei (M, d_M) ebenfalls ein metrischer Raum und $h : E \rightarrow M$ \mathcal{P}_X -fast-überall stetig. Gehen die Zufallsvariablen $X_n \xrightarrow[n \rightarrow \infty]{d} X$ so gehen auch:

$$h(X_n) \xrightarrow[n \rightarrow \infty]{d} h(X)$$

(vgl. analogen Satz 2.1.3 über Maße). Analog, ist μ ein W-Maß auf $(M, \mathcal{B}(M))$ und $X_n \xrightarrow[n \rightarrow \infty]{d} \mu$, dann gehen auch

$$h(X_n) \xrightarrow[n \rightarrow \infty]{d} \mu \circ h^{-1}$$

2.2.2 Lemma: Charakterisierung von Konvergenz in Verteilung

Seien $(X_n), X$ \mathbb{R}^m -wertige Zufallsvariablen. Dann sind äquivalent:

1. X_n konvergiert in Verteilung gegen X : $X_n \xrightarrow[n \rightarrow \infty]{d} X$.
2. Sind \tilde{X}_n, \tilde{X} die Fourier-transformierten von X_n bzw. X , so konvergieren die \tilde{X}_n punktweise gegen \tilde{X} :

$$\lim_{n \rightarrow \infty} \tilde{X}_n(\xi) = \tilde{X}(\xi) \quad \forall \xi \in \mathbb{R}^m$$

3. Im Fall $m = 1$: Für die Verteilungsfunktionen F_n, F von X_n bzw. X gilt

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

in allen Stetigkeitspunkten von F .

2.2.3 Satz: Übertragung der Konvergenz in Verteilung

Sei (M, d_M) metrisch, separabel und X_n, Y_n, X M -wertige Zufallsvariablen mit $X_n \xrightarrow[n \rightarrow \infty]{d} X$ und $d(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{d} 0$ fast sicher. Dann gehen auch

$$Y_n \xrightarrow[n \rightarrow \infty]{d} X$$

Für einen Beweis siehe [5].

Folgerung: Sind X_n, X \mathbb{R}^n -wertig und $X_n \xrightarrow[n \rightarrow \infty]{d} X$, $a_n \xrightarrow[n \rightarrow \infty]{d} 0$, so geht auch

$$X_n + a_n \xrightarrow[n \rightarrow \infty]{d} X$$

2.2.4 Der zentrale Grenzwertsatz (CLT)

Sei X_1, X_2, \dots eine Folge von unabhängig, identisch verteilten, reellen Zufallsvariablen auf $(\Omega, \mathfrak{G}, \mathcal{P})$ mit $\mathbb{E}|X_j|^2 < \infty$ und

$$\mu := \mathbb{E}X_j, \quad \sigma := \sqrt{\text{Var}(X_1)} .$$

Dann konvergiert die Folge von Zufallsvariablen

$$Z_n := \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^n [X_j - \mu] \quad , \quad n \in \mathbb{N}$$

in Verteilung gegen die Standardnormalverteilung $\mathcal{N}_{0,1}$: $Z_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}_{0,1}$. Insbesondere heißt dies: Ist Φ die Verteilungsfunktion von $\mathcal{N}_{0,1}$, so geht

$$\lim_{n \rightarrow \infty} \mathcal{P}(Z_n \leq t) = \Phi(t)$$

für jedes $t \in \mathbb{R}$.

2.2.5 Der mehrdimensionale zentrale Grenzwertsatz (CLT)

Sei X, X_1, X_2, \dots eine Folge von unabhängig, identisch verteilten, \mathbb{R}^m -Zufallsvektoren mit $\mathbb{E}\|X\|^2 < \infty$ und Kovarianzmatrix $\text{Cov}(X)$. Dann gilt:

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n [X_k - \mathbb{E}X] \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}_{0, \text{Cov}(X)}$$

(vgl. 3.2.6 unten).

2.3 Konvergenz stochastischer Prozesse

2.3.1 Definition: Konvergenz stochastischer Prozesse

Seien $(X_t^n)_{t \in T}$, $n \in \mathbb{N}$ und $(X_t)_{t \in T}$ \mathbb{R}^m -wertige stochastische Prozesse. Dann konvergieren $X^n \xrightarrow[n \rightarrow \infty]{d} X$ in endlich dimensionalen Verteilungen \Leftrightarrow

$$\forall t_1, \dots, t_k \in T : (X_{t_1}^n, \dots, X_{t_k}^n) \xrightarrow[n \rightarrow \infty]{d} (X_{t_1}, \dots, X_{t_k})$$

$$\Leftrightarrow \forall t_1, \dots, t_k \in T : \forall f \in \mathcal{C}_b(\mathbb{R}^{mk}) : \mathbb{E}f(X_{t_1}^n, \dots, X_{t_k}^n) \xrightarrow[n \rightarrow \infty]{} \mathbb{E}f(X_{t_1}, \dots, X_{t_k})$$

$$\stackrel{(2.2.2)}{\Leftrightarrow} \forall t_1, \dots, t_k \in T, \forall \alpha_1, \dots, \alpha_k \in \mathbb{R}^m : \mathbb{E} \exp \left[i \sum_{j=1}^k \alpha_j X_{t_j}^n \right] \xrightarrow[n \rightarrow \infty]{} \mathbb{E} \exp \left[i \sum_{j=1}^k \alpha_j X_{t_j} \right]$$

Man schreibt $X^n \xrightarrow[n \rightarrow \infty]{\text{f.d.d.}} X$.

Bemerkung: Diese Art der Konvergenz ist meist einfach nachweisbar, leider aber oft zu schwach. So folgt z.B. aus ihr nicht

$$\sup_{t \in T} |X_n(t)| \xrightarrow[n \rightarrow \infty]{d} \sup_{t \in T} |X(t)|$$

(im Falle der Messbarkeit).

2.3.2 Theorem von Kratowski

Sei (E, d) ein vollständiger, separabler metrischer Raum und $\mathfrak{B} \subset \mathcal{B}(E)$ separiere die Punkte in E , das heißt:

$$\forall x \neq y \in E : \exists A \in \mathfrak{B} : x \notin A \ni y$$

Dann gilt $\sigma(\mathfrak{B}) = \mathcal{B}(E)$.

2.3.3 Bemerkung über Borel- σ -Algebren in \mathcal{C}

Sei (M, d_M) ein separabler metrischer Raum¹⁰, $T \neq \emptyset$ eine Indexmenge und $E \subseteq M^T$. Mit der Metrik d_E werde (E, d_E) zu einem vollständigen, separablen metrischen Raum, so dass die Abbildungen

$$\zeta_t : (E, \mathcal{B}(E)) \rightarrow (M, \mathcal{B}(M)), \quad \zeta_t : f \mapsto f(t)$$

messbar sind. Dann gilt

$$\mathcal{B}(E) = \underbrace{\mathcal{C}(M^T) \cap E}_{\text{Schnitt-}\sigma\text{-Algebra}}$$

Beweis: Offensichtlich ist

$$\mathcal{C}_0(M^T) \cap E = \left\{ \underbrace{(\zeta_{t_1}, \dots, \zeta_{t_m})^{-1}(B)}_{\in \mathcal{B}(E)} : B \in \mathcal{B}(M), t_1, \dots, t_m \in T \right\} =: \mathfrak{B}$$

und $\mathfrak{B} \subseteq \mathcal{B}(E)$ separiert die Punkte in E , denn: Für $f \neq g \in E$ existiert ein $t_0 \in T$ mit $f(t_0) \neq g(t_0)$, also

$$g \notin \underbrace{\zeta_{t_0}^{-1} \left(\underbrace{B_{d_M(f(t_0), g(t_0))/2}^o}_{\in \mathcal{B}(M)}(f(t_0)) \right)}_{\in \mathfrak{B}} \ni f$$

Demnach

$$\mathcal{C}(M^T) \cap E = \sigma(\mathcal{C}_0(M^T)) \cap E = \sigma_E(\mathfrak{B}) \stackrel{(2.3.2)}{=} \mathcal{B}(E)$$

Bemerkung: Die Messbarkeit von $(\zeta_{t_1}, \dots, \zeta_{t_m})$ setzt die Separabilität von M voraus. \square

2.3.4 Satz über Messbarkeit von Pfaden

Sei (M, d_M) ein separabler, metrischer Raum und $(X_t)_{t \in T}$ ein M -wertiger stochastischer Prozess auf $(\Omega, \mathfrak{S}, \mathcal{P})$ mit Pfaden in $E \subseteq (M)^T$. Dabei sei (E, d_E) ein vollständiger, separabler metrischer Raum und die Abbildungen

$$\zeta_t : f \mapsto f(t), \quad f \in M$$

für beliebiges $t \in T$ messbar in $(E, \mathcal{B}(E))$. Dann ist auch $X : \Omega \rightarrow (E, \mathcal{B}(E))$ messbar.

Beweis: Nach Lemma 1.3.4 ist $X : \Omega \rightarrow (M, \mathcal{C}(M^T))$ messbar. Nach Bemerkung 2.3.3 ist dann auch $X : \Omega \rightarrow (E, \mathcal{B}(E))$ messbar.

\square

Folgerung: Hat X Pfade in E , so erzeugt X ein Verteilungsgesetz auf $(E, \mathcal{B}(E))$.

Beispiel: Für kompakten, metrischen Raum (T, d_T) und normierten, separablen Raum $(M, \|\cdot\|_M)$ wird $(\mathcal{C}(T, M), \|\cdot\|_\infty) \subseteq M^T$ zu einem separablen Banachraum, mit ζ_t messbar (sogar stetig) von $\mathcal{C}(T, M)$ nach M . Daher ist jeder M -wertige Prozess $(X_t)_{t \in T}$ mit stetigen Pfaden auch messbar nach $(\mathcal{C}(T, M), \|\cdot\|_\infty)$.

2.3.5 Beispiel: Schwache Konvergenz von Pfad-Suprema

Für kompakten metrischen Raum (T, d_T) und $E = (\mathcal{C}(T), \|\cdot\|_\infty)$ ist $\zeta_t : f \mapsto f(t)$ messbar auf $(E, \mathcal{B}(E))$. Nach Bemerkung 2.3.3 gilt dann:

$$\mathcal{B}(\mathcal{C}(T)) = \mathcal{C}(\mathbb{R}^T) \cap \mathcal{C}(T)$$

Seien nun X, X^n \mathbb{R} -wertige stochastische Prozesse auf $(\Omega, \mathfrak{S}, \mathcal{P})$ mit stetigen Pfaden und $X^n \xrightarrow[\text{d}]{n \rightarrow \infty} X$ ¹¹. Da $\|\cdot\|_\infty : E \rightarrow \mathbb{R}$ stetig ist, folgt dann nach 2.1.3 auch¹²

$$\|X^n\|_\infty \xrightarrow[\text{d}]{n \rightarrow \infty} \|X\|_\infty$$

¹⁰Ausgestattet mit der Borel- σ -Algebra $\mathcal{B}(M)$ induziert durch die Metrik d_M .

¹¹Betrachtet als Zufallsvariablen $X, X_n : \Omega \rightarrow (E, \|\cdot\|_\infty)$

¹²Als Zufallsvariablen $\|X^n\|_\infty, \|X\|_\infty : \Omega \rightarrow \mathbb{R}$.

Insbesondere: Ist $\mathcal{P}(\|X\|_\infty = \alpha) = 0$ für ein $\alpha \in \mathbb{R}$, das heißt die Verteilungsfunktion von $\|X\|_\infty$ ist stetig in α , so folgt nach Charakterisierungssatz 2.2.2 (3)

$$\mathcal{P}(\|X^n\|_\infty \leq \alpha) \xrightarrow{n \rightarrow \infty} \mathcal{P}(\|X\|_\infty \leq \alpha)$$

Ähnlich: Für endliches Borelmaß μ auf $(T, \mathcal{B}(T))$ ist die Abbildung $f \mapsto \int_T f d\mu$ stetig auf $(\mathcal{C}(T), \|\cdot\|_\infty)$, also geht auch

$$\int_T X_t^n \mu(dt) \xrightarrow[n]{d} \int_T X_t \mu(dt)$$

2.3.6 Definition: Radon-Maß

Sei $(T, \mathcal{O}(T))$ ein Hausdorff Raum. Ein lokal-endliches¹³ Maß μ auf $(T, \mathcal{B}(T))$ heißt *Radonsch* (*straff*, engl. *tight*) falls μ *inner-regulär* ist, das heißt für $A \in \mathcal{B}(T)$ gilt:

$$\mu(A) = \sup \{ \mu(K) : K \subseteq A \text{ kompakt} \}$$

Spezialfall: Ein Wahrscheinlichkeitsmaß μ auf einem metrischen Raum (E, d) ist genau dann Radonsch wenn zu jedem $\varepsilon > 0$ eine kompakte Menge $K \subseteq E$ existiert mit

$$\mu(K) \geq 1 - \varepsilon$$

Bemerkung: Jedes Borel-Wahrscheinlichkeitsmaß auf einem vollständigen, separablen, metrischen Raum ist Radonsch.

2.3.7 Definition: Gleichmäßig Radonsch

Zu messbarem Raum (Ω, \mathfrak{S}) sei $\mathfrak{M}_1(\Omega)$ die Menge aller Wahrscheinlichkeitsmaße auf (Ω, \mathfrak{S}) .

Ist $(T, \mathcal{O}(T))$ Hausdorffsch, so heißt eine Untermenge $\mathcal{M} \subset \mathfrak{M}_1(T)$ (*gleichmäßig Radonsch*) \Leftrightarrow

$$\forall \varepsilon > 0 : \exists K \subseteq T \text{ kompakt} : \mu(K) \geq 1 - \varepsilon \quad \forall \mu \in \mathcal{M}$$

Offensichtlich sind dann alle $\mu \in \mathcal{M}$ radonsch.

Beispiel: Die Menge von Wahrscheinlichkeitsmaßen

$$\{ \delta_n : n \in \mathbb{N} \} \subseteq \mathfrak{M}_1(\mathbb{R})$$

ist nicht gleichmäßig radonsch, obwohl jedes δ_n radonsch ist.

Bemerkung:

- (i) Ist (E, d) vollständig, separabel, metrisch und $\mathcal{M} \subseteq \mathfrak{M}_1(E)$ endlich, so ist \mathcal{M} stets radonsch.
- (ii) Ist (E, d) metrisch, so existiert auf $\mathfrak{M}_1(E)$ eine Metrik π (*Prokhorov Metrik*), definiert durch

$$\pi(\mu, \nu) := \inf \{ \varepsilon > 0 : \mu(A^\varepsilon) \leq \nu(A) + \varepsilon \quad \forall A \in \mathcal{B}(E) \}$$

Ist E sogar separabel, vollständig, so gilt für Wahrscheinlichkeitsmaße $\mu, (\mu_n)_{n \in \mathbb{N}} \in \mathfrak{M}_1(E)$:

$$\mu_n \xrightarrow[n]{w} \mu \Leftrightarrow \pi(\mu_n, \mu) \xrightarrow{n \rightarrow \infty} 0$$

das heißt π induziert auf $\mathfrak{M}_1(E)$ die gleiche Topologie wie die schwache Konvergenz.

¹³Ein Borel-Maß μ auf einem Hausdorff-Raum $(T, \mathcal{O}(T))$ heißt *lokal-endlich*, falls es zu jedem $x \in T$ eine offene Menge $U \in \mathcal{O}(T)$ gibt mit $\mu(U) < \infty$.

2.3.8 Definition: Relative Kompaktheit

Eine Teilmenge $A \subseteq T$ eines topologischen Raumes $(T, \mathcal{O}(T))$ heißt *relativ kompakt* falls der Abschluss $\text{cl}(A)$ kompakt ist.¹⁴

Bemerkung: Eine Teilmenge $A \subseteq E$ eines metrischen Raumes E ist genau dann relativ kompakt, falls jede Folge in A eine in E konvergente Teilfolge hat.

2.3.9 Theorem von Prokhorov (1956)

Sei (E, d) ein vollständiger, separabler metrischer Raum und $\mathcal{M} \subseteq \mathfrak{M}_1(E)$. Dann gilt: \mathcal{M} relativ kompakt in (\mathfrak{M}_1, π) ¹⁵ \Leftrightarrow \mathcal{M} ist gleichmäßig Radonsch.

2.3.10 Satz: Charakterisierung der Konvergenz stochastischer Prozesse

Sei $T \neq \emptyset$ eine Indexmenge und $E \subseteq \mathbb{R}^T$. Sei (E, d) metrisch, vollständig, separabel und $\zeta_t : f \mapsto f(t)$ stetig von E nach \mathbb{R} . Seien $(X^n), X$ reellwertige stochastische Prozesse mit Pfaden in $E \subseteq \mathbb{R}^T$.

Dann gilt $X^n \xrightarrow[\text{d}]{n \rightarrow \infty} X$ (als Zufallsvariablen $X, X^n : \Omega \rightarrow (E, \mathcal{B}(E))$) genau dann wenn:

(a) $X^n \xrightarrow[\text{f.d.d.}]{n \rightarrow \infty} X$

(b) $\{\mathcal{P}_{X^n} : n \in \mathbb{N}\}$ ist relativ kompakt in $(\mathfrak{M}_1(E), \pi)$.

Beweis: Richtung "⇒": Wegen Stetigkeit von ζ_t ist für $t_1, \dots, t_m \in T$ auch

$$f \mapsto (f(t_1), \dots, f(t_m)) \quad , \quad f \in E$$

stetig von E nach \mathbb{R}^m . Nach Satz 2.1.3 gehen deshalb auch

$$(X_{t_1}^n, \dots, X_{t_m}^n) \xrightarrow[\text{d}]{n \rightarrow \infty} (X_{t_1}, \dots, X_{t_m})$$

Definitionsgemäß geht auch $\mathcal{P}_{X^n} \xrightarrow[\text{w}]{n \rightarrow \infty} \mathcal{P}_X$, das heißt $\{\mathcal{P}_{X^n}\}_{n \in \mathbb{N}}$ ist lediglich eine konvergente Folge in $\mathfrak{M}_1(E)$ und als solche relativ kompakt.

Richtung "⇐": Beweis durch Widerspruch. Unter der Annahme $\mathcal{P}_{X^n} \not\xrightarrow[\text{w}]{n \rightarrow \infty} \mathcal{P}_X$ existiert ein $\varepsilon_0 > 0$ und Teilfolge (n_k) mit

$$\pi(\mathcal{P}_{X^{n_k}}, \mathcal{P}_X) \geq \varepsilon_0 \quad , \quad k \in \mathbb{N}$$

Nach Voraussetzung (b) existiert jedoch auch eine Teilfolge $(n'_k) \subset (n_k)$ und Wahrscheinlichkeitsmaß $Q \in \mathfrak{M}_1(E)$ mit

$$\mathcal{P}_{X^{n'_k}} \xrightarrow[\text{w}]{n \rightarrow \infty} Q$$

Insbesondere nach Satz 2.1.3

$$\underbrace{\mathcal{P}_{X^{n'_k}} \circ \overbrace{(\zeta_{t_1}, \dots, \zeta_{t_m})^{-1}}^{\text{stetig}}}_{\mathcal{P}_{t_1, \dots, t_m}^{X^{n'_k}}} \xrightarrow[\text{w}]{n \rightarrow \infty} \underbrace{Q \circ (\zeta_{t_1}, \dots, \zeta_{t_m})^{-1}}_{Q_{t_1, \dots, t_m}}$$

Nach Voraussetzung (a) gehen auch

$$\mathcal{P}_{t_1, \dots, t_m}^{X^{n'_k}} \xrightarrow[\text{w}]{n \rightarrow \infty} \mathcal{P}_{t_1, \dots, t_m}^X$$

Aufgrund der Eindeutigkeit des Grenzwertes (vgl. Satz 2.1.5) muss $Q_{t_1, \dots, t_m} = \mathcal{P}_{t_1, \dots, t_m}^X$ sein für beliebige $t_1, \dots, t_m \in T$, das heißt

$$\mathcal{P}_X = Q$$

¹⁴In vollständigen metrischen Räumen ist relative Kompaktheit äquivalent zu Präkompaktheit.

¹⁵Vergleiche mit Bemerkung (ii) in 2.3.7.

auf $\mathcal{C}_0(\mathbb{R}^T) \cap E$. Da $\mathcal{C}_0(\mathbb{R}^T) \cap E$ ein durchschnittsstabiler Erzeuger von $\overbrace{\mathcal{C}(\mathbb{R}^T) \cap E}^{\sigma(\mathcal{C}_0(\mathbb{R}^T) \cap E)}$ ist, sind $\mathcal{P}_X = Q$ sogar auf ganz $\underbrace{\mathcal{C}(\mathbb{R}^T) \cap E}_{\mathcal{B}(E)}$. Insbesondere
nach (2.3.3)

$$\mathcal{P}_{X^{n_k}} \xrightarrow[\text{w}]{n \rightarrow \infty} \mathcal{P}_X$$

Doch dies ist im direkten Widerspruch zu

$$\pi(\mathcal{P}_{X^{n_k}}, \mathcal{P}_X) \geq \varepsilon_0$$

□

3 Gaußprozesse

3.1 Charakteristische Funktion

3.1.1 Konventionen

Schreiben für $x, y \in \mathbb{C}^n$:

$$\begin{aligned} x &= (x^1, \dots, x^n) \\ \langle x, y \rangle &:= xy := \sum_{i=1}^n x^{i*} y^i \\ \|x\| &:= |x| := \sqrt{\langle x, x \rangle} \end{aligned}$$

3.1.2 Definition: Charakteristische Funktion

Zu Wahrscheinlichkeitsmaß \mathcal{P} auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ definieren die *charakteristische Funktion* (*Fourier-Transformierte*) von \mathcal{P} :

$$\tilde{\mathcal{P}} : \mathbb{R}^n \rightarrow \mathbb{C}, \quad \tilde{\mathcal{P}}(k) := \mathcal{F}(\mathcal{P})(k) := \int_{\mathbb{R}^n} e^{ikx} \mathcal{P}(dx), \quad k \in \mathbb{R}^n$$

Zu zufälligen Vektor $X : \Omega \rightarrow \mathbb{R}^n$ auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$ ist die *charakteristische Funktion* von X gemäß

$$\tilde{X} : \mathbb{R}^n \rightarrow \mathbb{C}, \quad \tilde{X}(k) := \mathcal{F}(X)(k) := \mathbb{E}(e^{ikX}) = \tilde{\mathcal{P}}_X(k)$$

definiert.

3.1.3 Satz: Eigenschaften der charakteristischen Funktion

Zu W-Maß \mathcal{P} auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ hat die charakteristische Funktion $\tilde{\mathcal{P}}$ folgende Eigenschaften:

1. $\tilde{\mathcal{P}}(0) = 1$
2. $\tilde{\mathcal{P}}$ ist stetig.
3. $\tilde{\mathcal{P}}$ ist nicht-negativ definit, das heißt die Matrix

$$\left(\tilde{\mathcal{P}}(k_i - k_j) \right)_{i,j=1}^m \in \mathbb{C}^{m \times m}$$

ist nicht-negativ definit¹⁶ in \mathbb{C}^m für beliebige $k_1, \dots, k_m \in \mathbb{R}^n$.¹⁷

3.1.4 Theorem von Bochner

Sei $\varphi : \mathbb{R}^n \rightarrow \mathbb{C}$ eine Abbildung mit allen 3 Eigenschaften in 3.1.3 erfüllt. Dann existiert ein eindeutig bestimmtes Wahrscheinlichkeitsmaß \mathcal{P} auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ mit

$$\tilde{\mathcal{P}} = \varphi$$

Insbesondere gilt

$$\mathcal{P} = \mathcal{Q} \Leftrightarrow \tilde{\mathcal{P}} = \tilde{\mathcal{Q}}$$

für W-Maße \mathcal{P}, \mathcal{Q} .

¹⁶Eine konjugiert-symmetrische Sesquilinearform $b : H \times H \rightarrow \mathbb{C}$ auf einem \mathbb{C} -Vektorraum H heißt *nicht-negativ definit*, falls gilt $b(v, v) \geq 0 \forall v \in H$. Ist ferner H ein Hilbertraum, so heißt ein selbstadjungierter, linearer Operator $A : H \rightarrow H$ *nicht-negativ definit* falls $\langle Hv, v \rangle \geq 0 \forall v \in H$ gilt. Im Falle einer hermiteschen Matrix, ist eine notwendige (jedoch nicht hinreichende!) Bedingung dass alle ihre Hauptminoren nicht-negativ sind.[6]

¹⁷Beachte dass obige Matrix hermitesch ist, denn $\tilde{\mathcal{P}}(-k) = \tilde{\mathcal{P}}^*(k) \forall k \in \mathbb{R}^n$.

3.1.5 Satz: Verknüpfungen charakteristischer Funktionen

1. Zu Wahrscheinlichkeitsmaßen \mathcal{P}, \mathcal{Q} auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ ist

$$\widetilde{(\mathcal{P} * \mathcal{Q})} = \widetilde{\mathcal{P}} \cdot \widetilde{\mathcal{Q}}$$

Die Zuordnung $\mathcal{P} \mapsto \widetilde{\mathcal{P}}$ ist also ein Monomorphismus¹⁸ zwischen den Monoiden $(\mathfrak{M}_1(\mathbb{R}^n), *)$ und $(\mathbb{C}^{\mathbb{R}^n}, \cdot)$. Analog gilt auch für zufällige, unabhängige Vektoren X, Y :

$$\widetilde{(X + Y)} = \widetilde{X} \cdot \widetilde{Y}$$

Insbesondere für zufälligen Vektor $X : \Omega \rightarrow \mathbb{R}^n$ und Konstante $X_0 \in \mathbb{R}^n$:

$$\widetilde{(X + X_0)}(k) = \widetilde{X}(k) \cdot e^{i\langle k, X_0 \rangle}$$

2. Sei \mathcal{P} ein W-Maß auf \mathbb{R}^n und \mathcal{Q} ein W-Maß auf \mathbb{R}^m , dazu das Produktmaß $\mathcal{P} \otimes \mathcal{Q}$ auf \mathbb{R}^{n+m} . Dann ist

$$\widetilde{(\mathcal{P} \otimes \mathcal{Q})}(a, b) = \widetilde{\mathcal{P}}(a) \cdot \widetilde{\mathcal{Q}}(b) \quad , \quad a \in \mathbb{R}^n, b \in \mathbb{R}^m$$

Analog für unabhängige, zufällige Vektoren $X : \Omega \rightarrow \mathbb{R}^n, Y : \Omega \rightarrow \mathbb{R}^m$:

$$\widetilde{(X, Y)}(a, b) = \widetilde{X}(a) \cdot \widetilde{Y}(b) \tag{3.1.5.1}$$

3. Sei \mathcal{P} ein W-Maß auf \mathbb{R}^n und $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear, dazu die adjungierte¹⁹ A^\dagger . Dann gilt:

$$\widetilde{\mathcal{P}_A} = \widetilde{\mathcal{P}} \circ A^\dagger \tag{3.1.5.2}$$

Beweis:

$$\widetilde{\mathcal{P}_A}(k) = \int_{\mathbb{R}^m} e^{iky} \mathcal{P}_A(dy) = \int_{\mathbb{R}^n} e^{ikA(x)} \mathcal{P}(dx) = \int_{\mathbb{R}^n} e^{iA^\dagger(k)x} \mathcal{P}(dx) = \widetilde{\mathcal{P}}(A^\dagger(k)) \quad , \quad k \in \mathbb{R}^m$$

4. Zu $x_0 \in \mathbb{R}^n$ ist

$$\widetilde{\delta_{x_0}}(k) = e^{ikx_0}$$

und somit nach (1) für W-Maß \mathcal{P} :

$$\widetilde{\mathcal{P} * \delta_{x_0}}(k) = e^{ikx_0} \widetilde{\mathcal{P}}(k) \quad , \quad k \in \mathbb{R}^n$$

Beachte dass $(\mathcal{P} * \delta_{x_0})(B) = \mathcal{P}(B - x_0)$.

5. Sei $X : \Omega \rightarrow \mathbb{R}^n$ ein zufälliger Vektor und $k \in \mathbb{R}^n$ fest, dazu die Zufallsvariable $\langle k, X \rangle : \Omega \rightarrow \mathbb{R}$. Dann ist

$$\widetilde{\langle k, X \rangle}(t) = \widetilde{X}(tk) = \mathbb{E}(e^{itkX}) \quad , \quad t \in \mathbb{R}$$

bzw.

$$\widetilde{X}(k) = \widetilde{\langle k, X \rangle}(1)$$

6. Sei \mathcal{P} ein W-Maß auf \mathbb{R}^n und $k \in \mathbb{R}^n$ fest, dazu die 1-Form $x \mapsto \langle k, x \rangle, x \in \mathbb{R}^n$. Dann ist

$$\widetilde{\mathcal{P}_k}(t) \stackrel{(3)}{=} \widetilde{\mathcal{P}}(k^\dagger(t)) = \widetilde{\mathcal{P}}(tk) \quad , \quad t \in \mathbb{R}$$

bzw.

$$\widetilde{\mathcal{P}}(k) = \widetilde{\mathcal{P}_k}(1)$$

¹⁸Ein Monomorphismus $f : G \rightarrow H$ zwischen zwei Halbgruppen G, H ist ein injektiver Homomorphismus.

¹⁹Zu linearem Operator $A : V \rightarrow W$ zwischen endlich dimensionalen Hilberträumen V, W ist der *adjungierte* $A^\dagger : W \rightarrow V$ der eindeutig bestimmte lineare Operator der erfüllt $\langle Av, w \rangle_W = \langle v, A^\dagger w \rangle_V$ für $v \in V, w \in W$. In Matrixdarstellung ist $(A^\dagger) = (A^*)^T$.

3.1.6 Satz von Cramér-Wold

Seien \mathcal{P} und \mathcal{Q} W-Maße auf \mathbb{R}^n mit

$$\mathcal{P}_k = \mathcal{Q}_k \quad \forall k \in \mathbb{R}^n \quad (3.1.6.1)$$

(dabei $x \mapsto \langle k, x \rangle$, $x \in \mathbb{R}^n$). Dann folgt $\mathcal{P} = \mathcal{Q}$.

Beweis: Nach Satz 3.1.5 (6) ist

$$\tilde{\mathcal{P}}(k) \stackrel{(3.1.5)}{=} \tilde{\mathcal{P}}_k(1) = \tilde{\mathcal{Q}}_k(1) \stackrel{(3.1.5)}{=} \tilde{\mathcal{Q}}(k) \quad \forall k \in \mathbb{R}^n$$

Nach Bochner (3.1.4) ist somit $\mathcal{P} = \mathcal{Q}$.

□

Interpretation: Äquivalent zu Gl. 3.1.6.1 (bzw. zu $\mathcal{P} = \mathcal{Q}$) ist

$$\forall k \in \mathbb{R}^n \quad \forall \alpha \in \mathbb{R} : \overbrace{\mathcal{P}(\{x \in \mathbb{R}^n : \langle x, k \rangle \leq \alpha\})}^{\mathcal{P}_k((-\infty, \alpha])} = \overbrace{\mathcal{Q}(\{x \in \mathbb{R}^n : \langle x, k \rangle \leq \alpha\})}^{\mathcal{Q}_k((-\infty, \alpha])}$$

Stimmen also die Maße \mathcal{P}, \mathcal{Q} auf allen Halbräumen $\{x \in \mathbb{R}^n : \langle k, x \rangle \leq \alpha\}$, $k \in \mathbb{R}^n, \alpha \in \mathbb{R}$ überein, so sind sie gleich.

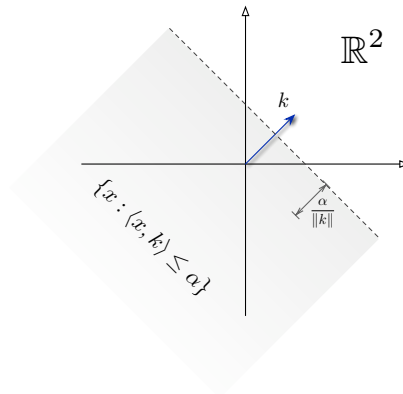


Abbildung 3.1: Zum Satz von Cramér-Wold über Charakterisierung von Maßen durch Halbräume.

Anwendung findet dieser Satz z.B. in der Computertomographie, in der die Bestimmung des Maßes (z.B. Konzentration bestimmter Stoffe) im untersuchten Objekt durch Projektionen (jedoch nur endlich vieler) Halbräume erfolgt.

3.1.7 Satz: Charakterisierung projektiver Systeme

Sei $T \neq \emptyset$ eine Indexmenge. Eine Familie

$$(Q_{t_1, \dots, t_m})_{\substack{m \in \mathbb{N} \\ t_i \in T \\ t_i \neq t_j \neq i}}$$

von W-Maßen auf $\mathbb{R}^{n \cdot m}$ ist genau dann ein projektives System auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ über T (vgl. Bedingungen 1.5.1.1 und 1.5.1.2) falls gilt:

(i) Zu $t_1, \dots, t_m \in T$, $t_i \neq t_j \neq i$ und Permutation $\pi \in \text{Sym}(m)$ ist

$$\tilde{Q}_{t_1, \dots, t_m}(k_1, \dots, k_m) = \tilde{Q}_{t_{\pi(1)}, \dots, t_{\pi(m)}}(k_{\pi(1)}, \dots, k_{\pi(m)}) \quad , \quad k_1, \dots, k_m \in \mathbb{R}^n \quad (3.1.7.1)$$

(ii) Zu $t_1, \dots, t_m \in T$, $t_i \neq t_j \neq i$ gilt

$$\tilde{Q}_{t_1, \dots, t_m}(k_1, \dots, k_{m-1}, 0) = \tilde{Q}_{t_1, \dots, t_{m-1}}(k_1, \dots, k_{m-1}) \quad , \quad k_1, \dots, k_{m-1} \in \mathbb{R}^n \quad (3.1.7.2)$$

Beweis:

1. Zu Permutation $\pi \in \text{Sym}(m)$ definiere den linearen Operator

$$T_\pi : \mathbb{R}^{mn} \rightarrow \mathbb{R}^{mn} \quad , \quad T_\pi \underbrace{(x_1, \dots, x_m)}_{x \in \mathbb{R}^{mn}} := (x_{\pi(1)}, \dots, x_{\pi(m)}) \quad , \quad x_1, \dots, x_m \in \mathbb{R}^n$$

Dann ist

$$\langle T_\pi x, y \rangle = \sum_{i=1}^m \langle (T_\pi(x))_i, y_i \rangle = \sum_{i=1}^m \langle x_{\pi(i)}, y_i \rangle \stackrel{j:=\pi(i)}{=} \sum_{j=1}^n \langle x_j, y_{\pi^{-1}(j)} \rangle = \langle x, T_{\pi^{-1}} y \rangle$$

das heißt $T_\pi^\dagger = T_{\pi^{-1}}$. Somit:

$$\begin{aligned} (1.5.1.1) \Leftrightarrow Q_{t_1, \dots, t_m} &= Q_{t_{\pi(1)}, \dots, t_{\pi(m)}} \circ T_\pi \\ &\stackrel{(3.1.5.2)}{=} \tilde{Q}_{t_{\pi(1)}, \dots, t_{\pi(m)}} \circ (T_\pi^{-1})^\dagger = \tilde{Q}_{t_{\pi(1)}, \dots, t_{\pi(m)}} \circ T_\pi \end{aligned}$$

2. Definieren den linearen Operator

$$P_m : \mathbb{R}^{mn} \rightarrow \mathbb{R}^{(m-1)n} \quad , \quad P_m(x_1, \dots, x_m) := (x_1, \dots, x_{m-1}) \quad , \quad x_1, \dots, x_m \in \mathbb{R}^n$$

Dann ist

$$P_m^\dagger(y_1, \dots, y_{m-1}) = (y_1, \dots, y_{m-1}, 0) \quad , \quad y_1, \dots, y_{m-1} \in \mathbb{R}^n$$

so dass folgende Äquivalenz gilt:

$$\begin{aligned} (1.5.1.2) \Leftrightarrow Q_{t_1, \dots, t_m} \circ P_m^{-1} &= Q_{t_1, \dots, t_{m-1}} \\ &\stackrel{(3.1.5.2)}{=} \tilde{Q}_{t_1, \dots, t_{m-1}} \circ P_m^\dagger \end{aligned}$$

□

3.1.8 Theorem: Existenz stochastischer Prozesse

Sei $T \neq \emptyset$ eine Indexmenge und

$$(\varphi_{t_1, \dots, t_m})_{\substack{m \in \mathbb{N} \\ t_i \in T \\ t_i \neq t_j \neq i}}$$

eine Familie von Abbildungen $\varphi_{t_1, \dots, t_m} : \mathbb{R}^{mn} \rightarrow \mathbb{C}$ mit:

- (i) $\varphi_{t_1, \dots, t_m}(0) = 1$.
- (ii) $\varphi_{t_1, \dots, t_m}$ stetig.
- (iii) $\varphi_{t_1, \dots, t_m}$ nicht-negativ definit.
- (iv) $\varphi_{t_1, \dots, t_m} = \varphi_{t_{\pi(1)}, \dots, t_{\pi(m)}} \circ T_\pi$ für $\pi \in \text{Sym}(m)$.
- (v) $\varphi_{t_1, \dots, t_m}(k_1, \dots, k_{m-1}, 0) = \varphi_{t_1, \dots, t_{m-1}}(k_1, \dots, k_{m-1})$ für $k_1, \dots, k_{m-1} \in \mathbb{R}^n$.

Dann existiert ein \mathbb{R}^n -wertiger stochastischer Prozess $(X_t)_{t \in T}$ mit

$$\tilde{\mathcal{P}}_{t_1, \dots, t_m}^X(k_1, \dots, k_m) = \mathbb{E} \left\{ \exp \left[i \sum_{i=1}^m \langle k_i, X_{t_i} \rangle \right] \right\} = \varphi_{t_1, \dots, t_m}(k_1, \dots, k_m) \quad , \quad k_1, \dots, k_m \in \mathbb{R}^n$$

für $t_1, \dots, t_m \in T$ paarweise verschieden.

Beweis: Nach Bochner (3.1.4) existieren Wahrscheinlichkeitsmaße Q_{t_1, \dots, t_m} auf \mathbb{R}^{mn} mit

$$\tilde{Q}_{t_1, \dots, t_m} = \varphi_{t_1, \dots, t_m}$$

Die (Q_{t_1, \dots, t_m}) erfüllen dabei nach Voraussetzung die Bedingungen (3.1.7.1) und (3.1.7.2), sind also ein projektives System auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Nach Existenztheorem 1.5.4 existiert somit ein \mathbb{R}^n -wertiger stochastischer Prozess $(X_t)_{t \in T}$ mit $\mathcal{P}_{t_1, \dots, t_m}^X = Q_{t_1, \dots, t_m}$ bzw.

$$\tilde{\mathcal{P}}_{(X_{t_1}, \dots, X_{t_m})} = \varphi_{t_1, \dots, t_m}$$

□

3.2 Gaußsche Vektoren

3.2.1 Definition: Normalverteilung

Es sei $\sigma \geq 0$ und $\mu \in \mathbb{R}$. Dann heißt das Maß $\mathcal{N}_{\mu, \sigma^2}$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ definiert durch

$$\mathcal{N}_{\mu, \sigma^2}(B) := \frac{1}{\sqrt{2\pi}\sigma} \int_B e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

für $\sigma^2 > 0$ bzw.

$$\mathcal{N}_{\mu, 0}(B) := \delta_\mu(B) := \begin{cases} 1 & : \mu \in B \\ 0 & : \text{sonst} \end{cases}$$

Normalverteilung auf \mathbb{R} , mit Mittelwert μ und Varianz σ^2 . Eine Zufallsvariable $\xi : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow \mathbb{R}$ heißt *Normalverteilt* falls ihr Verteilungsgesetz die Normalverteilung ist, das heißt

$$\mathcal{P}(\{\xi \in B\}) = \mathcal{N}_{\mu, \sigma^2}(B) \quad , \quad B \in \mathcal{B}(\mathbb{R})$$

Bemerkung: Die Fourier-Transformierte von $\mathcal{N}_{\mu, \sigma^2}$ ist gegeben durch

$$\tilde{\mathcal{N}}_{\mu, \sigma^2}(k) = \exp \left[ik\mu - \frac{\sigma^2 k^2}{2} \right] \quad , \quad k \in \mathbb{R} \quad (3.2.1.1)$$

3.2.2 Definition: Gaußvektor

Ein Zufälliger Vektor $Z = (Z_1, \dots, Z_n) : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow \mathbb{R}^n$ heißt *normalverteilt* oder *Gaußvektor* falls $\forall k \in \mathbb{R}^n$ die Zufallsvariable

$$\langle Z, k \rangle : \Omega \rightarrow \mathbb{R} \quad , \quad \omega \mapsto \langle Z(\omega), k \rangle$$

Normalverteilt ist.

