
Gerhard Keller

Wahrscheinlichkeitstheorie

Vorlesung im Wintersemester 2013/14

Department Mathematik, Universität Erlangen-Nürnberg

Letzte Änderung: 6. Februar 2014

Inhaltsverzeichnis

Vorbemerkungen und Literatur	i
I Maß- und Wahrscheinlichkeitsräume	2
1 Mengensysteme, σ -Algebren	2
2 Maße, Wahrscheinlichkeitsmaße	7
3 Messbare Funktionen und Abbildungen	17
4 Zufallsvariablen, Unabhängigkeit	21
II Integral, Erwartungswert und Konvergenzsätze	25
5 Integral und Erwartungswert	25
6 Konvergenzsätze und gleichgradige Integrierbarkeit	31
7 Minkowski, Hölder und Jensen	39
III Maße auf Produkträumen	44
8 Produkträume	44
9 Produktmaße	48
10 Maße auf unendlichen Produkträumen	54
IV Bedingte Erwartung, bedingte Wahrscheinlichkeit	59
11 Zerlegungssätze für Maße, Satz von Radon-Nikodym	59
12 Bedingte Erwartung	64
13 Bedingte Wahrscheinlichkeitsverteilungen	70

V	Gesetz der großen Zahl	76
14	Stationäre Prozesse	76
15	Der Ergodensatz	80
VI	Verteilungskonvergenz und zentraler Grenzwertsatz	83
16	Verteilungskonvergenz	83
17	Charakteristische Funktionen und der Zentrale Grenzwertsatz	89
VII	Martingale	97
18	Martingale, Submartingale, Stoppzeiten	97
19	Der Martingal-Konvergenzsatz	103
	Anhang	109
A	Maßtheorie	109
B	Integrationstheorie	117
C	Diverse Ergänzungen	126

Vorbemerkungen und Literatur

Die fundamentalen Begriffe der Wahrscheinlichkeitstheorie sind *Ereignisse*, *Zufallsvariablen*, *Wahrscheinlichkeiten* und *Erwartungswerte*. Die präzise Grundlage für diese Objekte liefert die Maß- und Integrationstheorie mit folgenden Entsprechungen:

- Ereignis – messbare Menge
- Zufallsvariable – messbare Funktion
- Wahrscheinlichkeiten – Wahrscheinlichkeitsmaß
- Erwartungswert – Integral

Daher gehe ich in dieser Vorlesung und beim Abfassen dieses Skripts davon aus, dass die Inhalte der Modul *Stochastische Modellbildung* und *Maßtheorie* im Wesentlichen bekannt und verstanden sind. Trotzdem werden die für die Wahrscheinlichkeitstheorie benötigten Begriffe und Sätze der Maßtheorie wiederholt - allerdings in der Regel ohne Beweise. Außerdem enthält der Anhang dieses Skripts (nicht die Vorlesung selbst!) viele Teile der Maßtheorie in ausführlicherer Form.

Eine Reihe der folgenden Bücher waren mir bei der Erstellung dieses Vorlesungsskripts hilfreich. Das Buch [5] schlägt die Brücke von der Stochastischen Modellbildung zur maßtheoretisch fundierten Wahrscheinlichkeitstheorie, um die es in dieser Vorlesung geht.

- [1] H. Bauer, *Maß- und Integrationstheorie*, de Gruyter (1990).
- [2] H. Bauer, *Wahrscheinlichkeitstheorie*, de Gruyter (1991).
- [3] P. Billingsley, *Probability and Measure*, Wiley (1986).
- [4] L. Breiman, *Probability*, Addison-Wesley (1968).
- [5] H.-O. Georgii, *Stochastik*, de Gruyter (2002).
- [6] A. Klenke, *Wahrscheinlichkeitstheorie*, Springer (2006).
- [7] S. Resnick, *A Probability Path*, Birkhäuser (1999).

Kapitel 1

Mengensysteme, σ -Algebren

In diesem Kapitel werden die mengentheoretischen Grundlagen der Wahrscheinlichkeitstheorie rekapituliert. Halbringe, Algebren, σ -Algebren und Dynkin-Systeme werden wiederholt und an einem für die Wahrscheinlichkeitstheorie fundamentalen Beispiel illustriert.

Notation: Sei Ω eine Menge. Statt mit 2^Ω (wie in der Maßtheorie) bezeichnen wir die Potenzmenge von Ω mit $\mathfrak{P}(\Omega)$.

1.1 Definition (Halbring, \cap -stabil) $\mathcal{S} \subseteq \mathfrak{P}(\Omega)$ ist ein Halbring (oder auch Semiring), falls

SR1) $\emptyset \in \mathcal{S}$,

SR2) $A, B \in \mathcal{S} \implies A \cap B \in \mathcal{S}$ (\cap -Stabilität),

SR3) $A, B \in \mathcal{S}$ und $A \subseteq B \implies \exists A_1, \dots, A_k \in \mathcal{S}$ paarweise disjunkt mit $B \setminus A = A_1 \cup \dots \cup A_k$.

1.2 Beispiel (Intervalle im \mathbb{R}^k)

$\Omega = \mathbb{R}^k$, $\mathcal{S} = \mathcal{I}^k := \{(\underline{a}, \underline{b}] : -\infty < \underline{a} \leq \underline{b} < +\infty\}$ mit $(\underline{a}, \underline{b}] := \{x \in \mathbb{R}^k : \underline{a} < x \leq \underline{b}\}$. Für Details siehe Beispiel A.1.

1.3 Beispiele (Folgenraum, Zylindermengen) Sei Σ eine endliche Menge, $\Omega = \Sigma^{\mathbb{N}} = \{\omega = (\omega_1, \omega_2, \omega_3, \dots) : \omega_n \in \Sigma \forall n \in \mathbb{N}\}$. Für $a_1, \dots, a_N \in \Sigma$ sei

$$[a_1, \dots, a_N] := \{\omega \in \Omega : \omega_n = a_n (n = 1, \dots, N)\}$$

der von a_1, \dots, a_N bestimmte Zylinder. Sei

$$\mathcal{Z}_N := \{[a_1, \dots, a_N] : a_1, \dots, a_N \in \Sigma\}$$

$$\mathcal{Z} := \bigcup_{N=1}^{\infty} \mathcal{Z}_N \cup \{\emptyset\}.$$

Dann ist \mathcal{Z} ein Halbring:

SR1) $\emptyset \in \mathcal{Z}$ ✓

SR2) $A = [a_1, \dots, a_M]$, $B = [b_1, \dots, b_N]$, o.B.d.A. $0 < M \leq N$.

$$A \cap B = \begin{cases} B, & \text{falls } a_j = b_j \forall j = 1, \dots, M \\ \emptyset & \text{sonst} \end{cases}$$

Außerdem: $A \cap \emptyset = \emptyset \forall A \in \mathcal{Z}$.

SR3) Seien $A, B \in \mathcal{Z}$. Ist $A = \emptyset$, so ist $B \setminus A = B \in \mathcal{Z}$, ist $B = \emptyset$, so ist $B \setminus A = \emptyset \in \mathcal{Z}$.
 Seien nun $A = [a_1, \dots, a_M], B = [b_1, \dots, b_N] \neq \emptyset$ und $A \subseteq B$. Dann ist $M \geq N$ und $a_j = b_j \forall j = 1, \dots, N$.

- ▷ Ist $M = N$, so ist $B = A$, also $B \setminus A = \emptyset$.
- ▷ Ist $M > N$, so ist

$$\begin{aligned} B \setminus A &= \{\omega \in \Omega : \omega_j = a_j \forall j = 1, \dots, N \text{ aber} \\ &\quad \exists j \in \{N + 1, \dots, M\} : \omega_j \neq a_j\} \\ &= \bigcup_{\substack{(c_{N+1}, \dots, c_M) \in \Sigma^{M-N} \\ \exists j \in \{N+1, \dots, M\} : c_j \neq a_j}} [a_1, \dots, a_N, c_{N+1}, \dots, c_M] \end{aligned}$$

1.4 Bemerkung Der Raum $\Omega = \Sigma^{\mathbb{N}}$ des letzten Beispiels wird mit folgender Metrik zu einem kompakten metrischen Raum:

$$d(\omega, \omega') := 2^{-\min\{n \in \mathbb{N} : \omega_n \neq \omega'_n\}}$$

Die Eigenschaften einer Metrik weist man leicht nach. (Übung!) Ebenfalls leicht ist der Beweis, dass alle Zylinder zugleich offen und abgeschlossen sind. (Übung!) Statt der Kompaktheit zeigen wir hier eine zunächst etwas schwächere Eigenschaft, aus der die Kompaktheit aber leicht gefolgert werden kann (Übung!). Diese Eigenschaft, auf die wir später noch einmal zurückgreifen werden, lautet:

Sind $A, A_k \in \mathcal{Z}$ und ist $A \subseteq \bigcup_{k=1}^{\infty} A_k$, so gibt es ein $m \in \mathbb{N}$ mit $A \subseteq \bigcup_{k=1}^m A_k$.

Sei dazu $\mathcal{E} := \{B \in \mathcal{Z} : \exists m \in \mathbb{N} \text{ s.d. } B \subseteq \bigcup_{k=1}^m A_k\}$. Zu zeigen ist also, dass $A \in \mathcal{E}$ ist. Da Σ endlich ist, existiert für jeden Zylinder $[b_1, \dots, b_n] \in \mathcal{Z} \setminus \mathcal{E}$ ein $b_{n+1} \in \Sigma$, so dass $[b_1, \dots, b_{n+1}] \in \mathcal{Z} \setminus \mathcal{E}$. Angenommen, der Zylinder $A = [a_1, \dots, a_N] \in \mathcal{Z} \setminus \mathcal{E}$. Dann folgt induktiv die Existenz von $a_{N+1}, a_{N+2}, \dots \in \Sigma$ derart, dass $[a_1, \dots, a_j] \in \mathcal{Z} \setminus \mathcal{E}$ für alle $j > N$. Andererseits ist $(a_1, a_2, a_3, \dots) \in A = \bigcup_{k=1}^{\infty} (A_k \cap A)$. Also gibt es ein $m \in \mathbb{N}$, für das $(a_1, a_2, a_3, \dots) \in A_m \cap A$. Aber $A_m \cap A$ ist ein Zylinder der Form $[a_1, \dots, a_N, a_{N+1}, \dots, a_{N+l}]$, und es folgt $[a_1, \dots, a_{N+l}] \subseteq A_m \subseteq \bigcup_{k=1}^m A_k$, im Widerspruch zu $[a_1, \dots, a_{N+l}] \in \mathcal{Z} \setminus \mathcal{E}$.

1.5 Beispiel Die Indexmenge \mathbb{N} im vorigen Beispiel kann man durch eine beliebige abzählbare Menge ersetzen, und die Eigenschaften bleiben erhalten. Beispiele:

- $\Omega = \{-1, +1\}^{\mathbb{Z}^d}$ ist die Menge aller $+/-$ -Konfigurationen auf dem d -dimensionalen Gitter.
- Sei $\tilde{\mathbb{Z}}^d$ die Menge aller „Kanten“ im \mathbb{Z}^d , die benachbarte Punkte verbinden. Interpretiert man diese Kanten als Verbindungen, die entweder offen oder geschlossen sein können, so ist $\{0, 1\}^{\tilde{\mathbb{Z}}^d}$ der Raum aller solcher Kantenkonfigurationen. In der Perkolationstheorie nimmt man an, dass jede Kante mit einer gewissen Wahrscheinlichkeit offen oder geschlossen ist, und man fragt z.B. nach der Wahrscheinlichkeit, dass es einen unendlich langen Pfad offener Kanten gibt.

1.6 Definition (Algebra, σ -Algebra) $\mathcal{A} \subseteq \mathfrak{P}(\Omega)$ ist eine Algebra, falls

- A1) $\Omega \in \mathcal{A}$,

$$A2) A \in \mathcal{A} \implies A^c := \Omega \setminus A \in \mathcal{A},$$

$$A3) A_1, \dots, A_n \in \mathcal{A} \implies \bigcup_{j=1}^n A_j \in \mathcal{A}.$$

\mathcal{A} ist eine σ -Algebra, falls statt (A3) das stärkere (A3 σ) gilt:

$$A3\sigma) A_n \in \mathcal{A} (n \in \mathbb{N}) \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Die Mengen $A \in \mathcal{A}$ heißen \mathcal{A} -messbar (oder einfach messbar).

Das Paar (Ω, \mathcal{A}) heißt ein messbarer Raum.

1.7 Definition (Erzeugte σ -Algebra) Sei $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$.

$$\sigma(\mathcal{C}) := \bigcap_{\substack{\mathcal{A} \supseteq \mathcal{C} \\ \mathcal{A} \text{ } \sigma\text{-Algebra}}} \mathcal{A}$$

heißt die von \mathcal{C} erzeugte σ -Algebra oder die kleinste \mathcal{C} enthaltende σ -Algebra.

(Beachte: 1. Der Durchschnitt einer beliebigen Familie von σ -Algebren ist wieder eine σ -algebra (Lemma A.2). 2. $\mathfrak{P}(\Omega)$ ist eine σ -Algebra, so dass der Durchschnitt immer über eine nicht leere Familie von σ -Algebren genommen wird.)

1.8 Bemerkungen

a) $\mathcal{C} \subseteq \sigma(\mathcal{C})$,

b) $\mathcal{C}_1 \subseteq \mathcal{C}_2 \implies \sigma(\mathcal{C}_1) \subseteq \sigma(\mathcal{C}_2)$,

c) \mathcal{A} σ -Algebra $\implies \sigma(\mathcal{A}) = \mathcal{A}$.

1.9 Beispiele (σ -Algebren) Die meisten interessanten σ -Algebren können nicht explizit angegeben werden, sondern sind nur durch $\mathcal{A} = \sigma(\mathcal{S})$ bestimmt, wo \mathcal{S} ein Halbring oder ein anderes einfaches Mengensystem ist. Für uns wichtig sind:

a) $\mathcal{B} := \sigma(\mathcal{I})$ heißt die σ -Algebra der Borelschen Mengen in \mathbb{R} . Analog: $\mathcal{B}^k := \sigma(\mathcal{I}^k)$.

b) Sei $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, $\bar{\mathcal{B}} := \{A \subseteq \bar{\mathbb{R}} : A \setminus \{-\infty, +\infty\} \in \mathcal{B}\}$. Dann ist $\bar{\mathcal{B}}$ eine σ -Algebra.

c) Allgemeiner: Sei (Ω, \mathcal{G}) ein topologischer Raum, d.h. \mathcal{G} ist die Familie aller offenen Teilmengen von Ω .¹ Dann heißt $\mathcal{B} = \sigma(\mathcal{G})$ die Borel- σ -Algebra von Ω . Es gilt auch $\mathcal{B} = \sigma(\{F \subseteq \Omega : F \text{ abgeschlossen}\})$ (Übung!).

d) $\mathcal{F} := \sigma(\mathcal{Z})$ ist die von den Zylindermengen erzeugte σ -Algebra auf $\Omega = \Sigma^{\mathbb{N}}$. Es ist gleichzeitig die Borel- σ -Algebra von $\Sigma^{\mathbb{N}}$. Um das einzusehen beachten wir zunächst:

$$\forall \omega \in \Sigma^{\mathbb{N}} \forall n \in \mathbb{N} : U_{2^{-n}}(\omega) = [\omega_1, \dots, \omega_n],$$

wobei $U_\delta(\omega) := \{\omega' \in \Sigma^{\mathbb{N}} : d(\omega', \omega) < \delta\}$, und stellen fest, dass die abzählbare (Übung!) Menge $\{(a_1, \dots, a_n, a, a, a, \dots) : n \in \mathbb{N}, a, a_1, \dots, a_n \in \Sigma\}$ dicht in $\Sigma^{\mathbb{N}}$ liegt. Die Behauptung folgt dann aus dem nachfolgenden Satz (angewandt mit $\delta_n = 2^{-n}$).

¹Egal, ob es sich um die offenen Teilmengen in einem metrischen Raum handelt, oder ob es um ein abstraktes System offener Mengen geht, es gilt immer: \mathcal{G} is abgeschlossen unter *endlicher* Durchschnittsbildung und unter *beliebiger* (auch überabzählbarer) Vereinigung.

Zur Erinnerung: Ein metrischer Raum ist separabel, falls er eine abzählbare dichte Teilmenge besitzt. Insbesondere ist jeder kompakte metrische Raum separabel.

1.10 Satz

Sei (Ω, d) ein separabler metrischer Raum, $D \subset \Omega$ eine abzählbare dichte Teilmenge und $(\delta_n)_{n \in \mathbb{N}}$ eine Nullfolge positiver Zahlen. Dann gilt für die Borel- σ -Algebra \mathcal{B} von Ω :

$$\mathcal{B} = \sigma(\mathcal{G}) = \sigma(\{U_{\delta_n}(\omega) : \omega \in D, n \in \mathbb{N}\}) .$$

Beweis: „ \supseteq “: Folgt aus Bemerkung 1.8b, da jedes $U_{\delta}(\omega)$ offen ist.

„ \subseteq “: Wegen Bemerkung 1.8b ist nur zu zeigen, dass $\mathcal{G} \subseteq \sigma(\{U_{\delta_n}(\omega) : \omega \in D, n \in \mathbb{N}\})$. Sei also $V \in \mathcal{G}$. Dann ist (Übung!)

$$V = \bigcup_{x \in D, n \in \mathbb{N}: U_{\delta_n}(x) \subseteq V} U_{\delta_n}(x) .$$

Das ist eine abzählbare Vereinigung, also $V \in \sigma(\{U_{\delta_n}(\omega) : \omega \in D, n \in \mathbb{N}\})$. □

1.11 Bemerkung Die Elemente einer σ -Algebra, also die speziellen Teilmengen von Ω , die in der σ -Algebra zusammengefasst werden, heißen in der Wahrscheinlichkeitstheorie auch *Ereignisse*. Die Punkte $\omega \in \Omega$ heißen dagegen *Elementarereignisse*. Schauen wir uns das einmal genauer im Fall des Folgenraums $\Omega = \{0, 1\}^{\mathbb{N}}$ mit der von den Zylindermengen erzeugten σ -Algebra \mathcal{F} an:

Ein Elementarereignis $\omega = (\omega_1, \omega_2, \dots) \in \{0, 1\}^{\mathbb{N}}$ interpretieren wir als eine unendliche Folge von Versuchsergebnissen, bei denen jeder Versuch zwei mögliche Ergebnisse hat, z.B. Erfolg/Misserfolg oder Kopf/Zahl. Hier sind einige Ereignisse:

- $[000] = \{\omega : \omega_1 = \omega_2 = \omega_3 = 0\}$ ist das Ereignis, dass die ersten drei Versuche das Ergebnis 0 haben. Nach Definition gehört $[000]$ zur σ -Algebra \mathcal{F} .
- $\{\omega : \sum_{n=1}^{\infty} \omega_n = \infty\}$ ist das Ereignis, dass das Ergebnis 1 unendlich oft auftritt. Es lässt sich auch schreiben als $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{\omega : \omega_n = 1\}$ und gehört deshalb zur σ -Algebra \mathcal{F} .
- $\{\omega : \lim_{n \rightarrow \infty} n^{-1} \sum_{k=1}^n \omega_k = \frac{1}{2}\}$ ist das Ereignis, dass asymptotisch gleich viele 0en wie 1en beobachtet werden. Es gehört ebenfalls zu \mathcal{F} (Übung!).

Betrachten wir noch die Menge $\Omega = \{0, 1\}^{\tilde{\mathbb{Z}}^d}$ der Perkulationskonfigurationen.

- $\{\omega : \{k \in \tilde{\mathbb{Z}}^d : \omega_k = 1\} \text{ enthält eine unendliche zusammenhängende Komponente}\}$ gehört zur σ -Algebra \mathcal{F} . (Das ist schon etwas schwieriger zu zeigen.)

Für einen flexiblen Umgang mit σ -Algebren benötigen wir noch eine weitere Struktur:

1.12 Definition (Dynkin-System) $\mathcal{D} \subseteq \mathfrak{P}(\Omega)$ heißt Dynkin-System, falls

D1) $\Omega \in \mathcal{D}$

D2) $A, B \in \mathcal{D}, A \subseteq B \implies B \setminus A \in \mathcal{D}$

D3) $A_n \in \mathcal{D} (n \in \mathbb{N})$ paarweise disjunkt $\implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$

(\bigcup bezeichnet eine disjunkte Vereinigung.)

1.13 Bemerkungen

- a) D1), D2) $\implies \emptyset \in \mathcal{D}$
- b) Jede σ -Algebra ist ein Dynkin-System.
- c) Der Durchschnitt beliebig vieler Dynkin-Systeme ist ein Dynkin-System (analog Lemma A.2).

1.14 Definition (Erzeugtes Dynkin-System) Für $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$ sei

$$\mathcal{D}(\mathcal{C}) := \bigcap_{\substack{\mathcal{D} \supseteq \mathcal{C} \\ \mathcal{D} \text{ Dynkin-System}}} \mathcal{D}$$

das von \mathcal{C} erzeugte Dynkin-System. (Es gelten die Bemerkungen 1.8 entsprechend.)

1.15 Lemma Ein \cap -stabiles Dynkin-System \mathcal{D} ist eine σ -Algebra. (Siehe Lemma A.3).

1.16 Satz (Dynkin-System / σ -Algebra)

Sei $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$ \cap -stabil. Dann ist $\sigma(\mathcal{C}) = \mathcal{D}(\mathcal{C})$. (Siehe Satz A.4.)