Bemerkungen:

- (i) Ist $Z = (Z_1, \dots, Z_n)$ Normalverteilt, so ist insbesondere jede Komponente Z_i Normalverteilt (setze $k := \mathbf{e}_i$).
- (ii) Die Angabe allein der Verteilungen der Z_i reicht allgemein nicht aus um Z eindeutig festzulegen. Insbesondere können Z_1, \dots, Z_n alle Normalverteilt und Z nicht normalverteilt sein.
- (iii) Da jeder einzelne Erwartungswert $\mathbb{E}|Z_i|^2 < \infty$ existiert, existiert auch der Erwartungswert

$$\mathbb{E}(Z) := (\mathbb{E}Z_1, \dots, \mathbb{E}Z_n)$$

- (iv) Sind ferner Z_1, \dots, Z_n unabhängig, so gilt

$$\widetilde{(Z_1, \dots, Z_n)}(k) \stackrel{\substack{\text{unabhängig} \\ (3.1.5.1)}}{=} \prod_{j=1}^n \tilde{Z}_j(k_j) = \exp \left[i \langle k, \mathbb{E}Z \rangle - \frac{1}{2} \sum_j \sigma_j^2 k_j^2 \right] \quad (3.2.2.1)$$

wobei

$$\sigma_j^2 := \mathbb{E}((Z_j - \mathbb{E}Z_j)^2) = \text{Var}(Z_j)$$

(v) Ist $Z = (Z_1, \dots, Z_n)$ Gaußsch und $\mu \in \mathbb{R}^n$, so ist auch $Z' := Z + \mu$ Gaußsch mit Erwartungswert

$$\mathbb{E}Z' = \mathbb{E}Z + \mu$$

(vi) Seien (Y_1, \dots, Y_n) , (Z_1, \dots, Z_m) von einander unabhängige Gaußvektoren mit Kovarianzmatrizen R_Y , R_Z . Dann ist auch

$$(Y_1, \dots, Y_n, Z_1, \dots, Z_m)$$

ein Gaußvektor mit Kovarianzmatrix

$$R_{(Y,Z)} = \left(\begin{array}{c|c} R_Y & 0 \\ \hline 0 & R_Z \end{array} \right)$$

3.2.3 Definition: Zentrierter Gaußvektor

Ein Gaußvektor $Z : \Omega \rightarrow \mathbb{R}^n$ heißt *zentriert* falls $\mathbb{E}(Z) = 0$ ist.

Bemerke:

(i) Zu beliebigen Gaußvektor $Z : \Omega \rightarrow \mathbb{R}^n$ ist

$$Z' := Z - \mathbb{E}(Z)$$

ein zentrierter Gaußvektor.

(ii) Z ist genau dann zentriert, falls für beliebige $k \in \mathbb{R}^n$ (oder äquivalent: beliebige $k \in \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$):

$$\underbrace{\mathbb{E} \langle Z, k \rangle}_{\langle \mathbb{E}Z, k \rangle} = 0$$

ist.

3.2.4 Definition: Standardnormalverteilter Vektor

Ein zufälliger Vektor $Z = (Z_1, \dots, Z_n) : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow \mathbb{R}^n$ heißt *standardnormalverteilt* falls Z_1, \dots, Z_n unabhängig und $\mathcal{N}_{0,1}$ -verteilt sind, das heißt insbesondere

$$\mathcal{P}(\{Z \in B\}) = \frac{1}{(2\pi)^{n/2}} \int_B e^{-\frac{\|x\|^2}{2}} dx$$

Bemerke:

(i) Ist $Z = (Z_1, \dots, Z_n)$ standardnormalverteilt, so ist $\langle Z, k \rangle \mathcal{N}_{0, \|k\|^2}$ -verteilt für beliebiges $k \in \mathbb{R}^n$, denn:

$$\langle Z, k \rangle = \sum_{j=1}^n k_j Z_j \stackrel{\text{unabhängig}}{\sim} \mathcal{N}_{0, \sum_{i=1}^n k_i^2} = \mathcal{N}_{0, \|k\|^2}, \quad a \in \mathbb{R}^n \quad (3.2.4.1)$$

(ii) Somit ist ersichtlich: Jeder standardnormalverteilte Zufallsvektor (Z_1, \dots, Z_n) ist normalverteilt, zentriert.

(iii) Nach Gl. (3.2.2.1) gilt für standardnormalverteilten (Z_1, \dots, Z_n) :

$$\tilde{Z}(k) = e^{-\frac{1}{2} \|k\|^2}, \quad k \in \mathbb{R}^n \quad (3.2.4.2)$$

3.2.5 Lemma: Existenz standardnormalverteilter Vektoren

Für beliebiges $n \in \mathbb{N}$ existiert ein standardnormalverteilter Vektor $(Z_1, \dots, Z_n) : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow \mathbb{R}^n$.

Beweis: Setzen

$$\Omega := \mathbb{R}^n, \quad \mathfrak{S} := \mathcal{B}(\mathbb{R}^n), \quad \mathcal{P} := \bigotimes_{i=1}^n \mathcal{N}_{0,1}$$

und dazu

$$Z : \Omega \rightarrow \mathbb{R}^n, \quad Z_i(x_1, \dots, x_n) := x_i, \quad i = 1, \dots, n$$

Dann ist Z standardnormalverteilt.

□

3.2.6 Satz: Charakterisierung zentrierter Gauß-Vektoren

Sei $Z = (Z_1, \dots, Z_n) : \Omega \rightarrow \mathbb{R}^n$ ein zufälliger Vektor. Dann sind äquivalent:

1. Z ist Gaußsch, zentriert.
2. Es existiert eine nicht-negativ-definite Matrix $R \in \mathbb{R}^{n \times n}$ ($R \geq 0$) mit

$$\tilde{Z}(k) = e^{-\frac{1}{2}\langle Rk, k \rangle}, \quad k \in \mathbb{R}^n \quad (3.2.6.1)$$

Gegebenfalls ist dies erfüllt durch die Kovarianzmatrix $R = (\mathbb{E}(Z_i Z_j))_{i,j=1}^n$.

3. Es existiert eine lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit

$$Z \stackrel{\text{d.}}{=} A(Y)$$

(d.h. $\mathcal{P}_Z = \mathcal{P}_{A(Y)}$) für irgendeinen standardnormalverteilten Zufallsvektor $Y : \Omega' \rightarrow \mathbb{R}^n$. Gegebenfalls ist dann $R = AA^\dagger$.

Beweis:

- 1 \rightarrow 2: Setzen $r_{ij} := \mathbb{E}(Z_i Z_j) = \text{Cov}(Z_i, Z_j)$ und $R := (r_{ij})_{i,j=1}^n$ (Kovarianzmatrix). Offensichtlich ist $R = R^\dagger$. Ferner gilt für $k \in \mathbb{R}^n$:

$$\langle Rk, k \rangle = \sum_{i,j=1}^n k_i k_j \mathbb{E}(Z_i Z_j) = \mathbb{E} \left[\sum_{i,j=1}^n k_i k_j Z_i Z_j \right] = \mathbb{E} \left| \sum_{i=1}^n k_i Z_i \right|^2 \geq 0$$

das heißt $R \geq 0$. Außerdem ist

$$\tilde{Z}(k) = \widetilde{\langle Z, k \rangle}(1) \stackrel{(3.2.1.1)}{=} e^{-\frac{\sigma^2}{2}}$$

mit

$$\sigma^2 = \mathbb{E} |\langle Z, k \rangle|^2 = \mathbb{E} \sum_{i,j=1}^n k_i k_j Z_i Z_j = \langle Rk, k \rangle$$

das heißt

$$\tilde{Z}(k) = e^{-\frac{1}{2}\langle Rk, k \rangle}$$

- 2 \rightarrow 3: Bekanntlich existiert zu $R \geq 0$ eine lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $R = AA^\dagger$. Zu standardnormalverteilten Vektor $Y : \Omega' \rightarrow \mathbb{R}^n$ (vgl. Existenzlemma 3.2.5) ist

$$\tilde{Y}(k) \stackrel{(3.1.5.1)}{=} \prod_{i=1}^n \tilde{Y}_i(k_i) \stackrel{(3.2.4.2)}{=} e^{-\frac{1}{2}\|k\|^2}, \quad k \in \mathbb{R}^n$$

Andererseits

$$\begin{aligned} \widetilde{A(Y)}(k) &= \tilde{Y}(A^\dagger k) = e^{-\frac{1}{2}\|A^\dagger k\|^2} = e^{-\frac{1}{2}\langle A^\dagger k, A^\dagger k \rangle} = e^{-\frac{1}{2}\langle AA^\dagger k, k \rangle} \\ &= e^{-\frac{1}{2}\langle Rk, k \rangle} = \tilde{Z}(k) \end{aligned}$$

Nach Bochner (3.1.4) sind somit $A(Y) \stackrel{\text{d.}}{=} Z$.

3 → 1: Aus $Z \stackrel{d.}{=} A(Y)$ folgt

$$\langle Z, k \rangle \stackrel{d.}{=} \langle A(Y), k \rangle = \langle Y, A^\dagger k \rangle \stackrel{(3.2.4.1)}{\sim} \mathcal{N}_{0, \|A^\dagger k\|^2}, \quad k \in \mathbb{R}^n$$

das heißt Z ist Gaußsch, zentriert.

□

Bemerkungen:

- (i) Aus dem Beweis wurde ersichtlich, dass die Kovarianzmatrix $R := (\mathbb{E}(Z_i Z_j))_{i,j=1}^n$ eines zentrierten Gaußvektors Z stets nicht-negativ definit ist. Ist sogar $R > 0$ (positiv definit), das heißt $\det(R) > 0$ so existiert die Inverse R^{-1} und es lässt sich zeigen:

$$\mathcal{P}(\{Z \in B\}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(R)}} \int_B e^{-\frac{1}{2} \langle R^{-1} x, x \rangle} dx, \quad B \in \mathcal{B}(\mathbb{R}^n)$$

Insbesondere ist die Zuordnung zwischen nicht-negativ definiten, symmetrischen Matrizen $R \in \mathbb{R}^{n \times n}$ und zentrierten Gaußvektoren (bzgl. Verteilung) eineindeutig (vgl. Darstellung 3.2.6.1), wobei die Existenz der Gaußvektoren aus dem Beweis 2 → 3 ersichtlich wurde. Man schreibt daher oft $X \sim \mathcal{N}_{0,R}$ und meint dass X Gaußsch, zentriert ist mit Kovarianzmatrix R .

- (ii) Aus Darstellung

$$\widetilde{Z}(k) = e^{-\frac{1}{2} \langle Rk, k \rangle}$$

für zentrierte Gaußvektoren Z folgt

$$\widetilde{\langle Z, k \rangle}(t) = \widetilde{Z}(tk) = e^{-\frac{t^2}{2} \langle Rk, k \rangle}$$

das heißt

$$\langle Z, k \rangle \sim \mathcal{N}_{0, \langle Rk, k \rangle}$$

(vgl. 3.2.1.1).

- (iii) Sei nun Z zentriert, Gaußverteilt mit Kovarianzmatrix $R = (\mathbb{E}(Z_i Z_j))_{i,j=1}^n$, dabei $\det(R) = 0$. Dazu sei $R = AA^\dagger$ für irgendeine $A \in \mathbb{R}^{n \times n}$, $E := A(\mathbb{R}^n)$ und $Z \stackrel{d.}{=} A(Y)$ für irgendeinen standardnormalverteilten Vektor Y . Dann ist $E \subset \mathbb{R}^n$ ein Unterraum von \mathbb{R}^n und

$$\mathcal{P}(\{Z \in E\}) = \mathcal{P}(\{A(Y) \in E\}) = 1$$

Z hat also fast-sicher Werte im (echten) Unterraum E mit $\dim(E) = \text{rank}(A) < n$.

Dabei ist

$$E = A(\mathbb{R}^n) = [\ker(A^\dagger)]^\perp$$

das heißt

$$x \in E^\perp \Leftrightarrow A^\dagger(x) = 0 \Leftrightarrow \underbrace{\langle A^\dagger x, A^\dagger x \rangle}_{\langle Rx, x \rangle} = 0$$

also

$$E = \{x : \langle Rx, x \rangle = 0\}^\perp$$

Beispiel: Für $Z = (Z_1, -Z_1)$ mit $Z_1 \sim \mathcal{N}_{0,1}$ ist

$$R = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

bzw.

$$\langle Rx, x \rangle = (x_1 - x_2)^2 = 0 \Leftrightarrow x_1 = x_2$$

das heißt

$$E = \{(y_1, y_2) : y_1 = -y_2\}$$

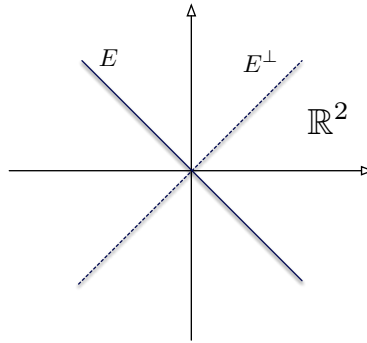


Abbildung 3.2: Zur Konzentration zentrierter Gaußvektoren. Gaußprozess $(Z_1, -Z_1)$ hier konzentriert auf Unterraum E .

- (iv) Sei $Z = (Z_1, \dots, Z_n)$ Gaußsch, zentriert, dazu Kovarianzmatrix R . Dann gilt:
 Z_1, \dots, Z_n sind unabhängig $\Leftrightarrow \mathbb{E}(Z_i Z_j) = 0 \quad \forall i \neq j$ (d.h. R diagonal).

Tatsächlich: Richtung \Rightarrow ist trivial, da unabhängige Zufallsvariablen stets unkorreliert sind. Sind umgekehrt die Z_1, \dots, Z_n unkorreliert, das heißt

$$R = \begin{pmatrix} r_{11} & \dots & 0 \\ & \ddots & \\ 0 & \dots & r_{nn} \end{pmatrix}$$

so existiert zu

$$A := \begin{pmatrix} \sqrt{r_{11}} & \dots & 0 \\ & \ddots & \\ 0 & \dots & \sqrt{r_{nn}} \end{pmatrix} \rightsquigarrow R = AA^\dagger$$

ein standardnormalverteilter Vektor Y mit

$$(Z_1, \dots, Z_n) = (\sqrt{r_{11}}Y_1, \dots, \sqrt{r_{nn}}Y_n)$$

(vgl. Beweis des obigen Satzes).

- (v) Mit analogen Überlegungen kann sogar allgemeiner gesagt werden: \mathbb{R}^n -Gaußvektoren Z_1, \dots, Z_m sind genau dann unkorreliert, das heißt $\underbrace{\text{Cov}(Z_i, Z_k)}_{\in \mathbb{R}^{n \times n}} = 0 \quad \forall i, k$, wenn sie unabhängig sind.
- (vi) Ist $Z = (Z_1, \dots, Z_n)$ Gaußsch, so sind insbesondere alle Z_i Normalverteilt. Die Umkehrung gilt im allgemeinen nicht! Sind jedoch Z_1, \dots, Z_n noch zusätzlich unabhängig, so ist Z Gaußsch.

3.2.7 Korollar für allgemeine Gaußvektoren

Sei $Z = (Z_1, \dots, Z_n) : \Omega \rightarrow \mathbb{R}^n$ ein zufälliger Vektor. Dann sind äquivalent:

1. Z ist Gaußsch.
2. Es existiert eine nicht-negativ definite, symmetrische Matrix $R \in \mathbb{R}^{n \times n}$ und Vektor $\mu \in \mathbb{R}^n$ mit

$$\tilde{Z}(k) = e^{-\frac{1}{2}\langle Rk, k \rangle + i\langle k, \mu \rangle}, \quad k \in \mathbb{R}^n \quad (3.2.7.1)$$

Gegebenfalls ist dann

$$R = (\text{Cov}(Z_i, Z_j))_{i,j=1}^n$$

genau die Kovarianzmatrix von Z und $\mu = \mathbb{E}(Z)$.

3. Es existiert eine affine Abbildung $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit

$$Z \stackrel{\text{d.}}{=} \mathcal{A}(Y)$$

für irgendeinen standardnormalverteilten Zufallsvektor $Y : \Omega' \rightarrow \mathbb{R}^n$.

Beweis:

1 \rightarrow 2: Da $Z' := Z - \mathbb{E}Z$ zentriert ist, ist nach 3.2.6:

$$R = (\mathbb{E} [(Z_i - \mathbb{E}Z_i)(Z_j - \mathbb{E}Z_j)])_{i,j=1}^n = (\mathbb{E} [Z'_i Z'_j])_{i,j=1}^n$$

nicht-negativ definit, symmetrisch und erfüllt

$$\tilde{Z}(k) = \mathcal{F}\{(Z - \mathbb{E}Z) + \mathbb{E}Z\}(k) \stackrel{(3.2.6)}{=} e^{-\frac{1}{2}\langle Rk, k \rangle} \cdot e^{i\langle k, \mathbb{E}Z \rangle}$$

2 \rightarrow 3: Der Vektor $Z' := Z - \mu$ erfüllt die Darstellung

$$\tilde{Z}'(k) = e^{-\frac{1}{2}\langle Rk, k \rangle}, \quad k \in \mathbb{R}^n$$

Nach 3.2.6 existiert ein standardnormalverteilter Vektor $Y : \Omega' \rightarrow \mathbb{R}^n$ und lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit

$$Z' \stackrel{\text{d.}}{=} L(Y)$$

also

$$Z = Z' + \mu \stackrel{\text{d.}}{=} \underbrace{L(Y) + \mu}_{\mathcal{A}(Y)}$$

3 \rightarrow 1: Sei $\mathcal{A}(x) = L(x) + \mu$ für irgendeine lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ und $\mu \in \mathbb{R}^n$. Dann erfüllt der Vektor $Z' = Z - \mu$:

$$Z' \stackrel{\text{d.}}{=} L(Y)$$

Somit ist Z' Gaußsch, zentriert, also

$$Z = Z' + \mu$$

Gaußsch.

□

Bemerkungen: Nach Korollar 3.2.7, ist ein Gaußvektor eindeutig durch seine Kovarianzmatrix R und Erwartungswert $\mathbb{E}(X)$ beschrieben. Umgekehrt, entspricht jedem Vektor $\mu \in \mathbb{R}^n$ und nicht-negativ definiten, symmetrischer Matrix $R \in \mathbb{R}^{n \times n}$ ein Gaußvektor X mit R als Kovarianzmatrix und Erwartungswert $\mathbb{E}(X) = \mu$.

3.2.8 Lemma: Rotationsinvarianz Gaußscher Vektoren

Sei $\vartheta \in [0, 2\pi]$ beliebiger Winkel, dazu

$$\mathcal{R}_\vartheta : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}, \quad \mathcal{R}_\vartheta(x, y) := (x \cos \vartheta + y \sin \vartheta, -x \sin \vartheta + y \cos \vartheta), \quad x, y \in \mathbb{R}^n \quad (3.2.8.1)$$

Seien $X, Y : \Omega \rightarrow \mathbb{R}^n$ unabhängige, identisch verteilte, zentrierte Gaußvektoren. Dann gilt:

$$\mathcal{R}_\vartheta(X, Y) \stackrel{\text{d.}}{=} (X, Y)$$

Beweis: Sei R die Kovarianzmatrix von X bzw. Y . Zum einen ist

$$\widetilde{(X, Y)}(a, b) = \widetilde{X}(a) \cdot \widetilde{Y}(b) \stackrel{(3.2.6)}{=} e^{-\frac{1}{2}\langle Ra, a \rangle - \frac{1}{2}\langle Rb, b \rangle}, \quad a, b \in \mathbb{R}^n$$

Andererseits ist

$$\mathcal{R}_\vartheta^\dagger(a, b) = (a \cos \vartheta - b \sin \vartheta, a \sin \vartheta + b \cos \vartheta)$$

und somit

$$\begin{aligned} \mathcal{F}\{\mathcal{R}_\vartheta(X, Y)\} &= \widetilde{(X, Y)}\left(\mathcal{R}_\vartheta^\dagger(a, b)\right) \\ &= \exp\left[-\frac{1}{2}\left(\cos^2 \vartheta \langle Ra, a \rangle - 2 \cos \vartheta \sin \vartheta \langle Ra, b \rangle + \sin^2 \vartheta \langle Rb, b \rangle + \sin^2 \vartheta \langle Ra, a \rangle + 2 \cos \vartheta \sin \vartheta \langle Ra, b \rangle + \cos^2 \vartheta \langle Rb, b \rangle\right)\right] \\ &= \exp\left[-\frac{1}{2}\langle Ra, a \rangle - \frac{1}{2}\langle Rb, b \rangle\right] = \widetilde{(X, Y)}(a, b) \end{aligned}$$

Nach Bochner 3.1.4 ist somit $(X, Y) \stackrel{\text{d.}}{=} \mathcal{R}_\vartheta(X, Y)$.

□

Spezialfall: Sind $X, Y : \Omega \rightarrow \mathbb{R}^n$ unabhängige, identisch verteilte, zentrierte Gaußvektoren, so gilt:

$$(X, Y) \stackrel{\text{d.}}{=} \frac{1}{\sqrt{2}}(X - Y, X + Y) \quad \Bigg| \quad \vartheta := -\frac{\pi}{4}$$

das heißt

$$\mathcal{P}(\{X \in A, Y \in B\}) = \mathcal{P}\left(\left\{\frac{X - Y}{\sqrt{2}} \in A, \frac{X + Y}{\sqrt{2}} \in B\right\}\right)$$

3.2.9 Lemma: Unitäre Transformationen von Gaußvektoren

Seien Z_1, \dots, Z_n unabhängig, identisch, zentriert Normalverteilt. Dann gilt für jede unitäre Transformation $U : \mathbb{R}^n \rightarrow \mathbb{R}^n$:

$$U(Z) \stackrel{\text{d.}}{=} Z$$

Beweis: Sei $\sigma^2 := \text{Var}(Z_i)$. Es genügt zu zeigen $\widetilde{U(Z)} = \widetilde{Z}$. Tatsächlich ist

$$\widetilde{U(Z)}(k) = \widetilde{Z}(U^\dagger k) \stackrel{(3.2.2.1)}{=} e^{-\frac{\sigma^2}{2}\|U^\dagger k\|^2} \stackrel{U \text{ unitär}}{=} e^{-\frac{\sigma^2}{2}\|k\|^2} = \widetilde{Z}(k)$$

□

3.3 Gaußprozesse

3.3.1 Definition: Gaußprozess

Sei $T \neq \emptyset$ eine beliebige Indexmenge. Ein \mathbb{R}^n -wertiger stochastischer Prozess $(X_t)_{t \in T}$ heißt *Gaußprozess*, wenn für alle $t_1, \dots, t_n \in T$ der Vektor $(X_{t_1}, \dots, X_{t_n})$ normalverteilt ist. Gegebenfalls heißt

$$R_X : T \times T \rightarrow \mathbb{R}^{n \times n}, \quad R_X(t, s) := \mathbb{E}[(X_t - \mathbb{E}X_t)^T \cdot (X_s - \mathbb{E}X_s)] = (\text{Cov}(X_t^i, X_s^j))_{i,j=1}^n$$

Kovarianzfunktion von X . Gilt stets $\mathbb{E}(X_t) = 0$, so heißt der Prozess *zentriert*.

Bemerkungen:

- (i) Ist $T = \{t_1, \dots, t_m\}$ so entspricht ein Gaußprozess genau einem Gaußvektor.
(ii) Ein Prozess $(X_t)_{t \in T}$ ist genau dann Gaußsch, wenn $\forall t_1, \dots, t_m \in T$, $a_1, \dots, a_m \in \mathbb{R}^n$ die Zufallsvariable

$$\sum_{i=1}^m \langle a_i, X_{t_i} \rangle$$

normalverteilt ist.

- (iii) Um Nachzuweisen dass $(X_t)_{t \in T}$ Gaußsch ist, genügt es zu zeigen dass $(X_{t_1}, \dots, X_{t_m})$ für $t_i \neq t_j \neq i$ Gaußsch sind. Dies wird ersichtlich wenn man sich klarmacht, dass für Gaußschen $(X_{t_1}, \dots, X_{t_m})$ auch $(X_{t_1}, \dots, X_{t_m}, X_{t_m})$ Gaußsch ist, denn für $a_i \in \mathbb{R}^n$ ist

$$\sum_{i=1}^{m+1} \langle a_i, X_{t_i} \rangle \stackrel{X_{t_{m+1}} := X_{t_m}}{=} \underbrace{\sum_{i=1}^{m-1} \langle a_i, X_{t_i} \rangle + \langle a_m + a_{m+1}, X_{t_m} \rangle}_{\text{Normalverteilt}}$$

(analog auch für mehrere *Duplikate*).

- (iv) Ist (X_t) Gaußsch, so ist insbesondere jedes X_t normalverteilt, das heißt insbesondere $\mathbb{E} \|X_t\|^2 < \infty$. Entsprechend ist dann

$$X'_t := X_t - \mathbb{E}(X_t), \quad t \in T$$

ein zentrierter Gaußprozess. Wir betrachten o.B.d.A im folgenden hauptsächlich zentrierte Gaußprozesse.

- (v) Die endlich-dimensionalen Verteilungen von X werden durch die Fourier-Transformationen

$$\tilde{\mathcal{P}}_{t_1, \dots, t_m}^X(k) \stackrel{(3.2.7.1)}{=} \exp \left[-\frac{1}{2} \langle Rk, k \rangle + i \langle k, \mathbb{E}(X_{t_1}, \dots, X_{t_m}) \rangle \right], \quad k \in \mathbb{R}^{nm}$$

beschrieben, wobei

$$R = \begin{pmatrix} \text{Cov}(X_{t_1}^1 X_{t_1}^1) & \text{Cov}(X_{t_1}^1 X_{t_1}^2) & \dots & \text{Cov}(X_{t_1}^1 X_{t_1}^n) & \text{Cov}(X_{t_1}^1 X_{t_2}^1) & \dots & \text{Cov}(X_{t_1}^1 X_{t_m}^n) \\ \text{Cov}(X_{t_1}^2 X_{t_1}^1) & \text{Cov}(X_{t_1}^2 X_{t_1}^2) & \dots & \text{Cov}(X_{t_1}^2 X_{t_1}^n) & \text{Cov}(X_{t_1}^2 X_{t_2}^1) & \dots & \text{Cov}(X_{t_1}^2 X_{t_m}^n) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \text{Cov}(X_{t_1}^n X_{t_1}^1) & \text{Cov}(X_{t_1}^n X_{t_1}^2) & \dots & \text{Cov}(X_{t_1}^n X_{t_1}^n) & \text{Cov}(X_{t_1}^n X_{t_2}^1) & \dots & \text{Cov}(X_{t_1}^n X_{t_m}^n) \\ \text{Cov}(X_{t_2}^1 X_{t_1}^1) & \text{Cov}(X_{t_2}^1 X_{t_1}^2) & \dots & \text{Cov}(X_{t_2}^1 X_{t_1}^n) & \text{Cov}(X_{t_2}^1 X_{t_2}^1) & \dots & \text{Cov}(X_{t_2}^1 X_{t_m}^n) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \text{Cov}(X_{t_m}^n X_{t_1}^1) & \text{Cov}(X_{t_m}^n X_{t_1}^2) & \dots & \text{Cov}(X_{t_m}^n X_{t_1}^n) & \text{Cov}(X_{t_m}^n X_{t_2}^1) & \dots & \text{Cov}(X_{t_m}^n X_{t_m}^n) \end{pmatrix}$$

$$= \begin{pmatrix} R_X(t_1, t_1) & R_X(t_1, t_2) & \dots & R_X(t_1, t_m) \\ R_X(t_2, t_1) & R_X(t_2, t_2) & \dots & R_X(t_2, t_m) \\ \vdots & \vdots & & \vdots \\ R_X(t_m, t_1) & R_X(t_m, t_2) & \dots & R_X(t_m, t_m) \end{pmatrix} = (R_X(t_i, t_j))_{i,j=1}^m$$

genau die Kovarianzmatrix von $(X_{t_1}, \dots, X_{t_m}) : \Omega \rightarrow \mathbb{R}^{nm}$ ist. Somit ist ersichtlich: Die Kovarianzfunktion R_X beschreibt zusammen mit den Erwartungswerten $\mathbb{E}(X_t)$, $t \in T$ komplett die endlich dimensionalen Verteilungen und bestimmen somit (bis auf Äquivalenz) den Gaußprozess.

3.3.2 Satz: Charakterisierung von Gaußprozessen durch Kovarianzfunktion

Zwei \mathbb{R}^n -wertige Gaußprozesse $(X_t)_{t \in T}$, $(Y_t)_{t \in T}$ sind genau dann äquivalent wenn sie gleiche Kovarianzfunktionen $R_X = R_Y$ und gleiche Erwartungswerte $\mathbb{E}(X_t) = \mathbb{E}(Y_t)$ haben.

Beweis: Siehe Bemerkung (v) in Definition 3.3.1.

3.3.3 Satz über die Kovarianzfunktion

Ist (X_t) ein \mathbb{R}^n -wertiger Gaußprozess, so erfüllt die Kovarianzfunktion $R_X : T \times T \rightarrow \mathbb{R}^{n \times n}$:

1. Quasi-Symmetrie: $(R_X(t, s)) = (R_X(s, t))^T$, $t, s \in T$
2. Nicht-negativ-Definitheit, das heißt für $t_1, \dots, t_m \in T$ ist $(R_X(t_i, t_j))_{i,j=1}^m$ nicht-negativ definit²⁰. Anders formuliert:

$$\sum_{i,j=1}^m \langle R_X(t_i, t_j) a_j, a_i \rangle \geq 0 \quad \forall a_1, \dots, a_m \in \mathbb{R}^n, t_1, \dots, t_m \in T \quad (3.3.3.1)$$

3. X_t, X_s sind unabhängig genau dann wenn $R_X(t, s) = 0$.

Beweis:

1. Trivial.
2. Da $(X_{t_1}, \dots, X_{t_m})$ nach Voraussetzung normalverteilt ist, ist seine Kovarianzmatrix $R = (R_X(t_i, t_j))_{i,j=1}^m$ nicht-negativ definit.
3. Es ist $R_X(t, s) = 0 \Leftrightarrow \text{Cov}(X_t^i, X_s^j) = 0$. Da für Gaußvektoren Unkorreliertheit äquivalent zu Unabhängigkeit ist, folgt die Behauptung.

□

3.3.4 Theorem: Existenz von Gaußprozessen

Sei $T \neq \emptyset$ eine Indexmenge, $(\mu_t)_{t \in T}$ eine Familie von Vektoren $\mu_t \in \mathbb{R}^n$ und $R : T \times T \rightarrow \mathbb{R}^{n \times n}$ mit folgenden Eigenschaften:

- (i) $(R(t, s))^T = (R(s, t))$
- (ii) R ist nicht-negativ definit (vgl. 3.3.3.1).

Dann existiert ein bis auf Äquivalenz eindeutig bestimmter \mathbb{R}^n -wertiger Gaußprozess $(X_t)_{t \in T}$ mit

$$R(t, s) = R_X(t, s) \stackrel{\text{def.}}{=} (\text{Cov}(X_t^i, X_s^j))_{i,j=1}^n, \quad \mathbb{E}(X_t) = \mu_t, \quad t, s \in T$$

Beweis: Sei zunächst $\mu_t = 0$, $t \in T$. Zu vorgegebenen $t_1, \dots, t_m \in T$, $t_i \neq t_j \neq i$ setzen

$$\tilde{R}_{t_1, \dots, t_m} := (R(t_i, t_j))_{i,j=1}^m \in \mathbb{R}^{nm \times nm}$$

(vgl. Bemerkung (v) in Definition 3.3.1). Dann sind $\tilde{R}_{t_1, \dots, t_m}$ nicht-negativ definit und symmetrisch, entsprechen also Normalverteilungen $Q_{t_1, \dots, t_m} \sim \mathcal{N}_{0, \tilde{R}_{t_1, \dots, t_m}}$. Dabei gilt für $k_1, \dots, k_m \in \mathbb{R}^n$ und Permutation $\pi \in S_m$:

$$\begin{aligned} \tilde{Q}_{t_{\pi(1)}, \dots, t_{\pi(m)}}(k_{\pi(1)}, \dots, k_{\pi(m)}) &\stackrel{(3.2.7.1)}{=} \exp \left[-\frac{1}{2} \sum_{i,j=1}^m \langle R(t_{\pi(i)}, t_{\pi(j)}) k_{\pi(j)}, k_{\pi(i)} \rangle \right] = \exp \left[-\frac{1}{2} \sum_{i,j=1}^m \langle R(t_i, t_j) k_j, k_i \rangle \right] \\ &= \tilde{Q}_{t_1, \dots, t_m}(k_1, \dots, k_m) \end{aligned}$$

(Eigenschaft 3.1.7.1). Ferner:

$$\begin{aligned} \tilde{Q}_{t_1, \dots, t_m}(k_1, \dots, k_{m-1}, 0) &\stackrel{k_m := 0}{=} \exp \left[-\frac{1}{2} \sum_{i,j=1}^m \langle R(t_i, t_j) k_j, k_i \rangle \right] = \exp \left[-\frac{1}{2} \sum_{i,j=1}^{m-1} \langle R(t_i, t_j) k_j, k_i \rangle \right] \\ &= \tilde{Q}_{t_1, \dots, t_{m-1}}(k_1, \dots, k_{m-1}) \end{aligned}$$

²⁰Nicht-negativ Definitheit der Kovarianzfunktion R_X ist nach Darstellung in Bemerkung (v), Definition 3.3.1 äquivalent zur nicht-negativ Definitheit der Kovarianzmatrix jedes Vektors $(X_{t_1}, \dots, X_{t_m})$.

(Eigenschaft 3.1.7.2). Nach Kolmogorov existiert also ein stochastischer Prozess $(X_t)_{t \in T}$ mit

$$\mathcal{P}_{t_1, \dots, t_m}^X = Q_{t_1, \dots, t_m}, \quad t_i \neq t_j \neq i$$

das heißt X ist Gaußsch (vgl. Bemerkung (iii) in Definition 3.3.1), zentriert. Für $t \neq s \in T$ besitzt insbesondere (X_t, X_s) die Kovarianzmatrix $\begin{pmatrix} R(t, t) & R(t, s) \\ R(s, t) & R(s, s) \end{pmatrix}$, das heißt

$$R_X(t, s) = (\text{Cov}(X_t^i, X_s^j))_{i,j=1}^n = R(t, s)$$

und

$$R_X(t, t) = (\text{Cov}(X_t^i, X_t^j))_{i,j=1}^n$$

(hier wurde angenommen $|T| \geq 2$, andernfalls ist das Theorem trivial). Im allgemeinen Fall $\mu \in \mathbb{R}^n$ erfüllt $X' := X + \mu$ die Behauptungen. Eindeutigkeit folgt nach Bemerkung (v) in Definition 3.3.1

□

3.3.5 Der Wiener-Prozess als Gauß-Prozess

Für $t, s \geq 0$ definieren $R(t, s) := \min\{t, s\}$. Dann ist zum einen $R : [0, \infty)^2 \rightarrow \mathbb{R}$ symmetrisch, und zum anderen nicht-negativ definit, denn

$$\sum_{i,j=1}^m R(t_i, t_j) a_i a_j = \sum_{i,j=1}^m a_i a_j \int_0^{\overbrace{\min\{t_i, t_j\}}^{\infty}} 1_{[0, t_i]}(x) \cdot 1_{[0, t_j]}(x) dx = \int_0^{\infty} \left| \sum_{i,j=1}^m a_i 1_{[0, t_i]} \right|^2 dx \geq 0$$

Dann existiert ein zentrierter Gaußprozess $(W_t)_{t \geq 0}$ mit Kovarianzfunktion R . Wegen $\mathbb{E}|W_0|^2 = \min\{0, 0\} = 0$ ist $W_0 = 0$ fast sicher.

Definieren

$$A_m : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad A(x_1, \dots, x_m) := (x_1, x_2 - x_1, \dots, x_m - x_{m-1}),$$

Wegen

$$\mathcal{F}\{(W_{t_1}, \dots, W_{t_m})\}(k) = \exp \left[-\frac{1}{2} \sum_{i,j=1}^m k_i k_j \min\{t_i, t_j\} \right]$$

für $0 \leq t_1 < \dots < t_m$, folgt dann

$$\begin{aligned} \mathcal{F}\{(A_m(W_{t_1}, \dots, W_{t_m}))\}(k) &= \mathcal{F}\{(W_{t_1}, \dots, W_{t_m})\}(A^\dagger k) \stackrel{a_{m+1} := 0}{=} \exp \left[-\frac{1}{2} \int_0^{\infty} \left| \langle A_m^\dagger k, (1_{[0, t_1]}(x), \dots, 1_{[0, t_m]}(x)) \rangle \right|^2 dx \right] \\ &= \exp \left[-\frac{1}{2} \int_0^{\infty} \left| \langle k, A(1_{[0, t_1]}(x), \dots, 1_{[0, t_m]}(x)) \rangle \right|^2 dx \right] \\ &= \exp \left[-\frac{1}{2} \int_0^{\infty} \left| \langle k, 1_{[0, t_1]}(x), \underbrace{1_{[0, t_2]} - 1_{[0, t_1]}}_{1_{(t_1, t_2]}}, \dots, \underbrace{1_{[0, t_m]} - 1_{[0, t_{m-1}]}}_{1_{(t_{m-1}, t_m]}} \rangle \right|^2 dx \right] \\ &\stackrel{t_0 := 0}{=} \exp \left[-\frac{1}{2} \int_0^{\infty} \underbrace{\left| \sum_{i=1}^m a_i 1_{(t_{i-1}, t_i]}(x) \right|^2}_{\text{Treppenfunktion}} dx \right] = \exp \left[-\frac{1}{2} \int_0^{\infty} \sum_{i=1}^m \left| a_i 1_{(t_{i-1}, t_i]}(x) \right|^2 dx \right] \\ &= \exp \left[-\frac{1}{2} \sum_{i=1}^m a_i^2 (t_i - t_{i-1}) \right] \end{aligned}$$

Somit ist

$$A_m(W_{t_1}, \dots, W_{t_m}) = (W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}})$$

Gaußsch, zentriert mit Kovarianzmatrix

$$\text{diag}(t_1, t_2 - t_1, \dots, t_m - t_{m-1})$$

Insbesondere sind die Zuwächse $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}}$ unabhängig mit

$$W_{t_j} - W_{t_{j-1}} \sim \mathcal{N}_{0, t_j - t_{j-1}}$$

das heißt W ist tatsächlich der Wiener-Prozess.

Zusammenfassung: Der Wiener-Prozess (W_t) ist der eindeutig bestimmte, zentrierte Gaußsche Prozess mit Kovarianzen

$$\text{Cov}(W_t, W_s) = \min\{t, s\} \quad .$$

Eigenschaften:

(i) Zu $c > 0$ setzen $X_t := W_{ct}$. Dann ist wegen

$$\sum_{i=1}^m a_i X_{t_j} = \sum_{i=1}^m a_i W_{ct_i} \quad , \quad a_i \in \mathbb{R}$$

normalverteilt
 $s_i := ct_i$

X ein zentrierter Gaußprozess, mit Kovarianzfunktion

$$R_X(t, s) = \mathbb{E}(X_t X_s) = \underbrace{\mathbb{E}(W_{ct} W_{cs})}_{R_W(ct, cs)} = c \cdot \min\{t, s\}$$

Setzt man andererseits $Y_t := \sqrt{c} \cdot W_t$ so erhält man ebenfalls einen zentrierten Gaußprozess, der erfüllt

$$R_Y(t, s) = c \cdot \min\{t, s\}$$

Nach 3.3.2 sind somit die beiden Prozesse äquivalent:

$$(W_{ct})_{t \geq 0} \cong (\sqrt{c} \cdot W_t)_{t \geq 0}$$

das heißt, *Zeittransformationen* entsprechen *Ortstransformationen*.

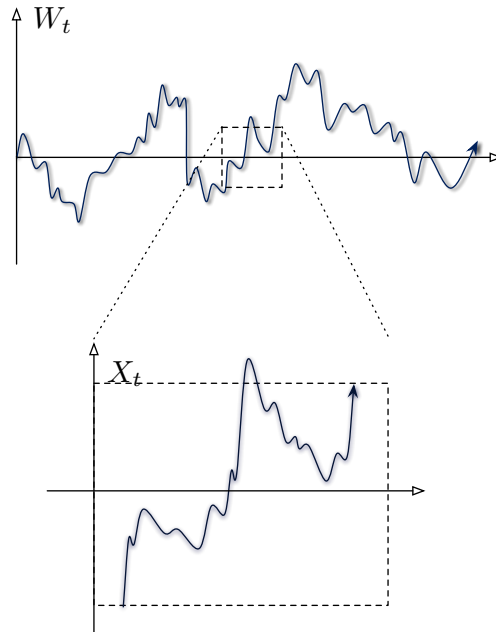


Abbildung 3.3: Zeitzoomen und gleichzeitige Ortsvergrößerung, führt dazu dass $(\frac{1}{\sqrt{c}}W_{ct})$ ähnlich zu (W_t) erscheint (hier $c \approx \frac{1}{6}$).

- (ii) Setzen $X_t := tW(\frac{1}{t})$, $t > 0$ und $X_0 := 0$. Dann ist $(X_t)_{t \geq 0}$ ein zentrierter Gaußprozess, mit Kovarianzfunktion

$$\mathbb{E}(X_t X_s) = t \cdot s \min \left\{ \frac{1}{t}, \frac{1}{s} \right\} = \min \{t, s\}$$

Fängt man also bei $t = \infty$ an und lässt den Prozess mit wachsender Vergrößerung *rückwärts* laufen, so erscheint dieser ähnlich zum ursprünglichen Wiener-Prozess.

- (iii) Setzt man ähnlich $X_t := W_{1-t} - W_1$, $0 \leq t \leq 1$ so ist X_t ebenfalls zentriert, Gaußsch und erfüllt

$$\mathbb{E}(X_t X_s) = \mathbb{E}[(W_{1-t} - W_1)(W_{1-s} - W_1)] = \min\{(1-t), (1-s)\} - (1-t) - (1-s) + 1 = \min\{t, s\}$$

Rückwärtsbeobachtung resultiert sozusagen auch in einem Wienerprozess.

Bemerkungen:

- (i) Es existiert eine Version mit Pfaden, die stetig aber in fast keinem Punkt differenzierbar sind.
- (ii) Die Pfade des Wiener-Prozesses sind fast-sicher α -Hölder²¹ für alle $\alpha < \frac{1}{2}$, aber nicht $\frac{1}{2}$ -Hölder.
- (iii) In jedem beliebigen Intervall $[a, b]$ haben fast alle Pfade unendliche Variation²²

$$\mathcal{P} \left(\left\{ \omega : \bigvee_a^b W(\omega) = \infty \right\} \right) = 1$$

²¹Eine Funktion $f : M \rightarrow N$ zwischen zwei metrische Räumen M, N heißt α -Hölder (Hölderstetig zum Exponenten α) \Leftrightarrow

$$\exists C, \alpha \geq 0 : d(f(x), f(y)) \leq C \cdot d(x, y)^\alpha \quad \forall x, y \in M$$

²²Eine Abbildung $f : [a, b] \rightarrow M$ vom Intervall $[a, b]$ in einem metrischen Raum M besitzt die *Variation*

$$\bigvee_a^b f := \sup \left\{ \sum_{i=1}^n d(f(t_i), f(t_{i-1})) : a \leq t_0 < t_1 < \dots < t_n = b \right\}$$

3.3.6 Die Brownsche Brücke

Zu $T = [0, 1]$ ist

$$R(t, s) := \min\{t, s\} - ts, \quad t, s \in T$$

symmetrisch, nicht-negativ definit. Der zur Kovarianzfunktion R gehörige, zentrierte Gaußsche Prozess ist (bis auf Äquivalenz) gegeben durch

$$B_t := W_t - t \cdot W_1, \quad 0 \leq t \leq 1 \quad (3.3.6.1)$$

Tatsächlich ist $\mathbb{E}(B_t) = 0$ und somit

$$\text{Cov}(B_t, B_s) = \mathbb{E}(B_t B_s) = \min\{t, s\} - t\mathbb{E}(W_s W_1) - s\mathbb{E}(W_t W_1) + ts\mathbb{E}(W_1 W_1) = \min\{t, s\} - ts$$

Der Prozess $(B_t)_{t \in T}$ heißt *Brownsche Brücke*. Wegen $\mathbb{E}|B_1|^2 = \min\{1, 1\} - 1 \cdot 1 = 0$ ist $B_1 = 0$ fast sicher.

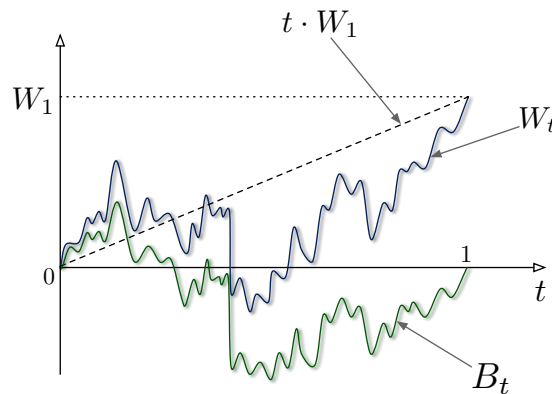


Abbildung 3.4: Typischer Pfad der Brownschen Brücke im Vergleich zum Wiener-Prozess.

3.3.7 Zufällige Gaußmaße

Sei (M, \mathcal{M}, μ) beliebiger Maßraum und $\mathcal{A} \subset \mathcal{M}$ (nicht unbedingt σ -algebra) mit $\mu|_{\mathcal{A}} < \infty$. Dann setzen für $A, B \in \mathcal{A}$:

$$R(A, B) := \mu(A \cap B)$$

Dann ist R symmetrisch und für $\alpha \in \mathbb{R}$ gilt

$$\sum_{i,j=1}^n a_i a_j R(A_i, A_j) = \sum_{i,j=1}^n a_i a_j \int_M 1_{A_i}(x) 1_{A_j}(x) \mu(dx) = \int_M \left| \sum_{j=1}^n a_j 1_{A_j}(x) \right|^2 \mu(dx) \geq 0$$

das heißt R ist nicht-negativ definit. Somit existiert ein zentrierter \mathbb{R} -wertiger Gaußprozess

$$(\Gamma_A)_{A \in \mathcal{A}}, \quad \Gamma_A : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow \mathbb{R}$$

mit

$$\mathbb{E}(\Gamma_A \Gamma_B) = \mu(A \cap B)$$

Insbesondere $\mathbb{E}(\Gamma_A^2) = \mu(A)$, das heißt $\Gamma_A \in L_2(\Omega, \mathcal{P})$. Man nennt $\Gamma : \mathcal{A} \times \Omega \rightarrow [0, \infty)$ *zufälliges Gaußsches Maß*, μ *Kontrollmaß* von Γ .

Eigenschaften:

1. Zu paarweise disjunkten $A_1, \dots, A_n \in \mathcal{M}$ mit $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$ sind $\Gamma_{A_1}, \dots, \Gamma_{A_n}$ unabhängig und

$$\sum_{j=1}^n \Gamma_{A_j} = \Gamma_{\bigcup_{j=1}^n A_j}$$

fast sicher.

Beweis: Per Konstruktion ist $(\Gamma_{A_1}, \dots, \Gamma_{A_n})$ zentriert, Gaußsch mit Kovarianzmatrix

$$\left(\mathbb{E}(\Gamma_{A_i} \Gamma_{A_j}) \right)_{i,j=1}^n = (\mu(A_i \cap A_j))_{i,j=1}^n = \text{diag}(\mu(A_1), \dots, \mu(A_n))$$

so dass nach Bemerkung (iv) in 3.2.6 die 1. Behauptung folgt.

Seien nun $A, B \in \mathcal{M}$ disjunkt mit $A \cup B \in \mathcal{M}$, dann besitzt $X := (\Gamma_A, \Gamma_B, \Gamma_{A \cup B})$ die Kovarianzmatrix

$$\begin{pmatrix} \mu(A) & 0 & \mu(A) \\ 0 & \mu(B) & \mu(B) \\ \mu(A) & \mu(B) & \mu(A) + \mu(B) \end{pmatrix} =: \hat{R}$$

Für $k := (1, 1, -1)$ ist $\langle \hat{R}k, k \rangle = 0$ das heißt

$$\langle \widetilde{X}, k \rangle(t) = e^{-\frac{t}{2} \langle \hat{R}k, k \rangle} = 1 \Rightarrow \langle X, k \rangle = 0 \text{ fast sicher}$$

bzw.

$$\mathcal{P}(\{\Gamma_A + \Gamma_B - \Gamma_{A \cup B}\} = 0) = 1$$

Somit gilt die fast-sichere Additivität von Γ auf \mathcal{M} . Es kann sogar gezeigt werden:

$$\Gamma_{\bigcup_{i=1}^{\infty} A_i} = \sum_{i=1}^{\infty} \Gamma_{A_i}$$

fast sicher für disjunkte A_i mit $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$. Für diese existiert also ein $N \in \mathfrak{C}$, $\mathcal{P}(N) = 0$ mit

$$\Gamma_{\bigcup_{i=1}^{\infty} A_i}(\omega) = \sum_{i=1}^{\infty} \Gamma_{A_i}(\omega) \quad \forall \omega \notin N$$

Das Problem liegt jedoch darin, dass N allgemein von den A_i abhängt und oft $|\mathcal{M}| = \infty$, somit kein gemeinsames N auffindbar ist

2. $\Gamma_{A \cup B} = \Gamma_A + \Gamma_B - \Gamma_{A \cap B}$ fast-sicher (wo sinnvoll).
3. Für disjunkte $A_1, \dots, A_n \in \mathcal{M}$ und $a_1, \dots, a_n \in \mathbb{R}$ ist:

$$\begin{aligned} \mathcal{F} \left\{ \sum_{j=1}^n a_j \Gamma_{A_j} \right\} (1) &= \prod_{j=1}^n \exp \left[-\frac{a_j^2}{2} \overbrace{\mu(A_j)}^{\mathbb{E}|\Gamma_{A_j}|^2} \right] = \exp \left[-\frac{1}{2} \sum_{j=1}^n a_j^2 \mu(A_j) \right] \\ &= \exp \left[-\frac{1}{2} \int_M \underbrace{\left| \sum_{j=1}^n a_j 1_{A_j} \right|^2}_{f \in L_2(M, \mu)} \mu(dx) \right] = \exp \left[-\frac{1}{2} \|f\|_{L_2(M, \mu)}^2 \right] \end{aligned}$$

Nach Darstellung (3.2.1.1) ist andererseits

$$\|f\|_{L_2(M, \mu)}^2 \stackrel{(3.2.1.1)}{=} \mathbb{E} \left| \sum_{j=1}^n a_j \Gamma_{A_j} \right|^2 = \left\| \sum_{j=1}^n a_j \Gamma_{A_j} \right\|_{L_2(\Omega, \mathcal{P})}^2$$

(da $\sum_{j=1}^n a_j \Gamma_{A_j} \sim \mathcal{N}_{0, \|f\|_{L_2(M, \mu)}^2}$). Diese Abbildung

$$f \mapsto \Gamma(f) := \sum_{j=1}^n a_j \Gamma_{A_j} \quad \text{für} \quad f = \sum_{j=1}^n a_j 1_{A_j}, \quad A_j \in \mathcal{M} : \text{ paarweise disjunkt}$$

von Treppenfunktionen f nach Zufallsvariablen $\Gamma(f)$, kann tatsächlich zu einer Isometrie von $L_2(M, \mu)$ nach $L_2(\Omega, \mathcal{P})$ fortgesetzt werden (ohne Beweis).

4. Für $A, B \in \mathcal{M}$ gilt stets

$$\mathbb{E} |\Gamma_A - \Gamma_B|^2 = \mu(A) - 2\mu(A \cap B) + \mu(B) = \mu(A \Delta B)$$

Beispiele:

(i) Für $M := [0, \infty)$ mit Lebesgue-Maß λ , $\mathcal{M} := \{[0, t] : t \geq 0\}$ ist

$$X_t := \Gamma_{[0,t]} \sim \mathcal{N}_{0,\lambda[0,t]} = \mathcal{N}_{0,t}$$

der zentrierte Gaußprozess mit $\mathbb{E}(X_t X_s) = \lambda([0, t] \cap [0, s]) = \min\{t, s\}$, also genau der Wiener-Prozess.

(ii) Für $M = [0, 1]^n$ mit Lebesgue-Maß λ und $\mathcal{M} := \{[0, t_1] \times \dots \times [0, t_n] : 0 \leq t_i \leq 1\}$ setzen

$$W_t^n := \Gamma_{[0,t_1] \times \dots \times [0,t_n]}, \quad t \in [0, 1]^n$$

Dann ist $(W_t)_t \in [0, 1]^n$ zentriert, Gaußsch mit

$$\mathbb{E}(W_t^n W_s^n) = \lambda([0, t_1] \times \dots \times [0, t_n] \cap [0, s_1] \times \dots \times [0, s_n]) = \prod_{j=1}^n \min\{t_j, s_j\}$$

und heißt *Brownian-Sheet* Prozess.

(iii) Zu *Hurst-Index* $0 < H < 1$ und $T := [0, \infty)$ setzen

$$R(t, s) := \frac{1}{2} \left\{ |t|^{2H} + |s|^{2H} - |t - s|^{2H} \right\}$$

Dann ist R nicht-negativ definit (ohne Beweis) und symmetrisch, entspricht also einem zentrierten Gaußprozess $(B^H)_{t \geq 0}$, der sogenannten *gebrochenen Brownschen Bewegung* (*gebrochener Wiener-Prozess*). So erhält man z.B. für $H = \frac{1}{2}$ genau den Wiener-Prozess W .

Die Pfade von B^H sind zwar α -Hölder für $\alpha < H$, jedoch nicht H -Hölder.

3.3.8 Gaußprozesse aus Integral-Kernen

Sei $T \neq \emptyset$ eine beliebige Indexmenge. Sei (M, \mathcal{M}, μ) ein Maßraum und $K : T \times M \rightarrow \mathbb{R}$, so dass $K_t : M \rightarrow \mathbb{R}$, $t \in T$ messbar, L_2 -integrierbar sind. Setzen

$$R(t_1, t_2) := \underbrace{\int_M K_{t_1}(s) K_{t_2}(s) \mu(ds)}_{< \infty}$$

Dann ist R symmetrisch und nicht-negativ-definit (ohne Beweis), es existiert also ein zentrierter, Gaußscher Prozess $(X_t)_{t \in T}$ mit

$$\mathbb{E}(X_t X_s) = \int_M K_{t_1}(s) K_{t_2}(s) \mu(ds)$$

Durch Wahl von K kann nun eine große Schar an Gaußprozessen konstruiert werden. So entspricht z.B. für $K_t(s) := 1_{[0,t]}(s)$ dem Wiener-Prozess.

Spezialfall: Zu $T = [0, \infty)$, $M = \mathbb{R}$ mit Lebesgue-Maß λ , setzen

$$K_t(s) := C_H \cdot \left[(\max\{t - s, 0\})^{H - \frac{1}{2}} - (\max\{-s, 0\})^{H - \frac{1}{2}} \right]$$

Dann erzeugt K wegen

$$\int_{\mathbb{R}} K_{t_1}(s) K_{t_2}(s) ds = \frac{1}{2} \left[|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H} \right] =: R(t_1, t_2)$$

genau die gebrochene Brownsche Bewegung.

3.3.9 Definition: Riemann-Liouville Operator

Zu $f : \mathbb{R} \rightarrow \mathbb{C}$ und $\alpha > 0$ definiert man den *Riemann-Liouville Operator*:

$$(\mathcal{R}_\alpha f)(t) := \frac{1}{\Gamma(\alpha)} \int_0^t (t-x)^{\alpha-1} f(x) dx$$

Bemerkungen:

- (i) Zu $\alpha, \beta > 0$ ist $\mathcal{R}_\alpha \circ \mathcal{R}_\beta = \mathcal{R}_{\alpha+\beta}$.
- (ii) $(R_1 f)' = f$ und allgemeiner $(R_n f)^{(n)} = f$, $n \in \mathbb{N}$. Demnach ist R_n dem n -ten Ableitungsoperator rechtsinvers:

$$\frac{d^n}{dx^n} \circ \mathcal{R}_n = \text{Id}$$

3.3.10 Der Riemann-Liouville-Prozess

Sei $(M, \mathcal{M}, \mu) = (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda)$, $T = \mathbb{R}_+$ und $\alpha > 0$. Setzen

$$K(t, s) := \frac{1}{\Gamma(\alpha)} (t-s)^{\alpha-1} \cdot 1_{[0,t]}(s) \quad , \quad t \in T, s \in M$$

Dann existiert ein reellwertiger, zentrierter Gaußprozess $(R_t^{(\alpha)})_{t \geq 0}$ mit

$$\mathbb{E} \left[R_{t_1}^{(\alpha)} R_{t_2}^{(\alpha)} \right] = \frac{1}{\Gamma(\alpha)^2} \int_0^{t_1 \wedge t_2} (t-s)^{\alpha-1} (t_2-s)^{\alpha-1} ds$$

(vgl. 3.3.8). Der Prozess R^α heißt *Riemann-Liouville-Prozess*. Von besonderem Interesse sind folgende Spezialfälle:

- $\alpha > \frac{1}{2}$: Der Prozess besitzt eine Version mit fast-sicher stetigen Pfaden.
- $\alpha > \frac{3}{2}$: Es existiert eine Version mit fast-sicher differenzierbaren Pfaden.
- $\alpha = 1$: $R^{(1)}$ ist genau der Wiener-Prozess.
- $\alpha = 2$: $R_t^{(2)} = \int_0^t W_t dx$ sind genau die *Schatten* der Pfade des Wiener-Prozesses.

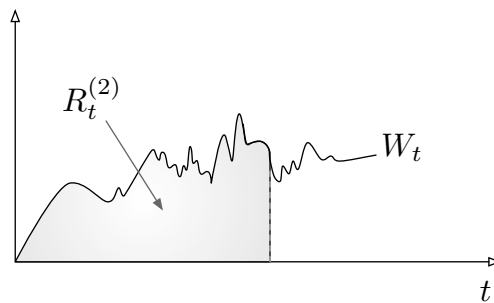


Abbildung 3.5: Zum Riemann-Liouville-Prozess $(R_t^{(2)})_{t \geq 0}$.

3.4 Integrierbarkeit von Gaußprozessen

3.4.1 Vorbetrachtung

Zu normalverteilter Zufallsvariable $\xi \sim \mathcal{N}_{0,\sigma^2}$ und $\beta > 0$ ist

$$\mathbb{E}(e^{\beta\xi^2}) = \frac{1}{\sqrt{2\pi\sigma}} \int_{\mathbb{R}} e^{\beta x^2} \cdot e^{-\frac{x^2}{2\sigma^2}} dx$$

Dabei ist $\mathbb{E}(e^{\beta\xi^2}) < \infty$ genau dann wenn $\beta < \frac{1}{2\sigma^2}$.

Sei nun $(X_t)_{t \in T}$ ein zentrierter, \mathbb{R} -wertiger Gaußprozess mit fast-sicher beschränkten Pfaden, dazu Abbildung

$$\eta := \sup_{t \in T} |X_t|$$

Es stellt sich die Frage: Gibt es ein $\beta > 0$ mit $\mathbb{E}(e^{\beta\eta^2}) < \infty$? Tatsächlich erwies sich 1970 die Aussage als gültig (Xavier Fernique) und 1976 wurde sogar der optimale Wert für β bestimmt (Michel Talagrand).

3.4.2 Definition: Zusammensetzung von Prozessen

Zu (S, \mathcal{S}) -wertigen stochastischen Prozessen $(X_t)_{t \in T}, (Y_t)_{t \in T}$ definiert man (X, Y) als den $(S^2, \mathcal{S} \otimes \mathcal{S})$ -wertigen stochastischen Prozess

$$(X, Y)_t := (X_t, Y_t) \quad , \quad t \in T$$

3.4.3 Lemma: Rotationsinvarianz Gaußscher Prozesse

Zu Winkel $\vartheta \in [0, 2\pi]$ und \mathbb{R}^n -wertigen Prozessen $(X_t)_{t \in T}, (Y_t)_{t \in T}$ setzen

$$(\mathcal{R}_\vartheta(X, Y))_t := [X_t \cdot \cos \vartheta + Y_t \cdot \sin \vartheta, -X_t \cdot \sin \vartheta + Y_t \cdot \cos \vartheta]$$

Sind X, Y identisch verteilt, unabhängig, zentriert, Gaußsch, so ist $\mathcal{R}_\vartheta(X, Y)$ ebenfalls zentriert, Gaußsch und sogar äquivalent zu (X, Y) .

Beweis: Seien zunächst $t_1, \dots, t_m \in T$ festgehalten. Setzen

$$\xi := (X_{t_1}, \dots, X_{t_m}) \quad , \quad \eta := (Y_{t_1}, \dots, Y_{t_m})$$

Dann sind ξ, η identisch verteilte, unabhängige, zentrierte Gaußsche Vektoren (in $\mathbb{R}^{n \times m}$). Nach Lemma 3.2.8 ist

$$\mathcal{R}_\vartheta(\xi, \eta) \stackrel{d.}{=} (\xi, \eta)$$

Für Mengen

$$A = A_1 \times \dots \times A_m, \quad B = B_1 \times \dots \times B_m, \quad A_j, B_j \in \mathcal{B}(\mathbb{R}^n)$$

ist dann

$$\mathcal{P} \{ \{ (X, Y)_{t_1} \in A_1 \times B_1, \dots, (X, Y)_{t_m} \in A_m \times B_m \} \} = \mathcal{P} \{ \{ \xi \in A, \eta \in B \} \}$$

$$\stackrel{(3.2.8)}{=} \mathcal{P} \{ \{ \xi \cos \vartheta + \eta \sin \vartheta \in A, -\xi \sin \vartheta + \eta \cos \vartheta \in B \} \}$$

$$= \mathcal{P} [X_{t_1} \cos \vartheta + Y_{t_1} \sin \vartheta \in A_1, X_{t_1} \sin \vartheta + Y_{t_1} \cos \vartheta \in B_1, \dots, X_{t_m} \cos \vartheta + Y_{t_m} \sin \vartheta \in A_m, -X_{t_m} \sin \vartheta + Y_{t_m} \cos \vartheta \in B_m]$$

$$= \mathcal{P} \{ \{ \mathcal{R}_\vartheta(X, Y)_{t_1} \in A_1 \times B_1, \dots, \mathcal{R}_\vartheta(X, Y)_{t_m} \in A_m \times B_m \} \}$$

Somit stimmen die Maße $\mathcal{P} \circ (X_{t_1}, Y_{t_1}, \dots, X_{t_m}, Y_{t_m})^{-1}$, $\mathcal{P} \circ (\mathcal{R}_\vartheta(X, Y)_{t_1}, \dots, \mathcal{R}_\vartheta(X, Y)_{t_m})^{-1}$ auf ganz $(\mathcal{B}(\mathbb{R}^{2n}))^m$ überein.

□

Folgerung: Speziell für $\vartheta = -\frac{\pi}{4}$ folgt

$$\mathcal{P} \left[\{(X, Y) \in B\} \right] = \mathcal{P} \left[\left\{ \left(\frac{X - Y}{\sqrt{2}}, \frac{X + Y}{\sqrt{2}} \right) \in B \right\} \right], \quad B \in \mathcal{C}((\mathbb{R}^{2n})^T)$$

(beachte Charakterisierungssatz 1.3.6 für äquivalente Prozesse).

3.4.4 Definition: Halbnorm

Sei (E, \mathcal{E}) ein \mathbb{K} -linearer, messbarer Raum ($\mathbb{K} = \mathbb{R}$ oder \mathbb{C}). Dann heißt eine Abbildung $\|\cdot\| : E \rightarrow [0, \infty]$ *Halbnorm*²³, falls gilt:

- $\|\alpha \cdot f\| = |\alpha| \cdot \|f\|$, $\alpha \in \mathbb{K}$, $f \in E$
- $\|f + g\| \leq \|f\| + \|g\|$, $f, g \in E$

Ist ferner jede Menge

$$\{f \in E : \|f\| \leq \rho\}, \quad \rho \in \mathbb{R}$$

messbar, so heißt $\|\cdot\|$ *messbar*.²⁴

Typisches Beispiel: Auf $(\mathbb{R}^T, \mathcal{C}(R^T))$ ist

$$\|f\|_{\text{Pf}} := \sup_{t \in T_0} |f(t)|, \quad f \in \mathbb{R}^T$$

($T_0 \subset T$ höchstens abzählbar) eine messbare Halbnorm (vgl. nächstes Lemma 3.4.5).

3.4.5 Lemma: Messbarkeit des Pfad-Supremums

Sei $T \neq \emptyset$ und $T_0 \subseteq T$ abzählbar. Sei ferner $(E, \|\cdot\|_E)$ ein normierter Raum. Dann ist

$$\|\cdot\|_{\text{Pf}} : (E^T, \mathcal{C}(E^T)) \rightarrow [0, \infty], \quad \|f\|_{\text{Pf}} := \sup_{t \in T_0} \|f(t)\|_E$$

eine messbare Halbnorm.

Beweis: Das $\|\cdot\|_{\text{Pf}}$ eine Halbnorm ist, ist klar. Zu zeigen wäre die Messbarkeit, wobei es genügt zu zeigen dass für $\lambda \in \mathbb{R}$ gilt $\|\cdot\|_{\text{Pf}}^{-1}((-\infty, \lambda]) \in \mathcal{C}(E^T)$ (offensichtlich ist $\|\cdot\|_{\text{Pf}}^{-1}(-\infty, \infty) = E^T$). Tatsächlich gilt

$$\|\cdot\|_{\text{Pf}}^{-1}((-\infty, \lambda]) \in \mathcal{C}(E^T) = \left\{ f \in E^T : \sup_{t \in T_0} \|f(t)\|_E \leq \lambda \right\} = \left\{ f \in E^T : \|f(t)\|_E \leq \lambda \forall t \in T_0 \right\}$$

$$= \bigcap_{t \in T_0} \left\{ f \in E^T : \|f(t)\|_E \leq \lambda \right\} = \underbrace{\bigcap_{\substack{t \in T_0 \\ \text{abzählbarer} \\ \text{Schnitt}}} \left\{ f \in E^T : f(t) \in \underbrace{\|\cdot\|_E^{-1}(-\infty, \lambda]}_{\substack{\text{messbar} \\ \text{da } \|\cdot\|_E \\ \text{stetig}}} \right\}}_{\text{Zylindermenge}} \in \mathcal{C}(E^T)$$

□

3.4.6 Korollar über die Messbarkeit des Pfad-Supremums

Sei (T, d_T) metrisch, separabel, $(E, \|\cdot\|_E)$ ein normierter Raum und $\mathcal{C}^*(T)$ der Raum aller, an jedem Punkt einseitig stetigen, Funktionen $f : T \rightarrow E$. Dann ist

$$\|\cdot\|_{\text{Pf}} : (\mathcal{C}^*(T), \mathcal{C}(E^T) \cap \mathcal{C}^*(T)) \rightarrow [0, \infty], \quad \|f\|_{\text{Pf}} := \sup_{t \in T} \|f(t)\|_E$$

eine messbare Halbnorm.

²³Typischerweise wird $\|\cdot\| : E \rightarrow [0, \infty]$ vorausgesetzt. Die Axiome sind im Fall $\|f\| = \infty$ entsprechend zu interpretieren.

²⁴Bemerke: Dies ist tatsächlich lediglich die Messbarkeit von $\|\cdot\| : (E, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Beweis: Da T separabel ist, existiert definitionsgemäß eine abzählbare, in T dichte Teilmenge $T_0 \subseteq T$. Doch dann ist für einseitig stetige $f : T \rightarrow E$:

$$\|f\|_{\text{Pf}} = \sup_{t \in T} \|f(t)\|_E = \sup_{t \in T_0} \|f(t)\|_E$$

Nach Lemma 3.4.5 ist dann $\|f\|_{\text{Pf}}$ messbar.

□

3.4.7 Hilfslemma

Sei $u_0 > 0$, dazu induktiv

$$u_{n+1} := \sqrt{2} \cdot u_n + u_0, \quad n \in \mathbb{N}_0$$

Dann gilt:

$$u_n = (\sqrt{2} + 1) \left(2^{\frac{n+1}{2}} - 1 \right) \cdot u_0$$

Beweis: Vollständige Induktion nach n .

3.4.8 Theorem von Ferniques (1970)

Sei $(X_t)_{t \in T}$ ein \mathbb{R}^n -wertiger, zentrierter Gaußprozess auf $(\Omega, \mathfrak{S}, \mathcal{P})$ und $\|\cdot\|$ eine messbare Halbnorm auf $(\mathbb{R}^n)^T$. Gilt für ein $u_0 > 0$:

$$\mathcal{P}(\{\omega \in \Omega : \|X(\omega)\| \leq u_0\}) =: \mathcal{P}(\{\|X\| \leq u_0\}) \geq \frac{1}{2}$$

so gilt

$$\mathcal{P}(\{\|X\| > u\}) \leq e^{-\frac{u^2}{24u_0^2}} \quad \forall u \geq u_0$$

Beweis: Betrachten zunächst zwei Elemente $f, g \in (\mathbb{R}^n)^T$. Dann gilt

$$\|f + g\| = \|g - f + 2f\| \leq \|g - f\| + 2\|f\| \rightsquigarrow \frac{\|f + g\|}{2} - \frac{\|g - f\|}{2} \leq \|f\| \quad (3.4.8.1)$$

Seien X_1, X_2 unabhängige Kopien von X (vgl. Existenzsatz 1.3.12) und $u \geq u_0$. Dann

$$2 \cdot \mathcal{P}(\|X\| \leq u_0) \cdot \mathcal{P}(\|X\| > u) \stackrel{X_i \stackrel{d}{=} X}{=} \mathcal{P}(\|X_1\| \leq u_0) \cdot \mathcal{P}(\|X_2\| > u) + \mathcal{P}(\|X_2\| \leq u_0) \cdot \mathcal{P}(\|X_1\| > u)$$

$$\stackrel{X_1, X_2 \text{ unabh.}}{=} \underbrace{\mathcal{P}(\|X_1\| \leq u_0, \|X_2\| > u)}_{\text{paarweise..}} + \underbrace{\mathcal{P}(\|X_2\| \leq u_0, \|X_1\| > u)}_{\text{..disjunkt}}$$

$$= \mathcal{P}(\{\|X_1\| \leq u_0, \|X_2\| > u\} \cup \{\|X_2\| \leq u_0, \|X_1\| > u\})$$

$$\stackrel{(3.4.3)}{=} \mathcal{P}\left(\underbrace{\left\{\frac{\|X_1 - X_2\|}{\sqrt{2}} \leq u_0, \frac{\|X_1 + X_2\|}{\sqrt{2}} > u\right\}}_{\substack{\subset \left\{\frac{u-u_0}{\sqrt{2}} < \|X_1\|\right\} \cap \left\{\frac{u-u_0}{\sqrt{2}} < \|X_2\|\right\} \\ \text{nach (3.4.8.1)}}}} \cup \underbrace{\left\{\frac{\|X_1 + X_2\|}{\sqrt{2}} \leq u_0, \frac{\|X_1 - X_2\|}{\sqrt{2}} > u\right\}}_{\substack{\subset \left\{\frac{u-u_0}{\sqrt{2}} < \|X_1\|\right\} \cap \left\{\frac{u-u_0}{\sqrt{2}} < \|X_2\|\right\} \\ \text{nach (3.4.8.1)}}}}\right)$$

$$\leq \mathcal{P}\left(\left\{\frac{u-u_0}{\sqrt{2}} < \|X_1\|\right\} \cap \left\{\frac{u-u_0}{\sqrt{2}} < \|X_2\|\right\}\right)$$

$$\stackrel{X_1, X_2 \text{ unabh.}}{\leq} \mathcal{P}\left(\left\{\|X_1\| > \frac{u-u_0}{\sqrt{2}}\right\}\right) \cdot \mathcal{P}\left(\left\{\|X_2\| > \frac{u-u_0}{\sqrt{2}}\right\}\right) \stackrel{X_i \stackrel{d}{=} X}{=} \mathcal{P}\left(\left\{\|X\| > \frac{u-u_0}{\sqrt{2}}\right\}\right)^2$$

Ferner also

$$\mathcal{P}(\|X\| > u) = 2 \cdot \frac{1}{2} \mathcal{P}(\|X\| > u) \leq 2 \cdot \mathcal{P}(\|X\| \leq u_0) \cdot \mathcal{P}(\|X\| > u) \leq \mathcal{P}\left(\|X\| > \frac{u - u_0}{\sqrt{2}}\right)^2, \quad u \geq u_0$$

Setzt man nun $u_1 := \sqrt{2}u_0 + u_0$ so ist

$$\mathcal{P}(\|X\| > u_1) \stackrel{u_1 > u_0}{\leq} \mathcal{P}(\|X\| > u_0)^2 \stackrel{\text{Voraussetzung}}{\leq} \left(\frac{1}{2}\right)^2$$

Setzt man nun iterativ $u_{n+1} := \sqrt{2}u_n + u_0$ so folgt per Induktion nach n :

$$\mathcal{P}(\|X\| > u_{n+1}) \stackrel{u_{n+1} > u_0}{\leq} \underbrace{\mathcal{P}(\|X\| > u_n)^2}_{\leq \left(\frac{1}{2}\right)^{2n} \text{ nach Induktionsvoraussetzung}} \leq \left(\frac{1}{2}\right)^{2^{n+1}}, \quad n \in \mathbb{N}_0 \quad (3.4.8.2)$$

Offensichtlich sind $u_0 < u_1 < u_2 < \dots$ und $u_n \xrightarrow{n \rightarrow \infty} \infty$, das heißt für beliebiges $v \geq u_0$ existiert ein $n \in \mathbb{N}_0$ mit $v \in [u_n, u_{n+1})$. Zusammen mit der Nebenbemerkung

$$\frac{v^2}{4u_0^2(\sqrt{2} + 1)^2} \leq \frac{u_{n+1}^2}{4u_0^2(\sqrt{2} + 1)^2} \stackrel{(3.4.7)}{=} \frac{1}{2} \left(2^{\frac{n+2}{2}} - 1\right)^2 \leq \frac{1}{4} \left(2^{\frac{n+2}{2}}\right)^2 = 2^n$$

folgt dann schließlich

$$\mathcal{P}(\|X\| > v) \stackrel{v \geq u_n}{\leq} \mathcal{P}(\|X\| > u_n) \leq \left(\frac{1}{2}\right)^{2^n} = e^{-2^n \ln 2} \leq \exp\left[-\underbrace{\frac{\ln 2}{4(\sqrt{2} + 1)^2}}_{\geq \frac{1}{24}} \cdot \frac{v^2}{u_0^2}\right] \leq \exp\left[-\frac{v^2}{24u_0^2}\right]$$

□

Beispiele:

(i) Spezialfall $T = \{1, \dots, n\}$. Für Gaußschen Vektor (X_1, \dots, X_n) und (Halb-)Norm

$$\|x\| := \sqrt{\sum_{j=1}^n x_j^2}, \quad x \in \mathbb{R}^n$$

gilt

$$\mathcal{P}(\{\|X\| \leq u_0\}) \geq \frac{1}{2} \Rightarrow \mathcal{P}(\{\|X\| > u\}) \leq e^{-\frac{u^2}{24u_0^2}} \quad \forall u \geq u_0$$

(ii) Sei $(E, \|\cdot\|)$ ein Banachraum und

$$\xi_1, \dots, \xi_n \sim \mathcal{N}_{0,1}$$

unabhängig, standardnormalverteilt. Dann gilt für $x_1, \dots, x_n \in E$:

$$\left[\mathbb{E}\left\|\sum_{j=1}^n \xi_j \cdot x_j\right\|^p\right]^{\frac{1}{p}} \leq C_{q,p} \cdot \left[\mathbb{E}\left\|\sum_{j=1}^n \xi_j \cdot x_j\right\|^q\right]^{\frac{1}{q}}, \quad \forall q \leq p$$

für geeignete, von E und n unabhängige, $C_{q,p}$.

Beispiel: $E = \mathcal{C}([0, 2\pi])$, $x_j(t) = \cos(jt)$ und Norm

$$\|g\| := \sup_{0 \leq t \leq 2\pi} |g(t)|, \quad g \in \mathcal{C}([0, 2\pi])$$

Siehe unten 3.4.13 für Erläuterungen.

3.4.9 Hilfslemma über Erwartungswerte

Sei $\eta : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow [0, \infty)$ eine nicht-negative, reelle Zufallsgröße. Dann:

1. Es gilt:

$$\mathbb{E}(\eta) = \int_0^\infty \mathcal{P}(\eta > t) dt$$

2. Ist $f : [0, \infty) \rightarrow [\alpha, \infty)$ ($0 \leq \alpha$) streng monoton wachsend²⁵ mit $[\alpha, \infty) \subset f([0, \infty))$ so gilt

$$\mathbb{E}(f(\eta)) = \alpha + \int_0^\infty f'(t) \mathcal{P}(\eta > t) dt$$

Beweis:

1. Schreiben

$$\mathbb{E}(\eta) = \int_0^\infty t \mathcal{P}_\eta(dt) = \int_0^\infty \int_0^t \lambda(dx) \mathcal{P}_\eta(dt) \stackrel{\text{vertausch. Integralreihenfolge}}{=} \int_0^\infty \underbrace{\int_x^\infty \mathcal{P}_\eta(dt)}_{\mathcal{P}(\eta \geq x)} \lambda(dx) = \int_0^\infty \mathcal{P}(\eta \geq t) \lambda(dt)$$

Da $\mathcal{P}(\eta \geq t)$ (als Funktion von t) monoton fallend ist, hat sie höchstens abzählbar viele Unstetigkeitsstellen, im restlichen Bereich gilt insbesondere $\mathcal{P}(\eta \geq t) = \mathcal{P}(\eta > t)$. Somit

$$\lambda(\{t \in \mathbb{R} : \mathcal{P}(\eta \geq t) \neq \mathcal{P}(\eta > t)\}) = 0$$

also

$$\mathbb{E}(\eta) = \int_0^\infty \mathcal{P}(\eta > t) \lambda(dt)$$

2. Nach (1) können wir schreiben:

$$\begin{aligned} \mathbb{E}(f(\eta)) &= \int_0^\infty \mathcal{P}(f(\eta) > t) dt \stackrel{\substack{\mathcal{P}(f(\eta) > t) = 1 \\ t \leq \alpha}}{=} \alpha + \int_\alpha^\infty \mathcal{P}(f(\eta) > t) dt \\ &= \alpha + \int_\alpha^\infty \mathcal{P}(\eta > f^{-1}(t)) dt \stackrel{s := f^{-1}(t)}{=} \int_{f^{-1}(\alpha)}^\infty f'(s) \mathcal{P}(\eta > s) ds \end{aligned}$$

(beachte dass $f : [0, \infty) \rightarrow f([\alpha, \infty))$ bijektiv ist und $f^{-1}(\alpha) = 0$).

□

3.4.10 Theorem über Gaußprozesse (Folgerung von Ferniques)

Sei $(X_t)_{t \in T}$ ein \mathbb{R}^n -wertiger, zentrierter Gaußprozess und $\|\cdot\|_{\text{Pf}}$ eine messbare Halbnorm auf $(\mathbb{R}^n)^T$. Gilt

$$\mathcal{P}(\{\|X\|_{\text{Pf}} < \infty\}) = 1$$

so existiert ein $\beta > 0$ mit

$$\mathbb{E} \left[e^{\beta \|X\|_{\text{Pf}}^2} \right] < \infty$$

²⁵Beachte dass streng monotone Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ fast-überall differenzierbar sind.