Kapitel 2

Maße, Wahrscheinlichkeitsmaße

In diesem Kapitel werden Maße eingeführt. Das sind Mengenfunktionen μ , die jeder Menge einer σ -Algebra eine Zahl zwischen 0 und $+\infty$ zuordnen und gewisse Additivitätseigenschaften besitzen. Werden die Werte im Intervall $[0, 1]$ angenommen, so spricht man von *Wahrscheinlichkeitsmaßen*. Wir wissen aus der Maßtheorie, dass man in der Regel ein Maß nicht für alle Elemente einer σ -Algebra explizit angeben muss (und auch nicht kann), sondern dass es reicht, es für Elemente eines erzeugenden Halbrings anzugeben. Wir beginnen mit einigen Beispielen von Wahrscheinlichkeitsmaßen, bevor wir die wesentlichen Aussagen der Theorie wiederholen. In diesen Beispielen werden zunächst nur *Inhalte* definiert.

2.1 Beispiel (Gleichverteilung auf $\Omega = [0, 1]$) Für $0 \leq a \leq b \leq 1$ setzen wir

$$\mu((a, b]) = b - a .$$

2.2 Beispiel (Lebesgue-Stieltjes Maß) Sei $\Omega = \mathbb{R}$ und $F : \Omega \rightarrow \mathbb{R}$ eine Verteilungsfunktion, d.h. es gelte

- (i) F ist monoton wachsend,
- (ii) F ist rechtsseitig stetig ($\lim_{x \searrow x_0} F(x) = F(x_0)$),
- (iii) $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow +\infty} F(x) = 1$.

Für $a \leq b$ setze

$$\mu(a, b] = F(b) - F(a) .$$

Mit $F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$ erhält man die Standardnormalverteilung (Maß mit Dichte).

2.3 Beispiel (Bernoulli-Maß auf $\Omega = \{0, 1\}^{\mathbb{N}}$) Sei $0 < p < 1$. Für $n \in \mathbb{N}$ und $a_1, \dots, a_n \in \{0, 1\}$ setzen wir

$$\begin{aligned} \mu([a_1, \dots, a_n]) &= p^{\sum_{k=1}^n a_k} \cdot (1-p)^{n-\sum_{k=1}^n a_k} \\ &= p^{\text{Anzahl der Erfolge}} \cdot (1-p)^{\text{Anzahl der Misserfolge}} . \end{aligned}$$

2.4 Beispiel (Perkolationswahrscheinlichkeiten auf $\Omega = \{0, 1\}^{\tilde{\mathbb{Z}}^d}$) Zur Erinnerung: $\tilde{\mathbb{Z}}^d$ ist die Menge aller Kanten, die benachbarte Punkte des \mathbb{Z}^d verbinden, d.h. Punkte, die sich in genau einer Koordinate um genau 1 unterscheiden. Nimmt man an, dass jede Kante unabhängig von

allen anderen mit Wahrscheinlichkeit $\in (0, 1)$ offen ist, so ist bei gegebenen $i_1, \dots, i_n \in \tilde{\mathbb{Z}}^d$ und $a_1, \dots, a_n \in \{0, 1\}$

$$A = \{\omega \in \{0, 1\}^{\tilde{\mathbb{Z}}^d} : \omega_{i_j} = a_j (j = 1, \dots, n)\}$$

das Ereignis, dass die Kanten i_1, \dots, i_n die Zustände a_1, \dots, a_n haben, wobei $a_j = 1$ bedeutet, dass die Kante i_j offen ist. Also wird man definieren

$$P(A) = p^{\sum_{j=1}^n a_j} \cdot (1 - p)^{n - \sum_{j=1}^n a_j} .$$

2.5 Definition (Mengenfunktionen) Sei $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$ und $\emptyset \in \mathcal{C}$. Eine Mengenfunktion $\mu : \mathcal{C} \rightarrow \mathbb{R}$ heißt

positiv, wenn $\mu(\emptyset) = 0$ und $\mu(A) \geq 0 \forall A \in \mathcal{C}$,

monoton, wenn $\mu(A) \leq \mu(B) \forall A, B \in \mathcal{C}$ mit $A \subseteq B$,

additiv, wenn $\mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$ für alle paarweise disjunkten $A_1, \dots, A_n \in \mathcal{C}$ mit $\bigcup_{i=1}^n A_i \in \mathcal{C}$,

σ -additiv, wenn $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ für alle paarweise disjunkten $A_1, A_2, \dots \in \mathcal{C}$ mit $\bigcup_{i=1}^{\infty} A_i \in \mathcal{C}$,

subadditiv, wenn $\mu(A) \leq \sum_{i=1}^n \mu(A_i)$ für alle $A, A_1, \dots, A_n \in \mathcal{C}$ mit $A \subseteq \bigcup_{i=1}^n A_i$,

σ -subadditiv, wenn $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$ für alle $A, A_1, A_2, \dots \in \mathcal{C}$ mit $A \subseteq \bigcup_{i=1}^{\infty} A_i$,

σ -endlich, wenn es paarweise disjunkte $B_1, B_2, \dots \in \mathcal{C}$ gibt mit $\Omega = \bigcup_{i=1}^{\infty} B_i$ und $\mu(B_i) < \infty \forall i \in \mathbb{N}$.

Notation: Seien $A, A_1, A_2, A_3, \dots \subseteq \Omega$. Wir schreiben

$$A_n \nearrow A, \text{ falls } A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \text{ und } \bigcup_{n=1}^{\infty} A_n = A$$

$$A_n \searrow A, \text{ falls } A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots \text{ und } \bigcap_{n=1}^{\infty} A_n = A$$

2.6 Definition (Inhalt)

a) Ein Inhalt μ ist eine positive und additive Mengenfunktion auf einer Algebra \mathcal{A} .

b) Ein Inhalt heißt σ -stetig wenn für alle $A, A_n \in \mathcal{A}$ mit $A_n \nearrow A$ gilt

$$\mu(A_n) \nearrow \mu(A) .$$

2.7 Bemerkung Ein Inhalt ist monoton, da für alle $A, B \in \mathcal{A}$ mit $A \subseteq B$ gilt:

$$\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) \geq \mu(A) .$$

2.8 Lemma Ein Inhalt μ auf einer Algebra \mathcal{A} ist σ -endlich genau dann, wenn es $A_k \in \mathcal{A}$ gibt mit $\mu(A_k) < \infty$ und $\Omega = \bigcup_{k=1}^{\infty} A_k$.

Beweis: Zu zeigen ist nur die „wenn“-Richtung: Seien die A_k wie im Lemma. Setze $B_1 := A_1$, $B_i := A_i \setminus \bigcup_{k=1}^{i-1} A_k$ ($i \geq 2$). Dann sind die B_i paarweise disjunkt, $\Omega = \bigcup_{i=1}^{\infty} B_i$, und es ist wegen der Monotonie $\mu(B_i) \leq \mu(A_i) < \infty$ für alle $i \in \mathbb{N}$. \square

Das folgende Lemma wurde in der Maßtheorie gezeigt, siehe Lemma A.6 für einen Beweis.

2.9 Lemma Sei \mathcal{S} ein Halbring, $\mu : \mathcal{S} \rightarrow [0, +\infty]$ additiv. Dann gilt:

i) $A, B \in \mathcal{S}, A \subseteq B \implies \mu(A) \leq \mu(B)$ (Monotonie)

ii) μ ist endlich subadditiv.

Beachte, dass die Subadditivität keine triviale Konsequenz der Additivität ist, da bei der Subadditivität nicht die paarweise Disjunktheit der B_i angenommen wird!

Im nächsten Satz zeigen wir einen engen Zusammenhang zwischen σ -Additivität und σ -Stetigkeit.

2.10 Satz (σ -Stetigkeit und σ -Additivität von Inhalten)

Sei μ ein Inhalt auf der Algebra \mathcal{A} . Äquivalent sind:

(i) μ ist σ -additiv

(ii) μ ist σ -subadditiv

(iii) μ ist σ -stetig

Außerdem folgt aus der σ -Stetigkeit für $A, A_n \in \mathcal{A}$:

(iv) $A_n \searrow A$ und $\mu(A_1) < \infty \implies \mu(A_n) \searrow \mu(A)$

(v) $A_n \searrow \emptyset$ und $\mu(A_1) < \infty \implies \mu(A_n) \searrow 0$

Ist μ ein endlicher Inhalt, so sind (i) - (v) äquivalent.

Beweis:

(i) \implies (iii): Seien $A, A_n \in \mathcal{A}, A_n \nearrow A$. Aus der Monotonie (Bemerkung 2.7) folgt

$$\mu(A_n) \nearrow s := \sup_{n \in \mathbb{N}} \mu(A_n).$$

Bleibt zu zeigen: $\mu(A) = s$. Sei dazu $B_1 := A_1, B_k := A_k \setminus A_{k-1}$ ($k \geq 2$). Dann ist $A_n = B_1 \cup \dots \cup B_n$, also $A = \bigcup_{n=1}^{\infty} A_n = \bigcup_{k=1}^{\infty} B_k$, und aus der σ -Additivität folgt

$$\mu(A) = \sum_{k=1}^{\infty} \mu(B_k) = \sup_{n \in \mathbb{N}} \sum_{k=1}^n \mu(B_k) = \sup_{n \in \mathbb{N}} \mu(A_n) = s.$$

(iii) \implies (ii): Seien $A, A_n \in \mathcal{A}, A \subseteq \bigcup_{n=1}^{\infty} A_n$. Setze $B_n := A \cap \bigcup_{k=1}^n A_k$. Dann ist $B_n \in \mathcal{A}$ und $B_n \nearrow A$. Aus der σ -Stetigkeit, der Monotonie (Bemerkung 2.7) und der endlichen Subadditivität (Lemma 2.9) folgt:

$$\mu(A) = \sup_{n > 0} \mu(B_n) \leq \sup_{n > 0} \mu \left(\bigcup_{k=1}^n A_k \right) \leq \sup_{n > 0} \sum_{k=1}^n \mu(A_k) = \sum_{k=1}^{\infty} \mu(A_k).$$

(ii) \Rightarrow (i): Seien $A, A_n \in \mathcal{A}$ mit $A = \bigcup_{n=1}^{\infty} A_n$. Dann folgt aus der Monotonie und endlichen Additivität von μ :

$$\sum_{k=1}^{\infty} \mu(A_k) = \sup_{n \in \mathbb{N}} \sum_{k=1}^n \mu(A_k) = \sup_{n \in \mathbb{N}} \mu \left(\bigcup_{k=1}^n A_k \right) \leq \mu(A),$$

und die Umkehrung folgt aus der σ -Subadditivität.

(iii) \Rightarrow (iv): $A_n \searrow A \implies (A_1 \setminus A_n) \nearrow (A_1 \setminus A)$. Unter Beachtung von $\mu(A_1) < \infty$ folgt nun aus der σ -Stetigkeit:

$$\mu(A_1) - \mu(A_n) = \mu(A_1 \setminus A_n) \nearrow \mu(A_1 \setminus A) = \mu(A_1) - \mu(A),$$

also $\mu(A_n) \searrow \mu(A)$.

(iv) \Rightarrow (v): Trivial

(v) \Rightarrow (iii) für endliches μ : Seien $A, A_n \in \mathcal{A}$, $A_n \nearrow A$. Dann geht $A \setminus A_n \searrow \emptyset$ und es ist $\mu(A \setminus A_1) < \infty$. Also folgt aus (v), dass $\mu(A) - \mu(A_n) = \mu(A \setminus A_n) \searrow 0$ und daher $\mu(A_n) \nearrow \mu(A)$. \square

2.11 Bemerkung Auf die Voraussetzung $\mu(A_1) < \infty$ in Satz 2.10 (iv) und (v) kann i.A. nicht verzichtet werden. Insbesondere sind (iii) und (iv) bei unendlichem μ nicht äquivalent. (Beispiel: Übung!)

2.12 Definition (Maß, Wahrscheinlichkeitsmaß) Sei $\mathcal{A} \subseteq \mathfrak{F}(\Omega)$ eine σ -Algebra.

a) Eine positive Mengenfunktion μ auf \mathcal{A} ist ein Maß, falls sie σ -additiv ist, d.h. falls für alle paarweise disjunkten $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$) gilt:

$$\mu \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu(A_n).$$

b) Ist μ ein Maß auf der σ -Algebra \mathcal{A} , so heißt das Tripel $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Ist außerdem $\mu(\Omega) < \infty$, so ist μ endlich, und ist $\mu(\Omega) = 1$, so heißt $(\Omega, \mathcal{A}, \mu)$ ein Wahrscheinlichkeitsraum und μ ein Wahrscheinlichkeitsmaß.

2.13 Beispiele (Maße auf abzählbaren Räumen) Sei Ω abzählbar, z.B. $\Omega = \mathbb{N}, \mathbb{Z}^d, \tilde{\mathbb{Z}}^d$. In diesem Fall betrachtet man fast immer die σ -Algebra $\mathcal{A} = \mathfrak{F}(\Omega) = \sigma(\{\{\omega\} : \omega \in \Omega\})$.

a) Sei ein „Wahrscheinlichkeitsvektor“ $(p_\omega)_{\omega \in \Omega}$ mit $p_\omega \geq 0$, $\sum_{\omega \in \Omega} p_\omega = 1$ gegeben. Dann wird durch $\mu(A) := \sum_{\omega \in A} p_\omega$ für $A \subseteq \Omega$ ein Wahrscheinlichkeitsmaß auf (Ω, \mathcal{A}) definiert. Die σ -Additivität zeigt man so: Seien $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt, $A = \bigcup_{n=1}^{\infty} A_n$. Dann ist $\omega \in A \Leftrightarrow \exists n \in \mathbb{N} : \omega \in A_n$, und falls ein solches n existiert, ist es wegen der Disjunktheit der A_j eindeutig. Also:

$$\mu(A) = \sum_{\omega \in A} p_\omega = \sum_{n \in \mathbb{N}} \sum_{\omega \in A_n} p_\omega = \sum_{n \in \mathbb{N}} \mu(A_n).$$

b) $\mu(A) = \text{card}(A)$. Die σ -Additivität zeigt man ähnlich wie vorher. (Dieses Beispiel ist dafür verantwortlich, dass man Sätze über die Summation von Reihen mit beliebiger Indexmenge als einfache Spezialfälle von Sätzen der Integrationstheorie auffassen kann.) Das Maß ist σ -endlich, da $\Omega = \bigcup_{\omega \in \Omega} \{\omega\}$ und $\mu(\{\omega\}) = 1$ für alle ω .

2.14 Beispiel (Lebesgue-Stieltjes-Maß auf \mathbb{R}) Sei $\mathcal{B} = \sigma(\mathcal{I}^1)$ die Borel- σ -Algebra auf \mathbb{R} und sei $F : \mathbb{R} \rightarrow \mathbb{R}$ monoton wachsend und rechtsseitig stetig. Ziel: Ein Maß λ_F auf \mathcal{B} mit $\lambda_F((a, b]) = F(b) - F(a)$, $\lambda_F(\emptyset) = 0$. Zunächst zeigt man nur:

Das so festgelegte λ_F ist additiv und σ -subadditiv auf \mathcal{I}^1 .

(Das wurde bereits in der Maßtheorie gezeigt, den Beweis finden Sie noch einmal in Satz A.5.)

Später folgt daraus die Fortsetzbarkeit zu einem Maß auf \mathcal{B} . Ist $F(x) = x$, so erhält man gerade das *Lebesgue-Maß* auf \mathbb{R} , während jedes F mit $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$ die Verteilungsfunktion eines Wahrscheinlichkeitsmaßes auf \mathbb{R} beschreibt.

2.15 Beispiele (zur σ -Endlichkeit)

a) Sei $F : \mathbb{R} \rightarrow \mathbb{R}$, λ_F wie in Beispiel 2.14. Da $F((-n, n]) = F(n) - F(-n) < \infty$ für $n \in \mathbb{N}$ und $(-n, n] \nearrow \mathbb{R}$, ist λ_F auf \mathcal{I} σ -endlich. Das gilt insbesondere für das Lebesgue-Maß λ .

b) Sei $\mu : \mathfrak{P}(\mathbb{R}) \rightarrow [0, +\infty]$, $\mu(A) = \text{card}(A)$. μ ist *nicht* σ -endlich, da für beliebige $C_n \subseteq \mathbb{R}$ gilt: $\mu(C_n) < \infty \implies C_n$ endlich $\implies \bigcup_{n=1}^{\infty} C_n$ abzählbar $\implies \mathbb{R} \neq \bigcup_{n=1}^{\infty} C_n$.

2.16 Beispiel (Bernoulli-Maß auf $\Sigma^{\mathbb{N}}$) Sei $\mathcal{F} = \sigma(\mathcal{Z})$, siehe Beispiel 1.9d), und sei $(p_i)_{i \in \Sigma}$ ein Wahrscheinlichkeitsvektor auf der endlichen Menge Σ . Wir wollen ein Wahrscheinlichkeitsmaß μ auf \mathcal{F} definieren, für das

$$\mu([a_1, \dots, a_N]) = \prod_{i=1}^N p_{a_i} \quad (a_1, \dots, a_N \in \Sigma),$$

also insbesondere $\mu([a]) = p_a$ für alle $a \in \Sigma$. Wie im letzten Beispiel zeigen wir zunächst nur, dass durch diese Festsetzung eine auf dem Halbring \mathcal{Z} additive und σ -subadditive Mengenfunktion gegeben ist.

2.17 Satz

Die Mengenfunktion $\mu : \mathcal{Z} \rightarrow [0, 1]$ aus Beispiel 2.16 ist additiv und σ -subadditiv.

Beweis: Additivität: Seien $A_1 = [a_1^1, \dots, a_{\ell_1}^1], \dots, A_n = [a_1^n, \dots, a_{\ell_n}^n] \in \mathcal{Z}$ paarweise disjunkt und sei $A := \bigcup_{j=1}^n A_j \in \mathcal{Z}$, also $A = [a_1, \dots, a_\ell]$ für geeignete $a_i \in \Sigma$. Beachte, dass dann

$$\ell_i > \ell \text{ und } a_j^i = a_j \text{ für alle } i = 1, \dots, n \text{ und } j = 1, \dots, \ell.$$

Wir zerlegen die A_j und auch A disjunkt in Teilzylinder gleicher Länge $L := \max\{\ell_1, \dots, \ell_n\}$:

$$A = \bigcup_{Z \in \mathcal{Z}_L, Z \subset A} Z, \quad A_j = \bigcup_{Z \in \mathcal{Z}_L, Z \subset A_j} Z.$$

Dann ist

$$\begin{aligned} \mu(A) &= \prod_{j=1}^{\ell} p_{a_j} = \prod_{j=1}^{\ell} p_{a_j} \cdot \left(\sum_{s \in \Sigma} p_s \right)^{L-\ell} \\ &= \sum_{s_{\ell+1}, \dots, s_L \in \Sigma} \left(\prod_{j=1}^{\ell} p_{a_j} \cdot \prod_{j=\ell+1}^L p_{s_j} \right) = \sum_{s_{\ell+1}, \dots, s_L \in \Sigma} \mu([a_1, \dots, a_\ell, s_{\ell+1}, \dots, s_L]) \\ &= \sum_{Z \in \mathcal{Z}_L, Z \subset A} \mu(Z), \end{aligned}$$

und ganz analog ist für alle A_1, \dots, A_n :

$$\mu(A_j) = \sum_{Z \in \mathcal{Z}_{\ell_j}, Z \subset A_j} \mu(Z).$$

Daher ist

$$\mu(A) = \sum_{Z \in \mathcal{Z}_\ell, Z \subset A} \mu(Z) = \sum_{j=1}^n \sum_{Z \in \mathcal{Z}_{\ell_j}, Z \subset A_j} \mu(Z) = \sum_{j=1}^n \mu(A_j).$$

σ -Subadditivität: μ additiv auf $\mathcal{Z} \xrightarrow{\text{Lemma 2.9}} \mu$ subadditiv auf \mathcal{Z} . Sei nun $A \subseteq \bigcup_{k=1}^{\infty} A_k$ mit $A, A_k \in \mathcal{Z}$. Nach Bemerkung 1.4 gibt es ein $m > 0$ derart, dass $A \subseteq \bigcup_{k=1}^m A_k$. Aus der Subadditivität und der Positivität von μ folgt nun: $\mu(A) \leq \sum_{k=1}^m \mu(A_k) \leq \sum_{k=1}^{\infty} \mu(A_k)$. \square

2.18 Bemerkung Sei μ ein Maß auf \mathcal{A} , $A \in \mathcal{A}$. Bezeichne

$$\mathcal{A}|_A := \{B \subseteq A : B \in \mathcal{A}\} = \{C \cap A : C \in \mathcal{A}\}.$$

- ▷ $\mathcal{A}|_A$ ist eine σ -Algebra. (Übung!)
- ▷ Die durch $B \mapsto \mu(B)$ definierte Mengenfunktion auf $\mathcal{A}|_A$ ist ein Maß – die Einschränkung von μ auf $\mathcal{A}|_A$. (Übung!) Wir bezeichnen sie mit $\mu|_A$ oder auch einfach wieder mit μ . Ein wichtiges Beispiel ist die Einschränkung des Lebesgue-Maßes von \mathbb{R} auf $[0, 1]$.

Bevor wir rekapitulieren, wie mit nicht ganz geringem Aufwand gezeigt wird, dass sich jede additive und σ -subadditive Mengenfunktion auf einem Halbring zu einem Maß auf der von dem Halbring erzeugten σ -Algebra fortsetzen lässt, zeigen wir hier schon einmal, dass es in den meisten Fällen höchstens eine solche Fortsetzung geben kann. Der Beweis wurde zwar schon in der Maßtheorie geführt, wird hier aber noch einmal detailliert wiederholt, da er in exemplarischer Weise die Argumentation mit Dynkin-Systemen illustriert.

2.19 Satz (Eindeutigkeit der Maßfortsetzung)

Sei $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$ \cap -stabil und seien μ_1, μ_2 Maße auf $\sigma(\mathcal{C})$. Sind die Einschränkungen $\mu_1|_{\mathcal{C}}$ und $\mu_2|_{\mathcal{C}}$ σ -endlich und stimmen sie überein, so ist $\mu_1 = \mu_2$.

Beweis: 1. Schritt: Für $C \in \mathcal{C}$ mit $\mu_1(C) = \mu_2(C) < \infty$ setze

$$\mathcal{D}_C := \{A \in \sigma(\mathcal{C}) : \mu_1(A \cap C) = \mu_2(A \cap C)\}.$$

Dann ist $\mathcal{C} \subseteq \mathcal{D}_C$ und \mathcal{D}_C ist ein Dynkin-System, denn:

D1) $\mu_1(\Omega \cap C) = \mu_1(C) = \mu_2(C) = \mu_2(\Omega \cap C)$, also ist $\Omega \in \mathcal{D}_C$.

D2) Seien $A, B \in \mathcal{D}_C$, $A \subseteq B$. Dann ist für $i = 1, 2$

$$\mu_i((B \setminus A) \cap C) = \mu_i((B \cap C) \setminus (A \cap C)) \stackrel{\mu_i(C) < \infty}{=} \mu_i(B \cap C) - \mu_i(A \cap C).$$

Da $A, B \in \mathcal{D}_C$, folgt: $\mu_1((B \setminus A) \cap C) = \mu_2((B \setminus A) \cap C)$, so dass $B \setminus A \in \mathcal{D}_C$.

D3) Seien $A_n \in \mathcal{D}_C$ paarweise disjunkt. Dann ist für $i = 1, 2$

$$\mu_i \left(\left(\bigcup_{n=1}^{\infty} A_n \right) \cap C \right) = \mu_i \left(\bigcup_{n=1}^{\infty} (A_n \cap C) \right) = \sum_{n=1}^{\infty} \underbrace{\mu_i(A_n \cap C)}_{\text{gleich für } i=1,2}$$

Also stimmt der erste dieser Ausdrücke auch für $i = 1$ und $i = 2$ überein, so dass $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}_C$.

2. Schritt: Aus Satz 1.16 folgt nun: $\sigma(\mathcal{C}) = \mathcal{D}(\mathcal{C}) \subseteq \mathcal{D}(\mathcal{D}_C) = \mathcal{D}_C$, d.h.

$$\mu_1(A \cap C) = \mu_2(A \cap C) \quad \text{für alle } A \in \sigma(\mathcal{C}),$$

und das gilt für jedes $C \in \mathcal{C}$ mit $\mu_i(C) < \infty$. Seien nun $C_k \in \mathcal{C}$ mit $\mu_i(C_k) < \infty$ und $\Omega = \bigcup_{k=1}^{\infty} C_k$. Für jedes $A \in \sigma(\mathcal{C})$ gilt dann

$$\mu_1(A) = \sum_{k=1}^{\infty} \mu_1(A \cap C_k) = \sum_{k=1}^{\infty} \mu_2(A \cap C_k) = \mu_2(A).$$

□

Nun rekapitulieren wir die wichtigsten Schritte der Konstruktion von Maßen nach Carathéodory. Die kompletten Beweise finden sich im Anhang.

2.20 Definition (Äußeres Maß) Die Mengenfunktion $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, +\infty]$ ist ein äußeres Maß, falls μ^* σ -subadditiv und $\mu^*(\emptyset) = 0$ ist.

2.21 Bemerkung Ein äußeres Maß ist *monoton*, denn ist $A \subseteq B$, also $A \subseteq B \cup \emptyset \cup \emptyset \cup \dots$, so folgt $\mu^*(A) \leq \mu^*(B)$.

2.22 Satz (Konstruktion äußerer Maße)

Sei $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$, $\emptyset \in \mathcal{C}$, und sei $\rho : \mathcal{C} \rightarrow [0, +\infty]$ eine Mengenfunktion mit $\rho(\emptyset) = 0$. (Z.B. kann \mathcal{C} ein Halbring sein.) Für $A \subseteq \Omega$ setze

$$\rho^*(A) := \inf \left\{ \sum_{n=1}^{\infty} \rho(C_n) : A \subseteq \bigcup_{n=1}^{\infty} C_n, C_n \in \mathcal{C} \forall n \right\}.$$

(Falls keine abzählbare \mathcal{C} -Überdeckung von A existiert, heißt das $\rho^*(A) = +\infty$.)

Dann ist ρ^* ein äußeres Maß.

Beispiel: $\Omega = \Sigma^{\mathbb{N}}$, $\mathcal{C} = \mathcal{Z}$ die Familie aller Zylindermengen, $\rho(A) = \mu(A)$ wie in Beispiel 2.16. Wir werden sehen, dass $\rho^*|_{\sigma(\mathcal{Z})}$ ein Wahrscheinlichkeitsmaß ist, das sogenannte *Bernoulli-Maß* zum Wahrscheinlichkeitsvektor $p = (p_a)_{a \in \Sigma}$.

2.23 Definition (μ^* -Messbarkeit) Sei μ^* ein äußeres Maß auf Ω . $A \subseteq \Omega$ heißt μ^* -messbar, falls

$$\mu^*(A \cap E) + \mu^*(A^c \cap E) = \mu^*(E) \quad \text{für alle } E \subseteq \Omega.$$

(Wegen der Subadditivität von μ^* ist das äquivalent zu

$$\mu^*(A \cap E) + \mu^*(A^c \cap E) \leq \mu^*(E) \quad \text{für alle } E \subseteq \Omega.)$$

Bezeichnung:

$$\mathcal{M}(\mu^*) := \{A \subseteq \Omega : A \mu^*\text{-messbar}\}$$

2.24 Satz (Carathéodory)

Ist μ^* ein äußeres Maß auf Ω , so ist $\mathcal{M}(\mu^*)$ eine σ -Algebra, und die Einschränkung von μ^* auf $\mathcal{M}(\mu^*)$ ist ein Maß.

Satz 2.24 ist der Schlüssel zum Hauptergebnis dieses Kapitels:

2.25 Satz (Fortsetzungssatz für Maße, klassische Form)

Jeder σ -stetige Inhalt μ auf einer Algebra \mathcal{A} lässt sich zu einem Maß auf $\sigma(\mathcal{A})$ fortsetzen. Ist μ σ -endlich auf \mathcal{A} , so ist diese Fortsetzung eindeutig.

Da jede Algebra insbesondere ein Halbring ist, ist dieser Satz enthalten in der folgenden, etwas technischeren, aber flexibler anwendbaren Aussage:

2.26 Satz (Fortsetzungssatz für Maße, flexible Form)

Sei μ eine positive, additive, σ -subadditive Mengenfunktion auf dem Halbring \mathcal{S} . Dann lässt sich μ zu einem Maß auf $\sigma(\mathcal{S})$ fortsetzen. Ist μ σ -endlich auf \mathcal{S} , so ist diese Fortsetzung eindeutig.

Beweis: Eindeutigkeit: Seien μ_1 und μ_2 Maße auf $\sigma(\mathcal{S})$, die μ fortsetzen. Dann ist $\mu_{1|_{\mathcal{S}}} = \mu = \mu_{2|_{\mathcal{S}}}$ σ -endlich, und aus Satz 2.19 folgt $\mu_1 = \mu_2$.

Existenz: Definiere ein äußeres Maß μ^* wie in Satz 2.22:

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(S_n) : A \subseteq \bigcup_{n=1}^{\infty} S_n, S_n \in \mathcal{S} \forall n \right\}$$

Aus dem Satz von Carathéodory folgt, dass die Einschränkung von μ^* auf die σ -Algebra $\mathcal{M}(\mu^*)$ ein Maß ist. Zu zeigen ist:

- (i) $\mu(A) = \mu^*(A)$ für alle $A \in \mathcal{S}$ und
- (ii) $\mathcal{S} \subseteq \mathcal{M}(\mu^*)$.

Denn dann ist auch $\sigma(\mathcal{S}) \subseteq \mathcal{M}(\mu^*)$, so dass $\mu^*|_{\sigma(\mathcal{S})}$ ein Maß ist, das μ fortsetzt.

Zu (i): Aus der σ -Subadditivität von μ auf \mathcal{S} folgt $\mu(A) \leq \sum_{n=1}^{\infty} \mu(S_n)$ wenn immer $A \subseteq \bigcup_{n=1}^{\infty} S_n$ mit $S_n \in \mathcal{S}$. Daher ist $\mu(A) \leq \mu^*(A)$. Da $A \subseteq A \cup \emptyset \cup \emptyset \cup \dots$, ist auch $\mu^*(A) \leq \mu(A) + \mu(\emptyset) + \mu(\emptyset) + \dots = \mu(A)$.

Zu (ii): Der Beweis ist identisch zu dem in der Maßtheorie. Sie finden ihn in Satz A.13. \square

2.27 Bemerkung In Zukunft wird das Maß $\mu^*|_{\sigma(\mathcal{S})}$ wieder mit μ bezeichnet.

2.28 Beispiele (Fortsetzung von 2.14 und 2.16) $\lambda_F : \mathcal{I} \rightarrow [0, +\infty]$ und $\mu : \mathcal{Z} \rightarrow [0, +\infty]$ erfüllen die Voraussetzungen von Satz 2.26. Damit ist auch der letzte Schritt zur Konstruktion von *Lebesgue-Stieltjes-Maßen* auf \mathbb{R} und von *Bernoulli-Maßen* auf $\Sigma^{\mathbb{N}}$ getan.

2.29 Satz (Approximationsatz für Maße)

Seien \mathcal{S} und $\mu : \mathcal{S} \rightarrow [0, \infty]$ wie in Satz 2.26, μ σ -endlich.

a) Zu $A \in \mathcal{M}(\mu^*)$ und $\epsilon > 0$ gibt es eine Folge $S_1, S_2, \dots \in \mathcal{S}$ paarweise disjunkter Mengen mit $A \subseteq \bigcup_{n=1}^{\infty} S_n$ und $\mu(\bigcup_{n=1}^{\infty} S_n \setminus A) < \epsilon$.

b) Zu $A \in \mathcal{M}(\mu^*)$ mit $\mu(A) < \infty$ und $\epsilon > 0$ gibt es eine endliche Folge $S_1, \dots, S_n \in \mathcal{S}$ paarweise disjunkter Mengen mit $\mu(A \Delta \bigcup_{k=1}^n S_k) < \epsilon$.

Beides gilt insbesondere für alle Mengen $A \in \sigma(\mathcal{S}) \subseteq \mathcal{M}(\mu^*)$.

Beweis: Wir führen den Beweis hier nur für endliches μ , der allgemeine Beweis ist ähnlich aber von der Notation her aufwändiger und findet sich im Anhang (Satz A.14).

a) Sei $A \in \mathcal{M}(\mu^*)$ und $\epsilon > 0$. Nach Konstruktion von μ^* gibt es Mengen $B_\ell \in \mathcal{S}$ mit

$$A \subseteq \bigcup_{\ell=1}^{\infty} B_\ell \quad \text{und} \quad \sum_{\ell=1}^{\infty} \mu(B_\ell) < \mu^*(A) + \epsilon = \mu(A) + \epsilon. \quad (*)$$

Dabei wurde die Annahme, dass $\mu(A) < \infty$ ist, benutzt.

Wir konstruieren aus den B_ℓ eine spezielle Überdeckung von A durch paarweise disjunkte Mengen aus \mathcal{S} : Sei $R_1 = B_1$, $R_2 = B_2 \setminus R_1$, und für allgemeine ℓ sei

$$R_{\ell+1} = B_{\ell+1} \setminus \bigcup_{k=1}^{\ell} R_k.$$

Die R_ℓ sind paarweise disjunkt, und man überlegt sich leicht (Induktion auf der Basis von Lemma 2.16 im Maßtheorie-Skript von Prof. Knauf), dass jedes R_ℓ endliche disjunkte Vereinigung von Mengen aus \mathcal{S} ist. Daher gibt es $S_1, S_2, \dots \in \mathcal{S}$ derart, dass

$$A \subseteq \bigcup_{\ell=1}^{\infty} B_\ell = \bigcup_{\ell=1}^{\infty} R_\ell = \bigcup_{j=1}^{\infty} S_j$$

und

$$\begin{aligned} \mu\left(\bigcup_{j=1}^{\infty} S_j \setminus A\right) &= \mu\left(\bigcup_{\ell=1}^{\infty} B_\ell \setminus A\right) = \mu\left(\bigcup_{\ell=1}^{\infty} B_\ell\right) - \mu(A) \\ &\leq \sum_{\ell=1}^{\infty} \mu(B_\ell) - \mu(A) \leq \epsilon, \end{aligned}$$

wobei (*) für die letzte Ungleichung benutzt wurde.

b) Da $\mu(A) < \infty$, ist auch $\mu\left(\bigcup_{j=1}^{\infty} S_j\right) < \mu(A) + \epsilon < \infty$, und es gibt ein $m > 0$ derart, dass $\sum_{j=m+1}^{\infty} \mu(S_j) < \epsilon$. Also:

$$\mu\left(A \Delta \bigcup_{j=1}^m S_j\right) \leq \mu\left(\bigcup_{j=m+1}^{\infty} S_j\right) + \mu\left(\bigcup_{j=1}^{\infty} S_j \setminus A\right) < 2\epsilon,$$

d.h. Aussage b).

□

2.30 Bemerkung (Regularität) Sei Ω ein topologischer Raum, $\mathcal{O} \subseteq \mathfrak{P}(\Omega)$ das System der offenen Mengen, $\mathcal{B} := \sigma(\mathcal{O})$ die Borel'sche σ -Algebra. Ein Maß μ auf (Ω, \mathcal{B}) heißt

- i) von außen regulär, falls zu jedem $A \in \mathcal{B}$ und $\epsilon > 0$ eine offene Menge $G \supseteq A$ mit $\mu(G \setminus A) < \epsilon$ existiert;
- ii) von innen regulär, falls zu jedem $A \in \mathcal{B}$ mit $\mu(A) < \infty$ und $\epsilon > 0$ eine kompakte Menge $K \subseteq A$ mit $\mu(A \setminus K) < \epsilon$ existiert;
- iii) regulär, falls es von außen und von innen regulär ist.

Aus Satz 2.29 folgt man leicht die *Regularität* aller Maße μ auf $(\mathbb{R}^k, \mathcal{B}^k)$, die auf allen Intervallen $(a, b] \in \mathcal{I}^k$ endlich sind (Beweis: Übung!). Insbesondere ist auch das Lebesgue-Stieltjes-Maß λ auf \mathbb{R} regulär, ebenso wie das später zu konstruierende Lebesgue-Maß auf \mathbb{R}^k .

Die Regularität des Bernoulli-Maßes P (bzw. seiner Vervollständigung, die ebenfalls mit P bezeichnet wird) ist in Satz 2.29 enthalten: Da alle $S \in \mathcal{S} = \mathcal{Z}$ offen sind, bedeutet Aussage 2.29a) die Regularität von außen. Da $\Omega = \Sigma^{\mathbb{N}}$ kompakt und da P endlich ist, sind Regularität von innen und Regularität von außen äquivalent.

Kapitel 3

Messbare Funktionen und Abbildungen

In ähnlicher Weise wie man auf topologischen Räumen stetige Abbildungen definiert, führen wir in diesem Kapitel den Begriff der *messbaren Abbildung* zwischen messbaren Räumen ein. Viele der Begriffe und Sätze sind aus der Maßtheorie bekannt. Im Rahmen der Wahrscheinlichkeitstheorie werden messbare Funktionen später als Zufallsvariablen interpretiert.

3.1 Definition (Messbare Abbildung) Seien (Ω, \mathcal{A}) , (Ω', \mathcal{A}') messbare Räume.

- Eine Abbildung $T : \Omega \rightarrow \Omega'$ heißt \mathcal{A} - \mathcal{A}' -messbar (oder einfach messbar, wenn keine Verwechslungsgefahr besteht), falls $T^{-1}A' \in \mathcal{A}$ für alle $A' \in \mathcal{A}'$ (kurz: $T^{-1}\mathcal{A}' \subseteq \mathcal{A}$). Erinnerung: $T^{-1}A' = \{x \in \Omega : Tx \in A'\}$ und $T^{-1}\mathcal{A}' = \{T^{-1}A' : A' \in \mathcal{A}'\}$.
- Eine Funktion $f : \Omega \rightarrow \mathbb{R}(\overline{\mathbb{R}})$ heißt \mathcal{A} -messbar (oder einfach messbar), falls sie \mathcal{A} - $\mathcal{B}(\overline{\mathbb{B}})$ -messbar ist.

3.2 Bemerkung $f : \Omega \rightarrow \overline{\mathbb{R}}$ ist messbar genau dann, wenn $f^{-1}\{-\infty\}$, $f^{-1}\{+\infty\} \in \mathcal{A}$ und $f^{-1}M \in \mathcal{A}$ für alle $M \in \mathcal{B}$.

3.3 Lemma Seien (Ω, \mathcal{A}) , (Ω', \mathcal{A}') messbare Räume und sei $T : \Omega \rightarrow \Omega'$. Dann sind $T^{-1}\mathcal{A}'$ und $\mathcal{A}'_0 := \{A' \in \mathcal{A}' : T^{-1}A' \in \mathcal{A}\}$ σ -Algebren.

Beweis: $T^{-1}\mathcal{A}'$: A1) $\emptyset \in \mathcal{A}'$, da \mathcal{A}' σ -Algebra, also $\emptyset = T^{-1}\emptyset \in T^{-1}\mathcal{A}'$. A2) Ist $T^{-1}A' \in T^{-1}\mathcal{A}'$, so ist auch $(T^{-1}A')^c = T^{-1}(A'^c) \in T^{-1}\mathcal{A}'$, da mit A' auch A'^c zu \mathcal{A}' gehört. A3 σ) Sind $T^{-1}A'_n \in T^{-1}\mathcal{A}'$ für $A'_n \in \mathcal{A}'$ ($n \in \mathbb{N}$), so ist auch $\bigcup_{n=1}^{\infty} T^{-1}A'_n = T^{-1}(\bigcup_{n=1}^{\infty} A'_n) \in T^{-1}\mathcal{A}'$, da $\bigcup_{n=1}^{\infty} A'_n \in \mathcal{A}'$.

\mathcal{A}'_0 : A1) $T^{-1}\emptyset = \emptyset \in \mathcal{A}$, da \mathcal{A} σ -Algebra; also $\emptyset \in \mathcal{A}'_0$. A2) Sei $A' \in \mathcal{A}'_0$. Dann ist $T^{-1}(A'^c) = (T^{-1}A')^c \in \mathcal{A}$, denn \mathcal{A} ist eine σ -Algebra, so dass auch $A'^c \in \mathcal{A}'_0$. A3 σ) Sind $A'_n \in \mathcal{A}'_0$ ($n \in \mathbb{N}$), so sind die $T^{-1}A'_n \in \mathcal{A}$, also auch $T^{-1}(\bigcup_{n=1}^{\infty} A'_n) = \bigcup_{n=1}^{\infty} T^{-1}A'_n \in \mathcal{A}$, so dass $\bigcup_{n=1}^{\infty} A'_n \in \mathcal{A}'_0$. \square

Folgender Satz hilft, die Messbarkeit einer Abbildung in konkreten Fällen zu beweisen (siehe auch Satz B.1):

3.4 Satz (Messbarkeit auf Erzeugern)

- a) Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$ messbare Räume, $\mathcal{M}' \subseteq \mathcal{A}'$, und sei $T : \Omega \rightarrow \Omega'$. Dann ist

$$\sigma(T^{-1}\mathcal{M}') = T^{-1}\sigma(\mathcal{M}')$$
- b) Ist außerdem $\sigma(\mathcal{M}') = \mathcal{A}'$ und ist $T^{-1}A' \in \mathcal{A}$ für alle $A' \in \mathcal{M}'$, so ist T messbar.

- 3.5 Beispiele** a) Betrachte $f : \Omega \rightarrow \mathbb{R}$ und $\mathcal{M} = \{(-\infty, x] : x \in \mathbb{R}\}$. Da $\mathcal{I}^1 \subseteq \sigma(\mathcal{M})$ und $\mathcal{M} \subseteq \mathcal{B}$, ist $\mathcal{B} = \sigma(\mathcal{I}^1) \subseteq \sigma(\mathcal{M}) \subseteq \mathcal{B}$, also $\mathcal{B} = \sigma(\mathcal{M})$. Daher ist f messbar, falls für alle $x \in \mathbb{R}$

$$\{f \leq x\} := \{\omega \in \Omega : f(\omega) \leq x\} = f^{-1}((-\infty, x]) \text{ messbar ist.}$$

Analoges gilt für die Mengen $\{f \geq x\}$, $\{f < x\}$ und $\{f > x\}$ (nicht aber für $\{f = x\}$!).

- b) Alle $\pi_i : \mathbb{R}^k \rightarrow \mathbb{R}, \underline{x} \mapsto x_i$ ($i = 1, \dots, k$) sind messbar, da

$$\pi_i^{-1}((-\infty, x]) = \mathbb{R}^{i-1} \times (-\infty, x] \times \mathbb{R}^{k-i} = \bigcup_{n=1}^{\infty} ((-n, n]^{i-1} \times (-n, x] \times (-n, n]^{k-i}) \in \mathcal{B}^k$$

- c) Betrachte nun $f : \Omega \rightarrow \mathbb{R}^k$ und $\mathcal{M} = \{(-\infty, \underline{x}] : \underline{x} \in \mathbb{R}^k\}$. Da $(-\infty, \underline{x}] = \bigcup_{n=1}^{\infty} (-n\underline{1}, \underline{x}] \in \sigma(\mathcal{I}^k)$, ist $\mathcal{M} \subseteq \sigma(\mathcal{I}^k)$. Andererseits lässt sich jedes $(\underline{a}, \underline{b}] \in \mathcal{I}^k$ als

$$(-\infty, \underline{b}] \setminus \left(\bigcup_{i=1}^k (-\infty, (b_1, \dots, b_{i-1}, a_i, b_{i+1}, \dots, b_k)] \right)$$

darstellen. Also ist $\mathcal{I}^k \subseteq \sigma(\mathcal{M})$ und daher $\mathcal{B}^k = \sigma(\mathcal{I}^k) = \sigma(\mathcal{M})$. Schreibe f als (f_1, \dots, f_k) , jedes $f_i : \Omega \rightarrow \mathbb{R}$. Für $\underline{x} \in \mathbb{R}^k$ ist dann

$$f^{-1}((-\infty, \underline{x}]) = \{\omega \in \Omega : f(\omega) \leq \underline{x}\} = \bigcap_{i=1}^k \{\omega \in \Omega : f_i(\omega) \leq x_i\}.$$

Daher ist f messbar, falls alle f_i messbar sind. (Die Umkehrung folgt aus Teil b und dem nächsten Satz, da $f_i = \pi_i \circ f$.)