Beweis: Nach Hilfslemma 3.4.9 (2) ist

$$\mathbb{E} \left[e^{\beta \|X\|_{\text{Pf}}^2} \right] = 1 + \int_0^\infty 2\beta s \cdot e^{\beta s^2} \mathcal{P}(\|X\|_{\text{Pf}} > s) ds$$

(setze $f(s) := e^{\beta s^2}$, $\alpha := 1$ und $\eta := \|X\|_{\text{Pf}}$). Wegen $\mathcal{P}(\|X\|_{\text{Pf}} < \infty) = 1$ existiert ein $u_0 > 0$ mit

$$\mathcal{P}(\|X\|_{\text{Pf}} \leq u_0) \geq \frac{1}{2}$$

das heißt

$$\mathbb{E} \left[e^{\beta \|X\|_{\text{Pf}}^2} \right] \leq 1 + 2\beta \int_0^{u_0} s e^{\beta s^2} ds + 2\beta \int_{u_0}^\infty s e^{\beta s^2} \mathcal{P}(\|X\|_{\text{Pf}} > s) ds$$

$$\stackrel{(3.4.8)}{\leq} 1 + 2\beta \int_0^{u_0} s e^{\beta s^2} ds + 2\beta \int_{u_0}^\infty s e^{\beta s^2} e^{-\frac{s^2}{24u_0^2}} ds < \infty \quad \text{für} \quad \beta < \frac{1}{24u_0^2}$$

□

Folgerungen: Sei $(X_t)_{t \in T}$ \mathbb{R}^n -wertig, Gaußsch, zentriert mit $\mathcal{P}(\|X\|_{\text{Pf}} < \infty) = 1$.

(i) Sei $\lambda \in \mathbb{R}$ beliebig. Wegen $e^{\lambda u} \leq e^{\beta u^2}$ für u groß genug ($u \geq \frac{\lambda}{\beta}$), ist

$$\mathbb{E} \left[e^{\lambda \|X\|_{\text{Pf}}} \right] < \infty$$

(ii) Sei $0 < p < \infty$ beliebig. Wegen $s^p \leq e^s$ für s groß genug, ist

$$\mathbb{E} \|X\|_{\text{Pf}}^p < \infty$$

3.4.11 Lemma über die p -Potenz von Pfadnormen

Sei $(X_t)_{t \in T}$ ein \mathbb{R}^n -wertiger, zentrierter Gaußprozess und $\|\cdot\|_{\text{Pf}}$ eine messbare Halbnorm auf $(\mathbb{R}^n)^T$. Gilt

$$\mathcal{P}(\|X\|_{\text{Pf}} \leq u_0) \geq \frac{1}{2}$$

für irgendein $u_0 \geq 0$, so folgt

$$(\mathbb{E} \|X\|_{\text{Pf}}^p)^{\frac{1}{p}} \leq C_p \cdot u_0, \quad 0 < p < \infty$$

für universelle, allein von p abhängige Konstanten $C_p > 1$.

Beweis:

$$\begin{aligned}
\mathbb{E} \|X\|_{\text{Pf}}^p &\stackrel{(3.4.9)}{=} \int_0^\infty \underbrace{\mathcal{P}(\|X\|_{\text{Pf}}^p > t)}_{\leq 1} dt \leq u_0^p + \int_{u_0^p}^\infty \underbrace{\mathcal{P}(\|X\|_{\text{Pf}}^p > t)}_{\mathcal{P}(\|X\|_{\text{Pf}} > t^{\frac{1}{p}})} dt \stackrel{s:=t^{\frac{1}{p}}}{=} u_0^p + \int_{u_0}^\infty p s^{p-1} \mathcal{P}(\|X\|_{\text{Pf}} > s) ds \\
&= u_0^p + \sum_{k=0}^\infty p \int_{u_k}^{u_{k+1}} \underbrace{s^{p-1} \mathcal{P}(\|X\|_{\text{Pf}} > s)}_{\substack{\leq 2^{-2^k} \\ (3.4.8.2)}} ds \quad \Bigg| \quad u_k := (\sqrt{2} + 1) \left(2^{\frac{k+1}{2}} - 1\right) u_0 \\
&\leq u_0^p + \sum_{k=0}^\infty 2^{-2^k} \underbrace{(u_{k+1}^p - u_k^p)}_{\leq u_{k+1}^p} \\
&\leq u_0^p + u_0^p \sum_{k=0}^\infty 2^{3p} 2^{\frac{kp}{2}} 2^{-2^k} \quad \Bigg| \quad u_{k+1}^p = (\sqrt{2} + 1)^p \cdot \left(2^{\frac{k+1}{2}} - 1\right)^p \cdot u_0^p \leq 2^{3p} 2^{\frac{kp}{2}} u_0^p \\
&= \underbrace{\left[1 + \sum_{k=0}^\infty 2^{3p} 2^{\frac{kp}{2}} 2^{-2^k}\right]}_{(C_p)^p < \infty} \cdot u_0^p
\end{aligned}$$

□

Beispiel: Aus $\mathcal{P}(\|X\|_{\text{Pf}} \leq u_0) \geq \frac{1}{2}$ folgt wegen $C_1 \leq \frac{5}{2}$:

$$\mathbb{E} \|X\|_{\text{Pf}} \leq \frac{5}{2} u_0$$

3.4.12 Theorem über die Monotonie der p -Potenzen von Pfadnormen

Sei $(X_t)_{t \in T}$ Gaußsch, zentriert, \mathbb{R}^n -wertig und $\|\cdot\|_{\text{Pf}}$ eine messbare Halbnorm auf $(\mathbb{R}^n)^T$. Dann existieren universelle Konstanten $C_{p,q} > 1$ so dass

$$(\mathbb{E} \|X\|_{\text{Pf}}^q)^{\frac{1}{q}} \leq C_{p,q} \cdot (\mathbb{E} \|X\|_{\text{Pf}}^p)^{\frac{1}{p}}, \quad 0 < p \leq q < \infty$$

Beweis: Allgemein gilt

$$\mathcal{P}(\|X\|_{\text{Pf}} \geq u) = \mathcal{P}(\|X\|_{\text{Pf}}^p \geq u^p) \stackrel{\text{Markov}}{\leq} \frac{\mathbb{E} \|X\|_{\text{Pf}}^p}{u^p}$$

Speziell für $u_0 := 3^{\frac{1}{p}} (\mathbb{E} \|X\|_{\text{Pf}}^p)^{\frac{1}{p}}$ gilt dann

$$\mathcal{P}(\|X\|_{\text{Pf}} > u_0) \leq \frac{\mathbb{E} \|X\|_{\text{Pf}}^p}{3 \mathbb{E} \|X\|_{\text{Pf}}^p} = \frac{1}{3} < \frac{1}{2} \Rightarrow$$

Nach Lemma 3.4.11 folgt

$$(\mathbb{E} \|X\|_{\text{Pf}}^q)^{\frac{1}{q}} \leq C_q \cdot u_0 = \underbrace{C_q 3^{\frac{1}{p}}}_{C_{p,q}} (\mathbb{E} \|X\|_{\text{Pf}}^p)^{\frac{1}{p}}$$

□

Bemerkungen:

(i) Allgemein gilt für beliebige Zufallsgröße η die Abschätzung

$$(\mathbb{E} |\eta|^p)^{\frac{1}{p}} \leq (\mathbb{E} |\eta|^q)^{\frac{1}{q}} \quad , \quad 0 < p \leq q < \infty \quad (3.4.12.1)$$

Ein umgekehrte Abschätzung existiert jedoch im allgemeinen nicht. Theorem 3.4.12 sichert diese lediglich für zentrierte Gaußprozesse.

(ii) Es kann gezeigt werden, dass die *optimalen* $C_{p,q}$ gegeben sind durch

$$C_{p,q} = \frac{\left[\frac{1}{\sqrt{\pi}} \Gamma\left(\frac{q+1}{2}\right) \right]^{\frac{1}{q}}}{\left[\frac{1}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right) \right]^{\frac{1}{p}}} \quad (3.4.12.2)$$

Als Beispiel dazu sei die standardnormalverteilte Zufallsvariable $\xi \sim \mathcal{N}_{0,1}$ betrachtet. Dann ist zum einen $(\mathbb{E} |\xi|^2)^{\frac{1}{2}} = 1$, zum anderen

$$\mathbb{E} |\xi| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x| e^{-\frac{x^2}{2}} dx = \sqrt{\frac{2}{\pi}} \underbrace{\int_0^{\infty} x e^{-\frac{x^2}{2}} dx}_1 = \sqrt{\frac{2}{\pi}}$$

so dass die Abschätzung

$$(\mathbb{E} |\xi|^2)^{\frac{1}{2}} \leq C_{1,2} \cdot \mathbb{E} |\xi|$$

den kleinstmöglichen Wert $C_{1,2} \geq \sqrt{\frac{\pi}{2}}$ liefert, was genau Darstellung (3.4.12.2) entspricht.

3.4.13 Gaußprozesse auf Banachräumen

Sei $(E, \|\cdot\|)$ ein reeller, separabler Banachraum mit $\|\cdot\|$, dazu $(E', \|\cdot\|)$ der Banachraum der beschränkten, linearen Funktionale auf E , mit der Norm

$$\|a\| := \sup_{\|x\| \leq 1} |a(x)| \quad , \quad a \in E'$$

Eine bzgl. der Borel- σ -Algebra $\mathcal{B}(E)$ messbare Abbildung (Zufallsgröße) $Y : \Omega \rightarrow E$ heißt *Gaußverteilt* falls $a(Y) : \Omega \rightarrow \mathbb{R}$ Gaußverteilt ist $\forall a \in E'$. Y heißt ferner *zentriert*, falls $\mathbb{E}(a(Y)) = 0$ für $a \in E'$.

Bemerkung: In \mathbb{R}^n ist jede 1-Form $a \in (\mathbb{R}^n)'$ eindeutig durch einen Vektor $(k_1, \dots, k_n) \in \mathbb{R}^n$ mit $a(x) := \langle k, x \rangle$ erzeugt. Somit erweist sich obige Definition tatsächlich als Verallgemeinerung von Gaußvektoren.

Mit $E'_1 := \{a \in E' : \|a\| \leq 1\}$ erhält man durch

$$X_a(\omega) := a(Y(\omega)) \quad , \quad a \in E'_1$$

einen \mathbb{R} -wertigen Gaußprozess $(X_a)_{a \in E'_1}$, denn

$$\langle (X_{a_1}, \dots, X_{a_n}), (k_1, \dots, k_n) \rangle = \sum_{j=1}^n k_j X_{a_j} = \sum_{j=1}^n k_j (a_j \circ Y) = \underbrace{\left[\sum_{j=1}^n k_j a_j \right]}_{b \in E'} \circ Y \sim \mathcal{N}_{\mu, \sigma^2} \quad , \quad a_1, \dots, a_n \in E'_1, k_1, \dots, k_n \in \mathbb{R}$$

Sei nun Y zentriert, dann ist offensichtlich auch (X_a) zentriert. Nach Hahn-Banach²⁶ gilt

$$\sup_{a \in B'_1} |X_a| = \|Y\| < \infty$$

²⁶Das Hahn-Banach-Theorem lautet:

Sei $(E, \|\cdot\|)$ ein normierter Raum und $E_0 \subset E$ ein Unterraum. Ist a_0 eine stetige Linearform auf E_0 , so existiert eine stetige Linearform a auf E mit $a|_{E_0} = a_0$ und $\|a\| = \|a_0\|$.

. Eine Folgerung des Theorems ist:

Für beliebiges $x \in E$ ist $\|x\| = \sup_{\|a\| \leq 1} |a(x)|$.

Genauer gesagt existiert eine *normierende* Folge von Linearformen $a_1, a_2, \dots \in E'_1$ so dass

$$\sup_{a \in B'_1} |X_a| = \sup_{i \in \mathbb{N}} |X_{a_i}| =: \|X\|_{\text{Pf}}$$

Insbesondere ist $\|\cdot\|_{\text{Pf}}$ eine messbare Halbnorm auf $(\mathbb{R}^{E'_1}, \mathcal{C}(\mathbb{R}^{E'_1}))$, also $\exists \beta > 0$ so dass

$$\mathbb{E} \left[e^{\beta \|Y\|_{\text{Pf}}^2} \right] < \infty$$

und ferner

$$(\mathbb{E} \|Y\|^q)^{\frac{1}{q}} \leq C_{p,q} \cdot (\mathbb{E} \|Y\|^p)^{\frac{1}{p}}, \quad 0 < p \leq q < \infty$$

für geeignete Konstanten $C_{p,q}$ (vgl. 3.4.12). Seien nun speziell ξ_1, \dots, ξ_n unabhängig, standardnormalverteilt und $x_1, \dots, x_n \in E$. Dann definiert

$$Y(\omega) := \sum_{j=1}^n \xi_j(\omega) \cdot x_j$$

einen zentrierten Gaußprozess, also

$$\left(\mathbb{E} \left\| \sum_{j=1}^n \xi_j x_j \right\|^q \right)^{\frac{1}{q}} \leq C_{p,q} \cdot \left(\mathbb{E} \left\| \sum_{j=1}^n \xi_j x_j \right\|^p \right)^{\frac{1}{p}}, \quad 0 < p \leq q < \infty \quad (3.4.13.1)$$

Ist ferner $E = H$ ein Hilbertraum, so gilt

$$\mathbb{E} \left\| \sum_{j=1}^n \xi_j x_j \right\|^2 = \mathbb{E} \left\langle \sum_{j=1}^n \xi_j x_j, \sum_{i=1}^n \xi_i x_i \right\rangle = \sum_{j,i=1}^n \langle x_j, x_i \rangle \underbrace{\mathbb{E}(\xi_j \xi_i)}_{\delta_{ij}} = \sum_{j=1}^n \langle x_j, x_j \rangle = \sum_{j=1}^n \|x_j\|^2$$

das heißt

$$\left(\mathbb{E} \left\| \sum_{j=1}^n \xi_j x_j \right\|^2 \right)^{\frac{1}{2}} = \left(\sum_{j=1}^n \|x_j\|^2 \right)^{\frac{1}{2}} \quad (3.4.13.2)$$

und somit

$$\frac{1}{C_{p,2}} \cdot \left(\sum_{j=1}^n \|x_j\|^2 \right)^{\frac{1}{2}} \stackrel{\substack{(3.4.12.1) \text{ falls } p \geq 2 \\ (3.4.13.1) \text{ falls } p < 2}}{\leq} \left(\mathbb{E} \left\| \sum_{j=1}^n \xi_j x_j \right\|^p \right)^{\frac{1}{p}} \stackrel{\substack{(3.4.13.1) \text{ falls } p \geq 2 \\ (3.4.12.1) \text{ falls } p < 2}}{\leq} C_{2,p} \cdot \left(\sum_{j=1}^n \|x_j\|^2 \right)^{\frac{1}{2}}$$

3.5 Reihendarstellung Gaußscher Prozesse

3.5.1 Einführung

Die Simulation stochastischer Prozesse am Rechner ist eine Aufgabe mit der immer häufiger so mancher Wissenschaftler konfrontiert wird. Dabei handelt es sich bei den simulierten Prozessen typischerweise um *Näherungen*, etwa als abgebrochene Reihen elementarer Prozesse.

Betrachtet sei z.B. der Prozess $(R_n)_{n \in \mathbb{N}_0}$ mit unabhängigen, identisch verteilten Zuwächsen

$$\mathcal{P}(Z_i = 1) = \frac{1}{2} = \mathcal{P}(Z_i = -1), \quad i = 1, 2, 3, \dots$$

(1-dimensionaler Random-Walker mit Schrittlänge ± 1), dazu die Folge gestauchter Prozesse

$$X_n^k := \frac{R_k}{\sqrt{n}}, \quad k = 0, \dots, n$$

($n \geq 1$). Dann konvergiert X mit $n \rightarrow \infty$ gegen den standard-Wiener-Prozess $W_{0 \leq t \leq 1}$ (Invarianzprinzip von Donsker).

Alternativer Zugang: Setzt man zu unabhängigen, standardnormalverteilten ξ_1, ξ_2, \dots und konstanten Funktionen $\psi_k(t)$:

$$X_t := \sum_{k=1}^{\infty} \xi_k \cdot \psi_k(t)$$

so definiert X (im Falle der Konvergenz) erneut einen stochastischen Prozess. So ergibt z.B. für $T = [0, 1]$ die Reihe

$$X_t := \xi_0 + \sqrt{2} \sum_{k=1}^{\infty} \xi_k \frac{\sin(k\pi t)}{k\pi}$$

(Überlagerung von Schwingungen mit *zufälligen Koeffizienten*) genau den Wiener-Prozess (siehe später 3.5.7).

Aus praktischen Gründen werden am Rechner allgemein nur endliche Summen anstelle von Reihen verwendet. Dabei stellt sich unvermeidbar die Frage der Konvergenzgeschwindigkeit und Darstellungsoptimierung.

3.5.2 Lemma: Reihenkonvergenz mit zufälligen Koeffizienten

Seien $\xi_1, \xi_2, \dots : \Omega \rightarrow \mathbb{R}$ standardnormalverteilt und $\alpha_1, \alpha_2, \dots \in \mathbb{R}$. Dann existiert:

$$\sum_{j=1}^{\infty} \xi_j \alpha_j$$

fast-überall genau dann wenn

$$\sum_{j=1}^{\infty} \alpha_j^2 < \infty$$

Gegebenfalls gilt dann

$$\sum_{j=1}^{\infty} \xi_j \alpha_j \sim \mathcal{N}_{0, \sum_{j=1}^{\infty} \alpha_j^2}$$

Beweis: Richtung "⇐": Sei $\sum_{j=1}^{\infty} \alpha_j^2 < \infty$. Dann ist analog zu 3.4.13 für $\varepsilon > 0$

$$\mathbb{E} \left| \sum_{j=n}^m \alpha_j \xi_j \right|^2 = \sum_{j=n}^m \alpha_j^2 < \varepsilon^2 \quad \forall n_\varepsilon < n < m$$

für genügend großes n_ε , das heißt

$$\sum_{j=1}^m \alpha_j \xi_j, \quad m \in \mathbb{N}$$

ist Cauchy (bzgl. $\|\cdot\|_{L_2}$), konvergiert also in L_2 (und somit auch L_1) gegen

$$X := \sum_{j=1}^{\infty} \alpha_j \xi_j \in L_2$$

Nach Levy konvergiert die Reihe auch Punktweise fast-überall. Ferner ist

$$\mathbb{E} e^{itX} = \lim_{n \rightarrow \infty} \mathbb{E} \left\{ \exp \left[it \sum_{j=1}^n \xi_j \alpha_j \right] \right\} \stackrel{(3.2.4.2)}{=} \lim_{n \rightarrow \infty} \exp \left[-\frac{t^2}{2} \sum_{j=1}^n \alpha_j^2 \right] = \exp \left[-\frac{t^2}{2} \sum_{j=1}^{\infty} \alpha_j^2 \right]$$

das heißt nach Gleichung (3.2.1.1):

$$X \sim \mathcal{N}_{0, \sum_{j=1}^{\infty} \alpha_j^2}$$

Richtung "⇒": Es existiere

$$X := \sum_{j=1}^{\infty} \alpha_j \xi_j$$

fast überall. Dann ist einerseits

$$\begin{aligned}\tilde{X}(t) &= \mathbb{E}e^{itX} \stackrel{\text{Lebesgue}}{=} \lim_{n \rightarrow \infty} \mathbb{E} \exp \left[it \sum_{j=1}^n \alpha_j \xi_j \right] = \lim_{n \rightarrow \infty} \prod_{j=1}^n \mathbb{E} \exp [it\alpha_j \xi_j] \\ &= \lim_{n \rightarrow \infty} \prod_{j=1}^n \exp \left[-\frac{t^2}{2} \alpha_j^2 \right] = \lim_{n \rightarrow \infty} \exp \left[-\frac{t^2}{2} \sum_{j=1}^n \alpha_j^2 \right]\end{aligned}$$

Andererseits ist \tilde{X} als charakteristische Funktion stetig und erfüllt $\tilde{X}(0) = 1$, insbesondere $\tilde{X}|_U > 0$ in einer Umgebung U von 0. Wäre

$$\sum_{j=1}^{\infty} \alpha_j^2 = \infty$$

so müsse $X(t) = 0 \forall t \neq 0$ sein, ein Widerspruch!

□

Folgerung: Für ξ_1, ξ_2, \dots unabhängig, identisch $\mathcal{N}_{0,1}$ -verteilt und

$$X := \sum_{j=1}^{\infty} \alpha_j \xi_j, \quad \sum_{j=1}^{\infty} \alpha_j^2 < \infty$$

ist dann

$$\mathbb{E}|X|^2 = \sum_{j=1}^{\infty} \alpha_j^2$$

und nach Satz 3.4.12, Darstellung 3.4.13.2 und Abschätzung 3.4.12.1 existieren $\forall p > 0$ Konstanten c_p, C_p so dass

$$c_p \left(\sum_{j=1}^{\infty} \alpha_j^2 \right)^{\frac{1}{2}} \leq \left(\mathbb{E} \left| \sum_{j=1}^{\infty} \alpha_j \xi_j \right|^p \right)^{\frac{1}{p}} \leq C_p \left(\sum_{j=1}^{\infty} \alpha_j^2 \right)^{\frac{1}{2}}$$

(vgl. 3.4.13).

3.5.3 Satz: Gaußsche Prozesse als Reihen

Seien $T \neq \emptyset$, $\psi_k : T \rightarrow \mathbb{R}$, $k \in \mathbb{N}$ mit

$$\sum_{k=1}^{\infty} |\psi_k(t)|^2 < \infty, \quad t \in T$$

und ξ_1, ξ_2, \dots identisch, unabhängig $\mathcal{N}_{0,1}$ -verteilt. Dann existiert nach Konvergenzsatz 3.5.2 für jedes feste $t \in T$ eine \mathcal{P} -Nullmenge N_t so dass

$$\sum_{l=1}^{\infty} \xi_l(\omega) \psi_k(t)$$

für $\omega \notin N_t$ existiert. Setzt man nun

$$X_t(\omega) := \begin{cases} \sum_{k=1}^{\infty} \xi_k(\omega) \psi_k(t) & : \omega \notin N_t \\ 0 & : \text{sonst} \end{cases}, \quad t \in T$$

so ist (X_t) ein zentrierter, reellwertiger Gaußprozess mit Kovarianzfunktion

$$R_X(t, s) = \mathbb{E}(X_t X_s) = \sum_{k=1}^{\infty} \psi_k(t) \psi_k(s)$$

Beweis: Seien $t_1, \dots, t_n \in T$ und $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Dann ist

$$\sum_{j=1}^n \alpha_j X_{t_j} = \sum_{j=1}^n \alpha_j \sum_{k=1}^{\infty} \xi_k \cdot \psi_k(t_j) = \sum_{k=1}^{\infty} \xi_k \underbrace{\left[\sum_{j=1}^n \alpha_j \psi_k(t_j) \right]}_{\beta_k} \stackrel{(3.5.2)}{\sim} \mathcal{N}_{0, \sum_{k=1}^{\infty} \beta_k^2} \quad \left| \quad \sum_{k=1}^{\infty} \beta_k^2 < \infty \right.$$

das heißt X ist zentriert, Gaußsch. Ferner folgt aus

$$\begin{aligned} \mathbb{E} |X_t \pm X_s|^2 &= \mathbb{E} \left| \sum_{k=1}^{\infty} \xi_k [\psi_k(t) \pm \psi_k(s)] \right|^2 \stackrel{(3.4.13.2)}{=} \sum_{k=1}^{\infty} [\psi_k(t) \pm \psi_k(s)]^2 \\ &= \sum_{k=1}^{\infty} \psi_k(t)^2 \pm 2 \sum_{k=1}^{\infty} \psi_k(t) \psi_k(s) + \sum_{k=1}^{\infty} \psi_k(s)^2 \end{aligned}$$

entsprechend

$$\mathbb{E}(X_t X_s) = \frac{1}{4} \left\{ \mathbb{E} |X_t + X_s|^2 - \mathbb{E} |X_t - X_s|^2 \right\} = \sum_{k=1}^{\infty} \psi_k(t) \psi_k(s)$$

□

3.5.4 Hilfslemma über Gaußsche Integrale

Für jedes $t_0 > 0$ existiert eine Konstante $C_{t_0} > 0$ mit

$$\int_t^{\infty} e^{-\frac{x^2}{2}} dx \leq C_{t_0} \cdot \frac{e^{-\frac{t^2}{2}}}{t} \quad \forall t \geq t_0$$

Beweis: Setze

$$\psi(t) := \int_t^{\infty} e^{-\frac{x^2}{2}} dx, \quad \varphi(x) := \frac{e^{-\frac{t^2}{2}}}{t}$$

dann gilt wegen $\psi(t), \varphi(t) \xrightarrow{t \rightarrow \infty} 0$ nach L'Hopital:

$$\lim_{t \rightarrow \infty} \frac{\psi(t)}{\varphi(t)} = \lim_{t \rightarrow \infty} \frac{\psi'(t)}{\varphi'(t)} = \lim_{t \rightarrow \infty} \frac{1}{1 + \frac{1}{t^2}} = 1$$

das heißt

$$\sup_{t \geq t_0} \frac{\psi(t)}{\varphi(t)} < \infty \quad \forall t \geq t_0$$

□

Insbesondere gilt

$$\int_t^{\infty} e^{-\frac{x^2}{2}} dx \leq \tilde{C}_{t_0} \cdot e^{-\frac{t^2}{2}} \quad \forall t \geq t_0 > 0 \quad (3.5.4.1)$$

3.5.5 Hilfslemma

Sei $0 \leq x \leq 1$ und $n \in \mathbb{N}$. Dann gilt $1 - (1-x)^n \leq nx$.

Beweis: Für $g(x) := 1 - (1-x)^n - nx$ gilt $g(0) = 0$ und $g'(x) = n(1-x)^{n-1} - n \leq 0$, das heißt g ist monoton fallend in $[0, 1]$. Demnach ist $g(x) \leq 0$ für $0 \leq x \leq 1$.

3.5.6 Lemma: Supremum normalverteilter Zufallsvariablen

Seien ξ_1, ξ_2, \dots unabhängig, identisch $\mathcal{N}_{0,1}$ -verteilt. Dann gilt

$$\mathbb{E} \left[\sup_{1 \leq j \leq n} |\xi_j| \right] \leq C \cdot \sqrt{\ln(n)} \quad , \quad n \geq 2$$

für irgendeine Konstante C .

Beweis: Es gilt:

$$\begin{aligned} \mathbb{E} \sup_{j \leq n} |\xi_j| &= \int_0^\infty \mathcal{P} \left(\sup_{j \leq n} |\xi_j| > t \right) dt = \int_0^\infty \underbrace{\left[1 - \mathcal{P} \left(\sup_{j \leq n} |\xi_j| \leq t \right) \right]}_{\mathcal{P}(|\xi_j| \leq t \ \forall j \leq n)} dt = \int_0^\infty \left[1 - \underbrace{\prod_{j=1}^n \mathcal{P}(|\xi_j| \leq t)}_{\substack{\mathcal{P}(|\xi| \leq t)^n \\ \xi \sim \mathcal{N}_{0,1}}} \right] dt \\ &= \int_0^\infty \left[1 - [1 - \mathcal{P}(|\xi| > t)]^n \right] dt = \int_0^{\sqrt{2 \ln n}} \underbrace{\left[1 - [1 - \mathcal{P}(|\xi| > t)]^n \right]}_{\leq 1} dt + \int_{\sqrt{2 \ln n}}^\infty \underbrace{\left[1 - [1 - \mathcal{P}(|\xi| > t)]^n \right]}_{\substack{\leq n \mathcal{P}(|\xi| > t) \\ (3.5.5)}} dt \\ &\leq \sqrt{2 \ln n} + n \int_{\sqrt{2 \ln n}}^\infty \mathcal{P}(|\xi| > t) dt \\ &\leq \sqrt{2 \ln n} + n \int_{\sqrt{2 \ln n}}^\infty C \cdot e^{-\frac{t^2}{2}} dt \quad \Bigg| \quad \mathcal{P}(|\xi| > t) = \sqrt{\frac{2}{\pi}} \int_t^\infty e^{-\frac{x^2}{2}} dx \stackrel{(3.5.4.1)}{\leq} \underbrace{C}_{\text{const}} \cdot e^{-\frac{t^2}{2}} \\ &\stackrel{(3.5.4.1)}{\leq} \sqrt{2 \ln n} + n C' \underbrace{e^{-\frac{(\sqrt{2 \ln n})^2}{2}}}_{\frac{1}{n}} \leq \sqrt{2 \ln n} + C' \leq \underbrace{C''}_{\text{const}} \cdot \sqrt{\ln n} \end{aligned}$$

□

3.5.7 Der Wiener-Prozess als zufällige Reihe

Sei $T = [0, 1]$ und $(g_k)_{k=1}^\infty$ eine Orthonormalbasis²⁷ in $L_2[0, 1]$ (reell). Setzt man

$$\psi_k(t) := \langle 1_{[0,t]}, g_k \rangle = \int_0^t g_k(x) dx$$

so ist nach Parseval

$$\sum_{k=1}^\infty |\psi_k(k)|^2 = \|1_{[0,t]}\|_{L_2}^2 = t < \infty$$

²⁷Eine Orthonormalbasis $(g_k)_{k \in \mathbb{N}} \subset \mathcal{H}$ eines Hilbertraumes \mathcal{H} ist ein vollständiges Orthonormalsystem, das heißt $\langle g_k, g_j \rangle = \delta_{kj}$ und $\text{span} \{x_1, x_2, \dots\}$ ist dicht in \mathcal{H} . Gegebenfalls gilt dann

$$x = \sum_{k=1}^\infty \langle x, x_k \rangle x_k \quad \forall x \in \mathcal{H}$$

und die Parsevalsche Gleichung

$$\sum_{k=1}^\infty |\langle x, x_k \rangle|^2 = \|x\|^2 \quad \forall x \in \mathcal{H}$$

Nach 3.5.3 ist

$$X_t := \sum_{k=1}^{\infty} \xi_k \psi_k(t) \quad , \quad t \in [0, 1]$$

ein zentrierter Gaußprozess mit

$$\mathbb{E}(X_t X_s) = \sum_{k=1}^{\infty} \psi_k(t) \psi_k(s) = \sum_{k=1}^{\infty} \langle 1_{[0,t]}, g_k \rangle \langle g_k, 1_{[0,s]} \rangle = \langle 1_{[0,t]}, 1_{[0,s]} \rangle = t \wedge s$$

Nach 3.3.5 ist somit X genau der Wiener-Prozess.

Beispiele:

(i) Das Orthonormalsystem

$$g_0 := 1 \quad , \quad g_k(x) = \sqrt{2} \cos(\pi k x) \quad , \quad k \geq 1$$

erzeugt mit

$$\psi_k(t) = \langle 1_{[0,t]}, g_k \rangle = \begin{cases} t & : k = 0 \\ \frac{\sqrt{2}}{k\pi} \sin(k\pi t) & : k \geq 1 \end{cases}$$

den Wiener-Prozess:

$$W_t = \xi_0 \cdot t + \sqrt{2} \cdot \sum_{k=1}^{\infty} \xi_k \cdot \frac{\sin(k\pi t)}{k\pi} \quad , \quad \xi_0, \xi_1, \dots \sim \mathcal{N}_{0,1} \tag{3.5.7.1}$$

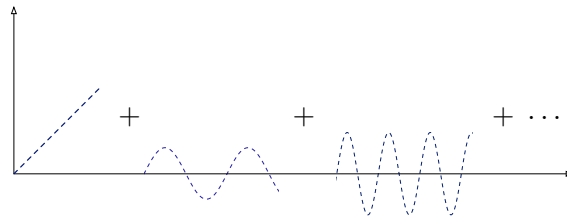


Abbildung 3.6: Zur Konstruktion des Wiener-Prozesses durch zufällige harmonische Reihen.

Mit Darstellung (3.3.6.1) erhält man ferner eine Reihendarstellung der Brownschen Brücke:

$$B_t = \sqrt{2} \cdot \sum_{k=1}^{\infty} \xi_k \cdot \frac{\sin(k\pi t)}{k\pi} \quad , \quad \xi_0, \xi_1, \dots \sim \mathcal{N}_{0,1} \tag{3.5.7.2}$$

(ii) Durch das Orthonormalsystem

$$h_{-1}(x) = 1 \quad , \quad h_{j,k} := \left\{ 1_{\left[\frac{2k}{2^{j+1}}, \frac{2k+1}{2^{j+1}}\right]} - 1_{\left[\frac{2k+1}{2^{j+1}}, \frac{2k+2}{2^{j+1}}\right]} \right\} \quad , \quad j \in \mathbb{N}_0, \quad k = 0, 1, \dots, 2^j - 1, \quad x \in [0, 1]$$

(Haarsche Basis) erhält man das sogenannte *Faber-Schader-System*

$$u_{-1}(x) := x \quad , \quad u_{j,k}(x) := \int_0^x h_{j,k}(s) ds \quad , \quad x \in [0, 1]$$

so dass für unabhängige $\xi_{-1}, \xi_{j,k} \sim \mathcal{N}_{0,1}$ die Reihe

$$\underbrace{\xi_{-1} \cdot t}_{X_t^0} + \underbrace{\sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \xi_{j,k} \cdot u_{j,k}(t)}_{X_t^j} \tag{3.5.7.3}$$

fast-sicher konvergiert und dem Wiener-Prozess entspricht.

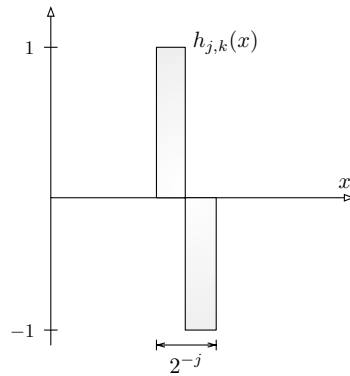


Abbildung 3.7: Zur Definition der Haarschen Basis.

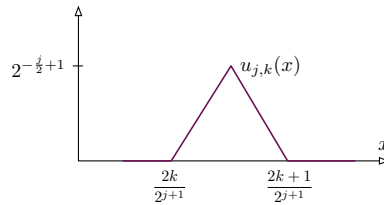


Abbildung 3.8: Zur Konstruktion des Wiener-Prozesses durch zufällige Faber-Schader-Dreiecke.

Die Teil-Prozesse $(X^j)_{j=-1}^\infty$ sind dabei unabhängig und besitzen stetige Pfade. Insbesondere ist

$$\|X^j\|_{\text{Pf}} := \|X^j\|_\infty := \sup_{0 \leq t \leq 1} |X_t^j| = 2^{-\frac{j}{2}+1} \sup_{k=0, \dots, 2^j-1} |\xi_{j,k}|$$

eine messbare Halbnorm auf $(\mathbb{R}^{[0,1]}, \mathcal{C}(\mathbb{R}^{[0,1]}))$ und nach Lemma 3.5.6 gilt

$$\mathbb{E} \|X^j\|_{\text{Pf}} = 2^{-\frac{j}{2}+1} \mathbb{E} \sup_{k \leq 2^j-1} |\xi_{j,k}| \stackrel{(3.5.6)}{\leq} \underbrace{C}_{\text{const}} 2^{-\frac{j}{2}} \sqrt{\ln 2^j} \leq \underbrace{C'}_{\text{const}} \cdot 2^{-\frac{j}{2}} \sqrt{j} \tag{3.5.7.4}$$

3.5.8 Vorbemerkungen zu separablen Banachräumen

Sei $(E, \|\cdot\|)$ ein separabler Banachraum und $(\Omega, \mathfrak{E}, \mathcal{P})$ ein Wahrscheinlichkeitsraum. Dann nennen

$$L_1(\Omega, \mathcal{P}, E) := \left\{ f : \Omega \rightarrow E \mid f : \Omega \rightarrow E \text{ messbar, } \underbrace{\int_{\Omega} \|f(\omega)\| \mathcal{P}(d\omega)}_{\|f\|_1} < \infty \right\}$$

Dabei ist $\|\cdot\|_1$ auf L_1 tatsächlich eine Norm, insofern man die Elemente in L_1 als Äquivalenzklassen von

$$f \sim g \Leftrightarrow f = g \pmod{0}$$

betrachtet. Für Funktionenfolge $(f_k)_{k \in \mathbb{N}} \subset L_1(\Omega, \mathcal{P}, E)$ mit

$$\sum_{k=1}^{\infty} \|f_k\|_1 < \infty$$

existiert

$$\sum_{k=1}^{\infty} f_k$$

als Grenzwert in $(L_1, \|\cdot\|_1)$ und somit auch in Wahrscheinlichkeit. Sind ferner $X^j : \Omega \rightarrow E$ unabhängige Zufallsvariablen, so gilt die Äquivalenz:

$$\sum_{j=1}^{\infty} X^j \text{ ex. } \mathcal{P}\text{-fast überall} \Leftrightarrow \sum_{j=1}^{\infty} X^j \text{ konvergent in Wahrscheinlichkeit}$$

3.5.9 Theorem: Reihenkonvergenz beim Wiener-Prozess

Sei $(\Omega, \mathfrak{G}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $\xi_{j,k} \sim \mathcal{N}_{0,1}$, $j \in \mathbb{N}_0$, $k = 0, \dots, 2^j - 1$ unabhängig, identisch, standardnormalverteilt auf Ω . Dazu seien X^j die Glieder in der Faber-Schader Reihendarstellung des Wiener-Prozesses wie in (3.5.7.3). Dann existiert eine Teilmenge $\Omega_0 \in \mathfrak{G}(\Omega)$ mit $\mathcal{P}(\Omega_0) = 1$ so dass

$$\sup_{0 \leq t \leq 1} \left| W_t(\omega) - \sum_{j=-1}^N X_t^j(\omega) \right| \xrightarrow{N \rightarrow \infty} 0 \quad , \quad \omega \in \Omega_0$$

das heißt die Pfade konvergieren fast-sicher gleichmäßig gegen die Pfade des Wiener-Prozesses

$$W_t = \sum_{j=-1}^{\infty} X_t^j$$

Beweis: Betrachten den separablen Banachraum $(\mathcal{C}[0, 1], \|\cdot\|_{\infty})$ und die Pfade $X^j : \Omega \rightarrow \mathcal{C}[0, 1]$. Dann ist nach (3.5.7.4)

$$\underbrace{\mathbb{E} \|X^j\|_{\infty}}_{\|X^j\|_1 \text{ vgl. (3.5.8)}} \leq \underbrace{C}_{\text{const}} \cdot 2^{-\frac{j}{2}} \sqrt{j}$$

und insbesondere

$$\sum_{j=-1}^{\infty} \|X^j\|_1 < \infty$$

Nach 3.5.8 geht also

$$\sum_{j=-1}^N X^j \xrightarrow{N \rightarrow \infty} W$$

in Wahrscheinlichkeit und somit \mathcal{P} -fast sicher in $\mathcal{C}[0, 1]$.

□

Folgerung: Nach obigem Theorem ist

$$\sum_{j=1}^{\infty} X_t^j(\omega)$$

für fast alle $\omega \in \Omega$ gleichmäßiger Grenzwert stetiger Funktionen, das heißt

$$t \mapsto \sum_{j=1}^{\infty} X_t^j(\omega) \quad , \quad t \in [0, 1]$$

ist fast sicher stetig. Haben also eine Version des Wiener-Prozesses mit fast-sicher stetigen Pfaden gefunden.