- d) Ähnlich zeigt man, dass $f : \Omega \rightarrow \bar{\mathbb{R}}^k$ genau dann messbar ist, falls alle $f_i : \Omega \rightarrow \bar{\mathbb{R}}$ messbar sind.

3.6 Satz

Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}'), (\Omega'', \mathcal{A}'')$ messbare Räume, $T_1 : \Omega \rightarrow \Omega', T_2 : \Omega' \rightarrow \Omega''$ messbar. Dann ist $T_2 \circ T_1 : \Omega \rightarrow \Omega''$ messbar.

Beweis: $(T_2 \circ T_1)^{-1}(\mathcal{A}'') = T_1^{-1}(T_2^{-1}\mathcal{A}'') \subseteq T_1^{-1}\mathcal{A}' \subseteq \mathcal{A}$. □

3.7 Satz

Jedes stetige $f : \mathbb{R}^k \rightarrow \mathbb{R}^n$ ist messbar.

Beweis: Wegen Beispiel 3.5c) reicht es, den Fall $n = 1$ zu betrachten. Dann gilt:

$$f \text{ stetig} \implies \{f < x\} \text{ offen f\u00fcr alle } x \in \mathbb{R} \implies \{f < x\} \in \mathcal{B}^k \text{ f\u00fcr alle } x \in \mathbb{R},$$

und die Messbarkeit von F folgt aus Satz 3.4. □

Im Folgenden wollen wir auch mit $\bar{\mathbb{R}}$ -wertigen Funktionen rechnen. Daher vereinbaren wir folgende Regeln f\u00fcr das Rechnen mit $\pm\infty$:

$$0 \cdot (\pm\infty) = 0, \quad x \cdot (+\infty) = +\infty \text{ f\u00fcr } x > 0, \quad +\infty + \infty = +\infty, \quad (+\infty) \cdot (+\infty) = +\infty, \quad \text{u.s.w.}$$

Verboten wird nat\u00fcrlich $+\infty - \infty$!

3.8 Satz

Sei (Ω, \mathcal{A}) ein messbarer Raum. Summen, Produkte, Maxima und Minima endlich vieler messbarer Funktionen von Ω nach \mathbb{R} oder $\bar{\mathbb{R}}$ sind messbar.

Beweis: Seien $f_1, \dots, f_n : \Omega \rightarrow \mathbb{R}$ messbar. Dann ist $F := (f_1, \dots, f_n) : \Omega \rightarrow \mathbb{R}^n$ messbar, und wegen Satz 3.7 und 3.6 ist $g \circ F$ f\u00fcr jedes stetige $g : \mathbb{R}^n \rightarrow \mathbb{R}$ messbar. W\u00e4hle nun f\u00fcr g :

$$g(\underline{x}) = \sum_{i=1}^n x_i, \quad \prod_{i=1}^n x_i, \quad \max_{i=1}^n x_i, \quad \min_{i=1}^n x_i.$$

Der Beweis f\u00fcr $\bar{\mathbb{R}}$ -wertige f_i verl\u00e4uft \u00e4hnlich. □

3.9 Satz

Sei f_1, f_2, \dots eine Folge messbarer $\bar{\mathbb{R}}$ -wertiger Funktionen auf (Ω, \mathcal{A}) .

- a) Die Funktionen $\sup_n f_n, \inf_n f_n, \limsup_{n \rightarrow \infty} f_n$ und $\liminf_{n \rightarrow \infty} f_n$ sind messbar.
- b) $E := \{\omega : \lim_{n \rightarrow \infty} f_n(\omega) \text{ existiert in } \bar{\mathbb{R}}\} \in \mathcal{A}$ und $1_E \cdot \lim_{n \rightarrow \infty} f_n$ ist messbar.

F\u00fcr den *Beweis* siehe Satz B.2.

Messbare Funktionen lassen sich durch *Elementarfunktionen* monoton approximieren:

3.10 Definition Sei (Ω, \mathcal{A}) ein messbarer Raum. $f : \Omega \rightarrow [0, \infty]$ hei\u00dft *Elementarfunktion*, falls $f = \sum_{i=1}^n \alpha_i 1_{A_i}$ f\u00fcr gewisse $\alpha_i \geq 0$ und paarweise disjunkte $A_i \in \mathcal{A}$. Die Menge aller Elementarfunktionen wird mit $\mathcal{E}(\Omega, \mathcal{A})$ bezeichnet.

3.11 Bemerkung Man kann immer annehmen, dass die A_1, \dots, A_n eine Partition von X bilden, d.h. dass $A_1 \cup \dots \cup A_n = \Omega$, denn ist das nicht der Fall, so erg\u00e4nze die n Mengen um $A_{n+1} := \Omega \setminus (A_1 \cup \dots \cup A_n)$ und setze $\alpha_{n+1} = 0$.

Ebenfalls bekannt aus der Ma\u00dftheorie ist folgender Satz (siehe auch Satz B.3):

3.12 Satz

Sei (Ω, \mathcal{A}) ein messbarer Raum und $f : \Omega \rightarrow [0, +\infty]$ messbar. Dann gibt es eine Folge von \mathcal{A} -Elementarfunktionen f_n derart, dass $f_n \nearrow f$.

Ma\u00dfe lassen sich, wie in der Ma\u00dftheorie gesehen, durch messbare Abbildungen transportieren:

3.13 Satz

Seien (Ω, \mathcal{A}) , (Ω', \mathcal{A}') messbare Räume, $T : \Omega \rightarrow \Omega'$ messbar, und sei μ ein Maß auf (Ω, \mathcal{A}) .
Definiere

$$\mu \circ T^{-1} : \mathcal{A}' \rightarrow [0, +\infty] \quad \text{durch} \quad \mu \circ T^{-1}(A') = \mu(T^{-1}A').$$

Dann ist $\mu \circ T^{-1}$ ein Maß auf \mathcal{A}' (das Bildmaß von μ unter T). Es ist ein Wahrscheinlichkeitsmaß genau dann, wenn μ eines ist.

Beweis: Elementar, siehe Satz B.4. □

Notation: In der Maßtheorie wurde das Bildmaß mit $T\mu$ bezeichnet.

3.14 Beispiel (Bernoulli-Verteilung und Binomial-Verteilung) Sei $\Omega = \{0, 1\}^n$, μ das Bernoullimaß zum Parameter p . Betrachte $T_n : \Omega \rightarrow \mathbb{R}$, $T_n(\omega) = \sum_{i=1}^n \omega_i$. Dann ist

$$\mu \circ T_n^{-1}(\{k\}) = \mu(T_n^{-1}\{k\}) = p^k(1-p)^{n-k} \cdot \text{card}\{\omega \in \Omega : \sum_{i=1}^n \omega_i = k\} = p^k(1-p)^{n-k} \binom{n}{k}$$

die Binomialverteilung $B(n, p)$.

Kapitel 4

Zufallsvariablen, Unabhängigkeit

In diesem Abschnitt sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum. Die Mengen $A \in \mathcal{A}$ heißen auch Ereignisse. Oft kann man nicht alle Ereignisse beobachten, sondern nur solche, die sich durch gegebene “Messinstrumente” erfassen lassen. Solche Situationen werden mit Hilfe messbarer Funktionen modelliert, die wir dann *Zufallsvariablen* nennen werden.

4.1 Definition a) Sei (M, \mathcal{M}) ein messbarer Raum. Eine \mathcal{A} - \mathcal{M} -messbare Abbildung $X : \Omega \rightarrow M$ heißt eine M -wertige Zufallsvariable. Ist $(M, \mathcal{M}) = (\mathbb{R}, \mathcal{B})$ oder $(\overline{\mathbb{R}}, \overline{\mathcal{B}})$, so sprechen wir einfach von einer Zufallsvariablen.

b) Das Wahrscheinlichkeitsmaß $P_X := P \circ X^{-1}$ auf (M, \mathcal{M}) heißt die Verteilung von X .

c) Ist μ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B})$, so heißt $F : \mathbb{R} \rightarrow [0, 1], x \mapsto \mu((-\infty, x])$ die Verteilungsfunktion von μ . Ist $\mu = P_X$ die Verteilung einer reellwertigen Zufallsvariablen, so spricht man auch von der Verteilungsfunktion von X , also $F(x) = P_X((-\infty, x]) = P\{X \leq x\}$.

d) Eine Familie $(X_i)_{i \in I}$ von Zufallsvariablen heißt identisch verteilt, falls P_{X_i} für alle i die gleiche Verteilung ist.

4.2 Bemerkung Sind F_1 und F_2 Verteilungsfunktionen von μ_1 und μ_2 und ist $F_1 = F_2$, so ist $\mu_1 = \mu_2$. Das folgt sofort aus dem Eindeutigkeitssatz für die Maßfortsetzung (Satz 2.19), denn bezeichnet \mathcal{C} den \cap -stabilen Erzeuger $\{(-\infty, x] : x \in \mathbb{R}\}$ von \mathcal{B} , so bedeutet $F_1 = F_2$ gerade, dass $\mu_1|_{\mathcal{C}} = \mu_2|_{\mathcal{C}}$.

4.3 Lemma Ist F die Verteilungsfunktion eines Wahrscheinlichkeitsmaßes μ , so gilt:

- 1) $x \leq y \implies F(x) \leq F(y)$.
- 2) $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$.
- 3) F ist rechtsseitig stetig, d.h. $\lim_{y \downarrow x} F(y) = F(x)$.

Beweis: Übung! Benutze die Monotonie und die monotone Stetigkeit von μ . □

4.4 Beispiel Sei $\Omega = \Sigma^{\mathbb{N}}$, P das Bernoulli-Maß zum Wahrscheinlichkeitsvektor $(p_i)_{i \in \Sigma}$ auf Σ . Betrachte Zufallsvariablen $X_n : \Omega \rightarrow \mathbb{R}, X_n(\omega) := \omega_n$ (Messbarkeit?). Dann ist P_{X_n} für alle n das durch den Wahrscheinlichkeitsvektor $(p_i)_{i \in \Sigma}$ auf Σ gegebene Wahrscheinlichkeitsmaß.

- a) Sei nun $\Sigma = \{0, 1\}$, $p_0 = p, p_1 = 1 - p$. Setze $S_n := \sum_{k=1}^n X_k$. Dann ist $S_n(\omega) = \sum_{k=1}^n \omega_k = T_n(\omega_1, \dots, \omega_n)$ aus Beispiel 3.14. Daher ist

$$P\{S_n = k\} = P_{S_n}(\{k\}) = P \circ S_n^{-1}(\{k\}) = \binom{n}{k} (1-p)^k p^{n-k}.$$

- b) Sei wieder $\Omega = \Sigma^{\mathbb{N}}$ und wähle ein $s \in \Sigma$ mit $p_s > 0$. Setze $Z(\omega) := \min\{n > 0 : \omega_n = s\}$, $Z(\omega) = +\infty$ falls kein solches ω_n existiert. Dann ist P_Z die *geometrische Verteilung* mit Parameter $1 - p_s$:

$$P\{Z = k\} = P\left(\bigcap_{i=1}^{k-1} \{\omega_i \neq s\} \cap \{\omega_k = s\}\right) = (1 - p_s)^{k-1} p_s.$$

Viele weitere Beispiele klassischer Verteilungen findet man in dem Textbuch von Georgii [5].

4.5 Definition (Unabhängige Familien von Mengensystemen) Eine Familie $(\mathcal{E}_i)_{i \in I}$, $\mathcal{E}_i \subseteq \mathcal{A}$ ($i \in I$) mit beliebiger Indexmenge I heißt unabhängig, falls für jede endliche Teilmenge $I_0 \subseteq I$ und jede Wahl von $E_i \in \mathcal{E}_i$ ($i \in I_0$) gilt

$$P\left(\bigcap_{i \in I_0} E_i\right) = \prod_{i \in I_0} P(E_i). \quad (*)$$

4.6 Satz (Unabhängige Familien von Mengensystemen)

- a) $(\mathcal{E}_i)_{i \in I}$ ist unabhängig genau dann, wenn $(\mathcal{E}_i)_{i \in I_0}$ für jede endliche Teilmenge $I_0 \subseteq I$ unabhängig ist.
- b) Ist I endlich und ist $\Omega \in \mathcal{E}_i$ für alle $i \in I$, so ist $(\mathcal{E}_i)_{i \in I}$ unabhängig genau dann, wenn (*) für $I_0 = I$ gilt.
- c) Ist $(\mathcal{E}_i)_{i \in I}$ unabhängig und sind alle $\mathcal{E}_i \cup \{\emptyset\}$ \cap -stabil, so ist auch $(\sigma(\mathcal{E}_i)_{i \in I})$ unabhängig.

Beweis:

- a) Klar, da jedes endliche I_0 Teilmenge von sich selbst ist.

b) \implies : Klar.

\impliedby : Sei $I_1 \subseteq I$, $E_i \in \mathcal{E}_i$ ($i \in I_1$), $E_i := \Omega$ ($i \notin I_1$). Dann ist

$$P\left(\bigcap_{i \in I_1} E_i\right) = P\left(\bigcap_{i \in I} E_i\right) = \prod_{i \in I} P(E_i) = \prod_{i \in I_1} P(E_i).$$

- c) Wegen a) kann man annehmen, dass I endlich ist, o.B.d.A. $I = \{1, \dots, n\}$. Da $\Omega \in \sigma(\mathcal{E}_i)$ für alle $i \in I$, ist wegen b) nur zu zeigen: (*) gilt für jede Wahl von $E_i \in \sigma(\mathcal{E}_i)$, $i \in I$. Daher zeigen wir durch Induktion nach $k = 1, \dots, n + 1$:

$$E_i \in \sigma(\mathcal{E}_i) \ (i = 1, \dots, k - 1) \text{ und } E_i \in \mathcal{E}_i \ (i = k, \dots, n) \implies P\left(\bigcap_{i=1}^n E_i\right) = \prod_{i=1}^n P(E_i).$$

Für $k = n + 1$ ist das die Behauptung.

„ $k = 1$ “: trivial.

„ $k \Rightarrow k + 1$ “: Für jede Wahl der $E_i \in \sigma(\mathcal{E}_i)$ ($i = 1, \dots, k - 1$) und $E_i \in \mathcal{E}_i$ ($i = k + 1, \dots, n$) betrachte die Mengenfunktionen

$$A \mapsto P \left(A \cap \bigcap_{\substack{i=1 \\ i \neq k}}^n E_i \right) \quad \text{und} \quad A \mapsto P(A) \cdot \prod_{\substack{i=1 \\ i \neq k}}^n P(E_i) . \quad (**)$$

Beides sind endliche Maße auf (Ω, \mathcal{A}) , die nach Induktionsannahme auf $\mathcal{E}_k \cup \{\emptyset\}$ übereinstimmen. Wegen des Eindeutigkeitsatzes 2.19 stimmen sie dann auch auf $\sigma(\mathcal{E}_k)$ überein. Das ist (***) für $k + 1$.

□

4.7 Definition (von einer Zufallsvariablen erzeugte σ -Algebra) Ist X eine Zufallsvariable, so heißt

$$\sigma(X) := X^{-1}\bar{\mathcal{B}}$$

die von X erzeugte σ -Algebra.

4.8 Definition (Unabhängigkeit von Zufallsvariablen)

Die Zufallsvariablen $(X_i)_{i \in I}$ sind unabhängig, falls die σ -Algebren $(\sigma(X_i))_{i \in I}$ unabhängig sind.

4.9 Satz

Seien X_1, \dots, X_N Zufallsvariablen. Es gelte:

$$P(\{X_i \leq z_i \forall i = 1, \dots, N\}) = \prod_{i=1}^N P(\{X_i \leq z_i\}) \quad \text{für alle } z_1, \dots, z_N \in (-\infty, \infty] . \quad (*)$$

Dann sind X_1, \dots, X_N unabhängig.

Beweis: Für $i = 1, \dots, N$ sei $\mathcal{E}_i = \{\{X_i \leq t\} : t \in \mathbb{R}\}$. (*) besagt in Anbetracht von Satz 4.6b), dass die \mathcal{E}_i unabhängige Mengensysteme sind. Außerdem sind die $\mathcal{E}_i \cup \{\emptyset\}$ Durchschnitts-stabil, so dass die erzeugten σ -Algebren $\sigma(\mathcal{E}_i)$ nach Satz 4.6c) unabhängig sind. Weiterhin gilt:

$$\begin{aligned} \sigma(\mathcal{E}_i) &= \sigma(\{X_i^{-1}((-\infty, t]) : t \in \mathbb{R}\}) \stackrel{3.4b)}{=} X_i^{-1}(\sigma(\{(-\infty, t] : t \in \mathbb{R}\})) \\ &\stackrel{3.5a)}{=} X_i^{-1}(\bar{\mathcal{B}}) = \sigma(X_i) . \end{aligned}$$

Also sind die X_i nach Definition 4.8 unabhängig. □

4.10 Übung Die Ereignisse $E_i \in \mathcal{A}$ ($i \in I$) heißen unabhängig, wenn für jede endliche Teilmenge $I_0 \subseteq I$ gilt $P(\bigcap_{i \in I_0} E_i) = \prod_{i \in I_0} P(E_i)$. Zeigen Sie, dass die folgenden Aussagen äquivalent sind:

- i) $(E_i)_{i \in I}$ ist unabhängig
- ii) $(\sigma(\{E_i\}))_{i \in I}$ ist unabhängig

iii) $(1_{E_i})_{i \in I}$ ist unabhängig

4.11 Bemerkung Sind $(X_i)_{i \in I}$ unabhängig, $f_i : \bar{\mathbb{R}} \rightarrow \bar{\mathbb{R}}$ messbar, so sind $(f_i(X_i))_{i \in I}$ unabhängig, denn $\sigma(f_i \circ X_i) = X_i^{-1}(f_i^{-1}\bar{\mathcal{B}}) \subseteq X_i^{-1}\bar{\mathcal{B}} = \sigma(X_i)$.

4.12 Beispiel (Unabhängigkeit des Bernoulli-Prozesses) Sei $\Omega = \Sigma^{\mathbb{N}}$, P ein Bernoulli-Maß auf Ω und $X_n(\omega) = \omega_n$ für alle $n \in \mathbb{N}$. Dann sind die $(X_n)_{n \in \mathbb{N}}$ unabhängig. (Übung!)

4.13 Bemerkung (Paarweise Unabhängigkeit impliziert nicht Unabhängigkeit) Sind in einer Familie $(X_i)_{i \in I}$ alle Paare (X_i, X_j) mit $i \neq j$ unabhängig, so muss $(X_i)_{i \in I}$ nicht unabhängig sein. Das gilt schon für eine Familie (X_1, X_2, X_3) . *Beispiel:* Das wohl einfachste Beispiel erhält man, wenn man die Gleichverteilung P auf $\Omega = \{0, 1, 2, 3\}$ und die Zufallsvariablen X_1, X_2, X_3 mit $X_i(0) = X_i(i) = 1$ und $X_i(j) = 0$ sonst betrachtet (Übung).

Kapitel 5

Integral und Erwartungswert

Aufbauend auf dem Maß- bzw. Wahrscheinlichkeitsbegriff führen wir jetzt die Begriffe *Integral* bzw. *Erwartungswert* ein und stellen einige elementare, aber für die Wahrscheinlichkeitstheorie zentrale Gleichungen und Ungleichungen mit Erwartungswerten bereit. Die maßtheoretischen Aspekte dieses Kapitels sind größtenteils aus dem Modul Maßtheorie bekannt.

Während des ganzen Kapitels sei $(\Omega, \mathcal{A}, \mu)$ ein *Maßraum*.

5.1 Definition (Integral) Sei $f : \Omega \rightarrow \bar{\mathbb{R}}$ messbar. Ist $f \geq 0$, so ist das Integral von f (bzgl. μ) definiert als

$$\int f d\mu = \sup \left\{ \sum_{i=1}^n \alpha_i \mu(A_i) : f \geq \sum_{i=1}^n \alpha_i 1_{A_i} \in \mathcal{E}(\Omega, \mathcal{A}) \right\} .$$

Für allgemeine f betrachte

$$f^+(\omega) := \max\{f(\omega), 0\} \quad \text{und} \quad f^-(\omega) := \max\{-f(\omega), 0\} .$$

f^+ und f^- sind messbar (Satz 3.8), nichtnegativ und $f = f^+ - f^-$. Man setzt

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu ,$$

vorausgesetzt, dass nicht $\int f^+ d\mu = \int f^- d\mu = +\infty$. In diesem Fall ist $\int f d\mu$ nicht definiert. Ist $f \geq 0$, so ist $f^+ = f$ und $f^- = 0$, und die Definition des Integrals ist mit der vorherigen für nichtnegative f konsistent.

Wenn es erforderlich ist, schreiben wir auch $\int f(\omega) d\mu(\omega)$ anstatt $\int f d\mu$. Weitere Schreibweise: $\int f(\omega) \mu(d\omega)$.

5.2 Lemma Seien f, g, f_n nichtnegative, messbare $\bar{\mathbb{R}}$ -wertige Funktionen und $A \in \mathcal{A}$.

- Ist $f = \sum_{i=1}^n \alpha_i 1_{A_i}$, $\alpha_i \in [0, +\infty]$, $A_i \in \mathcal{A}$, so ist $\int f d\mu = \sum_{i=1}^n \alpha_i \mu(A_i)$. Insbesondere ist $\int 1_A d\mu = \mu(A)$.
- Aus $0 \leq f \leq g$ folgt $0 \leq \int f d\mu \leq \int g d\mu$.
- Aus $0 \leq f_n \nearrow f$ folgt $0 \leq \int f_n d\mu \nearrow \int f d\mu$.
- Sind $\alpha, \beta \in \bar{\mathbb{R}}$, $\alpha, \beta \geq 0$, so gilt

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu .$$

Beweis: Siehe Lemma B.5.

5.3 Definition (Nullmenge, fast sicher)

- a) $N \in \mathcal{A}$ heißt Nullmenge, falls $\mu(N) = 0$.
- b) Eine Eigenschaft $E(\omega)$ die Punkten in Ω zukommen kann, gilt fast sicher (f.s.), falls es eine Nullmenge N gibt, so dass $E(\omega)$ für alle $\omega \in \Omega \setminus N$.

5.4 Satz
 Seien $f, g : \Omega \rightarrow \bar{\mathbb{R}}$ messbar, $f \geq 0$.

- a) $f(\omega) = 0$ f.s. $\iff \int f d\mu = 0$
- b) $\int f d\mu < \infty \implies f(\omega) < \infty$ f.s.
- c) $f(\omega) \leq g(\omega)$ f.s. $\implies \int f d\mu \leq \int g d\mu$

Beweis: Siehe Satz B.6.

5.5 Definition (Integrierbarkeit) Eine messbare Funktion $f : \Omega \rightarrow \bar{\mathbb{R}}$ heißt integrierbar (oder genauer μ -integrierbar), falls

$$\int f^+ d\mu < \infty \quad \text{und} \quad \int f^- d\mu < \infty .$$

Äquivalent dazu ist die Forderung $\int |f| d\mu < \infty$, da $|f| = f^+ + f^-$. Sei $\mathcal{L}_\mu^1 := \{f : \int |f| d\mu < \infty\}$ der Raum der μ -integrierbaren Funktionen.

- 5.6 Lemma**
- a) $|f| \leq |g|$ und $g \in \mathcal{L}_\mu^1 \implies f \in \mathcal{L}_\mu^1$
 - b) $\mu(\Omega) < \infty$ und f beschränkt $\implies f \in \mathcal{L}_\mu^1$ (Das gilt insbesondere in Wahrscheinlichkeitsräumen!)

Beweis:

- a) $\int |f| d\mu \leq \int |g| d\mu < \infty$ wegen Lemma 5.2b).
- b) Sei $|f| \leq M$. Dann ist

$$\int |f| d\mu \stackrel{\text{Lemma 5.2b)}}{\leq} \int M d\mu \stackrel{\text{Lemma 5.2a)}}{=} M \cdot \mu(\Omega) < \infty$$

□

5.7 Satz (Monotonie Linearität und Positivität des Integrals)

 Seien $f, g \in \mathcal{L}_\mu^1$.

a) $f \leq g$ f.s. $\implies \int f d\mu \leq \int g d\mu$

 b) Für alle $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g \in \mathcal{L}_\mu^1$ und

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu \quad (*)$$

 c) $|\int f d\mu| \leq \int |f| d\mu$ (Insbesondere: $f \geq 0 \implies \int f d\mu \geq 0$)

Beweis: Siehe Satz B.7.

5.8 Bemerkung Das Integral einer \mathbb{C} -wertigen Funktion wird folgendermaßen definiert: $f : \Omega \rightarrow \mathbb{C}$ ist messbar, falls $\Re f$ und $\Im f$ messbar sind. f ist integrierbar, falls $\Re f$ und $\Im f$ integrierbar sind. Für ein solches integrierbares f sei

$$\int f d\mu := \int \Re f d\mu + i \int \Im f d\mu.$$

Dann gelten Satz 5.7b) und c) analog, wobei auch die Koeffizienten α und β komplex sein können.

5.9 Definition Für $1 \leq p < \infty$ sei

$$\mathcal{L}_\mu^p := \left\{ f : \Omega \rightarrow \bar{\mathbb{R}} \mid f \text{ messbar, } \int |f|^p d\mu < \infty \right\}$$

5.10 Bemerkung Ist $\mu(\Omega) < \infty$, so ist $\mathcal{L}_\mu^p \subseteq \mathcal{L}_\mu^{p'}$ falls $p' \leq p$. Zum Beweis vergleiche Bemerkung 7.2.

5.11 Bemerkung Folgende Beobachtung ist oft nützlich: Ist $0 \leq f : \Omega \rightarrow \mathbb{R}$ messbar, so ist

$$\sum_{k=1}^{\infty} \mu\{f \geq k\} \leq \int f d\mu \leq \sum_{k=0}^{\infty} \mu\{f > k\},$$

denn für $n < f(\omega) < n + 1$ gilt

$$\sum_{k=1}^{\infty} 1_{\{f \geq k\}}(\omega) = n \leq f(\omega) \leq n + 1 = \sum_{k=0}^{\infty} 1_{\{f > k\}}(\omega),$$

und falls $f(\omega) = n + 1$ kann man das erste n durch $n + 1$ ersetzen.

Notation: Ist $f : \Omega \rightarrow \mathbb{R}$ integrierbar oder ≥ 0 und ist $A \in \mathcal{A}$, so ist

$$\int_A f d\mu := \int 1_A \cdot f d\mu.$$

Nun wenden wir den Integralbegriff auf Zufallsvariablen an. Sei daher (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum.

5.12 Definition (Erwartungswert, Varianz, Momente) Sei X eine Zufallsvariable.

- a) $E[X] := \int X dP$ (falls wohldefiniert) heißt der Erwartungswert von X . Wir schreiben auch EX .
- b) Ist $X \in \mathcal{L}_P^r$ oder $X^r \geq 0$, so ist das r -te Moment $E[X^r]$ von X wohldefiniert ($r = 1, 2, 3, \dots$).
- c) Sei $X \in \mathcal{L}_P^1$. $V(X) := E[(X - EX)^2]$ heißt die Varianz von X , $\sigma(X) := \sqrt{V(X)}$ die Streuung.

5.13 Lemma Für $X \in \mathcal{L}_P^1$ gilt:

- a) $V(X) = E[X^2] - (EX)^2$,
- b) $V(X) = 0 \iff X = EX$ f.s. $\iff X = \text{const}$ f.s.,
- c) $V(X) < \infty \iff X \in \mathcal{L}_P^2$
- d) Sei $X \in \mathcal{L}_P^2$. Dann ist

$$V(X) \leq E[(X - a)^2] \quad \text{für alle } a \in \mathbb{R}$$

mit Gleichheit genau dann, wenn $a = EX$.

Beweis:

a)

$$\begin{aligned} V(X) + \underbrace{2(E|X|)^2}_{< \infty} &= E[(X - EX)^2 + 2|X| \cdot E|X|] = E[X^2 - \underbrace{2X \cdot EX}_{\geq 0} + 2|X| \cdot E|X| + (EX)^2] \\ &= E[X^2] - 2EX EX + 2E|X| E|X| + (EX)^2 = E[X^2] - (EX)^2 + 2(E|X|)^2. \end{aligned}$$

b) Folgt aus Satz 5.4.

c) Folgt aus Teil a).

d)

$$f(a) := E[(X - a)^2] = E[X^2] - 2a EX + a^2 = (a - EX)^2 + V(X).$$

f hat ein eindeutiges Minimum bei $a = EX$ und $f(EX) = V(X)$.

□

5.14 Satz (Markov- und Chebyshev-Ungleichung)

Sei X eine Zufallsvariable, $\epsilon > 0$. Dann ist $P(\{|X| \geq \epsilon\}) \leq \epsilon^{-1} E|X|$ (Markov-Ungleichung) und daher auch (Chebyshev-Ungleichung)

$$P(\{|X| \geq \epsilon\}) = P(\{X^2 \geq \epsilon^2\}) \leq \epsilon^{-2} E[X^2]$$

Beweis:

$$P(|X| \geq \epsilon) = \int 1_{\{|X| \geq \epsilon\}} dP \leq \int 1_{\{|X| \geq \epsilon\}} \epsilon^{-1} |X| dP \leq \int \epsilon^{-1} |X| dP = \epsilon^{-1} E|X|$$

□

5.15 Definition Für $X, Y \in \mathcal{L}_P^2$ heißt $Cov(X, Y) := E[(X - EX)(Y - EY)]$ die Kovarianz von X und Y . X und Y heißen unkorreliert, falls $Cov(X, Y) = 0$. (Dass die Kovarianz für $X, Y \in \mathcal{L}_P^2$ immer wohldefiniert und endlich ist, wird man in Satz 7.1 sehen.)

5.16 Bemerkung $Cov(X, Y) = E[XY] - EX \cdot EY$. Insbesondere: Sind X, Y unkorreliert, so ist $E[XY] = EX \cdot EY$.

5.17 Satz
 Seien $X_i, Y_j \in \mathcal{L}_P^2$ ($i = 1, \dots, m; j = 1, \dots, n$) und $a_i, b_j, c, d \in \mathbb{R}$. Dann ist

$$Cov \left(\sum_{i=1}^m a_i X_i + c, \sum_{j=1}^n b_j Y_j + d \right) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j Cov(X_i, Y_j)$$

Beweis:

$$\begin{aligned} Cov \left(\sum_{i=1}^m a_i X_i + c, \sum_{j=1}^n b_j Y_j + d \right) &= E \left[\sum_{i=1}^m a_i (X_i - EX_i) \cdot \sum_{j=1}^n b_j (Y_j - EY_j) \right] \\ &= \sum_{i=1}^m \sum_{j=1}^n E [a_i b_j (X_i - EX_i)(Y_j - EY_j)] \\ &= \sum_{i=1}^m \sum_{j=1}^n a_i b_j Cov(X_i, Y_j) \end{aligned}$$

□

5.18 Korollar Seien $X_i \in \mathcal{L}_P^2$. Dann gilt

$$V \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n V(X_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^n Cov(X_i, X_j)$$

Insbesondere: Sind die X_i paarweise unkorreliert, so ist $V(\sum_{i=1}^n X_i) = \sum_{i=1}^n V(X_i)$.

5.19 Satz
 Seien $X_1, \dots, X_n \in \mathcal{L}_P^1$ unabhängig. Dann ist $\prod_{i=1}^n X_i \in \mathcal{L}_P^1$ und

$$E \left[\prod_{i=1}^n X_i \right] = \prod_{i=1}^n EX_i \quad (*)$$

Beweis: (Der Beweis verläuft ähnlich wie der einfachere von Satz 5.21.) Ist $X_i = 1_{A_i}$ ($i = 1, \dots, n$), so gilt (*) wegen der Unabhängigkeit. Da beide Seiten von (*) multilinear in X_1, \dots, X_n sind, gilt (*) auch für endliche Linearkombinationen von Indikatorfunktionen, und aus Satz 3.12 zusammen mit Lemma 5.2c) folgt, dass (*) für nichtnegative $X_i \in \mathcal{L}_P^1$ gilt. Aufgrund der Zerlegung $X_i = X_i^+ - X_i^-$ folgt (*) dann durch nochmalige Ausnutzung der

Multilinearität auch für beliebige $X_i \in \mathcal{L}_P^1$. (Dieses Beweisschema wird auch “algebraische Induktion” genannt.) Einen davon unabhängigen Beweis erhalten wir später als Korollar zum Satz von Fubini. \square

5.20 Korollar Sind $X_i \in \mathcal{L}_\mu^2$ unabhängig ($i \in I$), so sind sie paarweise unkorreliert.

5.21 Satz (Integration mit transformierten Maßen)
 Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$ messbare Räume, $T : \Omega \rightarrow \Omega'$ messbar, und sei μ ein Maß auf (Ω, \mathcal{A}) . Ist $f : \Omega' \rightarrow \mathbb{R}$ messbar und ist f nichtnegativ oder $f \circ T \in \mathcal{L}_\mu^1$, so gilt

$$\int_{\Omega} f \circ T \, d\mu = \int_{\Omega'} f \, d(\mu \circ T^{-1}). \quad (**)$$

(Ist $\mu = P$ ein Wahrscheinlichkeitsmaß und $T = X$ eine Ω' -wertige Zufallsvariable, so schreibt sich das als $E_P[f(X)] = \int f \, dP_X$.)
 Insbesondere gilt im Fall $\Omega = \Omega'$: Ist $\mu \circ T^{-1} = \mu$, so ist $\int f \circ T \, d\mu = \int f \, d\mu$.

Beweis: Wie oben folgt die Behauptung durch “algebraische Induktion”:

- i) (**) gilt nach Definition von $\mu \circ T^{-1}$ falls $f = 1_A$.
- ii) Da beide Seiten von (**) als Funktion von f linear sind, überträgt sich (**) auf endliche Linearkombinationen von Indikatorfunktionen, insbesondere auf $f \in \mathcal{E}(\Omega, \mathcal{A})$.
- iii) Zur Übertragung von (**) auf nichtnegative messbare Funktionen kann man entweder direkt Definition 5.1 anwenden (man muss dabei etwas aufpassen!), oder man kann Satz 3.12 kombiniert mit Lemma 5.2c anwenden.
- iv) Beliebige $f \in \mathcal{L}_{\mu \circ T^{-1}}^1$ zerlegt als $f = f^+ - f^-$ und wendet Satz 5.7b an.

\square

5.22 Bemerkung *Achtung:* Satz 5.21 stellt keine Integraltransformationsformel bereit, wie man sie aus der Analysis kennt. Dazu müsste auf der rechten Seite das transportierte Maß $\mu \circ T^{-1}$ auf Ω' durch ein dort vorgegebenes Maß μ' mal einer “Jacobischen Determinante” ersetzt werden. Das wäre ein viel tiefer liegendes Ergebnis.

Kapitel 6

Konvergenzsätze und gleichgradige Integrierbarkeit

In diesem Kapitel sei $(\Omega, \mathcal{A}, \mu)$ wieder ein Maßraum. Im Zentrum unserer Überlegungen steht die Frage, unter welchen Voraussetzungen Grenzwertbildung von Folgen messbarer Funktionen und Integration in der Reihenfolge vertauscht werden können. Dabei werden verschiedene Konvergenzbegriffe untersucht, und mit dem Borel-Cantelli Lemma und dem schwachen Gesetz der großen Zahl zwei einfache, aber für die Wahrscheinlichkeitstheorie wichtige Aussagen bewiesen.

6.1 Satz (Monotone Konvergenz)

Seien $f, f_n : \Omega \rightarrow \bar{\mathbb{R}}$ messbar.

- a) Aus $0 \leq f_n \nearrow f$ f.s. folgt $\int f_n d\mu \nearrow \int f d\mu$.
- b) Sind die $f_n \geq 0$, so gilt $\int \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int f_n d\mu$.
- c) Sind die $f_n \in \mathcal{L}_{\mu}^1$, gilt $f_n \nearrow f$ f.s. und ist $\sup_n \int f_n d\mu < \infty$, so ist $f \in \mathcal{L}_{\mu}^1$ und $\int f_n d\mu \nearrow \int f d\mu$. (Diese Aussage heißt auch Satz von Beppo Levi.)

Beweis: Siehe Satz B.8.

6.2 Bemerkung In Satz 6.1c) sowie in vielen anderen Konvergenzaussagen dieses Kapitels kann man auf die Voraussetzung, dass f messbar ist, verzichten, falls $(\Omega, \mathcal{A}, \mu)$ vollständig ist, denn dann folgt die Messbarkeit von f aus der fast sicheren Konvergenz von f_n gegen f .

6.3 Korollar (Maße mit Dichten) Ist $0 \leq f$ messbar, so wird durch $\nu(A) := \int_A f d\mu$ ein Maß auf (Ω, \mathcal{A}) definiert. ν ist endlich genau dann, wenn $f \in \mathcal{L}_{\mu}^1$, und ν ist ein Wahrscheinlichkeitsmaß genau dann, wenn $\int f d\mu = 1$. Schreibweise: $\nu = f \cdot \mu$ oder $\nu = f\mu$.

Beweis: Wir zeigen die σ -Additivität, der Rest ist trivial: Seien $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt. Mit $f_n := 1_{A_n} f$ folgt aus Satz 6.1b)

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \int \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \nu(A_n)$$

□

Wir führen folgende Bezeichnung ein:

$$\limsup_{n \rightarrow \infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{\omega : \omega \in A_k \text{ für unendlich viele } k\}$$

6.4 Korollar (Borel-Cantelli Lemma) Seien $A_n \in \mathcal{A}$, $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Dann ist

$$\mu\left(\limsup_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} \mu\left(\bigcup_{k=n}^{\infty} A_k\right) = 0,$$

d.h. $\{n \in \mathbb{N} : \omega \in A_n\}$ ist für μ -fast alle ω endlich.

Beweis: Aus Satz 6.1b) folgt für $f := \sum_{n=1}^{\infty} 1_{A_n}$, dass $\int f d\mu = \sum_{n=1}^{\infty} \mu(A_n) < \infty$, so dass $f < \infty$ f.s. nach Satz 5.4b). Das ist gleichbedeutend mit $\text{card}\{n \in \mathbb{N} : \omega \in A_n\} < \infty$ f.s. \square

6.5 Bemerkung Sind die Ereignisse A_n unabhängig, so gilt auch folgende ‘‘Umkehrung’’: Ist $\sum_{n=1}^{\infty} \mu(A_n) = \infty$, so ist $\mu(\limsup_{n \rightarrow \infty} A_n) = 1$, d.h. $\{n \in \mathbb{N} : \omega \in A_n\}$ ist für μ -fast alle ω unendlich.

6.6 Beispiel Sei $(X_n)_n$ eine Folge integrierbarer, identisch verteilter Zufallsvariablen auf (Ω, \mathcal{A}, P) . Dann ist $\lim_{n \rightarrow \infty} n^{-1} X_n = 0$ f.s., denn

$$\begin{aligned} \omega \in \{n^{-1} X_n \not\rightarrow 0\} &\Leftrightarrow \exists k \in \mathbb{N} : \text{card} \left\{ n \in \mathbb{N} : \frac{1}{n} |X_n(\omega)| > \frac{1}{k} \right\} = \infty \\ &\Leftrightarrow \exists k \in \mathbb{N} : \text{card} \{n \in \mathbb{N} : k|X_n(\omega)| > n\} = \infty \\ &\Leftrightarrow \omega \in \bigcup_{k=1}^{\infty} \limsup_{n \rightarrow \infty} \{k|X_n| > n\}, \end{aligned}$$

so dass

$$P(\{n^{-1} X_n \not\rightarrow 0\}) \leq \sum_{k=1}^{\infty} P\left(\limsup_{n \rightarrow \infty} \{k|X_n| > n\}\right),$$

und für jedes k ist $P(\limsup_{n \rightarrow \infty} \{k|X_n| > n\}) = 0$, da nach Bemerkung 5.11

$$\sum_{n=1}^{\infty} P(\{k|X_n| > n\}) = \sum_{n=1}^{\infty} P(\{k|X_1| > n\}) \leq k \int |X_1| dP < \infty.$$

Dabei haben wir benutzt, dass, da alle X_n identisch verteilt sind,

$$\begin{aligned} P(\{k|X_n| > n\}) &= P(X_n^{-1}(\{x \in \mathbb{R} : k|x| > n\})) = P_{X_n}(\{x \in \mathbb{R} : k|x| > n\}) \\ &= P_{X_1}(\{x \in \mathbb{R} : k|x| > n\}) = \dots = P(\{k|X_1| > n\}). \end{aligned}$$

6.7 Satz (Lemma von Fatou)

Seien $0 \leq f_n : \Omega \rightarrow \bar{\mathbb{R}}$ messbar. Dann ist

$$\int \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu.$$

Beweis siehe Satz B.9.

6.8 Satz (Majorisierte Konvergenz, auch Konvergenzsatz von Lebesgue)

Seien $f, f_n, g : \Omega \rightarrow \overline{\mathbb{R}}$ messbar, $|f_n| \leq g$ f.s. und $g \in \mathcal{L}_\mu^1$. Existiert $f = \lim_{n \rightarrow \infty} f_n$ f.s., so ist $f \in \mathcal{L}_\mu^1$ und

$$\lim_{n \rightarrow \infty} \int |f - f_n| d\mu = 0, \quad \text{insbesondere} \quad \int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Beweis siehe Satz B.10.

6.9 Bemerkung Ist $\mu(\Omega) < \infty$ und sind die f_n im vorherigen Satz gleichmäßig beschränkt, so kann der Satz angewandt werden.

6.10 Korollar Seien $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \subseteq \Omega$ messbar, $A = \bigcup_{n=1}^{\infty} A_n$ und $f \in \mathcal{L}_\mu^1$. Dann ist

$$\lim_{n \rightarrow \infty} \int_{A_n} f d\mu = \int_A f d\mu.$$

Beweis: Wähle $f_n := f 1_{A_n}$ und $g := |f|$, und berücksichtige, dass $\int_{A_n} f d\mu = \int f 1_{A_n} d\mu$.
□

6.11 Korollar (Riemann- und Lebesgue-Integral) Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist f auch integrierbar bzgl. der Einschränkung des Lebesgue-Maßes auf $[a, b]$, und die beiden Integrale stimmen überein, d.h.

$$\mathbb{R}\text{-} \int_a^b f(x) dx = \int_{[a,b]} f d\lambda.$$

Beweis siehe Korollar B.11.