4 Empirische Prozesse

4.0.10 Aufgabenstellung und Definition

Betrachtet seien die zufälligen, unabhängigen und identisch verteilten reellen Zufallsgrößen X_1, X_2, \dots auf $(\Omega, \mathfrak{G}, \mathcal{P})$. Gesucht sei nun das Verteilungsgesetz der X_j , sprich die Wahrscheinlichkeiten

$$\mathcal{P}(X_j \in [a, b]) \quad , \quad a < b \in \mathbb{R}$$

bzw. die Verteilungsfunktion²⁸

$$F : \mathbb{R} \rightarrow [0, 1] \quad , \quad F(t) := \underbrace{\mathcal{P}(\{X_j \leq t\})}_{\mathcal{P}_X((-\infty, t])}$$

Man nennt dabei

$$F_n^*(t, \omega) := \frac{1}{n} \# \{j \leq n : X_j(\omega) \leq t\} \quad , \quad t \in \mathbb{R}, \omega \in \Omega \quad (4.0.10.1)$$

empirische Verteilungsfunktion n -ter Ordnung bzgl. F . Sind z.B. $X_1(\omega) \leq X_2(\omega) \leq \dots$ so gilt

$$F_n^*(t, \omega) = \frac{k}{n} \quad \text{mit } t \in [X_k(\omega), X_{k+1}(\omega))$$

Beachte dass $(F_n^*(t))_{t \in \mathbb{R}}$ ein stochastischer Prozess ist mit Werten $0, \frac{1}{n}, \frac{2}{n}, \dots, 1$, dessen Pfade Verteilungsfunktionen sind. Das zugehörige Wahrscheinlichkeitsmaß

$$\mathcal{P}_n^*(\omega) := \frac{1}{n} \sum_{j=1}^n \delta_{X_j(\omega)} \quad , \quad \omega \in \Omega \quad (4.0.10.2)$$

auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ heißt (für ein festes ω) *empirisches Verteilungsgesetz*²⁹. So kann man schreiben

$$F_n^*(t, \omega) = \mathcal{P}_n^*(\omega)((-\infty, t])$$

Sei nun t festgehalten, dazu die Folge unabhängiger, identisch verteilter Zufallsvariablen

$$Y_j := 1_{(-\infty, t]}(X_j) \quad , \quad j \in \mathbb{N}$$

Dann gilt:

$$\mathcal{P}(Y_j = 0) = \mathcal{P}(X_j > t) = 1 - F(t)$$

$$\mathcal{P}(Y_j = 1) = F(t)$$

das heißt die $Y_j \sim \mathcal{B}_{1, F(t)}$ sind Binomialverteilt mit Parametern $1, F(t)$. Wegen

$$F_n^*(t) = \frac{1}{n} \sum_{j=1}^n Y_j$$

ist $nF_n^*(t) \sim \mathcal{B}_{n, F(t)}$ Binomialverteilt mit Parametern $n, F(t)$ und es gilt

$$\mathbb{E}F_n^*(t) = \frac{1}{n} (n \cdot F(t)) = F(t) \quad , \quad n \in \mathbb{N} \quad (4.0.10.3)$$

(empirische Verteilungsfunktion ist *im Mittel richtig*). Insbesondere ist

$$\mathbb{E}Y_j = F(t)$$

Ähnlich ergibt sich die Varianz:

$$\text{Var}(F_n^*(t)) = \frac{1}{n^2} \sum_{j=1}^n \text{Var}(Y_j) = \frac{1}{n^2} n F(t)(1 - F(t)) = \frac{1}{n} F(t)(1 - F(t))$$

²⁸Eine rechtsseitig stetige, monoton wachsende Funktion $F : \mathbb{R} \rightarrow [0, 1]$ mit

$$\lim_{t \rightarrow -\infty} F(t) = 0, \quad \lim_{t \rightarrow \infty} F(t) = 1$$

heißt *Verteilungsfunktion*. Zu jeder Verteilungsfunktion F existiert ein eindeutig bestimmtes Wahrscheinlichkeitsmaß \mathcal{P}_X auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ so dass $\mathcal{P}_X((-\infty, t]) = F(t)$. Umgekehrt definiert jedes Wahrscheinlichkeitsmaß auf diese Weise eine Verteilungsfunktion.

²⁹Keine Zufallsvariable!

4.0.11 Satz: Verteilungsgesetz der empirischen Verteilungsfunktion

Seien X_1, X_2, \dots identisch, unabhängig verteilte reelle Zufallsvariablen auf $(\Omega, \mathfrak{E}, \mathcal{P})$ mit der Verteilungsfunktion F und empirischer Verteilungsfunktion F_n^* , $n \in \mathbb{N}$. Dann gilt:

1. Äquivalenzinvarianz: Die endlich-dimensionalen Verteilungen von $(F_n^*)_{t \in \mathbb{R}}$ hängen nur von der Verteilung der X_j ab³⁰.
2. Für festes $n \in \mathbb{N}$ und $t \in \mathbb{R}$ gilt:

$$\mathcal{P}\left(\left\{F_n^*(t) = \frac{k}{n}\right\}\right) = \binom{n}{k} \cdot F(t)^k \cdot (1 - F(t))^{n-k}$$

3. Für festes $t \in \mathbb{R}$ geht

$$F_n^*(t) \xrightarrow{n \rightarrow \infty} F(t)$$

punktweise \mathcal{P} -fast-überall.

4. Für festes $t \in \mathbb{R}$ geht

$$\sqrt{n} (F_n^*(t) - F(t)) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}_{0, F(t)(1-F(t))}$$

in Verteilung.

Beweis:

1. Setzen

$$S_t : \mathbb{R}^n \rightarrow \mathbb{R}, (x_1, \dots, x_n) \mapsto \sum_{j=1}^n \mathbf{1}_{(-\infty, t]}(x_j)$$

Für $A \in \mathcal{B}(\mathbb{R}^n)$ und $t_1, \dots, t_m \in \mathbb{R}$ ist dann

$$\begin{aligned} \mathcal{P}_{(F_n^*(t_1), \dots, F_n^*(t_m))}(A) &= \mathcal{P}\left(\left\{\omega : \left[\overbrace{S_{t_1}(X_1(\omega), \dots, X_n(\omega))}^{F_n^*(t_1)}, \dots, \overbrace{S_{t_m}(X_1(\omega), \dots, X_n(\omega))}^{F_n^*(t_m)}\right] \in nA\right\}\right) \\ &= \mathcal{P}_{(S_{t_1}, \dots, S_{t_m}) \circ (X_1, \dots, X_n)}(nA) = \mathcal{P}_{(X_1, \dots, X_n)} \circ (S_{t_1}, \dots, S_{t_m})^{-1}(nA) \end{aligned}$$

2. Klar.

3. Die Zufallsvariablen $Y_j := \mathbf{1}_{(-\infty, t]}(X_j)$ sind unabhängig, identisch verteilt. Nach dem starken Gesetz der großen Zahlen (SLLN)³¹ gilt also

$$F_n^*(t) = \frac{1}{n} \sum_{j=1}^n Y_j \xrightarrow[n \rightarrow \infty]{\text{SLLN}} \underbrace{\mathbb{E}Y_j}_{F(t)}$$

fast sicher.

4. Die Behauptung folgt nach dem zentralen Grenzwertsatz 2.2.4:

$$\sqrt{n} (F_n^*(t) - F(t)) = \frac{1}{\sqrt{n}} \sum_{j=1}^n (Y_j - \overbrace{\mathbb{E}Y_j}^{F(t)}) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}_{0, \text{Var}(Y_j)} = \mathcal{N}_{0, F(t)(1-F(t))}$$

□

³⁰Insbesondere macht es Sinn von der empirischen Verteilungsfunktion von F , ohne Angabe von Zufallsvariablen X_1, X_2, \dots zu reden, insofern nur die endlich-dimensionalen Verteilungen bzw. die Verteilung der Pfade von F_n^* betrachtet werden.

³¹Für unabhängig, identisch verteilte reelle Zufallsvariablen Y_1, Y_2, \dots mit $\mathbb{E}Y_j < \infty$ konvergiert

$$\bar{Y}_n := \frac{1}{n} \sum_{j=1}^n Y_j$$

fast sicher gegen $\mathbb{E}Y_j$:

$$\bar{Y}_n \xrightarrow[n \rightarrow \infty]{\text{fast überall}} \mathbb{E}Y_j$$

4.0.12 Definition: Empirischer Prozess

Seien X_1, X_2, \dots unabhängig, identisch verteilte reelle Zufallsgrößen auf $(\Omega, \mathfrak{E}, \mathcal{P})$ mit Verteilungsfunktion F und empirischer Verteilungsfunktion F_n^* , $n \in \mathbb{N}$. Dann heißt der zufällige Prozess

$$E_n^*(t) := \sqrt{n}(F_n^*(t) - F(t)) \quad , \quad t \in \mathbb{R}$$

empirischer Prozess zur Verteilungsfunktion F von n -ter Ordnung.

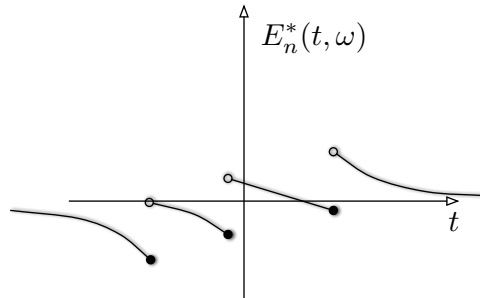


Abbildung 4.1: Typischer Verlauf eines Pfades eines empirischen Prozesses.

Bemerkung: Nach Satz 4.0.11 (4) geht für beliebige $u, t \in \mathbb{R}$:

$$\mathcal{P}(E_n^*(t) \leq u) \xrightarrow{n \rightarrow \infty} \underbrace{\frac{1}{\sqrt{2\pi\sigma_t}} \int_{-\infty}^u e^{-\frac{x^2}{2\sigma_t^2}} dx}_{\mathcal{N}_{0, \sigma_t^2}((-\infty, u])} \quad , \quad \sigma_t := \sqrt{F(t)(1-F(t))}$$

Beispiel: Betrachtet sei die Verteilungsfunktion der Gleichverteilung auf $[0, 1]$:

$$G(t) := \begin{cases} 0 & : t \leq 0 \\ t & : 0 \leq t \leq 1 \\ 1 & : t \geq 1 \end{cases}$$

Dazu seien u_1, u_2, \dots Verteilt gemäß G . Dann ist

$$U_n^*(t) := \sqrt{n}(G_n^*(t) - G(t)) \quad , \quad t \in \mathbb{R}$$

genau der empirische Prozess der Gleichverteilung auf $[0, 1]$.

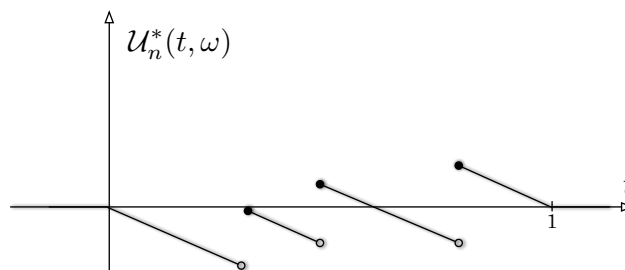


Abbildung 4.2: Typischer Verlauf eines Pfades des zur Gleichverteilung gehörigen empirischen Prozesses ($n = 3$).

4.0.13 Endlich dimensionale Verteilungen empirischer Prozesse

Gegeben sei die Verteilungsfunktion $F : \mathbb{R} \rightarrow [0, 1]$ und X_1, \dots, X_m unabhängig, identisch verteilt gemäß F , dazu die empirische Verteilungsfunktion F_n^* und der empirische Prozess E_n^* . Zu $-\infty < t_1 < t_2 < \dots < t_m < \infty$ ist nun das Verteilungsgesetz der $(E_n^*(t_1), \dots, E_n^*(t_m))$ gesucht. Dabei nimmt $E_n^*(t_r)$ prinzipiell die Werte

$$\underbrace{\sqrt{n} \cdot \left[\frac{k}{n} - F(t_r) \right]}_{\gamma_{k,r}}, \quad k = 0, \dots, n$$

Für $k_1, \dots, k_m \in \{0, \dots, n\}$ gilt dann

$$\begin{aligned} \mathcal{P}(\{E_n^*(t_1) = \gamma_{k_1,1}, \dots, E_n^*(t_m) = \gamma_{k_m,m}\}) &= \mathcal{P}\left(\underbrace{\left\{F_n^*(t_1) = \frac{k_1}{n}, \dots, F_n^*(t_m) = \frac{k_m}{n}\right\}}_{\substack{0 \text{ falls nicht} \\ 0 \leq k_1 \leq k_2 \leq \dots \leq k_m \leq n \\ (\text{da } t_1 \leq \dots \leq t_m)}}\right) \\ &= \mathcal{P}(\{\#\{j : X_j \leq t_1\} = k_1 \wedge \#\{j : t_{i-1} < X_j \leq t_i\} = k_i - k_{i-1}, i = 2, \dots, m \wedge \#\{j : t_m < X_j\} = n - k_m\}) \end{aligned}$$

$$\stackrel{\text{Multinomial-}}{=} \frac{n!}{k_1!(k_2 - k_1)! \dots (n - k_m)!} \cdot F(t_1)^{k_1} (F(t_2) - F(t_1))^{k_2 - k_1} \dots (1 - F(t_m))^{n - k_m}$$

Fazit: Die endlich-dimensionalen Verteilungen von E_n^* hängen nur von F ab! Insbesondere kann somit jeder Verteilungsfunktion F ein empirischer Prozess $(E_n^*)_{t \in \mathbb{R}}$ zugeordnet werden, der bis auf Äquivalenz eindeutig bestimmt ist.

Anders gesagt: Je zwei Folgen unabhängiger, gemäß F identisch verteilter reellwertiger Zufallsgrößen X_1, X_2, \dots und $\tilde{X}_1, \tilde{X}_2, \dots$, erzeugen jeweils empirische Prozesse E_n^*, \tilde{E}_n^* die gleiche endlich-dimensionale Verteilungen besitzen, das heißt $E_n^* \stackrel{d.}{=} \tilde{E}_n^*$.

4.0.14 Definition: Kanonische Darstellung

Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion. Dann setzen zu $x \in (0, 1)$:

$$F^-(x) := \inf \{t \in \mathbb{R} : F(t) \geq x\}$$

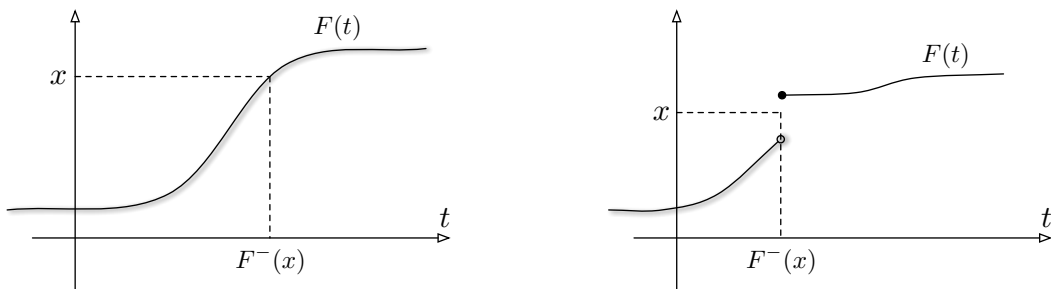


Abbildung 4.3: Zur Definition von F^- für zwei typische Verteilungsfunktionen F .

4.0.15 Lemma über F^-

Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion, dazu F^- wie in 4.0.14 definiert. Sei ferner \mathcal{U} eine auf $[0, 1]$ gleichverteilte Zufallsvariable und

$$X := F^-(\mathcal{U})$$

Dann gilt:

$$\mathcal{P}(X \leq t) = F(t)$$

das heißt $X \sim F$.

Beweis: Sei $\mathcal{U} : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow \mathbb{R}$ gleichverteilt auf $[0, 1]$. Dann gilt für $\omega \in \Omega, s \in \mathbb{R}$:

$$X(\omega) \leq s \Leftrightarrow F^-(\mathcal{U}(\omega)) \leq s \Leftrightarrow \inf \{t \in \mathbb{R} : F(t) \geq \mathcal{U}(\omega)\} \leq s$$

Offensichtlich folgt aus $F(s) \geq \mathcal{U}(\omega)$ stets $\inf \{t \in \mathbb{R} : F(t) \geq \mathcal{U}(\omega)\} \leq s$. Doch auch die Umkehrung gilt: Ist nämlich $F^-(\mathcal{U}(\omega)) \leq s$ so existiert eine Folge $(t_n) \subset \mathbb{R}$ mit $F(t_n) \geq \mathcal{U}(\omega)$ und $t_n \searrow F^-(\mathcal{U}(\omega))$. Da F rechtsseitig stetig ist, folgt dann

$$F(s) \stackrel{\substack{F \\ \text{monoton} \\ \text{wachsend}}}{\geq} F(F^-(\mathcal{U}(\omega))) = \lim_{n \rightarrow \infty} F(t_n) \geq \mathcal{U}(\omega)$$

das heißt

$$X(\omega) \leq s \Leftrightarrow F(s) \geq \mathcal{U}(\omega)$$

Somit ist

$$\mathcal{P}(X \leq s) = \mathcal{P}(\mathcal{U} \leq F(s)) = F(s)$$

□

4.0.16 Satz: Erzeugung empirischer Prozesse durch Gleichverteilungen

Sei F eine Verteilungsfunktion und \mathcal{U}_n^* der empirische Prozess der Gleichverteilung auf $[0, 1]$. Dann ist

$$(\mathcal{U}_n^*(F(t)))_{t \in \mathbb{R}}$$

ein von F erzeugter empirischer Prozess n -ter Ordnung.

Beweis: $\mathcal{U}_1, \mathcal{U}_2$ seien unabhängig, gleichmäßig verteilt auf $[0, 1]$, dazu

$$X_j := F^-(\mathcal{U}_j)$$

Dann sind nach Lemma 4.0.15 die Zufallsgrößen X_1, X_2 unabhängig mit Verteilungsfunktion F . Weiterhin

$$\begin{aligned} \mathcal{U}_n^*(F(t)) &= \sqrt{n} \left[\frac{1}{n} \# \{j \leq n : \mathcal{U}_j \leq F(t)\} - F(t) \right] \\ &= \sqrt{n} \left[\frac{1}{n} \# \{j \leq n : X_j \leq t\} - F(t) \right] \quad \left| \quad \{ \omega : \mathcal{U}_j(\omega) \leq F(t) \} = \left\{ \omega : \underbrace{F^-(\mathcal{U}_j(\omega))}_{X_j(\omega)} \leq t \right\} \right. \\ &= \sqrt{n} [F_n^*(t) - F(t)] =: E_n^*(t) \end{aligned}$$

□

Folgerung: Ist F stetig so ist $F(\mathbb{R}) = [0, 1]$ und es gilt

$$\sup_{t \in \mathbb{R}} \sqrt{n} \cdot |F_n^*(t) - F(t)| = \sup_{t \in \mathbb{R}} |E_n^*(t)| \stackrel{\text{d.}}{=} \sup_{0 \leq s \leq 1} |\mathcal{U}_n^*(s)|$$

4.1 Konvergenz der endlich dimensionalen Verteilungen

4.1.1 Satz: Momente von empirischen Prozessen

Sei F eine Verteilungsfunktion, dazu der empirische Prozess E_n^* . Für $t_1, \dots, t_m \in \mathbb{R}$ gilt:

1. $\mathbb{E}(F_n^*(t_1), \dots, F_n^*(t_m)) = (F(t_1), \dots, F(t_m))$
2. $\text{Cov}(E_n^*(t), E_n^*(s)) = F(t \wedge s) - F(t) \cdot F(s)$ für $t, s \in \mathbb{R}$.

Beweis:

1. Folgt direkt aus $\mathbb{E}F_n^*(t) = F(t)$ (4.0.10.3).
2. Seien X_1, X_2, \dots unabhängig, identisch verteilt gemäß F . Dann gilt zunächst:

$$\begin{aligned} \mathbb{E}(F_n^*(t) \cdot F_n^*(s)) &= \frac{1}{n^2} \sum_{i,j=1}^n \underbrace{\mathbb{E}(1_{(-\infty,t]}(X_i) \cdot 1_{(-\infty,s]}(X_j))}_{\begin{cases} F(t)F(s) & : i \neq j \\ F(t \wedge s) & : i = j \end{cases}} \\ &= \frac{1}{n^2} [nF(t \wedge s) + (n^2 - n)F(t)F(s)] = \frac{1}{n}F(t \wedge s) + \left(1 - \frac{1}{n}\right) \cdot F(t)F(s) \end{aligned}$$

Somit folgt

$$\text{Cov}(E_n^*(t), E_n^*(s)) \stackrel{\mathbb{E}E_n^*=0}{=} \mathbb{E}(E_n^*(t) \cdot E_n^*(s)) = n [\mathbb{E}(F_n^*(t)F_n^*(s)) - F(t) \cdot F(s)] = F(t \wedge s) - F(t) \cdot F(s)$$

□

Folgerungen:

1. Für die Kovarianzmatrix gilt:

$$\text{Cov}[(E_n^*(t_1), \dots, E_n^*(t_m))] = [F(t_i \wedge t_j) - F(t_i) \cdot F(t_j)]_{i,j=1}^m, \quad t_1, \dots, t_m \in \mathbb{R}$$

2. Für auf $[0, 1]$ gleichverteilte Zufallsgröße \mathcal{U} gilt:

$$\text{Cov}(\mathcal{U}_n^*(t), \mathcal{U}_n^*(s)) = t \wedge s - ts, \quad 0 \leq t \leq s \leq 1$$

4.1.2 Satz über die Brownsche Brücke

Sei $(B_t)_{0 \leq t \leq s}$ der Prozess der Brownschen Brücke (vgl. Abschnitt 3.3.6), F irgendeine Verteilungsfunktion mit empirischem Prozess E_n^* und $Y_t := B_{F(t)}$, $t \in \mathbb{R}$. Dann geht:

$$E_n^* \xrightarrow[n \rightarrow \infty]{\text{f.d.d.}} Y$$

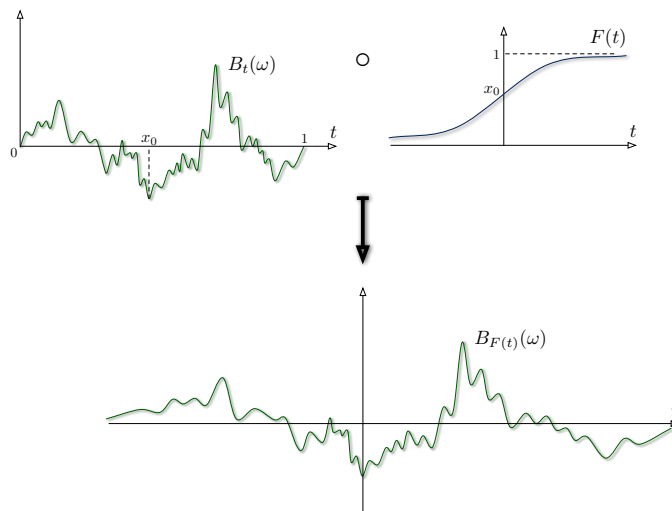


Abbildung 4.4: Brownsche Brücke unter Zeittransformation $t \mapsto F(t)$. Abgebildet ist typischer Pfad des entsprechenden Prozesses.

Beweis: Seien $t_1, \dots, t_m \in \mathbb{R}$ beliebig und X, X_1, X_2, \dots unabhängig, identisch verteilt gemäß F . Es gilt zunächst

$$E_n^*(t) = \frac{1}{\sqrt{n}} [nF_n^*(t) - nF(t)] = \frac{1}{\sqrt{n}} \sum_{k=1}^m [1_{(-\infty, t]}(X_k) - \mathbb{E}1_{(-\infty, t]}(X_k)] \quad , \quad t \in \mathbb{R}$$

und somit

$$(E_n^*(t_1), \dots, E_n^*(t_m)) = \frac{1}{\sqrt{n}} \sum_{k=1}^m [Z_k - \mathbb{E}Z_1] \xrightarrow[\text{d}]{n \rightarrow \infty} \mathcal{N}_{0, \text{Cov}(Z)} \quad \left| \quad Z_k := \underbrace{(1_{(-\infty, t_1]}(X_k), \dots, 1_{(-\infty, t_m]}(X_k))}_{\text{identisch verteilte Vektoren in } \mathbb{R}^m}$$

wobei

$$Z := (1_{(-\infty, t_1]}(X), \dots, 1_{(-\infty, t_m]}(X))$$

$$\begin{aligned} \text{Cov}(Z^i, Z^j) &= \frac{\text{Cov}[1_{(-\infty, t_i]}(X), 1_{(-\infty, t_j]}(X)]}{\mathbb{E}(1_{(-\infty, t_i]}(X) - \mathbb{E}1_{(-\infty, t_i]}(X)) \cdot (1_{(-\infty, t_j]}(X) - \mathbb{E}1_{(-\infty, t_j]}(X))} \\ &= F(t_i \wedge t_j) - F(t_i) \cdot F(t_j) \end{aligned}$$

Andererseits ist der Vektor $(Y_{t_1}, \dots, Y_{t_m})$ offensichtlich Gaußsch, zentriert und besitzt Kovarianzmatrix

$$\text{Cov}\left(\underbrace{Y_{t_i}}_{B_{F(t_i)}}, \underbrace{Y_{t_j}}_{B_{F(t_j)}}\right) \stackrel{(3.3.6)}{=} \underbrace{F(t_i) \wedge F(t_j)}_{\substack{F(t_i \wedge t_j) \\ \text{da } F \\ \text{monoton wachsend}}} - F(t_i) \cdot F(t_j)$$

spricht $(Y_{t_1}, \dots, Y_{t_m}) \sim \mathcal{N}_{0, \text{Cov}(Z)}$. Somit geht insbesondere

$$(E_n^*(t_1), \dots, E_n^*(t_m)) \xrightarrow[\text{f.d.d.}]{n \rightarrow \infty} (Y_{t_1}, \dots, Y_{t_m})$$

□

4.2 Das Theorem von Glivenko und Cantelli (1933)

4.2.1 Vorbetrachtung

Zu Verteilungsfunktion F sei $F_n^* : \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ die empirische Verteilungsfunktion und

$$\mathcal{F}_n(\omega) := \sup_{t \in \mathbb{R}} |F_n^*(t, \omega) - F(t)| \quad (4.2.1.1)$$

Dann ist $\mathcal{F}_n : \Omega \rightarrow \mathbb{R}$ tatsächlich messbar, denn aufgrund der rechtsseitigen Stetigkeit von $F_n^*(\cdot, \omega)$ ist

$$F_n^*(s, \omega) = \inf \{F_n^*(t, \omega) : t \geq s, t \in \mathbb{Q}\}$$

und somit

$$\mathcal{F}_n(\omega) = \sup_{t \in \mathbb{Q}} \underbrace{|F_n^*(t, \omega) - F(t)|}_{\text{messbar}}$$

Nach 4.0.11 geht dabei $F_n^*(t) \xrightarrow{n \rightarrow \infty} F(t)$ fast sicher. Nach folgendem Theorem 4.2.2 ist sogar fast sicher eine gleichmäßige Konvergenz vorhanden!

4.2.2 Theorem von Glivenko-Cantelli

Sei F eine Verteilungsfunktion, dazu die empirische Verteilungsfunktion F_n^* und \mathcal{F}_n wie in (4.2.1.1) definiert. Dann gilt

$$\mathcal{P}(\mathcal{F}_n \xrightarrow{n \rightarrow \infty} 0) = 1$$

das heißt $F_n^*(\cdot, \omega)$ konvergiert gleichmäßig gegen $F(\cdot)$ fast-sicher.

Beweis: Seien X, X_1, X_2, \dots unabhängig, identisch verteilt gemäß F , dazu $F_n^* : \mathbb{R} \times (\Omega, \mathfrak{C}, \mathcal{P}) \rightarrow \mathbb{R}$. Zu $t \in \mathbb{R}$ existiert nach 4.0.11 eine Nullmenge $N_t \in \mathfrak{C}$, $\mathcal{P}(N_t) = 0$ mit

$$F_n^*(t, \omega) \xrightarrow{n \rightarrow \infty} F(t) \quad , \quad \omega \notin N_t \tag{4.2.2.1}$$

Eigentlich existiert sogar eine Nullmenge $M_t \in \mathfrak{C}$, $\mathcal{P}(M_t) = 0$ mit

$$\underbrace{F_n^*(t-0, \omega)}_{\lim_{\tau \nearrow t} F_n^*(\tau, \omega)} \xrightarrow{n \rightarrow \infty} F(t-0) \quad , \quad \omega \notin M_t \tag{4.2.2.2}$$

denn

$$F_n^*(t-0) = \frac{1}{n} \# \{j \leq n : X_j < t\} = \frac{1}{n} \sum_{j=1}^n \underbrace{1_{(-\infty, t)}(X_j)}_{\substack{\text{unabhängig} \\ \text{identisch} \\ \text{verteilt}}} \xrightarrow[n \text{ fast sicher (SLLN)}]{n \rightarrow \infty} \mathbb{E}1_{(-\infty, t)}(X) = \mathcal{P}(X < t) = F(t-0)$$

Betrachten nun

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

Für $s := F^-(x)$ gilt stets $F(s) \geq x$, denn

$$\exists t_n \searrow s \Rightarrow F(t_n) \geq x \xrightarrow[F \text{ rechtsseitig stetig}]{\implies} F(s) \geq x$$

Ist andererseits $t < \overbrace{F^-(x)}^s$, so gilt per Konstruktion $F(t) < x$, also

$$\lim_{t \nearrow s} F(t) \leq x$$

Somit

$$\underbrace{\lim_{t \nearrow s} F(t)}_{F(s-0)} \leq x \leq F(s) \quad , \quad s := F^-(x) \quad , \quad x \in (0, 1) \tag{4.2.2.3}$$

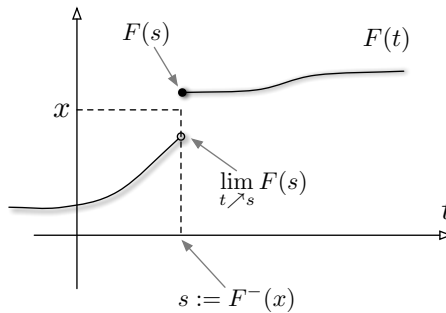


Abbildung 4.5: Zum Beweis von Glivenko-Cantelli.

Sei nun $m \geq 1$ festgehalten, und

$$s_{k,m} := F^-\left(\frac{k}{m}\right) \quad , \quad k = 1, \dots, m-1$$

dann gilt nach (4.2.2.3)

$$F(s_{k,m} - 0) \leq \frac{k}{m} \leq F(s_{k,m})$$

Speziell:

- **Fall** $k = 1$: Dann folgt

$$\underbrace{\mathcal{P}(X < s_{1,m})}_{F(s_{1,m}-0)} \leq \frac{1}{m} \quad (4.2.2.4)$$

- **Fall** $k = m - 1$, dann gilt

$$F(s_{m-1,m}) \geq 1 - \frac{1}{m} \quad (4.2.2.5)$$

- **Fall** $k = 2, \dots, m - 1$. Dann erhält man

$$F(s_{k,m} - 0) \leq \frac{k}{m} = \frac{k-1}{m} + \frac{1}{m} \leq F(s_{k-1,m}) + \frac{1}{m}$$

das heißt

$$\underbrace{F(s_{k,m} - 0) - F(s_{k-1,m})}_{\mathcal{P}(s_{k-1,m} < X < s_{k,m})} \leq \frac{1}{m} \quad (4.2.2.6)$$

Außerdem: Mit

$$\mathcal{F}_{m,n} := \sup_{1 \leq k \leq m-1} \{|F_n^*(s_{k,m}) - F(s_{k,m})| \vee |F_n^*(s_{k,m} - 0) - F(s_{k,m} - 0)|\} \quad , \quad m, n \in \mathbb{N}$$

und

$$N := \bigcup_{m=1}^{\infty} \bigcup_{k=1}^{m-1} [N_{s_{k,m}} \cup M_{s_{k,m}}]$$

gilt nach (4.2.2.1) und (4.2.2.2)

$$\lim_{n \rightarrow \infty} D_{m,n}(\omega) = 0 \quad \forall \omega \notin N \quad (4.2.2.7)$$

wobei $\mathcal{P}(N) = 0$. Sei nun $s \in \mathbb{R}$ beliebig und $m \in \mathbb{N}$.

- **Fall:** $s < s_{1,m}$, dann gilt:

$$F_n^*(s) \leq F_n^*(s_{1,m} - 0) \leq F(s_{1,m} - 0) + \mathcal{F}_{m,n} \stackrel{(4.2.2.4)}{\leq} F(s) + \frac{1}{m} + \overbrace{\mathcal{F}_{m,n}}^{\geq 0}$$

und

$$F_n^*(s) \geq 0 \stackrel{(4.2.2.4)}{\geq} \underbrace{F(s_{1,m} - 0)}_{\geq F(s)} - \frac{1}{m} \geq F(s) - \frac{1}{m} - \overbrace{\mathcal{F}_{m,n}}^{\geq 0}$$

- **Fall:** $s > s_{m-1,m}$, dann gilt:

$$F_n^*(s) \leq 1 \stackrel{(4.2.2.5)}{\leq} \underbrace{F(s_{m-1,m})}_{\leq F(s)} + \frac{1}{m} \leq F(s) + \frac{1}{m} + \mathcal{F}_{m,n}$$

und

$$F_n^*(s) \geq F_n^*(s_{m-1,m}) \geq \underbrace{F(s_{m-1,m})}_{\geq 1 - \frac{1}{m}} - \mathcal{F}_{m,n} \geq \overbrace{F(s)}^{\leq 1} - \frac{1}{m} - \mathcal{F}_{m,n} \stackrel{(4.2.2.5)}{\geq}$$

- **Fall:** $s_{k-1,m} \leq s < s_{k,m}$ für irgendein $k \in \{2, \dots, m-1\}$, dann gilt

$$F_n^*(s) \leq F_n^*(s_{k,m} - 0) \leq \underbrace{F(s_{k,m} - 0)}_{\substack{\leq F(s_{k-1,m}) + \frac{1}{m} \\ (4.2.2.6)}} + \mathcal{F}_{m,n} \leq F(s) + \frac{1}{m} + \mathcal{F}_{m,n}$$

und

$$F_n^*(s) \geq F_n^*(s_{k-1,m}) \geq \underbrace{F(s_{k-1,m})}_{\substack{\geq F(s_{k,m} - 0) - \frac{1}{m} \\ (4.2.2.6)}} - \mathcal{F}_{m,n} \geq F(s) - \frac{1}{m} - \mathcal{F}_{m,n}$$

In allen drei Fällen folgt dann

$$|F_n^*(s) - F(s)| \leq \frac{1}{m} + \mathcal{F}_{m,n} \quad , \quad s \in \mathbb{R}, \quad m \in \mathbb{N}$$

bzw.

$$\underbrace{\sup_{s \in \mathbb{R}} |F_n^*(s) - F(s)|}_{\mathcal{F}_n} \leq \frac{1}{m} + \mathcal{F}_{m,n} \quad , \quad m \in \mathbb{N} \quad (4.2.2.8)$$

Insbesondere gilt wegen (4.2.2.7) für $\omega \in N^c$:

$$\limsup_{n \rightarrow \infty} \mathcal{F}_n(\omega) \leq \frac{1}{m} \quad \forall m \in \mathbb{N}$$

also

$$\lim_{n \rightarrow \infty} \mathcal{F}_n(\omega) = 0 \quad , \quad \omega \in N^c$$

□

Bemerkung: Die Funktionenmenge

$$\mathcal{F} := \{1_{(-\infty, t]} : t \in \mathbb{R}\} \subseteq L_\infty(\mathbb{R})$$

erfüllt

$$\sup_{f \in \mathcal{F}} \left| \underbrace{\int_{\mathbb{R}} f d\mathcal{P}_n^*}_{f=1_{(-\infty, t]}} - \underbrace{\int_{\mathbb{R}} f d\mathcal{P}}_{F(t)} \right| \xrightarrow[n \rightarrow \infty]{\text{fast sicher}} 0$$

(vgl. Definition (4.0.10.2)). Man sagt: \mathcal{F} ist eine *Vapnik-Cernovenkis (VC) Klasse* [4].

4.3 Konvergenz der empirischen Prozesse

4.3.1 Vorbetrachtung

Gegeben sei die Verteilungsfunktion F mit empirischer Verteilungsfunktion F_n^* , dazu der empirische Prozess

$$E_n^*(t) := \sqrt{n} \cdot [F_n^*(t) - F(t)]$$

In 4.1.2 wurde gezeigt, dass

$$E_n^* \xrightarrow[n \rightarrow \infty]{\text{f.d.d.}} B_{F(\cdot)}$$

geht, wobei $(B_t)_{0 \leq t \leq 1}$ die Brownsche Brücke sei.

Ziel ist es nun zu zeigen:

$$E_n^* \xrightarrow[n \rightarrow \infty]{\text{d}} B_{F(\cdot)}$$

in irgendeinem metrischen $(E, d) \subseteq \mathbb{R}^T$, $T := \mathbb{R}$. Angesichts der Tatsache dass E_n^* typischerweise keine stetigen Pfade hat, stellt sich die Frage welche Mengen $E \subset \mathbb{R}^T$ in Frage kommen. So besitzt z.B. E_n^* beschränkte Pfade, doch ist

$$E := \{f \in \mathbb{R}^T : f \text{ beschränkt}\}$$

mit der Metrik $d(f, g) := \|f - g\|_\infty$ nicht separabel!

Die **Idee** ist nun

$$E_n^*(t) = \mathcal{U}_n^*(F(t))$$

(vgl. 4.0.16) in irgendeiner Weise durch einen Prozess mit stetigen Pfaden zu ersetzen. Ist F stetig, so reicht es, \mathcal{U}_n^* durch einen Prozess mit stetigen Pfaden zu *approximieren*.

4.3.2 Stetiger Ersatz zu \mathcal{U}_n^*

Sei G die Verteilungsfunktion der Gleichverteilung auf $[0, 1]$ mit empirischer Verteilungsfunktion

$$G_n^*(t) := \frac{1}{n} \# \{j \leq n : \mathcal{U}_j(\omega) \leq t\}$$

(mit $\mathcal{U}_1, \mathcal{U}_2, \dots \sim G$, iid auf $(\Omega, \mathfrak{S}, \mathcal{P})$, o.B.d.A. $\mathcal{U}_j : \Omega \rightarrow [0, 1]$) und empirischem Prozess

$$\mathcal{U}_n^*(t) := \sqrt{n} \cdot (G_n^*(t) - t) \quad , \quad 0 \leq t \leq 1$$

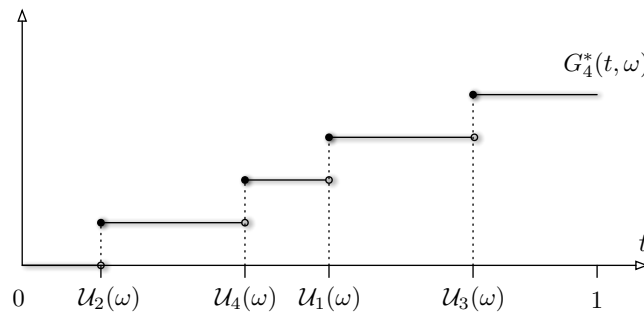


Abbildung 4.6: Typischer Pfad der empirischen Verteilungsfunktion G_4^* der Gleichverteilung.

Zu jedem $\omega \in \Omega$, $n \in \mathbb{N}$ seien o.B.d.A. $0 \leq \mathcal{U}_1(\omega) \leq \dots \leq \mathcal{U}_n(\omega) \leq 1$ umgeordnet. Definieren nun das (zufällige) Maß auf $[0, 1]$:

$$Q_n^*(B, \omega) := \frac{1}{n+1} \sum_{j=1}^{n+1} \frac{\lambda(B \cap [\mathcal{U}_{j-1}(\omega), \mathcal{U}_j(\omega)])}{\mathcal{U}_j(\omega) - \mathcal{U}_{j-1}(\omega)} \quad , \quad B \in \mathcal{B}([0, 1]) \quad \left| \quad \mathcal{U}_0(\omega) := 0, \mathcal{U}_{n+1}(\omega) := 1$$

mit der (zufälligen) Verteilungsfunktion

$$H_n^*(t, \omega) := Q_n^*([0, t], \omega) \quad , \quad 0 \leq t \leq 1 \tag{4.3.2.1}$$

Dabei nimmt $H_n^*(\cdot, \omega)$ an den Punkten $\mathcal{U}_j^*(\omega)$ die Werte $\frac{j}{n+1}$ an und verbindet diese linear (siehe Abbildung 4.7).

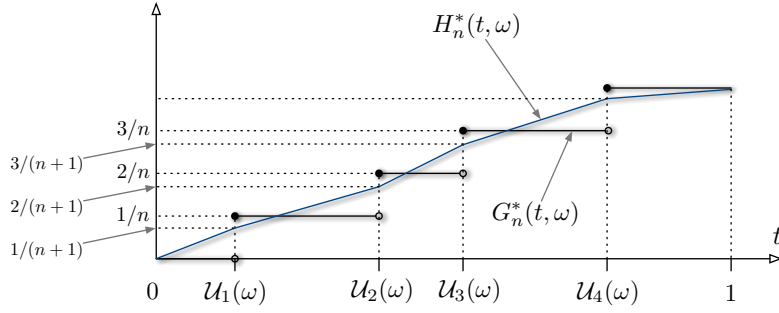


Abbildung 4.7: Typischer Pfad der empirischen Verteilungsfunktion $G_n^*(\cdot, \omega)$, dazu die entsprechende stetige Approximation $H_n^*(\cdot, \omega)$ für $n = 4$. Die Werte $U_j(\omega)$ sind hier o.B.d.A. geordnet.

Offensichtlich unterscheidet sich H_n^* von G_n^* an den Unstetigkeitsstellen von G_n^* am meisten, mit

$$G_n^*(U_j(\omega) - 0, \omega) \leq H_n^*(U_j(\omega), \omega) \leq G_n^*(U_j(\omega), \omega) \quad , \quad j = 1, \dots, n$$

Insbesondere gilt die Abschätzung

$$\left\{ \begin{array}{l} G_n^*(U_j) - H_n^*(U_j) \\ H_n^*(U_j) - G_n^*(U_j - 0) \end{array} \right\} = \left\{ \begin{array}{l} \frac{(n+1)j-jn}{n(n+1)} \\ \frac{nj-(n+1)(j-1)}{n(n+1)} \end{array} \right\} = \left\{ \begin{array}{l} \frac{j}{n(n+1)} \\ \frac{(n+1)-j}{n(n+1)} \end{array} \right\} \leq \frac{1}{n}$$

das heißt

$$\sup_{0 \leq t \leq 1} |H_n^*(t) - G_n^*(t)| \leq \frac{1}{n} \tag{4.3.2.2}$$

Für den *approximierten empirischen Prozess*

$$\mathcal{V}_n^*(t) := \sqrt{n} \cdot (H_n^*(t) - t)$$

gilt dann

$$\sup_{0 \leq t \leq 1} |\mathcal{V}_n^*(t) - \mathcal{U}_n^*(t)| \leq \frac{1}{\sqrt{n}} \quad , \quad n \in \mathbb{N}$$

Dabei ist zu beachten dass $(\mathcal{V}_n^*)_{0 \leq t \leq 1}$ stets stetige Pfade besitzt!

4.3.3 Satz: Konvergenz des approximierten empirischen Prozesses der Gleichverteilung

Sei $(\mathcal{U}_n^*)_{0 \leq t \leq 1}$ der empirische Prozess n -ter Ordnung der Gleichverteilung auf $[0, 1]$ und $(\mathcal{V}_n^*)_{0 \leq t \leq 1}$ die oben (4.3.2) eingeführte stetige Approximation. Dann geht letztere in f.d.d. gegen die Brownsche Brücke:

$$(\mathcal{V}_n^*)_{0 \leq t \leq 1} \xrightarrow[\text{f.d.d.}]{n \rightarrow \infty} (B_t)_{0 \leq t \leq 1}$$

Beweis: Seien $t_1, \dots, t_m \in [0, 1]$ beliebig, dazu die zufälligen Vektoren

$$Y_n := (\mathcal{V}_n^*(t_1), \dots, \mathcal{V}_n^*(t_m)) \quad , \quad Z_n := (\mathcal{U}_n^*(t_1), \dots, \mathcal{U}_n^*(t_m)) \quad , \quad n \in \mathbb{N}$$

Dann ist

$$\|Y_n - Z_n\|^2 = \sum_{j=1}^m |\mathcal{V}_n^*(t_j) - \mathcal{U}_n^*(t_j)|^2 \leq \frac{m}{n} \quad , \quad n \in \mathbb{N}$$

das heißt

$$\|Y_n - Z_n\| \xrightarrow{n \rightarrow \infty} 0$$

sicher. Zusammen mit $Z_n \xrightarrow[\text{d}]{n \rightarrow \infty} (B_{t_1}, \dots, B_{t_m})$ folgt dann nach 2.2.3 auch

$$Y_n \xrightarrow[\text{d}]{n \rightarrow \infty} (B_{t_1}, \dots, B_{t_m})$$

□

Ziel: Folgender Abschnitt wird als Ziel haben, die Konvergenz

$$(\mathcal{V}_n^*) \xrightarrow[n \rightarrow \infty]{d} B$$

in $(\mathcal{C}[0, 1], \|\cdot\|_\infty)$ zu beweisen. Nach 2.3.10 genügt es zu zeigen, dass

$$\{\mathcal{P}_{\mathcal{V}_n^*}\}_{n \in \mathbb{N}} \subseteq \mathfrak{M}_1[(\mathcal{C}[0, 1], \|\cdot\|_\infty)]$$

schwach-relativ kompakt ist.

4.3.4 Definition: Stetigkeitsmodul

Sei (T, d) ein metrischer Raum und $f \in \mathcal{C}(T)$. Dann heißt die Abbildung $\omega_f : (0, \infty) \rightarrow \mathbb{R}$ definiert durch

$$\omega_f(\delta) := \sup \{|f(t) - f(s)| : d(t, s) < \delta, t, s \in T\}, \quad \delta > 0$$

Stetigkeitsmodul von f auf T .

4.3.5 Eigenschaften des Stetigkeitsmoduls

Sei (T, d) ein metrischer Raum und $f \in \mathcal{C}(T)$, dazu das Stetigkeitsmodul ω_f . Dann:

1. $\omega_f : (0, \infty) \rightarrow \mathbb{R}$ ist monoton wachsend.
2. ω_f ist subadditiv, das heißt für $\delta, \delta' > 0$ gilt

$$\omega_f(\delta + \delta') \leq \omega_f(\delta) + \omega_f(\delta')$$

3. Ist $T_0 \subseteq T$ abzählbar und dicht in T , so gilt:

$$\omega_f(\delta) = \sup \{|f(t) - f(s)| : d(t, s) < \delta, t, s \in T_0\}, \quad \delta > 0$$

Insbesondere ist $\omega_f(\delta) : f \mapsto \omega_f(\delta)$ messbar von $\mathcal{C}(T)$ nach \mathbb{R} .³²

4. $f \in \mathcal{C}(T)$ ist genau dann gleichmäßig stetig, falls $\omega_f(\delta) \xrightarrow{\delta \searrow 0} 0$.

4.3.6 Satz von Arzela-Ascoli

Sei (T, d) ein kompakter, metrischer Raum. Eine Untermenge $K \subseteq \mathcal{C}(T)$ ist genau dann relativ kompakt, wenn:

- (a) K ist beschränkt, das heißt

$$\sup_{f \in K} \|f\|_\infty < \infty$$

- (b) K ist gleichgradig stetig, das heißt

$$\forall \varepsilon > 0 : \exists \delta > 0 : |f(t) - f(s)| < \varepsilon \quad \forall d(t, s) < \delta, f \in K$$

Bemerkung: Bedingung (b) ist genau dann erfüllt wenn

$$\sup_{f \in K} \omega_f(\delta) \xrightarrow{\delta \searrow 0} 0$$

³²Beachte dass $\zeta_t : \mathcal{C}(T) \rightarrow \mathbb{R}$, $\zeta_t : f \mapsto f(t)$ für jedes $t \in T$ stetig ist.

Beispiel: Typisches Beispiel einer beschränkten, aber nicht relativ kompakten Menge in $\mathcal{C}[0, 1]$ ist die Einheitskugel B_1 . So besitzt z.B. die Folge

$$(x_n)_{n \in \mathbb{N}} \subseteq B_1, \quad x_n(t) := \begin{cases} 0 & : t \leq \frac{1}{2} \\ (2t - 1)n/2 & : \frac{1}{2} \leq t \leq \frac{1}{2} + \frac{1}{n} \\ 1 & : \text{sonst} \end{cases}$$

keine in $\mathcal{C}[0, 1]$ konvergente Teilfolge.

4.3.7 Lemma über Wegbeschränktheit

Sei (E, d) ein metrischer Raum. Dann existiert ein $x_0 \in E$ mit

$$\forall \delta > 0 : \exists n \in \mathbb{N} : \forall x \in E : \exists x_1, \dots, x_n = x : d(x_i, x_{i-1}) < \delta, \quad i = 1, \dots, n \tag{4.3.7.1}$$

genau dann wenn

$$\forall \delta > 0 : \exists n \in \mathbb{N}, x_1, \dots, x_n \in E : \left[E \subseteq \bigcup_{i=1}^n B_\delta^o(x_i) \wedge \forall i_0 \in \{1, \dots, n-1\} : \exists i > i_0 : B_\delta^o(x_{i_0}) \cap B_\delta^o(x_i) \neq \emptyset \right]$$

Gegebenfalls nennen wir (E, d) *wegbeschränkt* und x_0 ein *Startpunkt* von E .

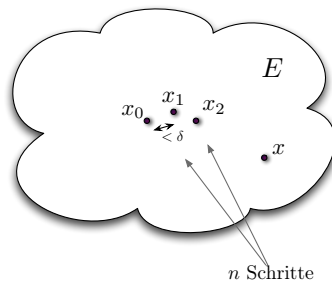


Abbildung 4.8: Zur Definition der Wegbeschränktheit.

Bemerke: Ist E wegbeschränkt, so ist jeder Punkt $x \in E$ auch Startpunkt.

4.3.8 Satz: Charakterisierung relativer Kompaktheit

Sei (T, d) ein wegbeschränkter (4.3.7), kompakter, metrischer Raum mit Startpunkt $t_0 \in T$. Dann ist $K \subseteq \mathcal{C}(T)$ genau dann relativ kompakt, wenn:

- (a) $\sup_{f \in K} |f(t_0)| < \infty$
- (b) $\sup_{f \in K} \omega_f(\delta) \xrightarrow{\delta \searrow 0} 0$

Beweis: Richtung "⇒". Nach Arzela-Ascoli 4.3.6 folgt unmittelbar Behauptung (a). Nach der Bemerkung im gleichen Satz, folgt auch Behauptung (b).

Richtung "⇐". Nach der Bemerkung in Arzela-Ascoli 4.3.6 genügt es zu zeigen: $\sup_{f \in K} \|f\|_\infty < \infty$.

Zu $\delta > 0$ sei $n_\delta \in \mathbb{N}$ wie in 4.3.7. Zu beliebigem $f \in K$ und $t \in T$, dazu $t_1, \dots, t_{n_\delta} := t$ mit $d(t_i, t_{i-1}) < \delta, i \in \{1, \dots, n_\delta\}$, gilt:

$$f(t) - f(t_0) = \sum_{j=1}^{n_\delta} [f(t_j) - f(t_{j-1})]$$

und somit

$$|f(t) - f(t_0)| \leq \sum_{j=1}^{n_\delta} |x(t_j) - x(t_{j-1})| \leq n_\delta \cdot \omega_f(\delta)$$

also

$$|f(t)| \leq |f(t_0)| + n_\delta \cdot \omega_f(\delta) \quad , \quad t \in T$$

Daher können wir allgemein abschätzen

$$\|f\|_\infty \leq |f(t_0)| + n_\delta \cdot \omega_f(\delta) \quad , \quad f \in K, \delta > 0$$

Nach Voraussetzung (b) existiert ein $\delta > 0$ mit $\sup_{f \in K} \omega_f(\delta) \leq 1$, also

$$\sup_{f \in K} \|f\|_\infty \leq \underbrace{\sup_{f \in K} |f(t_0)|}_{< \infty} + n_\delta \cdot \underbrace{\sup_{f \in K} \omega_f(\delta)}_{< \infty} < \infty$$

□

Bemerkung: Die Wegbeschränktheit von T ist tatsächlich notwendig. Als Beispiel sei betrachtet $T = [0, 1] \cup \{2\}$ (nicht wegbeschränkt) und die Familie stetiger Funktionen

$$f_n(t) := \begin{cases} 1 & : t \in [0, 1] \\ n & : t = 2 \end{cases} \quad , \quad n \in \mathbb{N}$$

Dann ist $\omega_{f_n}(\delta) = 0$ für $\delta < 1$ und $\sup_n |f_n(0)| < \infty$, das heißt die Voraussetzungen (a) und (b) des Satzes sind erfüllt. Doch die Menge $\{f_n\}_{n \in \mathbb{N}}$ ist trotzdem nicht beschränkt bzw. relativ kompakt!

4.3.9 Charakterisierung relativer Kompaktheit von Maß-Mengen

Sei (T, d) ein kompakter, metrischer Raum, dazu $(\mathcal{C}(T), \|\cdot\|_\infty)$. Dann ist $\mathcal{M} \subseteq \mathfrak{M}_1(\mathcal{C}(T))$ genau dann schwach-relativ kompakt, wenn:

$$(a) \quad \forall \varepsilon > 0 : \exists \rho > 0 : \mathcal{P}(\|f\|_\infty \geq \rho) \leq \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}.$$

$$(b) \quad \forall \eta > 0 : \sup_{\mathcal{P} \in \mathcal{M}} \mathcal{P}(\{f : \omega_f(\delta) \geq \eta\}) \xrightarrow{\delta \searrow 0} 0, \text{ das heißt } \omega(\delta) \xrightarrow[\mathcal{P}]{\delta \searrow 0} 0 \text{ gleichmäßig in } \mathcal{M}.$$

Beweis: Richtung "⇒": Sei \mathcal{M} schwach-relativ kompakt, dann ist nach Prokhorov 2.3.9 \mathcal{M} auch gleichmäßig Radonsch³³. Für beliebiges $\varepsilon > 0$ gilt also:

$$\exists K \subseteq \mathcal{C}(T) \text{ kompakt} : \mathcal{P}(K) \geq 1 - \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}$$

Dabei $\exists \rho > 0$ mit³⁴

$$K \subseteq \{f \in \mathcal{C}(T) : \|f\|_\infty < \rho\}$$

also

$$\mathcal{P}(\|f\|_\infty \geq \rho) = 1 - \mathcal{P}(\|f\|_\infty < \rho) \leq 1 - \mathcal{P}(K) \leq \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}$$

(Behauptung (a)). Ähnlich existiert zu $\eta > 0$ ein $\delta > 0$ mit

$$K \subseteq \{f \in \mathcal{C}(T) : \omega_f(\delta) < \eta\}$$

also

$$\mathcal{P}(\{f \in \mathcal{C}(T) : \omega_f(\delta) \geq \eta\}) \leq 1 - \mathcal{P}(K) \leq \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}$$

$$\Rightarrow \sup_{\mathcal{P} \in \mathcal{M}} \mathcal{P}(\{f \in \mathcal{C}(T) : \omega_f(\delta) \geq \eta\}) \leq \varepsilon$$

³³Beachte dass $(\mathcal{C}(T), \|\cdot\|_\infty)$ vollständig, separabel ist.

³⁴Erinnerung: Stetige Bilder kompakter Mengen sind wieder kompakt.

(Behauptung (b)).

Richtung "⇐": Nach Prokhorov 2.3.9 genügt es zu zeigen: \mathcal{M} ist gleichmäßig Radonsch. Sei also $\varepsilon > 0$. Nach Voraussetzung (b) existiert eine Folge $\delta_n \searrow 0$ so dass

$$\mathcal{P}\left(\underbrace{\left\{f \in \mathcal{C}(T) : \omega_f(\delta_n) \leq \frac{1}{n}\right\}}_{A_n}\right) \geq 1 - \frac{\varepsilon}{2^{n+1}} \quad \forall \mathcal{P} \in \mathcal{M}$$

Nach Voraussetzung (a) $\exists \rho > 0$ mit $\mathcal{P}(\|f\|_\infty \geq \rho) \leq \frac{\varepsilon}{2}$. Dabei ist die Menge

$$K_0 := \{f \in \mathcal{C}(T) : \|f\|_\infty < \rho\} \cap \bigcap_{n \in \mathbb{N}} A_n$$

relativ kompakt in $\mathcal{C}(T)$, denn

- K_0 ist beschränkt.
- Zu $f \in K_0$ ist $\omega_f(\delta_n) \leq \frac{1}{n}$, $n \in \mathbb{N}$ und somit

$$\sup_{f \in K_0} \omega_f(\delta_n) \xrightarrow{\delta_n \searrow 0} 0 \quad \Rightarrow \quad \sup_{f \in K_0} \omega_f(\delta) \xrightarrow{\delta \searrow 0} 0$$

(vgl. Arcela-Ascoli 4.3.6). Insbesondere ist $\text{cl}(K_0)$ kompakt und

$$\mathcal{P}(K_0^c) \leq \underbrace{\mathcal{P}(\|f\|_\infty \geq \rho)}_{\leq \frac{\varepsilon}{2}} + \sum_{n=1}^{\infty} \mathcal{P}(A_n^c) \leq \frac{\varepsilon}{2} + \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} = \varepsilon$$

$$\Rightarrow \mathcal{P}(K_0) \geq 1 - \varepsilon \quad \Rightarrow \quad \mathcal{P}(\text{cl}(K_0)) \geq 1 - \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}$$

das heißt \mathcal{M} ist tatsächlich gleichmäßig Radonsch.

□

Bemerkungen:

(i) Ist (T, d) wegbeschränkt mit Startpunkt $t_0 \in T$, so kann die Bedingung (a) ersetzt werden durch

$$\forall \varepsilon > 0 : \exists \rho > 0 : \mathcal{P}(\{f \in \mathcal{C}(T) : |f(t_0)| \geq \rho\}) \leq \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}$$

(ii) Aufgrund der Kompaktheit von T ist jedes $f \in \mathcal{C}(T)$ gleichmäßig stetig. Es gilt also $\omega_f(\delta) \xrightarrow{\delta \searrow 0} 0$ für jedes $f \in \mathcal{C}(T)$. Deshalb geht zwar

$$\mathcal{P}(f \in \mathcal{C}(T) : \omega_f(\delta) \geq \eta) \xrightarrow{\delta \searrow 0} 0 \quad \forall \eta > 0$$

für jedes $\mathcal{P} \in \mathcal{M}$, doch ist die Gleichmäßigkeit der Konvergenz in \mathcal{M} (Bedingung (b)) nicht gesichert! Nicht destotrotz, ist Bedingung (b) im Fall $\mathcal{M} = \{\mathcal{P}_1, \mathcal{P}_2, \dots\}$ äquivalent zu

$$\forall \eta > 0 \forall \varepsilon > 0 : \exists n_0 \in \mathbb{N}, \delta > 0 : \mathcal{P}_n(\{f \in \mathcal{C}(T) : \omega_f(\delta) \geq \eta\}) \leq \varepsilon \quad \forall n \geq n_0 \quad (\spadesuit)$$

Beweis: Richtung "⇒" ist klar. Sei nun (\spadesuit) erfüllt und $\eta, \varepsilon > 0$ vorgegeben, dazu δ, n_0 wie in (\spadesuit) . Dann

$$\exists \delta_1, \dots, \delta_{n_0-1} : \mathcal{P}_k(f : \omega_f(\delta_k) \geq \eta) \leq \varepsilon \quad \forall k \in \{1, \dots, n_0 - 1\}$$

Für $\tilde{\delta} := \min\{\delta_1, \dots, \delta_{n_0-1}, \delta\}$ folgt dann

$$\mathcal{P}_n(f : \omega_f(\tilde{\delta}) \geq \eta) \leq \varepsilon$$

4.3.10 Charakterisierung relativer Kompaktheit für stochastische Prozesse

Sei (T, d) ein kompakter, metrischer Raum und $(X_t^n)_{t \in T}$ stochastische Prozesse mit Pfaden in $\mathcal{C}(T)$. Dann ist $\{\mathcal{P}_{X^n}\}_{n \in \mathbb{N}} \subseteq \mathfrak{M}_1(\mathcal{C}(T))$ schwach-relativ kompakt, genau dann wenn:

- (a) $\forall \varepsilon > 0 : \exists \rho > 0 : \mathcal{P}\left(\|X^n\|_\infty \geq \rho\right) \leq \varepsilon \quad \forall n \in \mathbb{N}$
 (b) $\forall \eta > 0, \varepsilon > 0 : \exists \delta > 0 : \mathcal{P}(\omega_{X^n}(\delta) \geq \eta) \leq \varepsilon \quad \forall n \in \mathbb{N}$

Beweis: Spezialfall des Charakterisierungssatzes 4.3.9. ³⁵

4.3.11 Definition: Überdeckungszahl

Sei (T, d) ein kompakter, metrischer Raum. Dann existieren zu jedem $\delta > 0$ Punkte $t_1, \dots, t_n \in T$ mit

$$T \subseteq \bigcup_{j=1}^n B_\delta^o(t_j)$$

Wir nennen

$$\mathcal{N}(T, \delta) := \min \left\{ n \in \mathbb{N} \mid \exists t_1, \dots, t_n \in T : T \subseteq \bigcup_{j=1}^n B_\delta^o(t_j) \right\}$$

Überdeckungszahl von T .

Beispiel: Zu $T := [0, 1]$ ist $\mathcal{N}(T, \delta) = \lfloor \delta^{-1} \rfloor + 1$.

4.3.12 Satz über Überdeckungszahlen und Stetigkeitsmodule

Sei (T, d) ein kompakter, metrischer Raum und $\mathcal{P} \in \mathfrak{M}_1(\mathcal{C}(T))$ mit folgender Eigenschaft:

$$\forall \eta > 0, \varepsilon > 0 : \exists \delta > 0 : \mathcal{N}(T, \delta) \cdot \sup_{t \in T} \mathcal{P}\left(f \in \mathcal{C}(T) : \sup_{s \in B_{2\delta}^o(t)} |f(t) - f(s)| \geq \eta\right) \leq \varepsilon \quad (4.3.12.1)$$

Dann gilt:

$$\forall \eta, \varepsilon > 0 : \exists \delta > 0 : \mathcal{P}(f \in \mathcal{C}(T) : \omega_f(\delta) \geq \eta) \leq \varepsilon$$

Beweis: Gegeben seien $\eta, \varepsilon > 0$, dazu $\delta > 0$ wie in Voraussetzung (4.3.12.1), die Überdeckungszahl $N := \mathcal{N}(T, \delta)$ und $t_1, \dots, t_N \in T$ mit

$$T \subseteq \bigcup_{j=1}^N B_\delta^o(t_j)$$

Setze

$$A_j := \left\{ f \in \mathcal{C}(T) : \sup_{s \in B_{2\delta}^o(t_j)} |f(t_j) - x(s)| < \eta \right\}, \quad A := \bigcap_{j=1}^N A_j$$

Zu $f \in A$ und $t, s \in T$ mit $d(t, s) < \delta$ existiert dann $j \in \{1, \dots, N\}$ mit $t \in B_{2\delta}^o(t_j)$, $s \in B_{2\delta}^o(t_j)$ und somit

$$|f(t_j) - f(t)| < \eta, \quad |f(t_j) - f(s)| < \eta$$

$$\Rightarrow |f(t) - f(s)| < 2\eta$$

Daher $\omega_f(\delta) \leq 2\eta$ und allgemein

$$A \subseteq \{f \in \mathcal{C}(T) : \omega_f(\delta) \leq 2\eta\}$$

Somit können wir abschätzen

$$\mathcal{P}(f : \omega_f(\delta) > 2\eta) \leq \mathcal{P}(A^c) \leq \sum_{j=1}^N \underbrace{\mathcal{P}(A_j^c)}_{\leq \frac{\varepsilon}{N}} \leq \varepsilon$$

□

³⁵Da $(X^n) : (\Omega, \mathfrak{G}, \mathcal{P}) \rightarrow \mathcal{C}(T)$ messbar sind, induzieren diese die Bildmaße \mathcal{P}_{X^n} auf $\mathcal{C}(T)$.

Folgerung: Aus dem Beweis wird ersichtlich: Ist $\mathcal{M} \subseteq \mathfrak{M}_1(\mathcal{C}(T))$ mit

$$\forall \varepsilon, \eta > 0 : \exists \delta > 0 : \mathcal{N}(T, \delta) \cdot \sup_{t \in T} \mathcal{P} \left(f \in \mathcal{C}(T) : \sup_{s \in B_{2\delta}^o(t)} |f(t) - f(s)| \geq \eta \right) \leq \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}$$

so folgt sogar die gleichmäßige Abschätzung

$$\forall \eta, \varepsilon > 0 : \exists \delta > 0 : \sup_{\mathcal{P} \in \mathcal{M}} \mathcal{P} (f \in \mathcal{C}(T) : \omega_f(\delta) \geq \eta) \leq \varepsilon$$

also Bedingung 4.3.9 (b). Würde auch Bedingung 4.3.9 (a) gelten, so wäre \mathcal{M} schwach-relativ kompakt.

4.3.13 Satz zur relativen Kompaktheit von Maß-Mengen auf $\mathcal{C}[0, 1]$

Sei $\mathcal{M} \subseteq \mathfrak{M}_1(\mathcal{C}[0, 1])$ mit

$$(a) \quad \forall \varepsilon > 0 : \exists \rho > 0 : \mathcal{P} (f \in \mathcal{C}[0, 1] : |f(0)| \geq \rho) \leq \varepsilon \quad \forall \mathcal{P} \in \mathcal{M}$$

$$(b) \quad \forall \eta, \varepsilon > 0 : \exists 0 < \delta \leq 1 : \mathcal{P} \left(f : \sup_{t < s < t + \delta} |f(t) - f(s)| \geq \eta \right) \leq \varepsilon \cdot \delta \quad \forall t \in [0, 1], \mathcal{P} \in \mathcal{M}$$

Dann ist \mathcal{M} schwach-relativ kompakt.

Beweis: Im Kontext dieses Beweises sei o.B.d.A. jedes $f \in \mathcal{C}[0, 1]$ fortgesetzt auf $(-\infty, 0)$ mit $f(0)$ und auf $(1, \infty)$ mit $f(1)$. Dies ändert nichts an der Voraussetzung (b).

Da $[0, 1]$ wegbeschränkt ist mit Startpunkt $t_0 := 0$, entspricht Bedingung (a) der Bedingung 4.3.9 (a) (vgl. entsprechende Bemerkung (i)). Andererseits gilt für $f \in \mathcal{C}[0, 1]$ und $t \in [0, 1]$ die Implikation

$$|f(t) - f(s)| < \eta \quad \forall s \in (t, t + 4\delta) \quad \Rightarrow \quad |f(t') - f(s)| < 2\eta \quad \forall s \in (t' - 2\delta, t' + 2\delta) \quad \Bigg| \quad t' := t + 2\delta$$

(vgl. Abbildung 4.9) das heißt

$$\mathcal{P} \left(f : \sup_{s \in (t' - 2\delta, t' + 2\delta)} |f(t') - f(s)| \geq 2\eta \right) \leq \mathcal{P} \left(f : \sup_{s \in (t, t + 4\delta)} |f(t) - f(s)| \geq \eta \right) \quad , \quad \eta > 0, t' \in [0, 1], \mathcal{P} \in \mathcal{M}$$

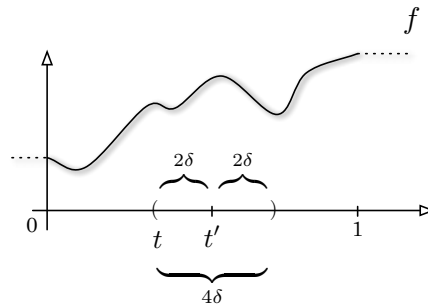


Abbildung 4.9: Zum Beweis des Satzes 4.3.13.

Nach Voraussetzung (b) existiert für beliebige $\frac{\eta}{2}, \frac{\varepsilon}{16} > 0$ ein $0 < 4\delta \leq 1$ so dass

$$\underbrace{\mathcal{P} \left(f : \sup_{s \in (t, t + 4\delta)} |f(t) - f(s)| \geq \frac{\eta}{2} \right)}_{\geq \mathcal{P} \left(f : \sup_{s \in B_{2\delta}^o(t')} |f(t') - f(s)| \geq \eta \right)} \leq \frac{\varepsilon}{16} \cdot (4\delta) = \frac{\varepsilon}{4} \delta \quad \forall t' \in [0, 1], \mathcal{P} \in \mathcal{M}$$

bzw.

$$\sup_{t' \in [0, 1]} \mathcal{P} \left(f : \sup_{s \in B_{2\delta}^o(t')} |f(t') - f(s)| \geq \eta \right) \leq \frac{\varepsilon}{4} \delta \quad \forall \mathcal{P} \in \mathcal{M}$$

Wegen $N_\delta := \mathcal{N}([0, 1], \delta) \leq \delta^{-1} + 1$ bzw. $\delta \leq \frac{1}{N_\delta - 1}$ ferner

$$\sup_{t' \in [0, 1]} \mathcal{P} \left(f : \sup_{s \in B_{2\delta}^0(t')} |f(t') - f(s)| \geq \eta \right) \leq \frac{\varepsilon}{4(N_\delta - 1)} \stackrel{N_\delta \geq 2}{\leq} \frac{\varepsilon}{N_\delta} \quad \forall \mathcal{P} \in \mathcal{M}$$

Nach der Folgerung in Satz 4.3.12 folgt dann auch Bedingung (b) in 4.3.9, und \mathcal{M} ist schwach-relativ kompakt. \square

Bemerkungen:

(i) Bedingung (b) kann im Fall $\mathcal{M} = \{\mathcal{P}_1, \mathcal{P}_2, \dots\}$ abgeschwächt werden zu

$$\forall \eta, \varepsilon > 0 : \exists n_0 \in \mathbb{N}, 0 < \delta \leq 1 : \mathcal{P}_n \left(f : \sup_{t < s < t + \delta} |f(t) - f(s)| \geq \eta \right) \leq \varepsilon \delta \quad \forall t \in [0, 1], n \geq n_0$$

(ii) Für Prozesse $(X_t^n)_{t \in [0, 1]}$, $n \in \mathbb{N}$ mit stetigen Pfaden und $\mathcal{M} := \{\mathcal{P}_{X^n}\}_{n \in \mathbb{N}}$ erhalten die Bedingungen des Satzes die Gestalt:

$$(a) \quad \forall \varepsilon > 0 : \exists \rho > 0 : \mathcal{P}(|X_0^n| \geq \rho) \leq \varepsilon \quad \forall n \in \mathbb{N}$$

$$(b) \quad \forall \eta, \varepsilon > 0 : \exists n_0 \in \mathbb{N}, 0 < \delta \leq 1 : \mathcal{P} \left(\sup_{t < s < t + \delta} |X_t^n - X_s^n| \geq \eta \right) \leq \varepsilon \delta \quad \forall t \in [0, 1], n \geq n_0$$

Beachte: Gilt $X_0^n = 0$ fast-sicher, so ist (a) automatisch erfüllt!

4.3.14 Hilfslemma über die Konvergenz der geglätteten Gleichverteilung

Sei \mathcal{U}_n^* der empirische Prozess der Gleichverteilung auf $[0, 1]$, dazu die stetige Approximation \mathcal{V}_n^* eingeführt in 4.3.2. Gilt

$$\forall \eta, \varepsilon > 0 : \exists n_{\eta, \varepsilon} \in \mathbb{N}, 0 < \delta \leq 1 : \mathcal{P} \left(\sup_{0 < s < \delta} |\mathcal{U}_n^*(s)| \geq \eta \right) \leq \varepsilon \delta \quad \forall t \in [0, 1], n \geq n_{\eta, \varepsilon}$$

so geht

$$(\mathcal{V}_n^*) \xrightarrow[\text{d}]{n \rightarrow \infty} B \quad (\text{Brownsche Brücke})$$

Beweis: Nach Charakterisierungssatz 2.3.10 und Satz 4.3.3 genügt es zu zeigen:

$$\{\mathcal{P}_{\mathcal{V}_n^*}\}_{n \in \mathbb{N}} \subseteq \mathfrak{M}_1[(\mathcal{C}[0, 1], \|\cdot\|_\infty)]$$

ist schwach-relativ kompakt. Nach Bemerkung 4.3.13 (ii) genügt es dafür zu zeigen:

$$\forall \eta, \varepsilon > 0 : \exists n_0 \in \mathbb{N}, 0 < \delta \leq 1 : \mathcal{P} \left(\sup_{t < s < t + \delta} |\mathcal{V}_n^*(t) - \mathcal{V}_n^*(s)| \geq \eta \right) \leq \varepsilon \delta \quad \forall t \in [0, 1], n \geq n_0 \quad (\clubsuit)$$

Dabei ist

$$\sup_{t < s < t + \delta} |\mathcal{V}_n^*(t) - \mathcal{V}_n^*(s)| \leq \sup_{t < s < t + \delta} |\mathcal{U}_n^*(t) - \mathcal{U}_n^*(s)| + \frac{2}{\sqrt{n}}, \quad \delta > 0, t \in [0, 1], n \in \mathbb{N}$$

das heißt

$$\mathcal{P} \left(\sup_{t < s < t + \delta} |\mathcal{V}_n^*(t) - \mathcal{V}_n^*(s)| \geq 2\eta \right) \leq \mathcal{P} \left(\sup_{t < s < t + \delta} |\mathcal{U}_n^*(t) - \mathcal{U}_n^*(s)| \geq 2\eta - \frac{2}{\sqrt{n}} \right)$$

$$\stackrel{\forall n \geq n_\eta}{\leq} \mathcal{P} \left(\sup_{t < s < t + \delta} |\mathcal{U}_n^*(t) - \mathcal{U}_n^*(s)| \geq \eta \right) \quad \Bigg| \quad 2\eta - \frac{2}{\sqrt{n_\eta}} \geq \eta$$

Nach Hilfslemmas A.0.15 und A.0.16 gilt

$$\sup_{t < s < t + \delta} |G_n^*(s) - G_n^*(t) - (s - t)| \stackrel{\text{d.}}{=} \sup_{0 < s < \delta} |G_n^*(s) - s| \quad , \quad t, \delta \in [0, 1]$$

bzw.

$$\sup_{t < s < t + \delta} |\mathcal{U}_n^*(s) - \mathcal{U}_n^*(t)| \stackrel{d.}{=} \sup_{0 < s < \delta} |\mathcal{U}_n^*(s)|$$

das heißt nach Voraussetzung:

$$\forall \eta, \varepsilon > 0 : \exists n_{\eta, \varepsilon} \in \mathbb{N}, 0 < \delta \leq 1 : \mathcal{P} \left(\sup_{t < s < t + \delta} |\mathcal{U}_n^*(t) - \mathcal{U}_n^*(s)| \geq \eta \right) \leq \varepsilon \delta \quad \forall t \in [0, 1], n \geq n_{\eta, \varepsilon} \quad (\spadesuit)$$

Seien nun $\eta, \varepsilon > 0$ beliebig, dazu $n_{\eta, \varepsilon}$ und $0 < \delta \leq 1$ wie in (\spadesuit) . Dann erfüllt $n_0 := \max \{n_{\eta, \varepsilon}, n_\eta\}$ die gewünschte Bedingung (\clubsuit) :

$$\mathcal{P} \left(\sup_{t < s < t + \delta} |\mathcal{V}_n^*(t) - \mathcal{V}_n^*(s)| \geq 2\eta \right) \leq \varepsilon \delta \quad \forall t \in [0, 1], n \geq n_0$$

□

4.3.15 Theorem von Donsker (1952)

Sei $(\mathcal{V}_n^*)_{0 \leq t \leq 1}$ die oben (4.3.2) eingeführte stetige Approximation des empirischen Prozesses \mathcal{U}_n^* n -ter Ordnung der Gleichverteilung auf $[0, 1]$. Dann gehen

$$\mathcal{V}_n^* \xrightarrow[n \rightarrow \infty]{d} B$$

(als Zufallsvariablen in $(\mathcal{C}[0, 1], \|\cdot\|_\infty)$).

Beweis: Sei $(\Omega, \mathfrak{E}, \mathcal{P})$ der zugrundeliegende Wahrscheinlichkeitsraum. Nach Hilfslemma 4.3.14 genügt es zu zeigen:

$$\forall \eta, \varepsilon > 0 : \exists n_0 \in \mathbb{N}, 0 < \delta \leq 1 : \mathcal{P} \left(\sup_{0 < s < \delta} |\mathcal{U}_n^*(s)| \geq \eta \right) \leq \varepsilon \delta \quad \forall n \geq n_0$$

Da \mathcal{U}_n^* rechtsseitig stetige Pfade besitzt, gilt für jede in $[0, \delta]$ dichte Teilmenge $D_\delta \subset [0, \delta]$:

$$\sup_{0 < s < \delta} |\mathcal{U}_n^*(s)| = \sup_{s \in D_\delta} |\mathcal{U}_n^*(s)| \quad .$$

So ist z.B.

$$D_\delta := \left\{ \frac{l}{2^k} \delta : k \in \mathbb{N}, l \in \{1, \dots, 2^k\} \right\} \quad , \quad \delta > 0$$

dicht in $[0, \delta]$. Für die monoton wachsende Folge

$$D_\delta^k := \left\{ \omega \in \Omega : \sup_{1 \leq l \leq 2^k} \left| \mathcal{U}_n^* \left(\frac{l\delta}{2^k} \right) \right| \geq \eta \right\} \quad , \quad k \in \mathbb{N}$$

gilt ferner

$$\bigcup_{k=1}^{\infty} D_\delta^k = \left\{ \omega \in \Omega : \sup_{s \in D_\delta} |\mathcal{U}_n^*(s)| \geq \eta \right\} \quad .$$

Daher ist

$$\underbrace{\mathcal{P} \left(\sup_{0 < s < \delta} |\mathcal{U}_n^*(s)| \geq \eta \right)}_{\lim_{k \rightarrow \infty} \mathcal{P}(D_\delta^k)} \leq \varepsilon \delta$$

äquivalent zu

$$\mathcal{P}(D_\delta^k) \leq \varepsilon \delta \quad \forall k \in \mathbb{N}$$

Gesucht ist nun für vorgegebene $\eta, \varepsilon > 0$ solch ein δ .

Behauptung: Es existiert eine Konstante $C \in \mathbb{R}$ die die Abschätzung

$$\mathcal{P} \left(\sup_{1 \leq l \leq m} \left| \mathcal{U}_n^* \left(\frac{l\delta}{m} \right) \right| \geq \eta \right) \leq C \cdot \frac{\delta^2}{\eta^4} + \mathcal{P} \left(|\mathcal{U}_n^*(\delta)| \geq \frac{\eta}{2} \right) \quad , \quad \eta, \delta > 0, m, n \in \mathbb{N} \quad (\clubsuit)$$

erfüllt.

Beweis:

Betrachten für feste $\delta \in (0, 1]$, $n, m \in \mathbb{N}$ die Zufallsgrößen

$$Y_l := \mathcal{U}_n^* \left(\frac{l\delta}{m} \right) - \mathcal{U}_n^* \left(\frac{(l-1)\delta}{m} \right), \quad l \in \{1, \dots, m\}$$

$$S_k := Y_1 + \dots + Y_k = \mathcal{U}_n^* \left(\frac{k\delta}{m} \right), \quad k \in \mathbb{N}, \quad S_0 := 0$$

$$M_m := \sup_{1 \leq k \leq m} |S_k|$$

$$M'_m := \sup_{1 \leq k \leq m} \{|S_k| \wedge |S_m - S_k|\}$$

Kontextgemäß wäre nun $\mathcal{P}(M_m \geq \eta)$ nach oben abzuschätzen durch einen Ausdruck der Form $\mathcal{P}(|S_m| \geq \tilde{\eta})$.³⁶ Wegen

$$|S_m - S_k| \geq |S_k| - |S_m|$$

ist

$$M'_m \geq \sup_{1 \leq k \leq m} \{|S_k| \wedge (|S_k| - |S_m|)\} = M_m - |S_m|$$

bzw.

$$\mathcal{P}(M_m \geq \eta) \leq \mathcal{P}(M'_m + |S_m| \geq \eta) \leq \mathcal{P}\left(M'_m \geq \frac{\eta}{2}\right) + \mathcal{P}\left(\underbrace{|S_m|}_{\mathcal{U}_n^*(\delta)} \geq \frac{\eta}{2}\right)$$

Daher genügt es zu zeigen

$$\mathcal{P}(M'_m \geq \eta) \leq \underbrace{\text{const}}_{\substack{\text{unabhängig} \\ \text{von} \\ \delta, \eta, m, n}} \cdot \frac{\delta^2}{\eta^4}$$

Durch die Abschätzung

$$\begin{aligned} \mathbb{E} \left\{ |S_j - S_i|^2 |S_k - S_j|^2 \right\} &= \mathbb{E} \left\{ \left| \mathcal{U}_n^* \left(\frac{j\delta}{m} \right) - \mathcal{U}_n^* \left(\frac{i\delta}{m} \right) \right|^2 \left| \mathcal{U}_n^* \left(\frac{k\delta}{m} \right) - \mathcal{U}_n^* \left(\frac{j\delta}{m} \right) \right|^2 \right\} \\ &= \mathbb{E} \left\{ |\mathcal{U}_n^*(s + \alpha) - \mathcal{U}_n^*(s)|^2 |\mathcal{U}_n^*(s + \alpha) - \mathcal{U}_n^*(s + \alpha + \beta)|^2 \right\} \\ &\stackrel{\text{(A.0.20)}}{\leq} 6\alpha\beta \quad \left| \quad s := \frac{i\delta}{m}, \alpha := \frac{(j-i)\delta}{m}, \beta := \frac{(k-l)\delta}{m} \right. \\ &= (j-i) \frac{\sqrt{6}\delta}{m} \cdot (k-j) \frac{\sqrt{6}\delta}{m} \\ &= \left(\sum_{l=i+1}^j \lambda_l \right) \left(\sum_{r=j+1}^k \lambda_r \right) \quad \left| \quad \lambda_1 := \dots := \lambda_m := \frac{\sqrt{6}\delta}{m} \right. \end{aligned}$$

folgt nach Satz A.0.19 tatsächlich die Abschätzung

$$\mathcal{P}(M'_m \geq \eta) \leq \frac{C}{\eta^4} (\lambda_1 + \dots + \lambda_m)^2 = 6C \cdot \frac{\delta^2}{\eta^4}$$

³⁶Zu beachten sei, dass gewissermaßen globale Extrema von Pfaden $(S_k)_{k=1}^n$ durch Werte an einem festen *Endpunkt* abzuschätzen sind! Tatsächlich stellen die Maximalungleichungen A.0.17 und A.0.18 auf den ersten Blick eine verlockende Möglichkeit dar, solch eine Abschätzung durchzuführen. Doch sind die hier vorhandenen Zufallsgrößen Y_1, \dots, Y_m **nicht** unabhängig, was eine Anwendung dieser beiden Maximalungleichungen ausschließt.