6.12 Satz (Differentiation parameterabhängiger Integrale)

Sei $G \subseteq \mathbb{R}$ offen, $f : G \times \Omega \rightarrow \mathbb{R}$ sei so, dass $\omega \mapsto f(t, \omega)$ für jedes $t \in G$ μ -integrierbar und $t \mapsto f(t, \omega)$ für jedes $\omega \in \Omega$ differenzierbar ist. Gibt es ein $\Psi \in \mathcal{L}_\mu^1$ derart, dass $|\frac{d}{dt} f(t, \omega)| \leq \Psi(\omega)$ für alle $(t, \omega) \in G \times \Omega$, so existiert $\frac{d}{dt} \int f(t, \omega) d\mu(\omega)$ für jedes $t \in G$ und es ist

$$\frac{d}{dt} \int f(t, \omega) d\mu(\omega) = \int \frac{d}{dt} f(t, \omega) d\mu(\omega).$$

(Entsprechendes gilt auch für \mathbb{C} -wertige f .)

Beweis siehe Satz B.12.

6.13 Beispiel Sei X eine Zufallsvariable auf (Ω, \mathcal{A}, P) . Dann ist $\exp(itX)$ für jedes $t \in \mathbb{R}$ messbar und da $|\exp(itX)| = 1$, ist $\exp(itX)$ integrierbar. Setze $\varphi(t) := E[\exp(itX)]$ (das ist die Fourier-Transformierte von X). Es ist $\frac{d}{dt} \exp(itX) = iX \exp(itX)$, also $|\frac{d}{dt} \exp(itX)| \leq |X|$ für alle $t \in \mathbb{R}$. Also gilt: Ist $X \in \mathcal{L}_P^1$, so ist

$$\varphi'(t) = i \cdot \int X \exp(itX) dP.$$

Im Rest dieses Kapitels beschränken wir uns auf den Fall endlicher Maße (im Hinblick auf unsere Hauptanwendung, die Wahrscheinlichkeitstheorie). Ziel ist es, die Voraussetzungen von Satz 6.8 abzuschwächen. Analoge Ergebnisse für allgemeine Maße findet man bei [1, §21].

6.14 Definition Sei μ ein endliches Maß auf (Ω, \mathcal{A}) . Eine Familie \mathcal{F} messbarer $\bar{\mathbb{R}}$ -wertiger Funktionen auf Ω heißt gleichgradig integrierbar, falls es zu jedem $\epsilon > 0$ ein $a > 0$ gibt, für das

$$\sup_{f \in \mathcal{F}} \int (|f| - a)^+ d\mu \leq \epsilon .$$

6.15 Bemerkung Sei $\mathcal{F} = \{f\}$ mit $f \in \mathcal{L}_\mu^1$. Dann ist \mathcal{F} gleichgradig integrierbar, denn $0 \leq (|f| - n)^+ \leq |f|$ und $\lim_{n \rightarrow \infty} (|f| - n)^+ = 0$ auf $\{|f| < \infty\}$, d.h. f.s., so dass aus dem Satz von der majorisierten Konvergenz folgt $\lim_{n \rightarrow \infty} \int (|f| - n)^+ d\mu = 0$.

6.16 Satz

Sei μ ein endliches Maß auf (Ω, \mathcal{A}) und \mathcal{F} eine Familie messbarer $\bar{\mathbb{R}}$ -wertiger Funktionen auf Ω . Äquivalent sind die folgenden Aussagen:

- i) \mathcal{F} ist gleichgradig integrierbar.
- ii) Für alle $\epsilon > 0$ gibt es ein $a > 0$ derart, dass

$$\sup_{f \in \mathcal{F}} \int_{\{|f| > a\}} |f| d\mu \leq \epsilon .$$

- iii) $\sup_{f \in \mathcal{F}} \int |f| d\mu =: C < \infty$ und

$$\forall \epsilon > 0 \exists \delta > 0 \forall A \in \mathcal{A} : \mu(A) \leq \delta \implies \sup_{f \in \mathcal{F}} \int_A |f| d\mu \leq \epsilon .$$

Beweis: Schema: iii) \implies ii) \implies i) \implies iii).

iii) \implies ii) Zu $\epsilon > 0$ wähle $\delta > 0$ wie in iii). Dann ist für $a = \frac{C}{\delta}$ und für jedes $f \in \mathcal{F}$ wegen der Markov-Ungleichung

$$\mu\{|f| > a\} \leq \frac{1}{a} \int |f| d\mu \leq \frac{C}{a} \leq \delta ,$$

also $\int_{\{|f| > a\}} |f| d\mu \leq \epsilon$ für jedes $f \in \mathcal{F}$.

ii) \implies i) Beachte dass $(|f| - a)^+ \leq |f| \cdot 1_{\{|f| > a\}}$.

i) \implies iii) Sei $\epsilon > 0$. Wähle $a = a(\frac{\epsilon}{2})$ wie in i), $\delta := \frac{\epsilon}{2a}$. Dann ist für $f \in \mathcal{F}$ und $A \in \mathcal{A}$

$$\int_A |f| d\mu \leq \int_A (|f| - a)^+ d\mu + \int_A a d\mu \leq \frac{\epsilon}{2} + a \cdot \mu(A) .$$

Für $A = \Omega$ folgt daraus

$$\sup_{f \in \mathcal{F}} \int |f| d\mu \leq \frac{\epsilon}{2} + a \cdot \mu(\Omega) < \infty ,$$

und für $\mu(A) < \delta$ erhält man

$$\sup_{f \in \mathcal{F}} \int_A |f| d\mu \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

6.17 Lemma Sei μ ein endliches Maß auf (Ω, \mathcal{A}) .

- a) \mathcal{F} gleichgradig integrierbar $\iff \{ |f| : f \in \mathcal{F} \}$ gleichgradig integrierbar
- b) \mathcal{F} gleichgradig integrierbar, $S > 0 \iff \left\{ \sum_{i=1}^N \alpha_i f_i : N \in \mathbb{N}, f_i \in \mathcal{F}, \sum_{i=1}^N |\alpha_i| \leq S \right\}$ gleichgradig integrierbar
- c) $\mathcal{F}_1, \mathcal{F}_2$ gleichgradig integrierbar $\implies \mathcal{F}_1 \cup \mathcal{F}_2$ gleichgradig integrierbar
- d) $\mathcal{F} \subseteq \mathcal{L}_\mu^1$ endlich $\implies \mathcal{F}$ gleichgradig integrierbar

Beweis:

a) trivial.

b) Sei $\epsilon > 0$ und seien C und $\delta = \delta(\frac{\epsilon}{S})$ wie in Satz 6.16iii) zu \mathcal{F} gewählt. Dann ist für $A \in \mathcal{A}$ mit $\mu(A) < \delta$, für $f_i \in \mathcal{F}$ und α_i mit $\sum_{i=1}^N |\alpha_i| \leq S$

$$\begin{aligned} \int \left| \sum_i \alpha_i f_i \right| d\mu &\leq \sum_i |\alpha_i| \cdot \int |f_i| d\mu \leq C \cdot S < \infty, \\ \int_A \left| \sum_i \alpha_i f_i \right| d\mu &\leq \sum_i |\alpha_i| \cdot \int_A |f_i| d\mu < \sum_i |\alpha_i| \frac{\epsilon}{S} \leq \epsilon. \end{aligned}$$

- c) Zu $\epsilon > 0$ seien δ_1 (für \mathcal{F}_1) und δ_2 (für \mathcal{F}_2) wie in Satz 6.16iii) gewählt. Benutze nun $\delta := \min\{\delta_1, \delta_2\} > 0$ für $\mathcal{F}_1 \cup \mathcal{F}_2$.
- d) folgt aus Bemerkung 6.15 und Aussage c).

□

Nun beweisen wir die angekündigte Verschärfung von Satz 6.8.

6.18 Satz (Konvergenzsatz für gleichgradig integrierbare Funktionen)

Sei μ ein endliches Maß auf (Ω, \mathcal{A}) , $(f_n)_{n \geq 1}$ eine gleichgradig integrierbare Folge messbarer \mathbb{R} -wertiger Funktionen auf Ω . Ist $f : \Omega \rightarrow \mathbb{R}$ messbar und $f = \lim_{n \rightarrow \infty} f_n$ f.s., so ist $f \in \mathcal{L}_\mu^1$ und

$$\lim_{n \rightarrow \infty} \left| \int f_n d\mu - \int f d\mu \right| = \lim_{n \rightarrow \infty} \int |f_n - f| d\mu = 0.$$

Beweis: Es ist $f \in \mathcal{L}_\mu^1$, da

$$\int |f| d\mu = \int \liminf_{n \rightarrow \infty} |f_n| d\mu \stackrel{\text{Lemma von Fatou}}{\leq} \liminf_{n \rightarrow \infty} \int |f_n| d\mu \leq \sup_n \int |f_n| d\mu \stackrel{\text{Satz 6.16iii)}}{<} \infty.$$

Setze $h_n = |f - f_n|$. Wegen Lemma 6.17 (erst d, dann c, dann b und dann a) ist $(h_n)_n$ gleichgradig integrierbar, so dass zu $\epsilon > 0$ ein $a > 0$ existiert mit

$$\sup_n \int_{\{h_n > a\}} h_n d\mu < \epsilon .$$

Daher:

$$\int |f - f_n| d\mu = \int h_n d\mu = \int_{\{h_n > a\}} h_n d\mu + \int h_n \cdot 1_{\{h_n \leq a\}} d\mu < \epsilon + \int h_n \cdot 1_{\{h_n \leq a\}} d\mu .$$

Da $|h_n \cdot 1_{\{h_n \leq a\}}| \leq a$ und $|h_n \cdot 1_{\{h_n \leq a\}}| \leq h_n = |f - f_n| \rightarrow 0$ f.s., folgt aus Bemerkung 6.9

$$\lim_{n \rightarrow \infty} \int h_n \cdot 1_{\{h_n \leq a\}} d\mu = 0 ,$$

also

$$\lim_{n \rightarrow \infty} \int |f - f_n| d\mu \leq \epsilon \quad \text{für beliebiges } \epsilon > 0 .$$

□

Wir diskutieren nun drei Konvergenzbegriffe für messbare Funktionen im Fall endlicher Maße. Für den allgemeinen Fall siehe [1, §20].

Seien f_n, f messbare \mathbb{R} -wertige Funktionen auf (Ω, \mathcal{A}) , μ ein endliches Maß auf \mathcal{A} .

▷ **Fast sichere Konvergenz** $f = \lim_{n \rightarrow \infty} f_n$ μ -f.s. genau dann, wenn

$$\mu \left(\left\{ \limsup_{k \rightarrow \infty} |f - f_k| > 0 \right\} \right) = \mu \left(\bigcup_{r=1}^{\infty} \left\{ \limsup_{k \rightarrow \infty} |f - f_k| > \frac{1}{r} \right\} \right) = 0$$

Das ist wegen der σ -Subadditivität von μ äquivalent zu

$$\forall r \in \mathbb{N} : \mu \left(\left\{ \limsup_{k \rightarrow \infty} |f - f_k| > \frac{1}{r} \right\} \right) = 0 ,$$

und da $\left\{ \limsup_{k \rightarrow \infty} |f - f_k| > \frac{1}{r} \right\} = \limsup_{k \rightarrow \infty} \left\{ |f - f_k| > \frac{1}{r} \right\}$ für jedes r , ist die fast sichere Konvergenz also äquivalent zu

$$\begin{aligned} \forall \epsilon > 0 : \mu \left(\limsup_{k \rightarrow \infty} \{ |f - f_k| > \epsilon \} \right) &= \mu \left(\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} \{ |f - f_k| > \epsilon \} \right) \\ &= \lim_{n \rightarrow \infty} \mu \left(\bigcup_{k \geq n} \{ |f - f_k| > \epsilon \} \right) = 0 \end{aligned}$$

▷ **Stochastische Konvergenz** $f_n \xrightarrow[\mu]{} f$ ($n \rightarrow \infty$) genau dann, wenn

$$\forall \epsilon > 0 : \lim_{n \rightarrow \infty} \mu(\{ |f - f_n| > \epsilon \}) = 0$$

▷ **L^1_μ -Konvergenz** $\lim_{n \rightarrow \infty} \int |f - f_n| d\mu = 0$

6.19 Lemma Seien f_n, f messbare \mathbb{R} -wertige Funktionen auf (Ω, \mathcal{A}) , μ ein endliches Maß auf \mathcal{A} .

$$a) f_n \rightarrow f \text{ } \mu\text{-f.s.} \implies f_n \xrightarrow{\mu} f$$

$$b) \int |f_n - f| d\mu \rightarrow 0 \implies f_n \xrightarrow{\mu} f$$

c) Für jeden der drei Konvergenzbegriffe gilt: Geht $f_n \rightarrow f$, so folgt

$$f_n \rightarrow g \iff f = g \text{ } \mu\text{-f.s.}$$

Beweis:

a) Folgt direkt aus obigen Charakterisierungen der Konvergenzbegriffe.

b) Folgt aus der Markov-Ungleichung $\mu\{|f_n - f| > \epsilon\} \leq \epsilon^{-1} \cdot \int |f_n - f| d\mu$.

c) “ \implies ” Wegen Teil a) und b) ist das nur für die stochastische Konvergenz zu zeigen: Für $\epsilon > 0$ und $n \in \mathbb{N}$ ist

$$\mu\{|f - g| > \epsilon\} \leq \mu\{|f - f_n| > \frac{\epsilon}{2}\} + \mu\{|f_n - g| > \frac{\epsilon}{2}\} \xrightarrow{n \rightarrow \infty} 0,$$

also $\mu\{|f - g| > \epsilon\} = 0$ und daher

$$\mu\{f \neq g\} = \mu\left(\bigcup_{r=1}^{\infty} \left\{|f - g| > \frac{1}{r}\right\}\right) \leq \sum_{r=1}^{\infty} \mu\left(\left\{|f - g| > \frac{1}{r}\right\}\right) = 0$$

“ \impliedby ” Ist $\mu\{f \neq g\} = 0$, so werden die definierenden Ausdrücke für die drei Konvergenzbegriffe beim Übergang von f zu g oder umgekehrt nur auf μ -Nullmengen geändert, d.h. ihre Werte ändern sich nicht.

□

6.20 Satz

Seien f_n, f wie vorher. Gilt $f_n \xrightarrow{\mu} f$, so gibt es eine Teilfolge $(f_{n_k})_k$ von $(f_n)_n$, die μ -f.s. gegen f konvergiert.

Beweis: Für alle $k > 0$ gibt es ein $n_k \in \mathbb{N}$ mit $\mu\{|f_{n_k} - f| > \frac{1}{k}\} \leq 2^{-k}$. Die n_k können so gewählt werden, dass $n_1 < n_2 < n_3 < \dots$. Aus dem Borel-Cantelli Lemma 6.4 folgt $\text{card}\{k \in \mathbb{N} : |f_{n_k}(\omega) - f(\omega)| > \frac{1}{k}\} < \infty$ für μ -fast alle ω und daher $\lim_{k \rightarrow \infty} f_{n_k} = f$ μ -f.s.

□

6.21 Satz

Seien f_n, f messbare \mathbb{R} -wertige Funktionen auf (Ω, \mathcal{A}) , μ ein endliches Maß auf \mathcal{A} . Folgende Aussagen sind äquivalent:

- i) $\int |f_n - f| d\mu < \infty$ für alle n , $\lim_{n \rightarrow \infty} \int |f_n - f| d\mu = 0$ und $f \in \mathcal{L}^1_{\mu}$.
- ii) $f_n \xrightarrow{\mu} f$ und $(f_n)_n$ ist gleichgradig integrierbar.

Beweis: “i) \implies ii)”: $f_n \xrightarrow{\mu} f$ folgt aus Lemma 6.19b). Sei nun $\epsilon > 0$. Wähle $a_1 > 0$ so, dass $\int (|f| - a_1)^+ d\mu < \frac{\epsilon}{3}$ (vergl. Bemerkung 6.15), und $n_0 = n_0(\epsilon) \in \mathbb{N}$ derart, dass

$$\int \||f_n| - |f|\| d\mu \leq \int |f_n - f| d\mu < \frac{\epsilon}{3} \text{ für alle } n \geq n_0.$$

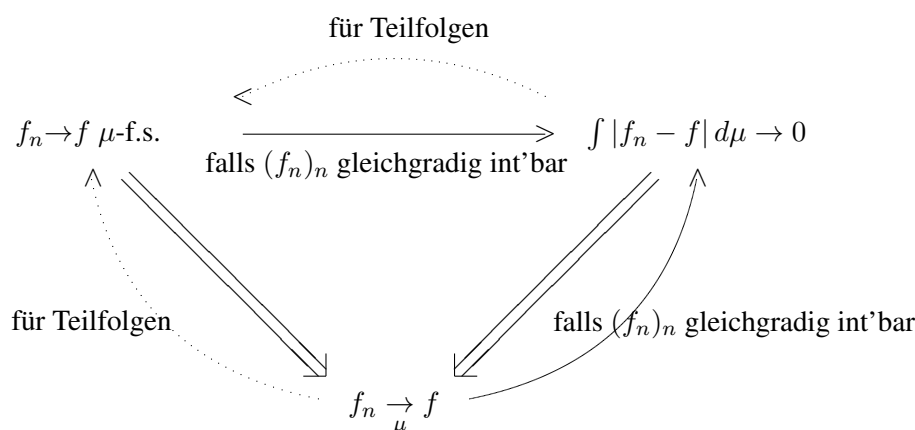
Es folgt

$$\int (|f_n| - a_1)^+ d\mu \leq \int (|f| - a_1)^+ d\mu + \int ||f_n| - |f|| d\mu < \frac{2}{3}\epsilon \quad \text{für alle } n \geq n_0.$$

Da alle f_n nach Voraussetzung integrierbar sind, ist die endliche Familie $\{f_1, \dots, f_{n_0}\}$ gleichgradig integrierbar (siehe Lemma 6.17), und es gibt ein $a_2 > 0$ derart, dass $\int (|f_i| - a_2)^+ d\mu < \epsilon$ für $i = 1, \dots, n_0$. Für $a = \max\{a_1, a_2\}$ folgt daher $\sup_{n>0} \int (|f_n| - a)^+ d\mu < \epsilon$.

“ii) \Rightarrow i)”: Sei $(f_{n_k})_{k>0}$ eine beliebige Teilfolge von $(f_n)_{n>0}$. Wegen Satz 6.20 gibt es eine Teilfolge $(f_{n_{k(i)}})_{i>0}$ von $(f_{n_k})_{k>0}$ mit $\lim_{i \rightarrow \infty} f_{n_{k(i)}} = f$ μ -f.s. Aus Satz 6.18 folgt: $f \in \mathcal{L}^1_\mu$ und $\lim_{i \rightarrow \infty} \int |f_{n_{k(i)}} - f| d\mu = 0$. Also hat jede Teilfolge der Folge $(\int |f_n - f| d\mu)_{n>0}$ eine Teilfolge, die gegen 0 konvergiert, so dass die Folge selbst auch gegen 0 konvergiert. \square

Wir fassen die Beziehungen zwischen den verschiedenen Konvergenzbegriffen bei endlichem Maß μ zusammen:



6.22 Bemerkung (Schwaches Gesetz der großen Zahl) Seien $X_1, X_2, \dots \in \mathcal{L}^2_P$ unkorreliert und $M := \sup_n V(X_n) < \infty$. Dann konvergiert

$$\frac{1}{n} \sum_{i=1}^n (X_i - EX_i) \xrightarrow{P} 0,$$

denn aus der Chebyshev-Ungleichung 5.14 folgt:

$$P \left(\left| \frac{1}{n} \sum_{i=1}^n (X_i - EX_i) \right| > \alpha \right) \leq \alpha^{-2} V \left(\frac{1}{n} \sum_{i=1}^n (X_i - EX_i) \right) = \alpha^{-2} \frac{1}{n^2} \sum_{i=1}^n V(X_i) \leq \alpha^{-2} M n^{-1}.$$

Kapitel 7

Minkowski, Hölder und Jensen

In diesem Kapitel stellen wir eine Reihe der wichtigsten Ungleichungen der Analysis zur Verfügung. Insbesondere zeigen wir, dass die sogenannten L^p -Räume Banach-Räume sind.

Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. Für messbare $f : \Omega \rightarrow \bar{\mathbb{R}}$ sei

$$\|f\|_p := \left(\int |f|^p d\mu \right)^{1/p}$$

Zur Erinnerung: Dann ist $\mathcal{L}_\mu^p = \{f : \Omega \rightarrow \bar{\mathbb{R}} \text{ mit } \|f\|_p < \infty\}$.

7.1 Satz (Minkowski- und Hölder-Ungleichung)

a) \mathcal{L}_μ^p ist ein Vektorraum.

b) $\|\cdot\|_p$ ist eine Semi-Norm auf \mathcal{L}_μ^p , d.h. für $f, g \in \mathcal{L}_\mu^p$ und $\alpha \in \mathbb{R}$ gilt

1) $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ (Minkowski-Ungleichung)

2) $\|\alpha f\|_p = |\alpha| \cdot \|f\|_p$

3) $\|f\|_p = 0 \iff f = 0$ μ -f.s.

c) $f \in \mathcal{L}_\mu^p, g \in \mathcal{L}_\mu^q \implies (fg) \in \mathcal{L}_\mu^1, \|fg\|_1 \leq \|f\|_p \cdot \|g\|_q$ (Hölder-Ungleichung)

Beweis:

b2)

$$\|\alpha f\|_p = \left(\int |\alpha f|^p d\mu \right)^{1/p} = |\alpha| \cdot \left(\int |f|^p d\mu \right)^{1/p} = |\alpha| \cdot \|f\|_p$$

b3)

$$\|f\|_p = 0 \iff \int |f|^p d\mu = 0 \iff |f|^p = 0 \text{ f.s.} \iff f = 0 \text{ f.s.}$$

c) Ist $\|f\|_p = 0$ oder $\|g\|_q = 0$, so ist $f = 0$ f.s. oder $g = 0$ f.s., also $fg = 0$ f.s. und daher $\|fg\|_1 = 0$.

Seien nun $\|f\|_p > 0$ und $\|g\|_q > 0$. Für beliebige $a, b > 0$ gilt

$$ab = e^{\frac{1}{p}(p \log a) + \frac{1}{q}(q \log b)} \leq \frac{1}{p} e^{p \log a} + \frac{1}{q} e^{q \log b} = \frac{a^p}{p} + \frac{b^q}{q},$$

da $x \mapsto e^x$ konvex ist. Ist $a = 0$ oder $b = 0$, so gilt diese Ungleichung ebenfalls. Also

$$\frac{|f(\omega)|}{\|f\|_p} \cdot \frac{|g(\omega)|}{\|g\|_q} \leq \frac{1}{p} \frac{|f(\omega)|^p}{\int |f|^p d\mu} + \frac{1}{q} \frac{|g(\omega)|^q}{\int |g|^q d\mu}$$

Integriert man beide Seiten dieser Ungleichung mit $d\mu$, so folgt

$$\frac{\int |fg| d\mu}{\|f\|_p \cdot \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1,$$

also die Behauptung.

- a) i) $0 \in \mathcal{L}_\mu^p$ ✓
 ii) Da $\|\alpha f\|_p = |\alpha| \cdot \|f\|_p$, ist mit f auch αf in \mathcal{L}_μ^p .
 iii) Seien $f, g \in \mathcal{L}_\mu^p$. Da

$$|f+g|^p \leq (|f|+|g|)^p \leq (2 \max\{|f|, |g|\})^p = 2^p \max\{|f|^p, |g|^p\} \leq 2^p (|f|^p + |g|^p), \quad (7.1)$$

ist

$$\int |f+g|^p \leq 2^p \left(\int |f|^p d\mu + \int |g|^p d\mu \right) < \infty,$$

also $f+g \in \mathcal{L}_\mu^p$.

- b1) Seien $f, g \in \mathcal{L}_\mu^p$. Es gilt

$$\int |f+g|^p d\mu \leq \int (|f|+|g|)^p d\mu, \quad (*)$$

und die Ungleichung folgt sofort für $p = 1$. Sei nun $p > 1$. Dann ist

$$\begin{aligned} \int (|f|+|g|)^p d\mu &= \int |f| (|f|+|g|)^{p-1} d\mu + \int |g| (|f|+|g|)^{p-1} d\mu \\ &\stackrel{\text{Hölder-Ungl.}}{\leq} \|f\|_p \left(\int (|f|+|g|)^{(p-1)q} d\mu \right)^{1/q} + \|g\|_p \left(\int (|f|+|g|)^{(p-1)q} d\mu \right)^{1/q}. \end{aligned}$$

Da aus $\frac{1}{p} + \frac{1}{q} = 1$ folgt, dass $p+q = pq$, ist $(p-1)q = p$, und wir erhalten

$$\int (|f|+|g|)^p d\mu \leq (\|f\|_p + \|g\|_p) \left(\int (|f|+|g|)^p d\mu \right)^{1/q},$$

woraus wegen $|f|+|g| \in \mathcal{L}_\mu^p$ folgt

$$\left(\int (|f|+|g|)^p d\mu \right)^{1/p} \leq \|f\|_p + \|g\|_p.$$

Mit (*) folgt daraus die behauptete Ungleichung. □

7.2 Bemerkungen

- a) Für $p = q = 2$ erhält man $\|fg\|_1 \leq \|f\|_2 \|g\|_2$ (Cauchy/Schwarz-Ungleichung)

b) Ist $\mu(\Omega) = 1$ und $1 \leq p < r$, so gilt für $f \in \mathcal{L}_\mu^r$

$$\|f\|_p = \left(\int |f|^p \cdot 1 \, d\mu \right)^{1/p} \stackrel{\text{Hölder-Ungl.}}{\leq} \left(\| |f|^p \|_{\frac{r}{p}} \cdot \|1\|_{\frac{r}{r-p}} \right)^{1/p} = \left(\int |f|^r \, d\mu \right)^{\frac{p}{r} \cdot \frac{1}{p}} = \|f\|_r$$

7.3 Bemerkung Sei

$$\mathcal{L}_\mu^\infty := \{f : \Omega \rightarrow \mathbb{R} \mid \exists M > 0 \text{ so dass } |f| \leq M \text{ } \mu\text{-f.s.}\} .$$

Für $f \in \mathcal{L}_\mu^\infty$ sei

$$\|f\|_\infty := \inf \{M > 0 : |f| \leq M \text{ } \mu\text{-f.s.}\} .$$

Dann gelten für $p = \infty$ und $q = 1$ die gleichen Aussagen wie in Satz 7.1 (Übung!).

7.4 Bemerkung Sei $\mathcal{F} \subseteq \mathcal{L}_\mu^1$ und $\sup_{f \in \mathcal{F}} \|f\|_r =: C < \infty$ für ein $r > 1$. Dann ist \mathcal{F} gleichgradig integrierbar, denn für $a > 0$ und $f \in \mathcal{F}$ gilt

$$\int_{\{|f|>a\}} |f| \, d\mu \leq a^{-(r-1)} \int_{\{|f|>a\}} |f|^r \, d\mu \leq a^{-(r-1)} C^r \rightarrow 0 \quad \text{für } a \rightarrow \infty .$$

7.5 Satz (Vollständigkeit von \mathcal{L}_μ^p)

Sei $1 \leq p \leq \infty$, $f_n \in \mathcal{L}_\mu^p$, und sei $\lim_{n \rightarrow \infty} \sup_{m \geq n} \|f_n - f_m\|_p = 0$. Dann existiert ein $f \in \mathcal{L}_\mu^p$ derart, dass $\lim_{n \rightarrow \infty} \|f_n - f\|_p = 0$.

Beweis: Der Beweis für $p = \infty$ bleibt zur Übung überlassen. Hier sei $1 \leq p < \infty$. Wähle eine Teilfolge $(f_{n_k})_k$ von $(f_n)_n$ mit $\|f_{n_{k+1}} - f_{n_k}\|_p \leq 2^{-k}$. Wir wollen zeigen, dass $f := \lim_{k \rightarrow \infty} f_{n_k}$ fast sicher existiert. Sei dazu $g_m := \sum_{k=1}^m |f_{n_{k+1}} - f_{n_k}|$. Dann folgt aus der Minkowski-Ungleichung, dass

$$\|g_m\|_p \leq \sum_{k=1}^m 2^{-k} \leq 1 .$$

Außerdem strebt $g_m \nearrow g := \sum_{k=1}^\infty |f_{n_{k+1}} - f_{n_k}|$, also auch $g_m^p \nearrow g^p$ und daher

$$\int g^p \, d\mu = \sup_m \int g_m^p \, d\mu = \sup_m \|g_m\|_p^p \leq 1 ,$$

also $g^p \in \mathcal{L}_\mu^1$, so dass $g < \infty$ μ -f.s. Insbesondere existiert

$$f := f_{n_1} + \sum_{k=1}^\infty (f_{n_{k+1}} - f_{n_k}) \quad \text{f.s., und} \quad |f| \leq |f_{n_1}| + g ,$$

so dass

$$\|f\|_p \leq \| |f_{n_1}| + g \|_p \leq \|f_{n_1}\|_p + \|g\|_p < \infty ,$$

also $f \in \mathcal{L}_\mu^p$. Aus dem Lemma von Fatou folgt schließlich

$$\begin{aligned} \|f_n - f\|_p^p &= \int |f_n - f|^p \, d\mu = \int \liminf_{m \rightarrow \infty} |f_n - f_{n_m}|^p \, d\mu \\ &\leq \liminf_{m \rightarrow \infty} \int |f_n - f_{n_m}|^p \, d\mu \leq \sup_{m \geq n} \|f_n - f_m\|_p^p =: \epsilon_n \end{aligned}$$

und $\epsilon_n \rightarrow 0$ nach Voraussetzung. □

7.6 Bemerkung (L^p -Räume) Sei L_μ^p die Menge der $\|\cdot\|_p$ -Äquivalenzklassen von \mathcal{L}_μ^p ($f \sim g$, falls $\|f - g\|_p = 0$, d.h. falls $f = g$ μ -f.s.). Aus dem vorhergehenden Satz folgt, dass L_μ^p ein *Banach-Raum* ist. Im Fall $p = 2$ ist es sogar ein *Hilbert-Raum* mit Skalarprodukt $\langle f, g \rangle = \int fg d\mu$. Das Skalarprodukt ist wohldefiniert, da $\int |fg| d\mu \leq \|f\|_2 \|g\|_2 < \infty$, und offensichtlich ist $\|f\|_2 = \langle f, f \rangle^{\frac{1}{2}}$.

Schon im Beweis der Hölder-Ungleichung spielte die Konvexität von $x \mapsto e^x$ eine entscheidende Rolle. Ist μ ein Wahrscheinlichkeitsmaß, so lässt sich eine viel allgemeinere Ungleichung für Integrale konvexer Funktionen beweisen.

7.7 Definition (Konvexität) Sei $I \subseteq \mathbb{R}$ ein Intervall. $\varphi : I \rightarrow \mathbb{R}$ ist konvex, falls

$$\varphi(ta + (1-t)b) \leq t\varphi(a) + (1-t)\varphi(b) \quad \text{für alle } a, b \in I \text{ und } t \in [0, 1].$$

φ ist strikt konvex, falls diese Ungleichung für $a \neq b$ und $t \in (0, 1)$ strikt ist.

7.8 Bemerkung $\varphi : I \rightarrow \mathbb{R}$ ist konvex genau dann, wenn

$$\frac{\varphi(v) - \varphi(u)}{v - u} \leq \frac{\varphi(w) - \varphi(u)}{w - u} \leq \frac{\varphi(w) - \varphi(v)}{w - v} \quad \text{für alle } u < v < w \text{ in } I. \quad (*)$$

(Setze $u = a, w = b, v = ta + (1-t)b$; dann einfache Übung.) Insbesondere ist jedes konvexe φ stetig im Inneren von I , also auch messbar, und für alle $x \in \overset{\circ}{I}$ existiert

$$\varphi^*(x) := \lim_{v \downarrow x} \frac{\varphi(v) - \varphi(x)}{v - x} = \lim_{n \rightarrow \infty} \frac{\varphi(x + \frac{1}{n}) - \varphi(x)}{\frac{1}{n}} \quad (\text{Monotonie!}).$$

Mit φ ist auch φ^* messbar, und wegen (*) ist

$$\frac{\varphi(x) - \varphi(u)}{x - u} \leq \varphi^*(x) \leq \frac{\varphi(v) - \varphi(x)}{v - x} \quad \text{für alle } u < x < v \text{ in } I. \quad (**)$$

Ist φ strikt konvex, so sind alle Ungleichungen in (*) und (**) strikt.

7.9 Satz (Jensen'sche Ungleichung)

Sei $I \subseteq \mathbb{R}$ ein Intervall (nicht notwendig endlich!), (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, X eine integrierbare Zufallsvariable mit Werten in I f.s., und sei $\varphi : I \rightarrow \mathbb{R}$ konvex. Dann ist

$$\varphi(EX) \leq E[\varphi(X)].$$

Ist φ strikt konvex, so gilt Gleichheit genau dann, wenn $X = EX$ f.s.

Beweis: Wegen (**) gilt für $x, y \in I$

$$\varphi^*(x) \cdot (y - x) + \varphi(x) \leq \varphi(y)$$

mit strikter Ungleichung falls φ strikt konvex und $x \neq y$ ist. Da $X(\omega) \in I$ f.s., ist auch $EX \in I$. Ist EX ein Endpunkt von I , z.B. $EX = \min I$, so ist $X \geq EX$ f.s. und daher $X = EX$ f.s., insbesondere $E[\varphi(X)] = \varphi(EX)$. Andernfalls ist $EX \in \overset{\circ}{I}$, und es folgt mit $x = EX$ und $y = X$

$$\varphi^*(EX) \cdot (X - EX) + \varphi(EX) \leq \varphi(X) \quad \text{f.s.} \quad (***)$$

Nimmt man auf beiden Seiten dieser Ungleichung den Erwartungswert, so folgt wegen $E[X - EX] = 0$ die Behauptung. Ist φ strikt konvex und X nicht fast sicher konstant, so herrscht mit positiver Wahrscheinlichkeit strikte Ungleichung in (***) \square

Die Hölder-Ungleichung kann direkt aus der Jensen-Ungleichung gefolgert werden. Das wird in Bemerkung B.13 gezeigt.

Kapitel 8

Produkt Räume

In diesem Abschnitt zeigen wir, wie man mehrere (auch unendlich viele) zufällige Einzelergebnisse (-messungen, -beobachtungen) auf einem einzigen großen Wahrscheinlichkeitsraum modelliert. Unter dem Blickwinkel der Modellbildung tun wir den ersten Schritt, um aus (einfachen) Modellen für einzelne zufällige Phänomene ein (komplizierteres) Modell für die gleichzeitige Beschreibung vieler solcher Phänomene zu konstruieren.

Sei I eine Indexmenge, z.B. $I = \{1, 2\}$, $I = \{1, \dots, n\}$ oder $I = \mathbb{N}$. Für das Verständnis dieses Kapitels sollten Sie zunächst nur an diese Indexmengen denken und sich erst beim zweiten oder dritten Lesen klar machen, dass alles auch für kompliziertere Indexmengen wie $I = \mathbb{Z}^d$ oder $I = \tilde{\mathbb{Z}}^d$ gilt, ja sogar für nicht abzählbare wie $I = [0, \infty)$ oder $I = \mathbb{R}^d$.

Notation: Sei $(A_i)_{i \in I}$ eine Familie von Mengen.

$$\prod_{i \in I} A_i := \{(x_i)_{i \in I} : x_i \in A_i \forall i \in I\}.$$

Also zum Beispiel $\prod_{i \in \{1,2\}} A_i = A_1 \times A_2$, $\prod_{i \in \{1, \dots, n\}} A_i = A_1 \times \dots \times A_n$.

Ist $A_i = A$ für alle $i \in I$, so ist

$$\prod_{i \in I} A_i = A^I = \{(x_i)_{i \in I} : x_i \in A \forall i \in I\} = \{x \mid x : I \rightarrow A \text{ Abbildung}\}.$$

Diesen Abbildungsstandpunkt nimmt man z.B. dann ein, wenn $I = [0, \infty)$ die positive Zeitachse sein soll und $x(t)$ den Wert einer zufälligen Größe zur Zeit t beschreibt.

8.1 Definition (Zylindermengen) Seien $(\Omega_i, \mathcal{A}_i)$, $i \in I$, messbare Räume. Wir bezeichnen mit

$$\mathcal{Z} = \mathcal{Z}(\mathcal{A}_i, i \in I) := \bigcup_{\substack{S \subseteq I \\ S \text{ endlich}}} \left\{ \prod_{i \in I} A_i : A_i \in \mathcal{A}_i (i \in S), A_i = \Omega_i (i \notin S) \right\}$$

die Zylindermengen (oder messbaren Rechtecke) zu $(\mathcal{A}_i, i \in I)$.

8.2 Beispiel Ist $I = \{1, \dots, n\}$, also jede Teilmenge S von I endlich, so ist einfach

$$\mathcal{Z} = \left\{ \prod_{i \in I} A_i : A_i \in \mathcal{A}_i \forall i \in I \right\}.$$

Aus der Maßtheorie bekannt ist folgender Satz:

8.3 Satz

$\mathcal{Z} = \mathcal{Z}(\mathcal{A}_i, i \in I)$ ist ein Halbring.

Beweis siehe Satz B.14.

8.4 Definition $\bigotimes_{i \in I} \mathcal{A}_i := \sigma(\mathcal{Z})$ ist die Produkt σ -Algebra der $\mathcal{A}_i, i \in I$. (Ist $I = \{1, 2\}$, so schreiben wir $\mathcal{A}_1 \otimes \mathcal{A}_2$.)

8.5 Bemerkung $\bigotimes_{i \in I} \mathcal{A}_i \neq \{\chi_{i \in I} A_i : A_i \in \mathcal{A}_i \forall i \in I\}$. Ganz allgemein gilt in keiner der beiden Richtungen Inklusion, aber für endliche oder abzählbare Indexmengen I gilt „ \supseteq “.

Oft kann man besser mit der folgenden äquivalenten Charakterisierung der Produkt σ -Algebra arbeiten.

Notation: Für $\emptyset \neq J \subseteq I$ sei die kanonische Projektion

$$\pi_J : \prod_{i \in I} \Omega_i \rightarrow \prod_{i \in J} \Omega_i, \quad (\omega_i)_{i \in I} \mapsto (\omega_i)_{i \in J}$$

Ist $J = \{j\}$, so schreiben wir π_j statt $\pi_{\{j\}}$. (Wenn nötig schreiben wir auch π_j^I statt π_j .)

8.6 Bemerkung Ist $I = \{1, 2, 3, 4\}$ und $\Omega_i = \mathbb{R}, \mathcal{A}_i = \mathcal{B}^1$ für alle $i \in I$, so ist $\pi_2^{-1} \mathcal{A}_2 = \{\mathbb{R} \times A \times \mathbb{R}^2 : A \in \mathcal{B}^1\} \subseteq \mathcal{Z}(\mathcal{A}_i, i \in I)$. Das gilt auch für allgemeine I und $(\Omega_i, \mathcal{A}_i)$:

$$\pi_j^{-1}(\mathcal{A}_j) \subseteq \mathcal{Z}(\mathcal{A}_i, i \in I) \quad \text{für alle } j \in I.$$

8.7 Satz

$$\bigotimes_{i \in I} \mathcal{A}_i = \sigma(\mathcal{Z}(\mathcal{A}_i, i \in I)) = \sigma\left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{A}_i\right).$$

D.h. $\bigotimes_{i \in I} \mathcal{A}_i$ ist die kleinste σ -Algebra auf $\prod_{i \in I} \Omega_i$, für die alle Projektionen π_i messbar sind.

Beweis: Die Inklusion „ \supseteq “ folgt aus der vorhergehenden Bemerkung. Für die Umkehrung betrachte $\chi_{i \in I} A_i \in \mathcal{Z}$ mit $A_i = \Omega_i$ für $i \in I \setminus S$. Es ist

$$\chi_{i \in I} A_i = \bigcap_{i \in I} \pi_i^{-1} A_i = \bigcap_{i \in S} \pi_i^{-1} A_i \in \sigma\left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{A}_i\right),$$

da S endlich ist und da $\pi_i^{-1} A_i \in \pi_i^{-1} \mathcal{A}_i$ für alle $i \in I$. Also ist $\sigma(\mathcal{Z}) \subseteq \sigma(\bigcup_{i \in I} \pi_i^{-1} \mathcal{A}_i)$. \square

8.8 Satz

Seien (Ω, \mathcal{A}) und $(\Omega_i, \mathcal{A}_i)$ ($i \in I$) messbare Räume. Dann gilt: $f : \Omega \rightarrow \prod_{i \in I} \Omega_i$ ist messbar genau dann, wenn $\pi_i \circ f : \Omega \rightarrow \Omega_i$ für alle $i \in I$ messbar ist.

Beweis: “ \implies ”: π_i ist $(\bigotimes_{j \in I} \mathcal{A}_j)$ - \mathcal{A}_i -messbar, also ist $\pi_i \circ f$ \mathcal{A} - \mathcal{A}_i -messbar.

“ \impliedby ”: $f^{-1}(\pi_i^{-1} \mathcal{A}_i) = (\pi_i \circ f)^{-1} \mathcal{A}_i \subseteq \mathcal{A}$ für alle $i \in I$. Also ist

$$\begin{aligned} f^{-1} \left(\bigotimes_{i \in I} \mathcal{A}_i \right) &= f^{-1} \left(\sigma \left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{A}_i \right) \right) \stackrel{\text{Satz 3.4}}{=} \sigma \left(f^{-1} \left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{A}_i \right) \right) \\ &= \sigma \left(\bigcup_{i \in I} f^{-1} (\pi_i^{-1} \mathcal{A}_i) \right) \subseteq \sigma(\mathcal{A}) = \mathcal{A} \end{aligned}$$

□

8.9 Korollar $\emptyset \neq J \subseteq I \implies \pi_J : \mathcal{X}_{i \in I} \Omega_i \rightarrow \mathcal{X}_{i \in J} \Omega_i$ ist $(\bigotimes_{i \in I} \mathcal{A}_i)$ - $(\bigotimes_{i \in J} \mathcal{A}_i)$ -messbar.