Behauptung: Zu $\eta, \varepsilon > 0$ existieren tatsächlich $n_{\eta, \varepsilon} \in \mathbb{N}$, $0 < \delta \leq 1$ mit:

$$\mathcal{P} \left(\sup_{1 \leq l \leq m} \left| \mathcal{U}_n^* \left(\frac{l\delta}{m} \right) \right| \geq \eta \right) \leq \varepsilon \delta \quad \forall m \in \mathbb{N}, n \geq n_{\eta, \varepsilon} \quad (\spadesuit)$$

Beweis:

Bekanntlich ist $\mathbb{E}(\mathcal{U}_n^*(\delta)) = 0$, $\text{Var}(\mathcal{U}_n^*(\delta)) = \delta(1-\delta)$ für jedes $0 < \delta \leq 1$. Nach 4.1.2 gehen außerdem

$$\mathcal{U}_n^*(\delta) \xrightarrow[n \rightarrow \infty]{d} B_\delta \stackrel{d.}{=} \sqrt{\delta(1-\delta)} \cdot \xi, \quad \xi \sim \mathcal{N}_{0,1}$$

das heißt nach Satz 4:

$$\mathcal{P} \left(|\mathcal{U}_n^*(\delta)| \geq \frac{\eta}{2} \right) \xrightarrow[n \rightarrow \infty]{} \mathcal{P} \left(|\xi| \geq \frac{\eta}{2} \frac{1}{\sqrt{\delta(1-\delta)}} \right) = \mathcal{P} \left(|\xi|^4 \geq \frac{\eta^4}{2^4 \delta^2 (1-\delta)^2} \right)$$

$$\stackrel{\text{Markov}}{\leq} \frac{2^4 \delta^2 (1-\delta)^2}{\eta^4} \cdot \underbrace{\mathbb{E} |\xi|^4}_3 \leq 3 \cdot 2^4 \frac{\delta^2}{\eta^4}$$

bzw.

$$\mathcal{P} \left(|\mathcal{U}_n^*(\delta)| \geq \frac{\eta}{2} \right) \leq \underbrace{C'}_{\text{const}} \cdot \frac{\delta^2}{\eta^4} \quad \forall n \geq n_{\eta, \delta} \quad \Bigg| \quad \text{z.B. } C' := 3 \cdot 2^4 + 1$$

für ein geeignetes $n_{\eta, \delta}$. Zusammen mit Abschätzung () lässt sich also schreiben

$$\mathcal{P} \left(\sup_{1 \leq l \leq m} \left| \mathcal{U}_n^* \left(\frac{l\delta}{m} \right) \right| \geq \eta \right) \leq C \cdot \frac{\delta^2}{\eta^4} + C' \cdot \frac{\delta^2}{\eta^4} = \underbrace{C''}_{\text{const}} \cdot \frac{\delta^2}{\eta^4}, \quad \delta, \eta > 0, m \in \mathbb{N}, n \geq n_{\eta, \delta}$$

Zu vorgegebenen $\eta, \varepsilon > 0$ erfüllt dann $\delta := \min \left\{ 1, \frac{\varepsilon \eta^4}{C''} \right\}$ die Ungleichung

$$\mathcal{P} \left(\sup_{1 \leq l \leq m} \left| \mathcal{U}_n^* \left(\frac{l\delta}{m} \right) \right| \geq \eta \right) \leq \varepsilon \delta \quad \forall m \in \mathbb{N}, n \geq n_{\eta, \varepsilon}$$

Aus Behauptung () und obiger Vorüberlegung folgt die Behauptung des Satzes.

□

Anwendungsbeispiele:

- Kolmogorov-Smirnov Test.
- Sichert dass $X_n : \Omega \rightarrow \mathcal{C}(T)$ (für (T, d) kompakt, metrisch) hinreichende Bedingung ist dass $\{\mathcal{P}_{X_n}\}_{n \in \mathbb{N}} \subseteq \mathfrak{M}_1(\mathbb{R})$ gleichmäßig radonsch ist.
- Weitere Anwendungen auf empirischen Prozessen.

Bemerkung: Die Frage nach der Konvergenz in Verteilung anderer stetiger Approximationen empirischer Prozesse lässt sich genauso stellen wie im Falle von \mathcal{U}_n^* und \mathcal{V}_n^* . Sind z.B. X_1, X_2, \dots unabhängig, identisch verteilt mit

$$\mathbb{E}X_i = 0, \quad \text{Var}(X_i) = 1$$

und dazu

$$Z^n \left(\frac{k}{n} \right) := \frac{S_k}{\sqrt{n}}, \quad S_k := X_1 + \dots + X_n, \quad k \in \{0, \dots, n\}$$

so kann man die lineare Verbindung Z^n ³⁷ der $Z^n(k/n)$ betrachten. Jedes $Z^n = (Z_t^n)_{t \in [0,1]}$ ist ein stochastischer Prozess mit stetigen Pfaden (vgl. Abbildung 4.10).

³⁷Das heißt linear in jedem Intervall $\left[\frac{k}{n}, \frac{k+1}{n} \right]$.

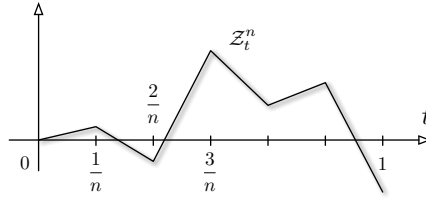


Abbildung 4.10: Typischer Pfad der linearen Verbindung Z^n der Partialsummen $\frac{S_k}{\sqrt{n}}$.

Es kann gezeigt werden:

$$Z^n \xrightarrow[\text{f.d.d.}]{n \rightarrow \infty} (W_t)_{t \in [0,1]} \quad (\text{Wiener-Prozess})$$

Der Beweis der Konvergenz

$$Z^n \xrightarrow[\text{d}]{n \rightarrow \infty} W$$

(als Abbildungen $Z^n, W : \Omega \rightarrow \mathcal{C}[0,1]$) läuft dann ganz analog zu vorhin ab. Insbesondere gilt:

$$\underbrace{\mathcal{P} \left(\frac{1}{\sqrt{n}} \sup_{1 \leq k \leq n} |S_k| \leq \alpha \right)}_{\mathcal{P} \left(\sup_{0 \leq t \leq 1} |Z_t^n| \leq \alpha \right)} \xrightarrow{n \rightarrow \infty} \mathcal{P} \left(\sup_{0 \leq t \leq 1} |W_t| \leq \alpha \right)$$

4.4 Anwendungen des Konvergenzsatzes von Donster

Zür $M \in \mathcal{B}(\mathbb{R})$ sei $\mathfrak{B}(M)$ der Raum aller messbaren, beschränkten Abbildungen $f : M \rightarrow \mathbb{R}$, ausgestattet mit der Norm

$$\|f\|_\infty := \sup_{t \in M} |f(t)|$$

4.4.1 Satz über den empirischen Prozess der Gleichverteilung und die Brownsche Brücke

Sei \mathcal{U}_n^* der empirische Prozess der Gleichverteilung auf $[0,1]$ und $h : (\mathfrak{B}[0,1], \|\cdot\|_\infty) \rightarrow \mathbb{R}$ stetig. Dann geht $h(\mathcal{U}_n^*)$ in Verteilung gegen die *transformierte* Brownsche Brücke

$$h(\mathcal{U}_n^*) \xrightarrow[\text{d}]{n \rightarrow \infty} h(B)$$

(als Abbildungen $h(B), h(\mathcal{U}_n^*) : \Omega \rightarrow \mathbb{R}$).

Beweis: Sei $(\mathcal{V}_n^*)_{0 \leq t \leq 1}$ die in 4.3.2 eingeführte stetige Approximation zu \mathcal{U}_n^* . Dabei wurde gezeigt dass

$$\|\mathcal{U}_n^* - \mathcal{V}_n^*\|_\infty = \sup_{0 \leq t \leq 1} |\mathcal{U}_n^*(t) - \mathcal{V}_n^*(t)| \leq \frac{1}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0$$

gilt, also auch

$$|h(\mathcal{U}_n^*) - h(\mathcal{V}_n^*)| \xrightarrow{n \rightarrow \infty} 0 \quad (\clubsuit)$$

Andererseits geht nach Theorem 4.3.15

$$\mathcal{V}_n^* \xrightarrow[\text{d}]{n \rightarrow \infty} B$$

und nach Satz 2.1.3

$$h(\mathcal{V}_n^*) \xrightarrow[\text{d}]{n \rightarrow \infty} h(B) \quad (\spadesuit)$$

(da $h : \mathcal{C}[0,1] \rightarrow \mathbb{R}$ stetig). Nach Übertragungssatz 2.2.3 folgt dann aus (\clubsuit) und (\spadesuit) auch

$$h(\mathcal{U}_n^*) \xrightarrow[\text{d}]{n \rightarrow \infty} h(B)$$

□

4.4.2 Korollar: Konvergenz der Betragssuprema empirischer Prozesse

Sei F eine stetige Verteilungsfunktion, dazu der empirische Prozess E_n^* . Dann geht:

$$\sup_{t \in \mathbb{R}} |E_n^*(t)| \xrightarrow[n \rightarrow \infty]{d} \sup_{0 \leq t \leq 1} |B_t|$$

mit der Brownschen Brücke $(B_s)_{0 \leq s \leq 1}$.

Beweis: Sei U_n^* der empirische Prozess der Gleichverteilung auf $[0, 1]$. Die Abbildung

$$h : (\mathcal{B}[0, 1], \|\cdot\|_\infty) \rightarrow \mathbb{R}, \quad h(f) := \sup_{0 \leq s \leq 1} |f(s)|$$

ist stetig, so dass nach Satz 4.4.1 folgt:

$$\underbrace{\sup_{t \in \mathbb{R}} |E_n^*(t)|}_{\substack{\text{messbar} \\ \text{nach (3.4.6)}}} \stackrel{(4.0.16)}{d} \sup_{t \in \mathbb{R}} |U_n^*(F(t))| = \sup_{0 \leq s \leq 1} |U_n^*(s)| \xrightarrow[n \rightarrow \infty]{d} \sup_{0 \leq s \leq 1} |B_s|$$

□

Bemerkung: Das Korollar stellt für stetige F eine Verschärfung des Theorems von Glivenko-Cantelli 4.2.2 dar, das besagte dass

$$\|F_n^* - F\|_\infty \xrightarrow[n \rightarrow \infty]{} 0$$

fast sicher, mit der empirischen Verteilungsfunktion $F_n^*(t) : (\Omega, \mathfrak{E}, \mathcal{P}) \rightarrow \mathbb{R}$. Obiges Korollar liefert nämlich eine Aussage über die tatsächliche *Konvergenzgeschwindigkeit*:

$$\mathcal{P} \left(\frac{\alpha}{\sqrt{n}} \leq \|F_n^* - F\|_\infty \leq \frac{\beta}{\sqrt{n}} \right) \xrightarrow[n \rightarrow \infty]{} \mathcal{P} \left(\alpha \leq \sup_{0 \leq s \leq 1} |B_s| \leq \beta \right) \quad \forall 0 \leq \alpha < \beta < \infty$$

(vgl. Charakterisierungssatz 2.1.2 (4))

4.4.3 Korollar: Konvergenz der Suprema empirischer Prozesse

Sei F eine stetige Verteilungsfunktion mit empirischem Prozess E_n^* . Dann geht

$$\sup_{t \in \mathbb{R}} E_n^*(t) \xrightarrow[n \rightarrow \infty]{d} \sup_{0 \leq s \leq 1} B_s$$

mit der Brownschen Brücke $(B_s)_{0 \leq s \leq 1}$.

Beweis: Analog zu 4.4.2, setze

$$h : (\mathcal{B}[0, 1], \|\cdot\|_\infty) \rightarrow \mathbb{R}, \quad h(f) := \sup_{0 \leq s \leq 1} f(t)$$

□

Folgerung: Ähnlich zur Bemerkung in Korollar 4.4.2 geht auch hier

$$\mathcal{P} \left(\frac{\alpha}{\sqrt{n}} \leq \sup_{t \in \mathbb{R}} (F_n^*(t) - F(t)) \leq \frac{\beta}{\sqrt{n}} \right) \xrightarrow[n \rightarrow \infty]{} \mathcal{P} \left(\alpha \leq \sup_{0 \leq s \leq 1} B_s \leq \beta \right) \quad \forall 0 \leq \alpha < \beta < \infty$$

wobei F_n^* die empirische Verteilungsfunktion von F auf $(\Omega, \mathfrak{E}, \mathcal{P})$ ist.

4.4.4 Lemma: Konvergenz der stetigen Approximation der Gleichverteilung

Seien \mathcal{V}_n^* , $\tilde{\mathcal{V}}_m^*$ zwei unabhängige stetige Approximationen des empirischen Prozesses der Gleichverteilung auf $[0, 1]$, definiert gemäß 4.3.2. Seien außerdem $m = m(n) =: m_n$ so dass

$$\lim_{n \rightarrow \infty} \frac{n}{m_n} = c : \text{const} \in (0, \infty)$$

Dann geht

$$\lim_{n \rightarrow \infty} \sqrt{\frac{n \cdot m_n}{n + m_n}} \cdot \left[\frac{\mathcal{V}_n^*}{\sqrt{n}}, \frac{\tilde{\mathcal{V}}_{m_n}^*}{\sqrt{m_n}} \right] \xrightarrow[n \rightarrow \infty]{d} \left[\frac{1}{\sqrt{1+c}} B^1, \sqrt{\frac{c}{1+c}} B^2 \right]$$

wobei B^1, B^2 zwei unabhängige Brownsche Brücken sind.

Beweis: Seien $\mathcal{V}_n^*, \tilde{\mathcal{V}}_m^* : (\Omega, \mathfrak{S}, \mathcal{P}) \rightarrow \mathcal{C}[0, 1]$. Per Konstruktion ist

$$\left[\frac{\mathcal{V}_n^*(\omega)}{\sqrt{n}}, \frac{\tilde{\mathcal{V}}_m^*(\omega)}{\sqrt{m}} \right] \in \mathcal{C}[0, 1] \times \mathcal{C}[0, 1] \quad , \quad \omega \in \Omega$$

wobei definiert sei

$$\|(x, y)\|_\infty := \max \{ \|x\|_\infty, \|y\|_\infty \} \quad , \quad (x, y) \in \mathcal{C}[0, 1] \times \mathcal{C}[0, 1] \quad (4.4.4.1)$$

Fernern gehen

$$\sqrt{\frac{m_n}{m_n + n}} \xrightarrow[n \rightarrow \infty]{} \frac{1}{\sqrt{1+c}} \quad , \quad \sqrt{\frac{n}{m_n + n}} \xrightarrow[n \rightarrow \infty]{} \frac{1}{\sqrt{1+\frac{1}{c}}}$$

Nach 4.3.15 und 2.2.1 (iv) gehen

$$\sqrt{\frac{m_n}{m_n + n}} \cdot \mathcal{V}_n^* \xrightarrow[n \rightarrow \infty]{d} \frac{B}{\sqrt{1+c}} \quad , \quad \sqrt{\frac{n}{m_n + n}} \cdot \mathcal{V}_{m_n}^* \xrightarrow[n \rightarrow \infty]{d} \frac{B}{\sqrt{1+\frac{1}{c}}}$$

wobei B^1, B^2 Brownsche Brücken sind. Setzt man nun

$$\mathcal{P}_n := \text{distr} \left(\sqrt{\frac{m_n}{m_n + n}} \cdot \mathcal{V}_n^* \right) \quad , \quad \mathcal{Q}_n := \text{distr} \left(\sqrt{\frac{n}{m_n + n}} \cdot \tilde{\mathcal{V}}_{m_n}^* \right)$$

so folgt nach 2.1.6 und 4.3.15 die Behauptung:

$$\text{distr} \left[\sqrt{\frac{m_n}{m_n + n}} \cdot \mathcal{V}_n^*, \sqrt{\frac{n}{m_n + n}} \cdot \mathcal{V}_{m_n}^* \right] \stackrel{\mathcal{V}_n^*, \tilde{\mathcal{V}}_{m_n}^* \text{ unabhängig}}{=} \mathcal{P}_n \otimes \mathcal{Q}_n \xrightarrow[n \rightarrow \infty]{d} \underbrace{\text{distr} \left(\frac{B}{\sqrt{1+c}} \right) \otimes \text{distr} \left(\frac{B}{\sqrt{1+\frac{1}{c}}} \right)}_{\substack{\text{distr} \left[\frac{B^1}{\sqrt{1+c}}, \frac{B^2}{\sqrt{1+\frac{1}{c}}} \right] \\ \text{mit } B^1, B^2 \text{ unabhängig}}}$$

□

4.4.5 Theorem von Smirnov

Seien F_n^*, \tilde{F}_m^* unabhängige empirische Verteilungsfunktionen der stetigen Verteilungsfunktion F . Seien $m = m_n$ so dass

$$\lim_{n \rightarrow \infty} \frac{n}{m_n} = c : \text{const} \in (0, \infty)$$

Dann gilt:

$$\sqrt{\frac{m_n \cdot n}{m_n + n}} \cdot \left\| F_n^* - \tilde{F}_{m_n}^* \right\|_\infty \xrightarrow[n \rightarrow \infty]{d} \sup_{0 \leq s \leq 1} |B_s|$$

Beweis: Seien $\mathcal{U}_1, \mathcal{U}_2, \dots, \tilde{\mathcal{U}}_1, \tilde{\mathcal{U}}_2, \dots$ unabhängig, gleichverteilt auf $[0, 1]$, dazu die (unabhängigen) empirischen Verteilungsfunktionen

$$G_n^*(t) := \frac{1}{n} \# \{j \leq n : \mathcal{U}_j \leq t\} \quad , \quad \tilde{G}_{m_n}^*(t) := \frac{1}{m_n} \# \{j \leq m_n : \tilde{\mathcal{U}}_j \leq t\} \quad , \quad 0 \leq t \leq 1$$

bzw. empirischen Prozesse

$$\mathcal{U}_n^*(t) := \sqrt{n} \cdot [G_n^*(t) - t] \quad , \quad \tilde{\mathcal{U}}_{m_n}^*(t) := \sqrt{m_n} \cdot [\tilde{G}_{m_n}^*(t) - t] \quad , \quad 0 \leq t \leq 1$$

und $H_n^*, \tilde{H}_{m_n}^*$ bzw. $\mathcal{V}_n^*, \tilde{\mathcal{V}}_{m_n}^*$ die in 4.3.2 eingeführten stetigen Approximationen der Verteilungsfunktionen bzw. der zugehörigen empirischen Prozesse.

Betrachten die auf $(\mathcal{C}[0, 1] \times \mathcal{C}[0, 1], \|\cdot\|_\infty)$ (vgl. (4.4.4.1)) stetige Abbildung

$$h : \mathcal{C}[0, 1] \times \mathcal{C}[0, 1] \rightarrow \mathbb{R} \quad , \quad h(x, y) := \sup_{0 \leq t \leq 1} |x(t) - y(t)| \quad , \quad x, y \in \mathcal{C}[0, 1]$$

Nach 2.2.1 (v) und vorigem Lemma 4.4.4 gehen

$$\sqrt{\frac{r_n}{m_n + n}} \cdot \sup_{0 \leq t \leq 1} \left| \frac{\mathcal{V}_n^*}{\sqrt{n}} - \frac{\tilde{\mathcal{V}}_{m_n}^*}{\sqrt{m_n}} \right| \xrightarrow[n \rightarrow \infty]{d} \underbrace{\sup_{0 \leq t \leq 1} \left| \frac{B_t^1}{\sqrt{1+c}} - \frac{B_t^2}{\sqrt{1+\frac{1}{c}}} \right|}_{\spadesuit}$$

wobei B^1, B^2 unabhängige Brownsche Brücken sind. Offensichtlich ist

$$X := \frac{B_t^1}{\sqrt{1+c}} - \frac{B_t^2}{\sqrt{1+\frac{1}{c}}}$$

ein zentrierter Gaußprozess, mit

$$\text{Cov}(X_t, X_s) \stackrel{B^1, B^2 \text{ unabhängig}}{=} \frac{1}{1+c} \underbrace{\mathbb{E}(B_t^1 B_s^1)}_{t \wedge s - ts} + \frac{1}{1+\frac{1}{c}} \underbrace{\mathbb{E}(B_t^2 B_s^2)}_{t \wedge s - ts} = t \wedge s - ts = \text{Cov}(B_t, B_s)$$

Nach Eindeutigkeitsatz 3.3.2 ist dann $X \cong B$, das heißt

$$(\spadesuit) = \sup_{0 \leq s \leq 1} |B_s|$$

Andererseits ist

$$\frac{\mathcal{V}_n^*}{\sqrt{n}} - \frac{\tilde{\mathcal{V}}_{m_n}^*}{\sqrt{m_n}} = H_n^* - \tilde{H}_{m_n}^*$$

und

$$\sup_{0 \leq t \leq 1} |H_n^*(t) - G_n^*(t)| \stackrel{(4.3.2.2)}{\leq} \frac{1}{n} \quad , \quad \sup_{0 \leq t \leq 1} |\tilde{H}_{m_n}^*(t) - \tilde{G}_{m_n}^*(t)| \stackrel{(4.3.2.2)}{\leq} \frac{1}{m_n}$$

bzw.

$$\begin{aligned} & \|G_n^* - \tilde{G}_{m_n}^*\|_\infty - \|H_n^* - \tilde{H}_{m_n}^*\|_\infty \leq \|G_n^* - H_n^*\|_\infty + \|\tilde{H}_{m_n}^* - \tilde{G}_{m_n}^*\|_\infty \\ & \|\tilde{H}_{m_n}^* - H_n^*\|_\infty - \|\tilde{G}_{m_n}^* - G_n^*\|_\infty \leq \|G_n^* - H_n^*\|_\infty + \|\tilde{H}_{m_n}^* - \tilde{G}_{m_n}^*\|_\infty \\ \Rightarrow & \left| r_n \|H_n^* - \tilde{H}_{m_n}^*\|_\infty - r_n \|G_n^* - \tilde{G}_{m_n}^*\|_\infty \right| \leq \underbrace{r_n \|G_n^* - H_n^*\|_\infty}_{\xrightarrow[n \rightarrow \infty]{} 0} + \underbrace{r_n \|\tilde{H}_{m_n}^* - \tilde{G}_{m_n}^*\|_\infty}_{\xrightarrow[n \rightarrow \infty]{} 0} \xrightarrow[n \rightarrow \infty]{} 0 \end{aligned}$$

Nach 2.2.3 konvergieren dann auch

$$r_n \left\| F_n^* - \tilde{F}_{m_n}^* \right\|_{\mathbb{R}, \infty} \stackrel{(4.0.16)}{=} r_n \left\| \frac{\mathcal{U}_n^*(F)}{\sqrt{n}} - \frac{\tilde{\mathcal{U}}_{m_n}^*(F)}{\sqrt{m_n}} \right\|_{\mathbb{R}, \infty} = r_n \|G_n^* - \tilde{G}_{m_n}^*\|_{[0,1], \infty} \xrightarrow[n \rightarrow \infty]{d} (\spadesuit) = \sup_{0 \leq s \leq 1} |B_s|$$

Bemerke: Das Verteilungsgesetz von $(F_n^* - \tilde{F}_{m_n}^*)$ ist unabhängig von jeglicher speziellen Version eindeutig durch F und die Unabhängigkeit der $F_n^*, \tilde{F}_{m_n}^*$ festgelegt. Somit ist die Behauptung gezeigt.

□

Bemerkungen:

- (i) Das Verblüffende am Theorem ist die Unabhängigkeit des Grenzwertes von dem speziellen c !
(ii) Der Satz gilt auch noch falls

$$0 < \liminf_{n \rightarrow \infty} \frac{n}{m_n} \leq \limsup_{n \rightarrow \infty} \frac{n}{m_n} < \infty$$

Es stellt sich nun die Frage: Wie ist überhaupt $\|B\|_{[0,1],\infty}$ verteilt?

4.4.6 Satz über die Supremum-Verteilung der Brownschen Brücke

Sei $(B_t)_{0 \leq t \leq 1}$ der Prozess der Brownschen Brücke auf $(\Omega, \mathfrak{S}, \mathcal{P})$. Dann gilt:

1.

$$\mathcal{P} \left(\sup_{0 \leq t \leq 1} |B_t| \leq u \right) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 u^2} =: K(u) \quad , \quad u > 0$$

2.

$$\mathcal{P} \left(\sup_{0 \leq t \leq 1} B_t \leq u \right) = 1 - e^{-2u^2} =: K^+(u)$$

Beweisidee: Gezeigt wird im Wesentlichen

$$\mathcal{P} \left(W \in A \mid 0 \leq W_1 \leq \varepsilon \right) \xrightarrow[\text{w}]{\varepsilon \searrow 0} \mathcal{P}_B \quad , \quad A \subseteq \mathcal{B}(\mathcal{C}[0,1])$$

wobei (W_t) der Wiener-Prozess ist.

4.4.7 Der Kolmogorov-Test als statistische Anwendung

Betrachtet seien die unabhängigen, identisch verteilten *Versuchsergebnisse* x_1, \dots, x_n , dazu die empirische Verteilungsfunktion

$$F_n^*(t) := F_n^*(t)(x_1, \dots, x_n) = \frac{1}{n} \# \{j \leq n : x_j \leq t\}$$

und die *Nullhypothese*, x_1, \dots, x_n seien gemäß der Verteilungsfunktion F verteilt. Seien nun

$$\mathcal{F}_n := \sup_{t \in \mathbb{R}} |F_n^*(t) - F(t)| \quad , \quad \mathcal{F}_n^+ := \sup_{t \in \mathbb{R}} (F_n^*(t) - F(t))$$

dann geht nach 4.4.2 & 4.4.3:

$$\sqrt{n} \cdot \mathcal{F}_n \xrightarrow[\text{d}]{n \rightarrow \infty} \sup_{0 \leq t \leq 1} |B_t| \quad , \quad \sqrt{n} \cdot \mathcal{F}_n^+ \xrightarrow[\text{d}]{n \rightarrow \infty} \sup_{0 \leq t \leq 1} B_t$$

und

$$\mathcal{P} \left(\sqrt{n} \cdot \mathcal{F}_n \leq u \right) \xrightarrow[\text{2.1.2(4)}]{n \rightarrow \infty} \mathcal{P} \left(\sup_{0 \leq t \leq 1} |B_t| \leq u \right) \stackrel{(4.4.6)}{=} K(u) := \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 u^2}$$

$$\mathcal{P} \left(\sqrt{n} \cdot \mathcal{F}_n^+ \leq u \right) \xrightarrow{n \rightarrow \infty} \mathcal{P} \left(\sup_{0 \leq t \leq 1} B_t \leq u \right) \stackrel{(4.4.6)}{=} K^+(u) := 1 - e^{-2u^2}$$

Ist nun $\alpha > 0$ (*Sicherheitswahrscheinlichkeit*) klein genug vorgegeben, dazu $u_\alpha, u_\alpha^+ > 0$ gegeben durch

$$K(u_\alpha) \stackrel{!}{=} 1 - \alpha \stackrel{!}{=} K^+(u_\alpha^+) \tag{4.4.7.1}$$

das heißt

$$\mathcal{P} \left(\mathcal{F}_n > \frac{u_\alpha}{\sqrt{n}} \right) \xrightarrow{n \rightarrow \infty} 1 - K(u_\alpha) = \alpha \xrightarrow{n \rightarrow \infty} \mathcal{P} \left(\mathcal{F}_n^+ > \frac{u}{\sqrt{n}} \right)$$

Gilt nun für großes $n \in \mathbb{N}$ doch (trotz *Erwartungen*)

$$\mathcal{F}_n(\omega) > \frac{u_\alpha}{\sqrt{n}}$$

so ist ein Ereignis mit eigentlicher Wahrscheinlichkeit $\lesssim \alpha$ eingetroffen! Die Hypothese ist dann natürlicherweise abzulehnen. Analog wird auch im Fall

$$\mathcal{F}_n^+(\omega) > \frac{u_\alpha^+}{\sqrt{n}} = \sqrt{\frac{-\ln \alpha}{2n}}$$

vorgegangen. In anderen Worten, kann der Wertebereich von (x_1, \dots, x_n) aufgeteilt werden gemäß

$$\mathbb{R}^n = \mathcal{H}_0^n \cup \mathcal{H}_1^n$$

mit dem *Annahme-Bereich*

$$\mathcal{H}_0^n := \left\{ \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : \mathcal{F}_n(\mathbf{x}) \leq \frac{u_\alpha}{\sqrt{n}} \right\}$$

bzw. dem *kritischen Bereich* $\mathcal{H}_1^n := (\mathcal{H}_0^n)^c$, und es gilt

$$\mathcal{P}((x_1, \dots, x_n) \in \mathcal{H}_0^n) \xrightarrow{n \rightarrow \infty} 1 - \alpha$$

Es stellt sich nun die Frage nach der Konsistenz dieses Tests. Das heißt, inwiefern geht $\mathcal{P}((x_1, \dots, x_n) \in \mathcal{H}_1^n) \xrightarrow{n \rightarrow \infty} 1$ im Falle einer falschen Nullhypothese?

4.4.8 Bemerkung: Interpretation des Kolmogorov-Tests

Sei F eine stetige Verteilungsfunktion. Zu iid Zufallsvariablen $\underbrace{Y_1, \dots, Y_n}_Y \sim F$ auf $(\Omega, \mathfrak{E}, \mathcal{P})$ seien für jedes $\omega \in \Omega$

die $\tilde{Y}_1(\omega) \leq \dots \leq \tilde{Y}_n(\omega)$ die geordneten $Y_i(\omega)$ (*Ordnungsstatistik*). Deren empirische Verteilungsfunktion ist gegeben ist durch

$$F_n^*(t, \omega) := F_n^*(t, Y) \stackrel{\text{def}}{=} \frac{1}{n} \# \{j \leq n : Y_j(\omega) \leq t\} = \frac{j}{n} \quad \text{wobei} \quad \tilde{Y}_j(\omega) \leq t < \tilde{Y}_{j+1}(\omega)$$

wobei im Kolmogorov-Test genau die Größen

$$\mathcal{F}_n^+(Y) := \sup_{t \in \mathbb{R}} [F_n^*(t, Y) - F(t)] = \sup_{1 \leq j \leq n} \left[\frac{j}{n} - F(\tilde{Y}_j) \right]$$

$$\mathcal{F}_n(Y) := \sup_{t \in \mathbb{R}} |F_n^*(t, Y) - F(t)|$$

getestet werden (vgl. Abbildung 4.11).

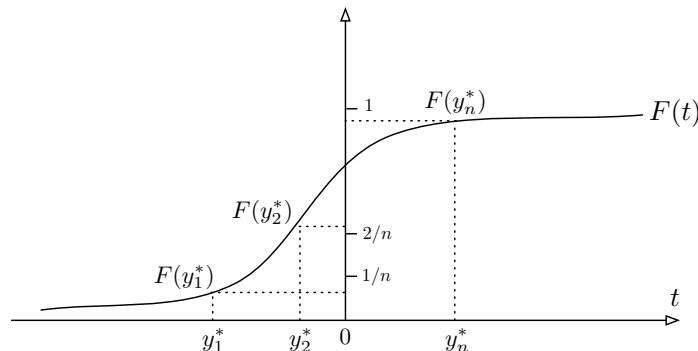


Abbildung 4.11: Zur Definition von \mathcal{F}_n und \mathcal{F}_n^+ .

Tatsächlich entsprechen nach Hilfslemma A.0.14 die Größen $F(\tilde{Y}_j)$ der Ordnungsstatistik auf $[0, 1]$ gleichverteilter Zufallsgrößen! Der Kolmogorov-Test stellt somit einen Vergleich zwischen den (im Idealfall *äquidistanten*) auf $[0, 1]$ verteilten $F(\tilde{Y}_j)$ und den Referenzwerten $\frac{1}{n}, \dots, \frac{n}{n}$ dar.

4.4.9 Hilfsaussage über punktweise konvergente Zufallsvariablen

Seien X_n, X reelle Zufallsvariablen auf $(\Omega, \mathfrak{G}, \mathcal{P})$ und $(a_n) \subseteq \mathbb{R}$ mit

$$\liminf_{n \rightarrow \infty} X_n > \limsup_{n \rightarrow \infty} a_n$$

\mathcal{P} -fast sicher. Dann gilt:

$$\mathcal{P}(\{X_n > a_n\}) \xrightarrow{n \rightarrow \infty} 1$$

Beweis:

$$\begin{aligned} \mathcal{P}(\{X_n > a_n\}) &\geq \mathcal{P}\left(\underbrace{\left\{\inf_{k \geq n} X_k > \sup_{k \geq n} a_k\right\}}_{\substack{\text{monoton wachsend} \\ \text{über } n}}\right) \xrightarrow{n \rightarrow \infty} \mathcal{P}\left[\bigcup_{n \in \mathbb{N}} \left\{\inf_{k \geq n} X_k > \sup_{k \geq n} a_k\right\}\right] \\ &\geq \mathcal{P}\left[\left\{\liminf_{n \rightarrow \infty} X_n > \limsup_{n \rightarrow \infty} a_n\right\}\right] = 1 \end{aligned}$$

□

4.4.10 Satz über die Konsistenz des Kolmogorov Tests

Seien $F \neq G$ stetige Verteilungsfunktionen mit empirischen Verteilungsfunktionen F_n^*, G_n^* auf $(\Omega, \mathfrak{G}, \mathcal{P})$. Sei $\alpha \in (0, 1)$ und u_α definiert wie oben in (4.4.7.1). Sei

$$\mathcal{H}_0^n := \left\{ \omega : \sup_{t \in \mathbb{R}} |G_n^*(t, \omega) - F(t)| \leq \frac{u_\alpha}{\sqrt{n}} \right\}$$

Dann geht

$$\mathcal{P}(\mathcal{H}_0^n) \xrightarrow{n \rightarrow \infty} 0$$

Beweis: Nach Voraussetzung $\exists t_0 \in \mathbb{R} : |G(t_0) - F(t_0)| =: \delta > 0$. Dann

$$\mathcal{P}((\mathcal{H}_0^n)^c) = \mathcal{P}\left(\sup_{t \in \mathbb{R}} |G_n^*(t) - F(t)| \geq \frac{u_\alpha}{\sqrt{n}}\right) \geq \mathcal{P}\left(|G_n^*(t_0) - F(t_0)| \geq \frac{u_\alpha}{\sqrt{n}}\right)$$

Wegen $G_n^*(t_0) \xrightarrow{n \rightarrow \infty} G(t_0)$ fast sicher, ist

$$\liminf_{n \rightarrow \infty} |G_n^*(t_0) - F(t_0)| = \delta > 0$$

fast sicher. Zusammen mit $\frac{u_\alpha}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0$ folgt nach Hilfsaussage 4.4.9:

$$\mathcal{P}((\mathcal{H}_0^n)^c) \xrightarrow{n \rightarrow \infty} 1$$

□

Interpretation: Gegeben sei eine Folge unabhängig, gemäß der stetigen Verteilungsfunktion G verteilter, zufälliger Messgrößen X_1, X_2, \dots , dazu (fälschlicherweise) vermutete stetige Verteilungsfunktion F (Nullhypothese). Gemäß des Kolmogorov-Tests sei

$$G_n^*(t) := \frac{1}{n} \# \{j \leq n : X_j \leq t\}$$

die durch die Messwerte definierte empirische Verteilungsfunktion. Ist $\alpha \in (0, 1)$ und u_α definiert wie in (4.4.7.1), dazu

$$\mathcal{H}_0^n := \left\{ \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : \sup_{t \in \mathbb{R}} |G_n^*(t)(x_1, \dots, x_n) - F(t)| \leq \frac{u_\alpha}{\sqrt{n}} \right\}$$

der Annahmehereich der Hypothese $X_i \sim F$, so müsse im Fall ihrer Richtigkeit $\mathcal{P}(\mathcal{H}_0^n) \xrightarrow{n \rightarrow \infty} 1 - \alpha$ gehen (vgl. Kolmogorov-Test 4.4.7). Tatsächlich geht jedoch wegen $G \neq F$:

$$\mathcal{P}(\mathcal{H}_0^n) \xrightarrow{n \rightarrow \infty} 0$$

(Konsistenz des Tests).

4.4.11 Verhältnis zwischen den beiden Kolmogorov-Test-Varianten

Seien nun $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n$ unabhängig, gleichverteilt auf $[0, 1]$, dazu die Ordnungsstatistik $\tilde{\mathcal{U}}_1, \dots, \tilde{\mathcal{U}}_n$, G_n^* die entsprechende empirische Verteilungsfunktion und

$$\mathcal{G}_n^+ := \sup_{0 \leq t \leq 1} [G_n^*(t) - t] = \sup_{1 \leq j \leq n} [G_n^*(\tilde{\mathcal{U}}_j) - \tilde{\mathcal{U}}_j] = \sup_{1 \leq j \leq n} \left[\frac{j}{n} - \tilde{\mathcal{U}}_j \right]$$

(vgl. Abbildung 4.12).

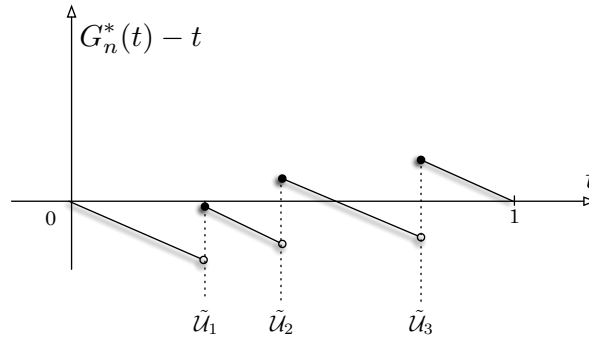


Abbildung 4.12: Typischer Verlauf eines Pfades von $(G_n^*(t) - t)$ ($n = 3$).

Beachte dass auch $(1 - \mathcal{U}_1), \dots, (1 - \mathcal{U}_n)$ auf $[0, 1]$ unabhängig, gleichverteilt sind, also

$$\begin{aligned} \mathcal{G}_n^+ &\stackrel{\text{d.}}{=} \sup_{1 \leq j \leq n} \left[\frac{j}{n} - \underbrace{(1 - \tilde{\mathcal{U}}_{n-j+1})}_{(1-\mathcal{U})_j} \right] = \sup_{1 \leq j \leq n} \left[\tilde{\mathcal{U}}_{n-j+1} - \frac{n-j}{n} \right] \\ &= \sup_{1 \leq j \leq n} \left[\tilde{\mathcal{U}}_j - \frac{j-1}{n} \right] = \sup_{1 \leq t \leq 1} [t - G_n^*(t)] \end{aligned} \quad (\clubsuit)$$

(vgl. Abbildung 4.13).

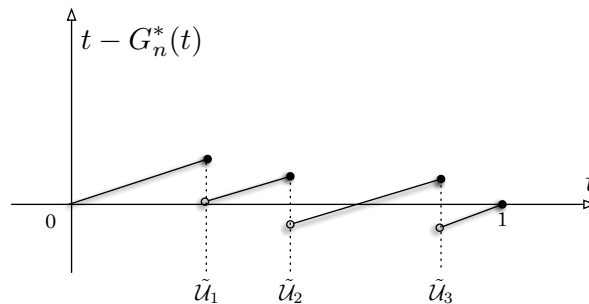


Abbildung 4.13: Typischer Verlauf eines Pfades von $(t - G_n^*(t))$ ($n = 3$).

Zu gegebenem $\alpha \in (0, 1)$ wähle nun u_α, u_α^+ so dass³⁸

$$\lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n > u_\alpha) = \alpha \quad , \quad \lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n^+ > u_\alpha^+) = \alpha$$

³⁸Es lässt sich zeigen:

$$\mathcal{P}(\mathcal{G}_n^+ \geq u) = \sum_{j=0}^{\lfloor n(1-u) \rfloor} u \cdot \binom{n}{j} \left[u + \frac{j}{n} \right]^{j-1} \left[1 - u - \frac{j}{n} \right]^{n-j} \quad , \quad u > 0$$

(vgl. Kolmogorov-Test 4.4.7). Dann:

$$\lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n > u_\alpha) = \alpha = \lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n^+ > u_\alpha^+) \stackrel{\mathcal{G}_n \geq \mathcal{G}_n^+}{\leq} \lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n > u_\alpha^+)$$

das heißt $u_\alpha^+ \leq u_\alpha$. Andererseits ist für $u \in \mathbb{R}$:

$$\begin{aligned} \mathcal{P}(\mathcal{G}_n > u) &\leq \mathcal{P}\left(\sup_{0 \leq t \leq 1} [G_n^*(t) - t] > u\right) + \mathcal{P}\left(\inf_{0 \leq t \leq 1} [G_n^*(t) - t] < -u\right) \\ &= \mathcal{P}(\mathcal{G}_n^+ > u) + \underbrace{\mathcal{P}\left(\sup_{0 \leq t \leq 1} [t - G_n^*(t)] > u\right)}_{\substack{\mathcal{P}(\mathcal{G}_n^+ > u) \\ \text{wegen } (\spadesuit)}} = 2\mathcal{P}(\mathcal{G}_n^+ > u) \quad , \quad u \in \mathbb{R}, n \in \mathbb{N} \quad (\spadesuit) \end{aligned}$$

bzw.

$$\underbrace{\lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n^+ > u_\alpha^+)}_{2 \lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n^+ > u_\alpha^+/2)} = \alpha = \lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n > u_\alpha) \stackrel{(\spadesuit)}{\leq} 2 \lim_{n \rightarrow \infty} \mathcal{P}(\sqrt{n} \cdot \mathcal{G}_n^+ > u_\alpha)$$

das heißt $u_\alpha \leq u_\alpha^+/2$. Zusammengefasst also

$$\boxed{u_\alpha^+ \leq u_\alpha \leq u_\alpha/2}$$

was genau einen Zusammenhang zwischen den beiden Testvarianten herstellt.

4.4.12 Kolmogorov-Smirnov-Test

Gegeben seien zwei Verteilungsfunktionen F, \tilde{F} , dazu jeweils die (unabhängigen) empirischen Verteilungsfunktionen F_n^*, \tilde{F}_n^* . Seien außerdem $m = m_n$ so dass

$$0 < \liminf_{n \rightarrow \infty} \frac{n}{m_n} \leq \limsup_{n \rightarrow \infty} \frac{n}{m_n} < \infty$$

Setzen

$$\mathcal{F}_{n,m}(\omega) := \sup_{t \in \mathbb{R}} |F_n^*(t, \mathbf{X}(\omega)) - F_m^*(t, \mathbf{Y}(\omega))| \quad \left| \quad X_1, \dots, X_n \text{ iid } \sim F, Y_1, \dots, Y_m \text{ iid } \sim \tilde{F} \right.$$

$$\stackrel{\text{d.}}{=} \sup_{t \in \mathbb{R}} |F_n^*(t) - \tilde{F}_m^*(t)|$$

Unter der Nullhypothese $F = \tilde{F}$ müsse nach Smirnov 4.4.5 gelten:

$$\sqrt{\frac{m_n \cdot n}{m_n + n}} \cdot \mathcal{F}_{n,m_n} \xrightarrow[n \rightarrow \infty]{\text{d}} \sup_{0 \leq s \leq 1} |B_s|$$

und somit nach Charakterisierungssatz 2.1.2 (4):

$$\mathcal{P}\left(\sqrt{\frac{m_n \cdot n}{m_n + n}} \cdot \mathcal{F}_{n,m_n} > u\right) \xrightarrow[n \rightarrow \infty]{\text{d}} \mathcal{P}\left(\sup_{0 \leq s \leq 1} |B_s| > u\right) \stackrel{(4.4.6)}{=} 1 - K(u) \quad , \quad u > 0$$

Für vorgegebenes $\alpha \in (0, 1)$ (klein), sei $u_\alpha > 0$ derart dass $K(u_\alpha) = 1 - \alpha$. Dann kann analog zum Kolmogorov-Test der Messraum zerlegt werden

$$\mathbb{R}^n \times \mathbb{R}^m = \underbrace{\mathcal{H}_0^{n,m}}_{(\mathcal{H}_1^{n,m})^c} \cup \mathcal{H}_1^{n,m} \quad , \quad \mathcal{H}_1^{n,m} := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^m : \sqrt{\frac{m \cdot n}{m + n}} \cdot \mathcal{F}_{n,m}(\mathbf{x}, \mathbf{y}) > u_\alpha \right\}$$

in den Annahme-Bereich $\mathcal{H}_0^{n,m}$ und den kritischen Bereich $\mathcal{H}_1^{n,m}$, wobei

$$\mathcal{P}((\mathbf{X}, \mathbf{Y}) \in \mathcal{H}_1^{n,m_n}) \xrightarrow{n \rightarrow \infty} \alpha$$

Für genügend große $n \in \mathbb{N}$ und genügend kleine $\alpha > 0$ kann somit das Eintreffen des Falls $(\mathbf{X}, \mathbf{Y}) \in \mathcal{H}_1^{n,m_n}$ als Indiz für die nicht-Gültigkeit der Nullhypothese gesehen werden.

4.4.13 Cramér von Mises Statistik

Gegeben sei die stetige, streng monoton wachsende Verteilungsfunktion F , dazu das Maß $\mathcal{P}_F \sim F$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$:

$$\mathcal{P}_F((a, b]) = F(b) - F(a)$$

und empirische Verteilungsfunktion F_n^* . Dann ist

$$\begin{aligned} I_n^2 &:= n \int_{\mathbb{R}} [F_n^*(t) - F(t)]^2 \mathcal{P}_F(dt) \stackrel{(4.0.16)}{=} \int_{\mathbb{R}} [G_n^*(F(t)) - F(t)]^2 \mathcal{P}_F(dt) \\ &\stackrel{F}{=} n \int_0^1 [G_n^*(s) - s]^2 (\mathcal{P}_F \circ F^{-1})(ds) = \int_0^1 \underbrace{n [G_n^*(s) - s]^2}_{[U_n^*(s)]^2} ds \end{aligned}$$

wobei U_n^* , G_n^* jeweils die empirische Verteilungsfunktion und der empirische Prozess der Gleichverteilung auf $[0, 1]$ seien. Da die Zuordnung

$$\mathfrak{B}[0, 1] \rightarrow \mathbb{R}, \quad f \mapsto \int_0^1 f^2(t) dt, \quad f \in \mathfrak{B}[0, 1]$$

stetig in $(\mathfrak{B}[0, 1], \|\cdot\|_\infty)$ ist, geht nach 4.4.1:

$$I_n^2 \xrightarrow[n \rightarrow \infty]{d} h(B) = \int_0^1 |B_t|^2 dt$$

Nach (3.5.7.2) besitzt B_t die Reihendarstellung

$$B_t = \sqrt{2} \cdot \sum_{k=1}^{\infty} \xi_k \frac{\sin(k\pi t)}{k\pi} \quad \Bigg| \quad \xi_1, \xi_2, \dots \sim \mathcal{N}_{0,1}, \text{ iid}$$

wobei die $\sqrt{2} \sin(k\pi \cdot)$ orthonormal sind in $L_2[0, 1]$. Nach Pythagoras folgt dann

$$I_n^2 \xrightarrow[n \rightarrow \infty]{d} \sum_{k=1}^{\infty} \frac{\xi_k^2}{(k\pi)^2}$$

und nach Charakterisierungssatz 2.1.2 (4):

$$\mathcal{P}(I_n^2 > u) \xrightarrow{n \rightarrow \infty} \mathcal{P}\left(\sum_{k=1}^{\infty} \frac{\xi_k^2}{(k\pi)^2} > u\right), \quad u \in \mathbb{R}$$

A Anhang

A.0.14 Hilfslemma über die Gleichverteilung

Sei F eine stetige Verteilungsfunktion und $X \sim F$. Dann ist $F(X)$ gleichverteilt auf $[0, 1]$.

Beweis: Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \rightarrow \mathbb{R}$ verteilt gemäß F und

$$\mathcal{F}^-(\alpha) := \inf \{t \in \mathbb{R} : F(t) = \alpha\} \quad , \quad \alpha \in (0, 1)$$

Offensichtlich ist $F(X) : \Omega \rightarrow [0, 1]$ und für $\alpha \in (0, 1)$ gilt

$$\mathcal{P}(X \leq F^-(\alpha)) \geq \mathcal{P}(F(X) < \alpha) \stackrel{\text{F stetig}}{=} \mathcal{P}(F(X) \leq \alpha) \geq \mathcal{P}(X \leq F^-(\alpha))$$

das heißt

$$\mathcal{P}(F(X) \leq \alpha) = \mathcal{P}(X \leq F^-(\alpha)) = F(F^-(\alpha)) = \alpha$$

Wegen Stetigkeit von oben, gilt

$$\mathcal{P}(F(X) = 0) = 0$$

das heißt tatsächlich $F(X)$ gleichverteilt auf $[0, 1]$. \square

A.0.15 Hilfslemma zur Verschiebungsinvarianz der empirischen Gleichverteilung

Für die empirische Gleichverteilung (G_n^*) auf $[0, 1]$ gilt:

$$[G_n^*(s) - G_n^*(t)]_{t \leq s \leq 1} \cong [G_n^*(s - t)]_{t \leq s \leq 1} \quad , \quad t \in [0, 1]$$

Beweis: Sei $G_n^* : [0, 1] \times (\Omega, \mathfrak{S}, \mathcal{P}) \rightarrow [0, 1]$ und G die Verteilungsfunktion der Gleichverteilung. Per Konstruktion nehmen $G_n^*(t)$ und $[G_n^*(s) - G_n^*(t)]$ nur die Werte $\frac{k}{n}$, $k = 0, \dots, n$. Gegeben seien nun $0 \leq k_1 \leq k_2 \leq \dots \leq k_m \leq n$ und $t \leq s_1 \leq \dots \leq s_m \leq 1$, dann entspricht

$$\mathcal{P} \left(G_n^*(s_1) - G_n^*(t) \leq \frac{k_1}{n}, \dots, G_n^*(s_m) - G_n^*(t) = \frac{k_m}{n} \right)$$

genau der Wahrscheinlichkeit, dass von n unabhängigen, auf $[0, 1]$ gleichverteilten $\mathcal{U}_1, \dots, \mathcal{U}_n$, k_1 Werte in $(t, s_1]$, $k_2 - k_1$ Werte in $(s_1, s_2]$, ..., $k_m - k_{m-1}$ Werte in $(s_{m-1}, s_m]$ und $n - k_m$ Werte in $[0, t] \cup (s_m, 1]$ annehmen (vgl. Abbildung A.1), und ist gegeben durch die Multinomialverteilung³⁹

$$n! \cdot \frac{[G(s_1) - G(t)]^{k_1} \cdot \dots \cdot [G(s_m) - G(s_{m-1})]^{k_m - k_{m-1}} \cdot [G(t) - G(0) + G(1) - G(s_m)]^{n - k_m}}{k_1! \cdot \dots \cdot k_m! \cdot (n - k_m)!} \quad (\clubsuit)$$

³⁹Gegeben eine $\{1, \dots, m\}$ -wertige Zufallsvariable X die mit dem Verteilungsgesetz $\mathcal{P}(X = k) = p_k$, $k \in \{1, \dots, m\}$. Zu N unabhängigen Kopien X_1, \dots, X_N bezeichne Z_k , $i = 1, \dots, m$ jeweils die Anzahl der Zufallsvariablen die den Wert k annehmen. Dann ist (Z_1, \dots, Z_m) Multinomialverteilt gemäß:

$$\mathcal{P}(Z_1 = z_1, \dots, Z_m = z_m) = \frac{N!}{z_1! \cdot \dots \cdot z_m!} \cdot p_1^{z_1} \cdot \dots \cdot p_m^{z_m} \quad , \quad z_k \in \{0, \dots, N\}$$

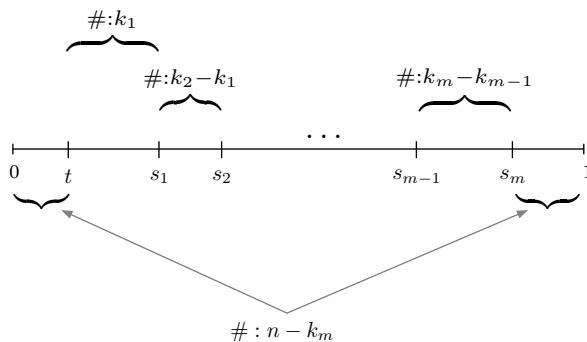


Abbildung A.1: Zur endlich-dimensionalen Verteilung von $[G_n^*(s) - G_n^*(t)]_{t \leq s \leq 1}$.

Andererseits entspricht

$$\mathcal{P} \left(G_n^*(s_1 - t) = \frac{k_1}{n}, \dots, G_n^*(s_m - t) = \frac{k_m}{m} \right)$$

genau der Wahrscheinlichkeit dass von den $\mathcal{U}_1, \dots, \mathcal{U}_n$, genau k_1 Werte in $[0, s_1 - t]$, $k_2 - k_1$ Werte in $(s_1 - t, s_2 - t]$, ..., $k_m - k_{m-1}$ Werte in $(s_{m-1} - t, s_m - t]$ und $n - k_m$ Werte in $(s_m - t, 1]$ annehmen (vgl. Abbildung A.2), und ist gegeben durch die Multinomialverteilung

$$n! \cdot \frac{G(s_1 - t)^{k_1} [G(s_2 - t) - G(s_1 - t)]^{k_2 - k_1} \dots [G(s_m - t) - G(s_{m-1} - t)]^{k_m - k_{m-1}} [1 - G(s_m - t)]^{n - k_m}}{k_1! \cdot (k_2 - k_1)! \cdot \dots \cdot (k_m - k_{m-1})! \cdot (n - k_m)!} \quad (\spadesuit)$$

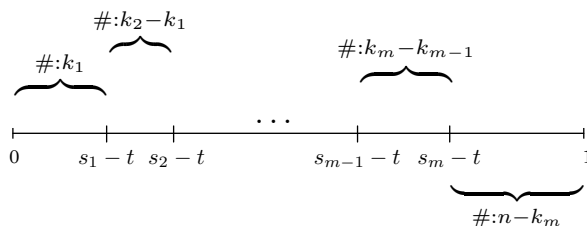


Abbildung A.2: Zur endlich-dimensionalen Verteilung von $[G_n^*(s - t)]_{t \leq s \leq 1}$.

Offensichtlich sind die Ausdrücke (\clubsuit) und (\spadesuit) identisch.

□

A.0.16 Hilfslemma: Eindeutigkeit der Pfad-Supremums-Verteilung

Sei (T, d) ein separabler, metrischer Raum, $(M, \|\cdot\|_M)$ ein separabler, normierter Raum und $(X_t)_{t \in T}$ ein M -wertiger, stochastischer Prozess $(\Omega, \mathfrak{S}, \mathcal{P})$ mit Pfaden in $\mathcal{C}(T, M) \subseteq M^T$. Zusammen mit der Norm $\|f\|_\infty := \sup_{t \in T} \|f(t)\|_M$ wird $(\mathcal{C}(T, M), \|\cdot\|_\infty)$ zu einem separablen Banachraum und

$$X : \Omega \rightarrow (\mathcal{C}(T, M), \mathcal{B}(\mathcal{C}(T, M)))$$

ist messbar (vgl. Satz 2.3.4). Es gilt nun: Die Verteilung von $\|X\|_\infty : \Omega \rightarrow [0, \infty)$ ist unabhängig von der Version von X , das heißt Äquivalenzinvariant.

Beweis: Da T separabel ist, existiert eine abzählbare, in T dichte Teilmenge $T_0 := \{t_1, t_2, \dots\} \subseteq T$. Daher:

$$\|f\|_\infty = \sup_{t \in T_0} \|f(t)\|_M \quad , \quad f \in \mathcal{C}(T, M)$$

und für $\varkappa \in \mathbb{R}$ gilt:

$$\begin{aligned} \mathcal{P}_{\|X\|_\infty}(-\infty, \varkappa) &\stackrel{\text{def.}}{=} \mathcal{P}(\|X\|_\infty \leq \varkappa) = \mathcal{P}\left(\bigcap_{t \in T_0} \{\|X_t\|_M \leq \varkappa\}\right) = \mathcal{P}\left(\underbrace{\bigcap_{k=1}^{\infty} \bigcap_{i=1}^k \{\|X_{t_i}\|_M \leq \varkappa\}}_{\text{monoton fallend in } k}\right) \\ &= \lim_{k \rightarrow \infty} \mathcal{P}\left(\underbrace{\bigcap_{i=1}^k \{\|X_{t_i}\|_M \leq \varkappa\}}_{\text{nur abhängig von endlich dim. Verteilungen}}\right) \end{aligned}$$

□

A.0.17 Maximalungleichung für unabhängige, symmetrische Zufallsgrößen

Seien Y_1, Y_2, \dots unabhängige, symmetrische⁴⁰ reelle Zufallsgrößen auf $(\Omega, \mathfrak{G}, \mathcal{P})$, dazu

$$S_k := Y_1 + \dots + Y_k \quad , \quad k \in \mathbb{N}_0$$

Dann gilt die Abschätzung:

$$\mathcal{P}\left(\sup_{1 \leq k \leq m} |S_k| \geq \eta\right) \leq 2 \cdot \mathcal{P}(|S_m| \geq \eta) \quad , \quad m \in \mathbb{N}, \eta \geq 0$$

A.0.18 Maximalungleichungen für unabhängige, zentrierte Zufallsgrößen

Seien Y_1, Y_2, \dots unabhängige, reelle Zufallsgrößen auf $(\Omega, \mathfrak{G}, \mathcal{P})$ mit

$$\mathbb{E}Y_j = 0 \quad , \quad \sigma_j^2 := \mathbb{E}Y_j^2 < \infty$$

dazu

$$S_k := Y_1 + \dots + Y_k \quad , \quad s_k^2 := \sigma_1^2 + \dots + \sigma_k^2 \quad , \quad k \in \mathbb{N}_0$$

Dann gilt die Abschätzung

$$\mathcal{P}\left(\sup_{1 \leq k \leq m} |S_k| \geq \eta s_m\right) \leq 2 \cdot \mathcal{P}\left(|S_m| \geq (\eta\sqrt{2})s_m\right) \quad , \quad m \in \mathbb{N}, \eta \geq 0$$

A.0.19 Lemma über 4. Momente und Zufallsgrößen

Seien Y_1, Y_2, \dots reelle Zufallsgrößen auf $(\Omega, \mathfrak{G}, \mathcal{P})$ mit 4. Momenten, dazu

$$S_k := Y_1 + \dots + Y_k \quad , \quad k \in \mathbb{N}_0$$

Gegeben seien Zahlen $\lambda_1, \dots, \lambda_m \geq 0$ mit

$$\mathbb{E}\left\{|S_j - S_i|^2 |S_k - S_j|^2\right\} \leq \left(\sum_{l=i+1}^j \lambda_l\right) \left(\sum_{r=j+1}^k \lambda_r\right) \quad , \quad 0 \leq i < j < k \leq m$$

Dann gilt die Abschätzung:

$$\mathcal{P}\left(\sup_{1 \leq k \leq m} \{|S_k| \wedge |S_m - S_k|\} \geq \eta\right) \leq \frac{\overbrace{C}^{\text{const}}}{\eta^4} (\lambda_1 + \dots + \lambda_m)^2 \quad , \quad m \in \mathbb{N}, \eta > 0$$

⁴⁰ $Y_j \stackrel{\text{d.}}{=} -Y_j$

Beispiel: Für Y_1, Y_2, Y_3 erfüllen $\lambda_1, \lambda_2, \lambda_3$ die Voraussetzungen des Lemmas falls:

$$\mathbb{E}(S_2 - S_1)^2(S_3 - S_2)^2 \leq \lambda_1 \lambda_3$$

$$\mathbb{E}(S_1)^2(S_3 - S_1)^2 \leq \lambda_1(\lambda_2 + \lambda_3)$$

$$\mathbb{E}(S_2 - S_1)^2 S_3^2 \leq (\lambda_1 + \lambda_2)\lambda_3$$

A.0.20 Satz über 4. Momente der empirischen Gleichverteilung

Sei U_n^* der empirische Prozess der Gleichverteilung auf $[0, 1]$. Dann gilt die Abschätzung

$$\mathbb{E} \left\{ [U_n^*(t) - U_n^*(t + \alpha)]^2 [U_n^*(t + \alpha) - U_n^*(t + \alpha + \beta)]^2 \right\} \leq 6\alpha\beta \quad , \quad \alpha, \beta > 0, \quad t \in [0, 1], \quad t + \alpha + \beta \leq 1$$

Beweis: Seien U_1, \dots, U_n unabhängige, in $[0, 1]$ gleichverteilte Zufallsvariablen auf $(\Omega, \mathfrak{C}, \mathcal{P})$, dazu

$$\xi_i(\omega) := \begin{cases} 1 - \alpha & : U_i(\omega) \in (t, t + \alpha] \\ -\alpha & : U_i(\omega) \notin (t, t + \alpha] \end{cases}$$

$$\eta_i(\omega) := \begin{cases} 1 - \beta & : U_i(\omega) \in (t + \alpha, t + \alpha + \beta] \\ -\beta & : U_i(\omega) \notin (t + \alpha, t + \alpha + \beta] \end{cases} \quad , \quad \omega \in \Omega, \quad i \in \{1, \dots, n\}$$

das heißt

$$\xi_i = (1 - \alpha) \cdot 1_{(t, t + \alpha]} \circ U_i - \alpha \cdot 1_{(t, t + \alpha]^c} \circ U_i$$

(analog auch für η_i). Per Konstruktion gilt dabei

$$\mathcal{P}(\xi_i = 1 - \alpha) = \alpha \qquad \mathcal{P}(\xi_i = -\alpha) = 1 - \alpha$$

$$\mathcal{P}(\eta_i = 1 - \beta) = \beta \qquad \mathcal{P}(\eta_i = -\beta) = 1 - \beta$$

das heißt $\mathbb{E}\xi_i = 0 = \mathbb{E}\eta_i$. Ferner ist

$$(\xi_i, \eta_i) = \begin{cases} (1 - \alpha, -\beta) & : \text{Wahrscheinlichkeit } \alpha \\ (-\alpha, 1 - \beta) & : \text{Wahrscheinlichkeit } \beta \\ (-\alpha, -\beta) & : \text{Wahrscheinlichkeit } 1 - \alpha - \beta \end{cases}$$

Insbesondere sind alle Paare (ξ_i, η_i) , $i \in \{1, \dots, n\}$ unabhängig, identisch verteilt.

Behauptung: Es gilt

$$\mathbb{E} \left\{ [U_n^*(t) - U_n^*(t + \alpha)]^2 [U_n^*(t + \alpha) - U_n^*(t + \alpha + \beta)]^2 \right\} = \frac{1}{n^2} \cdot \mathbb{E} \left\{ \left(\sum_{i=1}^n \xi_i \right)^2 \left(\sum_{j=1}^n \eta_j \right)^2 \right\} \quad (\spadesuit)$$

Beweis:

Einerseits ist

$$\sum_{i=1}^n \xi_i = \varkappa(1 - \alpha) - (n - \varkappa)\alpha = \varkappa - n\alpha$$

wobei

$$\varkappa := \# \{j \leq n : U_j(\omega) \in (t, t + \alpha]\}$$

Andererseits:

$$\varkappa = n [G_n^*(t + \alpha) - G_n^*(t)]$$

mit der entsprechenden empirischen Verteilungsfunktion G_n^* der Gleichverteilung auf $[0, 1]$. Somit

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i &= \sqrt{n} [G_n^*(t + \alpha) - G_n^*(t) - \alpha] = \sqrt{n} [G_n^*(t + \alpha) - (t + \alpha) - (G_n^*(t) - t)] \\ &= \mathcal{U}_n^*(t + \alpha) - \mathcal{U}_n^*(t) \end{aligned}$$

Ähnlich zeigt man

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n \eta_j = \mathcal{U}_n^*(t + \alpha + \beta) - \mathcal{U}_n^*(t + \alpha)$$

Zusammengefasst folgt die Behauptung.

Behauptung:

$$\mathbb{E} \left\{ \left(\sum_{i=1}^n \xi_i \right)^2 \left(\sum_{j=1}^n \eta_j \right)^2 \right\} \leq 6n^2 \alpha \beta \quad (\clubsuit)$$

Beweis:

Zu klären wäre zunächst, welche Summanden des Ausdrucks

$$\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \mathbb{E}(\xi_i \xi_j \eta_k \eta_l)$$

überhaupt einen Beitrag liefern. So liefern z.B. alle Kombinationen $i = j = k = l$ den Gesamtbeitrag

$$\begin{aligned} \sum_{i=1}^n \mathbb{E}(\xi_i^2 \eta_i^2) &= n \mathbb{E}(\xi_1^2 \eta_1^2) = n [\alpha(1 - \alpha)^2 \beta^2 + \beta \alpha^2 (1 - \beta)^2 + (1 - \alpha - \beta) \alpha^2 \beta^2] \\ &\leq n \cdot [\alpha \beta + \alpha \beta + \alpha \beta] = 3n \alpha \beta \end{aligned}$$

Andererseits liefern die Summanden mit paarweise unterschiedlichen i, j, k, l keinen Beitrag, da die $\xi_i, \xi_j, \eta_k, \eta_l$ in diesen Fällen unabhängig sind:

$$\mathbb{E}(\xi_i \xi_j \eta_k \eta_l) = \underbrace{\mathbb{E}(\xi_i) \mathbb{E}(\xi_j)}_0 \mathbb{E}(\eta_k) \mathbb{E}(\eta_l) = 0$$

Daher sind nur Kombinationen mit mindestens zwei identischen Indizes zu betrachten. Die Kombinationen mit $i = j, k \neq l$, das heißt $k \neq i$ oder $l \neq i$, liefern auch keinen Beitrag, denn z.B. für $k \neq i$ wäre η_k unabhängig von $\xi_i^2 \eta_l$ und

$$\mathbb{E}(\xi_i^2 \eta_l \eta_k) = \mathbb{E}(\xi_i^2 \eta_l) \cdot \underbrace{\mathbb{E}(\eta_k)}_0 = 0$$

Es verbleiben also nur noch 3 Fälle:

- $i = j \neq k = l$. Liefern den Beitrag

$$\begin{aligned} \sum_{i=1}^n \sum_{\substack{k=1 \\ k \neq i}}^n \mathbb{E}(\xi_i^2 \eta_k^2) &= n(n - 1) \mathbb{E}(\xi_1^2) \mathbb{E}(\eta_1^2) = n(n - 1) [\alpha(1 - \alpha)^2 + (1 - \alpha) \alpha^2] \cdot [\beta(1 - \beta)^2 + (1 - \beta) \beta^2] \\ &= n(n - 1) \alpha(1 - \alpha) \beta(1 - \beta) \leq n^2 \alpha \beta \end{aligned}$$

- $i = k \neq j = l$. Liefern den Beitrag

$$\sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \mathbb{E}(\xi_i \eta_i) \mathbb{E}(\xi_j \eta_j) = n(n - 1) (\mathbb{E}(\xi_1 \eta_1))^2 = (\dots) = n(n - 1) \alpha^2 \beta^2 \leq n^2 \alpha \beta$$

- $i = l \neq k = j$. Liefern analog zu oben einen Beitrag $\leq n^2\alpha\beta$.

Zusammengefasst ergibt sich also

$$\mathbb{E} \left\{ \left(\sum_{i=1}^n \xi_i \right)^2 \left(\sum_{j=1}^n \eta_j \right)^2 \right\} \leq 3n\alpha\beta + n^2\alpha\beta + n^2\alpha\beta + n^2\alpha\beta = 6n^2\alpha\beta$$

Aus (\spadesuit) und der Abschätzung (\clubsuit) folgt dann die Behauptung des Satzes.

□

Symbol-Referenz

\mathbb{R}_+ : $[0, \infty)$.

\mathbb{C}_+ : $\mathbb{R}_+ + i\mathbb{R}_+$.

\mathbf{e}_k : Standard-Einheitsvektor in \mathbb{R}^n .

iid: **I**dentically, **I**ntependently **D**istributed. Siehe 1.1.1.

z^* : Komplex-Konjugierte zu $z \in \mathbb{C}$.

$\mathcal{C}(M)$: Raum aller stetigen Funktionen $M \rightarrow \mathbb{C}$.

$\mathcal{C}(M, N)$: Raum aller stetigen Funktionen $M \rightarrow N$.

$\mathcal{C}_b(M)$: Raum aller stetigen, beschränkten Funktionen $M \rightarrow \mathbb{C}$. Konventionell ausgestattet mit der Norm $\|f\|_\infty := \sup_{t \in M} |f(t)|$.

$\mathcal{C}_b(M, N)$: Raum aller stetigen, beschränkten Funktionen $M \rightarrow N$.

1_A : Indikatorfunktion für Menge A .

ζ_t : Auswertefunktion auf S^T : $\zeta_t : f \mapsto f(t)$, $f \in S^T$, $t \in T$.

$\mathcal{P}(M)$: Potenzmenge von M .

$\sigma(\mathcal{A})$: Für $\mathcal{A} \subset \mathcal{P}(M)$, kleinste σ -Algebra die \mathcal{A} enthält.

$\mathfrak{S} \otimes \mathfrak{M}$: Für σ -Algebren $\mathfrak{S}, \mathfrak{M}$, produkt- σ -Algebra: $\mathfrak{S} \otimes \mathfrak{M} := \sigma(\{A \times B : A \in \mathfrak{S}, B \in \mathfrak{M}\})$.

\mathfrak{S}^n : Für σ -Algebra \mathfrak{S} : $\mathfrak{S}^n := \underbrace{\mathfrak{S} \otimes \mathfrak{S} \otimes \dots \otimes \mathfrak{S}}_{\times n}$.

$(T, \mathcal{O}(T))$: Topologischer Raum mit Topologie $\mathcal{O}(T)$.

$\mathcal{B}(T)$: Borel- σ -Algebra von T : $\mathcal{B}(T) := \sigma(\mathcal{O}(T))$.

$(T, \mathcal{B}(T))$: Messbarer Raum über Grundmenge T (topologischen Raum) und Borel- σ -Algebra $\mathcal{B}(T)$.

$\mathfrak{M}(\Omega)$: Für messbaren Raum (Ω, \mathfrak{S}) , die Menge aller Maße auf (Ω, \mathfrak{S}) .

$\mathfrak{M}_1(\Omega)$: Für messbaren Raum (Ω, \mathfrak{S}) die Menge aller Wahrscheinlichkeitsmaße auf (Ω, \mathfrak{S}) .

$(\mathfrak{M}_1(E), \pi)$: $\mathfrak{M}_1(E)$ für metrischen Raum (E, d) mit Prokhorov-Metrik π , siehe 2.3.7 (ii).

$\mu * \nu$: Faltung zweier Maße μ, ν .

$(\Omega, \mathfrak{S}, \mu)$: Maßraum über Grundmenge Ω , mit σ -Algebra \mathfrak{S} und Maß μ .

μ_X : Bildmaß von Maß μ unter Abbildung X : $\mu_X := \mu \circ X^{-1}$.

$\mu * \nu$: Faltung der Maße μ, ν auf $(\Omega, \mathfrak{S}, \mu)$: $\mu * \nu(B) := \int_{\Omega} \mu(B - x) d\nu(x)$.

λ : Lebesgue-Maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

λ^n : Lebesgue-Maß auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

\mathcal{U} : Auf $[0, 1]$ gleichverteilte Zufallsvariable.

$\mathcal{N}_{x_0, s}$: Normalverteilung mit Erwartungswert x_0 und Varianz s .

$\mathcal{N}_{\mu, R}$: Verteilung eines Gaußvektors $X : \Omega \rightarrow \mathbb{R}^n$ mit Erwartungswert $\mathbb{E}(X) = \mu$ und Kovarianzmatrix $R = (\text{Cov}(X_i, X_j))_{i, j=1}^n$.

$\mathcal{B}_{n, p}$: Binomialverteilung mit Versuchs-Erfolgswahrscheinlichkeit p , über n Versuche.

π_λ : Poissonverteilung mit Parameter (Intensität) $\lambda > 0$.

Exp_λ : Exponentialverteilung mit Parameter λ : $\text{Exp}_\lambda((t, \infty)) = e^{-\lambda t}$, $t \geq 0$.

$(X_t)_{t \in T}$: Stochastischer Prozess parametrisiert durch $T \neq \emptyset$. Siehe 1.1.3.

$Z_{t,s}$: Zuwachs eines stochastischen Prozesses $(X_t)_{t \in T}$: $Z_{t,s} := X_s - X_t$.

$\mathcal{P}_{t_1, \dots, t_n}^X$: Endlich-dimensionale Verteilung des stochastischen Prozesses $(X_t)_{t \in T}$, siehe 1.2.1.

\mathcal{P}_X : Verteilungsgesetz der Pfade eines S -wertigen stochastischen Prozesses $(X_t)_{t \in T}$ in $\mathcal{C}(S^T)$, siehe 1.3.5.

$L_p(\Omega)$: Funktionenraum der p -integrierbaren, komplexwertigen Funktionen im Maßraum $(\Omega, \mathfrak{S}, \mu)$.

\bar{A} : Für Untermenge $A \subset X$ eines topologischen Raumes X : Abschluss $\text{cl}(A) = \bar{A}$ von A .

$B_r(x)$: Abgeschlossene Kugel um x mit Radius r .

$B_r^o(x)$: Offene Kugel um x mit Radius r .

B_1^n : n -dimensionale Einheitskugel: $B_1^n := B_1(0)$.

$d(A, B)$: Für Untermengen $A, B \subset X$ eines metrischen Raumes X : $d(A, B) := \inf \{d(a, b) : a \in A, b \in B\}$.

$d(x, A)$: Für Untermenge $A \subset X$ eines metrischen Raumes und $x \in X$: $d(x, A) := d(\{x\}, A)$.

\mathcal{F} : Fouriertransformation.

$\text{supp}(f) := \text{cl} \{x \in X : f(x) \neq 0\}$: Träger von $f : X \rightarrow \mathbb{C}$.

$A_n^{x_0}$: Abbildung $A_n^{x_0} : G^n \rightarrow G^n$, siehe 1.6.1.1.

S^T : Für Mengen $S, T \neq \emptyset$: $S^T := \{f : T \rightarrow S\}$.

$\mathcal{C}_0(S^T)$: Für messbaren Raum (S, \mathcal{S}) und Indexmenge $T \neq \emptyset$: Menge aller Zylindermengen. Siehe 1.3.1.

$\mathcal{C}(S^T) := \sigma(\mathcal{C}_0(S^T))$. Siehe 1.3.3.

$\mathfrak{B}(\Omega)$: Raum aller beschränkten, komplexwertigen, gegebenenfalls messbaren, Funktionen auf Ω . Dabei ist $\mathfrak{B}(\Omega)$ mit der $\|\cdot\|_\infty$ -Norm ausgestattet.

T_π : Für Permutation $\pi \in \text{Sym}(n)$, Koordinaten-Permutation. Siehe 1.3.1.1.

$X \cong Y$: Äquivalenz stochastischer Prozesse $(X_t)_{t \in T}$, $(Y_t)_{t \in T}$. Siehe 1.2.2.

$X \stackrel{\text{d.}}{=} Y$: Zufallsvariablen X, Y identisch verteilt, das heißt $\mathcal{P}_X = \mathcal{P}_Y$.

A^\dagger : Adjungierte Operator zu A . Für Matrix $A \in \mathbb{C}^{n \times n}$ ist $A^\dagger = (A^T)^*$.

\mathcal{R}_ϑ : Rotation in \mathbb{R}^{2n} um Winkel ϑ . Siehe (3.2.8.1).

$\mathbb{E}(X)$: Erwartungswert der Zufallsvariable X .

$\text{Cov}(X, Y)$: Kovarianz zweier Zufallsvariablen X, Y , $\text{Cov}(XY) := \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$.

$\tilde{\mathcal{P}}$: (auch $\mathcal{F}(\mathcal{P})$) Fouriertransformierte des Maßes \mathcal{P} auf $\mathcal{B}(\mathbb{R}^n)$, siehe 3.1.2.

\tilde{X} : (auch $\mathcal{F}(X)$) Fouriertransformierte der Zufallsvariable X , $\tilde{X} := \widetilde{\mathcal{P}_X}$. Siehe 3.1.2.

$\mu_n \xrightarrow[\mathbf{w}]{n \rightarrow \infty} \mu$: Die Maß-Folge (μ_n) strebt schwach gegen μ . Siehe 2.1.1.

$X^n \xrightarrow[\mathbf{p}]{n \rightarrow \infty} X$: Die Zufallsvariablen $(X^n)_{n \in \mathbb{N}}$ gehen stochastisch (in Wahrscheinlichkeit) gegen X , siehe 2.2.1.

$X^n \xrightarrow[\mathbf{d}]{n \rightarrow \infty} X$: Die Zufallsvariablen $(X^n)_{n \in \mathbb{N}}$ gehen in Verteilung (schwach) gegen X , siehe 2.2.1.

$X^n \xrightarrow[\text{f.d.d.}]{n \rightarrow \infty} X$: Die Prozesse $(X_t^n)_{t \in T}$ gehen in endlich-dimensionalen Verteilungen gegen $(X_t)_{t \in T}$, siehe 2.3.1.

F_n^* : Zu Verteilungsfunktion F die empirische Verteilungsfunktion. Siehe (4.0.10.1).

\mathcal{F}_n : Für Verteilungsfunktion F mit empirischer Verteilungsfunktion F_n^* : $\mathcal{F}_n := \sup_{t \in \mathbb{R}} |F_n^*(t) - F(t)|$. Siehe (4.2.1.1).

$(W_t)_{t \geq 0}$: Wiener-Prozess der Molekularbewegung, siehe 1.7.7 Beispiel (i).

$(B_t)_{0 \leq t \leq 1}$: Brownsche Brücke, siehe 3.3.6.

A^ε : Für Untermenge A eines metrischen Raumes (X, d) und $\varepsilon \geq 0$: $A^\varepsilon := \bigcup_{x \in A} B_\varepsilon^o(x)$.

$\mathcal{N}(E, \delta)$: Überdeckungszahl des metrischen Raumes (E, d) bzgl. δ -Kugeln. Siehe 4.3.11.

Literatur

- [1] *Wahrscheinlichkeitstheorie*, H. Bauer
Walter de Gruyter, 2001
- [2] *Markov Chains & Stochastic Stability*, S.P. Meyn, R.L. Tweedie
Springer, 1993
<http://probability.ca/MT/BOOK.pdf> (16 Mai, 2009)
- [3] *Brownian Motion and Stochastic Calculus*, I. Karatzas, S. E. Shreve
Springer, 2004
- [4] *The Nature of Statistical Learning Theory*, V.N. Vapnik
Springer, 2000
- [5] *Weak convergence of measures: applications in probability*, P. Billingsley
SIAM, 1971
- [6] *The Principal Minor Test for Semidefinite Matrices*, J.E. Prussing
Journal of Guidance, Control and Dynamics, Vol. 9, No. 1, 1986