Beweis: Für alle $i \in J$ ist $\pi_i^J \circ \pi_J = \pi_i$ messbar. □

8.10 Satz

Seien $(\Omega_i, \mathcal{A}_i)$, $i \in I$, messbare Räume, $\mathcal{A}_i = \sigma(\mathcal{C}_i)$ für alle $i \in I$. Dann ist

$$\bigotimes_{i \in I} \mathcal{A}_i = \sigma \left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{C}_i \right)$$

Beweis: “ \supseteq ”: klar ✓

“ \subseteq ”: Für jedes $j \in I$ ist $\pi_j^{-1} \mathcal{C}_j \subseteq \sigma \left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{C}_i \right)$, also nach Satz 3.4

$$\pi_j^{-1} \mathcal{A}_j = \pi_j^{-1} \sigma(\mathcal{C}_j) = \sigma(\pi_j^{-1} \mathcal{C}_j) \subseteq \sigma \left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{C}_i \right).$$

Es folgt

$$\bigotimes_{j \in I} \mathcal{A}_j = \sigma \left(\bigcup_{j \in I} \pi_j^{-1} \mathcal{A}_j \right) \subseteq \sigma \left(\bigcup_{i \in I} \pi_i^{-1} \mathcal{C}_i \right).$$

□

8.11 Korollar Sind I_1, I_2 disjunkte Indextmengen, $(\Omega_i, \mathcal{A}_i)_{i \in I_1}$ und $(\Omega_i, \mathcal{A}_i)_{i \in I_2}$ Familien messbarer Räume, so ist

$$\left(\bigotimes_{i \in I_1} \mathcal{A}_i \right) \otimes \left(\bigotimes_{i \in I_2} \mathcal{A}_i \right) = \left(\bigotimes_{i \in I_1 \cup I_2} \mathcal{A}_i \right)$$

Beweis: Wir haben die folgenden beiden Inklusionen:

$$\mathcal{Z}(\mathcal{A}_i, i \in I_1 \cup I_2) \subseteq \mathcal{Z} \left(\bigotimes_{i \in I_1} \mathcal{A}_i, \bigotimes_{i \in I_2} \mathcal{A}_i \right) \subseteq \bigotimes_{i \in I_1 \cup I_2} \mathcal{A}_i.$$

Die erste ist klar, da jede Menge $\mathcal{X}_{i \in I_1 \cup I_2} A_i$ mit $A_i \in \mathcal{A}_i$ und nur endlich vielen $A_i \neq \Omega_i$ als $(\mathcal{X}_{i \in I_1} A_i) \times (\mathcal{X}_{i \in I_2} A_i)$, d.h. als Element von $\mathcal{Z} \left(\bigotimes_{i \in I_1} \mathcal{A}_i, \bigotimes_{i \in I_2} \mathcal{A}_i \right)$ aufgefasst werden kann. (Dabei wird die gleiche Identifikation wie bei der Gleichsetzung $(A \times B) \times C = A \times$

($B \times C$) vorgenommen.) Für die zweite Inklusion seien $I := I_1 \cup I_2$, $\Omega^{(k)} := \bigtimes_{i \in I_k} \Omega_i$ und $B^{(k)} \in \bigotimes_{i \in I_k} \mathcal{A}_i$ beliebig ($k = 1, 2$). Dann ist

$$B^{(1)} \times B^{(2)} = (B^{(1)} \times \Omega^{(2)}) \cap (\Omega^{(1)} \times B^{(2)}) \in \sigma \left(\pi_{I_1}^{-1}(\bigotimes_{i \in I_1} \mathcal{A}_i) \cup \pi_{I_2}^{-1} \left(\bigotimes_{i \in I_2} \mathcal{A}_i \right) \right),$$

und das ist gerade die zweite Inklusion. Die Behauptung folgt nun durch Übergang zu den erzeugten σ -Algebren. \square

Für späteren Gebrauch merken wir noch an:

8.12 Lemma $\bigcup_{I \in \mathcal{E}(T)} \pi_I^{-1}(\bigotimes_{i \in I} \mathcal{A}_i)$ ist eine Algebra.

Beweis: Ist $A \in \bigotimes_{i \in I} \mathcal{A}_i$, $B \in \bigotimes_{i \in J} \mathcal{A}_i$ mit $I, J \in \mathcal{E}(T)$, so sind $A, B \in \bigotimes_{i \in I \cup J} \mathcal{A}_i$, so dass $\pi_I^{-1}(A)$ und $\pi_J^{-1}(B)$ und alle daraus erzeugten Mengen zur σ -Algebra $\pi_{I \cup J}^{-1}(\bigotimes_{i \in I \cup J} \mathcal{A}_i)$ gehören. \square

Kapitel 9

Produktmaße

Das Produkt zweier Maße ist aus der Analysis bekannt. So gilt z.B. für das m - und das n -dimensionale Lebesguemaß $\lambda^m \otimes \lambda^n = \lambda^{m+n}$. Mit solchen Produkten lassen sich in der Wahrscheinlichkeitstheorie gemeinsame Verteilungen unabhängiger Zufallsvariablen konstruieren.

In diesem ganzen Kapitel eien $(\Omega_1, \mathcal{A}_1, \mu_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2)$ σ -endliche Maßräume.

9.1 Satz

Durch die Forderung

$$\mu_1 \otimes \mu_2 (A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \quad (A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2) \quad (*)$$

wird eindeutig ein σ -endliches Maß $\mu_1 \otimes \mu_2$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ festgelegt, das Produktmaß zu μ_1 und μ_2 .

Beweis: Durch (*) wird eine positive Mengenfunktion auf dem Halbring $\mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$ definiert, der ja die σ -Algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ erzeugt. Es gilt:

- (i) $\mu_1 \otimes \mu_2 (\emptyset) = \mu_1 \otimes \mu_2 (\emptyset \times \emptyset) = \mu_1(\emptyset) \cdot \mu_2(\emptyset) = 0 \cdot 0 = 0$.
- (ii) $\mu_1 \otimes \mu_2$ ist σ -additiv auf $\mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$: Sei $A \times B = \bigcup_{n=1}^{\infty} A_n \times B_n$ für $A, A_n \in \mathcal{A}_1$ und $B, B_n \in \mathcal{A}_2$. Dann ist

$$1_A(\omega_1) \cdot 1_B(\omega_2) = 1_{A \times B}(\omega_1, \omega_2) = \sum_{n=1}^{\infty} 1_{A_n \times B_n}(\omega_1, \omega_2) = \sum_{n=1}^{\infty} 1_{A_n}(\omega_1) \cdot 1_{B_n}(\omega_2).$$

Für festes ω_1 sind der erste und der letzte Ausdruck offensichtlich messbare Funktionen von ω_2 . Durch Integration mit $d\mu_2(\omega_2)$ und unter Beachtung von Satz 6.1b folgt:

$$\begin{aligned} 1_A(\omega_1) \cdot \mu_2(B) &= \int_{\Omega_2} 1_A(\omega_1) \cdot 1_B(\omega_2) d\mu_2(\omega_2) \\ &= \int_{\Omega_2} \sum_{n=1}^{\infty} 1_{A_n}(\omega_1) \cdot 1_{B_n}(\omega_2) d\mu_2(\omega_2) \\ &= \sum_{n=1}^{\infty} 1_{A_n}(\omega_1) \int_{\Omega_2} 1_{B_n}(\omega_2) d\mu_2(\omega_2) = \sum_{n=1}^{\infty} 1_{A_n}(\omega_1) \cdot \mu_2(B_n). \end{aligned}$$

Durch Integration mit $d\mu_1(\omega_1)$ folgt daraus mit gleicher Begründung:

$$\mu_1(A) \cdot \mu_2(B) = \sum_{n=1}^{\infty} \mu_1(A_n) \cdot \mu_2(B_n),$$

also

$$\mu_1 \otimes \mu_2(A \times B) = \sum_{n=1}^{\infty} \mu_1 \otimes \mu_2(A_n \times B_n).$$

(iii) $\mu_1 \otimes \mu_2$ ist σ -endlich auf $\mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$: Sei $\Omega_1 = \bigcup_{k=1}^{\infty} A_k$, $\Omega_2 = \bigcup_{n=1}^{\infty} B_n$ mit $A_k \in \mathcal{A}_1$, $\mu_1(A_k) < \infty$ und $B_n \in \mathcal{A}_2$, $\mu_2(B_n) < \infty$. Dann ist

$$\Omega_1 \times \Omega_2 = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} A_k \times B_n$$

und $\mu_1 \otimes \mu_2(A_k \times B_n) = \mu_1(A_k) \cdot \mu_2(B_n) < \infty$.

$\mu_1 \otimes \mu_2$ ist also eine positive, σ -additive und σ -endliche Mengenfunktion auf dem Halbring $\mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$ und kann nach Satz 2.26 eindeutig zu einem Maß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ fortgesetzt werden. \square

9.2 Beispiel (Lebesgue-Maß auf \mathbb{R}^2) $(\Omega_i, \mathcal{A}_i) = (\mathbb{R}, \mathcal{B}^1)$, $\mu_1 = \mu_2 = \lambda^1$ (Lebesgue-Maß auf \mathbb{R}). $\lambda^2 := \lambda^1 \otimes \lambda^1$ ist das Lebesgue-Maß auf \mathbb{R}^2 . Es ist eindeutig bestimmt durch $\lambda^2(A \times C) = \lambda^1(A) \times \lambda^1(C)$ für $A, C \in \mathcal{B}^1$ (sogar durch $\lambda^2((a, b] \times (c, d]) = (b - a) \cdot (d - c)$, Satz 8.10).

Zur Vorbereitung des Satzes von Fubini benötigen wir ein paar Lemmata.

9.3 Lemma Sei $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $f : \Omega_1 \times \Omega_2 \rightarrow \bar{\mathbb{R}}$ $\mathcal{A}_1 \otimes \mathcal{A}_2$ -messbar. Dann sind für alle $\tilde{\omega}_1 \in \Omega_1$, $\tilde{\omega}_2 \in \Omega_2$

$$\begin{aligned} A_{\tilde{\omega}_1} &:= \{\omega_2 \in \Omega_2 : (\tilde{\omega}_1, \omega_2) \in A\} \in \mathcal{A}_2, \\ A_{\tilde{\omega}_2} &:= \{\omega_1 \in \Omega_1 : (\omega_1, \tilde{\omega}_2) \in A\} \in \mathcal{A}_1, \\ f_{\tilde{\omega}_1} : \Omega_2 &\rightarrow \bar{\mathbb{R}}, \omega_2 \mapsto f(\tilde{\omega}_1, \omega_2) \quad \mathcal{A}_2\text{-messbar, und} \\ f_{\tilde{\omega}_2} : \Omega_1 &\rightarrow \bar{\mathbb{R}}, \omega_1 \mapsto f(\omega_1, \tilde{\omega}_2) \quad \mathcal{A}_1\text{-messbar.} \end{aligned}$$

Beweis: Für $\tilde{\omega}_1$ definiere $i : \Omega_2 \rightarrow \Omega$ durch $i(\omega_2) = (\tilde{\omega}_1, \omega_2)$. Da $\pi_1 \circ i = \tilde{\omega}_1 = \text{const}$ und $\pi_2 \circ i = \text{Id}_{\Omega_2}$ messbar sind, ist i nach Satz 8.8 \mathcal{A}_2 - $\mathcal{A}_1 \otimes \mathcal{A}_2$ -messbar, so dass $A_{\tilde{\omega}_1} = i^{-1}(A) \in \mathcal{A}_2$ und $f_{\tilde{\omega}_1} = f \circ i$ \mathcal{A}_2 - $\bar{\mathbb{B}}$ -messbar ist. Analog für $A_{\tilde{\omega}_2}$ und $f_{\tilde{\omega}_2}$. \square

9.4 Satz (Cavalieri-Prinzip)

Für $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ sind $\omega_1 \mapsto \mu_2(A_{\omega_1})$ und $\omega_2 \mapsto \mu_1(A_{\omega_2})$ messbar, und es gilt

$$\int_{\Omega_1} \mu_2(A_{\omega_1}) d\mu_1(\omega_1) = \mu_1 \otimes \mu_2(A) = \int_{\Omega_2} \mu_1(A_{\omega_2}) d\mu_2(\omega_2). \quad (**)$$

Beweis: Wir beweisen nur die erste Gleichheit (der Beweis der zweiten verläuft ganz analog). Zunächst betrachten wir den Fall dass die Maße μ_1 und μ_2 endlich sind.

Sei $\mathcal{D} := \{A \in \mathcal{A}_1 \otimes \mathcal{A}_2 : \omega_1 \mapsto \mu_2(A_{\omega_1}) \text{ ist messbar und (**) gilt}\}$. Für $A_i \in \mathcal{A}_i$ ist $A := A_1 \times A_2 \in \mathcal{D}$, denn dann ist $\mu_2(A_{\omega_1}) = 1_{A_1}(\omega_1) \cdot \mu_2(A_2)$ messbar und

$$\int_{\Omega_1} \mu_2(A_2) d\mu_1(\omega_1) = \mu_1(A_1) \cdot \mu_2(A_2) = \mu_1 \otimes \mu_2(A).$$

Also enthält \mathcal{D} einen \cap -stabilen Erzeuger von $\mathcal{A}_1 \otimes \mathcal{A}_2$, siehe Satz 8.3. Außerdem ist \mathcal{D} ein Dynkin-System:

D1) $\Omega_1 \times \Omega_2 \in \mathcal{D}$ wie eben gesehen.

D2) $A, B \in \mathcal{D}, A \subseteq B \implies \mu_2((B \setminus A)_{\omega_1}) = \mu_2(B_{\omega_1} \setminus A_{\omega_1}) = \mu_2(B_{\omega_1}) - \mu_2(A_{\omega_1})$ ist messbar als Funktion von ω_1 . (Hier wird die Endlichkeit des Maßes μ_2 benötigt!) Es folgt, da auch μ_1 endlich ist,

$$\begin{aligned} \int_{\Omega_1} \mu_2((B \setminus A)_{\omega_1}) d\mu_1(\omega_1) &= \int_{\Omega_1} \mu_2(B_{\omega_1}) d\mu_1(\omega_1) - \int_{\Omega_1} \mu_2(A_{\omega_1}) d\mu_1(\omega_1) \\ &= \mu_1 \otimes \mu_2(B) - \mu_1 \otimes \mu_2(A) = \mu_1 \otimes \mu_2(B \setminus A) \end{aligned}$$

D3) $A_n \in \mathcal{D} (n \geq 1)$ paarweise disjunkt $\implies \mu_2((\bigcup_n A_n)_{\omega_1}) = \mu_2(\bigcup_n (A_n)_{\omega_1}) = \sum_n \mu_2((A_n)_{\omega_1})$ ist messbar, und es ist

$$\int_{\Omega_1} \mu_2 \left(\left(\bigcup_n A_n \right)_{\omega_1} \right) d\mu_1(\omega_1) = \sum_n \int_{\Omega_1} \mu_2(A_n)_{\omega_1} d\mu_1(\omega_1) = \sum_n \mu_1 \otimes \mu_2(A_n).$$

Also ist $\mathcal{D} = \mathcal{A}_1 \otimes \mathcal{A}_2$ (Satz 1.16).

Sind μ_1 und μ_2 nur σ -endlich, So ist $\Omega_i = \bigcup_{n=1}^{\infty} A_n^i$ mit $\mu_i(A_n^i) < \infty$. Dann gilt (**) für jedes Maß $\mu_1(\cdot \cap A_m^1) \otimes \mu_2(\cdot \cap A_n^2)$ und überträgt sich durch monotone Konvergenz auf

$$\mu_1 \otimes \mu_2 = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mu_1(\cdot \cap A_m^1) \otimes \mu_2(\cdot \cap A_n^2).$$

□

9.5 Korollar Sei $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Ist $\mu_1 \otimes \mu_2(A) = 0$, so ist $\mu_1((A)_{\omega_2}) = 0$ für μ_2 -f.a. $\omega_2 \in \Omega_2$ und $\mu_2((A)_{\omega_1}) = 0$ für μ_1 -f.a. $\omega_1 \in \Omega_1$.

Beweis: (**) und Satz 5.4a. □

9.6 Korollar (Integral = “Fläche unter dem Funktionsgraphen”) Sei (Ω, \mathcal{A}) messbarer Raum, μ σ -endlich auf \mathcal{A} , $f : \Omega \rightarrow [0, +\infty]$ \mathcal{A} -messbar. Sei $A_f := \{(\omega, x) \in \Omega \times \mathbb{R} : 0 \leq x \leq f(\omega)\}$. Dann ist

$$\int_{\Omega} f d\mu = (\mu \otimes \lambda)(A_f) = \int_0^{\infty} \mu\{f \geq x\} d\lambda(x).$$

Beweis: Die Funktion $h := \pi_2 - f \circ \pi_1 : \Omega \times [0, \infty) \rightarrow \mathbb{R}$ ist messbar, so dass auch die Menge $A_f = \{(\omega, x) : h(\omega, x) = x - f(\omega) \leq 0\}$ messbar ist. Da $(A_f)_{\omega} = [0, f(\omega)]$, ist nach Cavalieri

$$(\mu \otimes \lambda)(A_f) = \int_{\Omega} \lambda((A_f)_{\omega}) d\mu(\omega) = \int_{\Omega} f(\omega) d\mu(\omega),$$

und aus $(A_f)_x = \{f \geq x\}$ folgt genau so

$$(\mu \otimes \lambda)(A_f) = \int_0^\infty \mu((A_f)_x) d\lambda(x) = \int_0^\infty \mu(\{f \geq x\}) d\lambda(x).$$

□

9.7 Satz (Satz von Fubini)

Sei $f : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$ $\mathcal{A}_1 \otimes \mathcal{A}_2$ -messbar. Ist $f^- \in \mathcal{L}_{\mu_1 \otimes \mu_2}^1$ oder ist $f^+ \in \mathcal{L}_{\mu_1 \otimes \mu_2}^1$, so sind $\omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} d\mu_2$ und $\omega_2 \mapsto \int_{\Omega_1} f_{\omega_2} d\mu_1$ fast sicher messbar und es ist

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right) d\mu_1(\omega_1) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2). \end{aligned} \quad (***)$$

(Beachte, dass für $f \geq 0$ oder $f \in \mathcal{L}_{\mu_1 \otimes \mu_2}^1$ gilt: $f^- \in \mathcal{L}_{\mu_1 \otimes \mu_2}^1$.)

Beweis: Wir beweisen wieder nur die erste Gleichheit. Für $f = 1_A$ mit $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ ist

$$\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) = \int_{\Omega_2} f_{\omega_1} d\mu_2 = \mu_2(A_{\omega_1}),$$

also ist (***) gerade (**) aus Satz 9.4.

Da die Abbildung $f \mapsto \left(\omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} d\mu_2 \right)$ linear in f ist, ist $\omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} d\mu_2$ auch für Elementarfunktionen messbar, und da die linke und die rechte Seite in (***) linear von f abhängen, folgt die Gültigkeit von (***) für Elementarfunktionen f .

Sind f_n Elementarfunktionen mit $0 \leq f_n \nearrow f$, so folgt aus dem Satz von der monotonen Konvergenz, dass $\int_{\Omega_2} (f_n)_{\omega_1} d\mu_2 \nearrow \int_{\Omega_2} f_{\omega_1} d\mu_2$ für jedes feste $\omega_1 \in \Omega_1$, woraus wieder die Messbarkeit von $\omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} d\mu_2$ folgt. Daher ist auch

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \otimes \mu_2) &= \sup_n \int_{\Omega_1 \times \Omega_2} f_n d(\mu_1 \otimes \mu_2) \\ &= \sup_n \int_{\Omega_1} \left(\int_{\Omega_2} (f_n)_{\omega_1} d\mu_2 \right) d\mu_1(\omega_1) \\ &= \int_{\Omega_1} \left(\sup_n \int_{\Omega_2} (f_n)_{\omega_1} d\mu_2 \right) d\mu_1(\omega_1) \\ &= \int_{\Omega_1} \left(\int_{\Omega_2} f_{\omega_1} d\mu_2 \right) d\mu_1(\omega_1) \end{aligned}$$

Ist $f^- \in \mathcal{L}_{\mu_1 \otimes \mu_2}^1$, $f = f^+ - f^-$, so gilt (**) für f^+ und f^- . Da $f^- \in \mathcal{L}_{\mu_1 \otimes \mu_2}^1$, ist die linke Seite von (***) für f^- endlich und damit auch die rechte, so dass auch $\int_{\Omega_2} f^-(\omega_1, \omega_2) d\mu_2(\omega_2)$ für μ_1 -f.a. $\omega_1 \in \Omega_1$ endlich ist (Satz 5.4b). Daher ist $\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2)$ für μ_1 -f.a. $\omega_1 \in \Omega_1$ wohldefiniert und $> -\infty$, und die Gleichheit in (***) für f folgt aus der Linearität des Integrals.

Ist $f^+ \in \mathcal{L}_{\mu_1 \otimes \mu_2}^1$, so argumentiert man ganz analog. □

9.8 Bemerkungen

a) Sind μ_1, \dots, μ_n σ -endliche Maße auf $(\Omega_1, \mathcal{A}_1), \dots, (\Omega_n, \mathcal{A}_n)$, so wird durch

$$\mu_1 \otimes \dots \otimes \mu_n := (\dots ((\mu_1 \otimes \mu_2) \otimes \mu_3) \dots \otimes \mu_n)$$

ein σ -endliches Maß auf $\mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_n$ definiert (beachte Korollar 8.11). Wegen der Festlegung von Produktmaßen durch ihre Werte auf Zylindermengen ist die Reihenfolge der Klammerung unwesentlich. Der Satz von Fubini überträgt sich analog. **Beispiel:** $\lambda^n := \lambda \otimes \dots \otimes \lambda$ (Lebesgue-Maß auf dem \mathbb{R}^n).

b) Sind in der Situation von Teil a) h_i Wahrscheinlichkeitsdichten auf $(\Omega_i, \mathcal{A}_i, \mu_i)$ ($i = 1, \dots, n$) und ist $h : \Omega_1 \times \dots \times \Omega_n \rightarrow [0, +\infty]$ definiert durch $h(\omega_1, \dots, \omega_n) = h_1(\omega_1) \dots h_n(\omega_n)$, so ist $(h_1 \mu_1) \otimes \dots \otimes (h_n \mu_n) = h \cdot (\mu_1 \otimes \dots \otimes \mu_n)$, denn beide Maße stimmen auf Zylindermengen $A_1 \times \dots \times A_n$ überein (Beweis für $n = 2$):

$$\begin{aligned} (h_1 \mu_1) \otimes (h_2 \mu_2)(A_1 \times A_2) &= (h_1 \mu_1)(A_1) \cdot (h_2 \mu_2)(A_2) \\ &= \int_{\Omega_1} 1_{A_1}(\omega_1) h_1(\omega_1) d\mu_1(\omega_1) \cdot \int_{\Omega_2} 1_{A_2}(\omega_2) h_2(\omega_2) d\mu_2(\omega_2) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} 1_{A_1}(\omega_1) h_1(\omega_1) 1_{A_2}(\omega_2) h_2(\omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2) \\ &= \int_{\Omega_1 \times \Omega_2} 1_{A_1 \times A_2}(\omega_1, \omega_2) h(\omega_1, \omega_2) d(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) \\ &= (h \cdot (\mu_1 \otimes \mu_2))(A_1 \times A_2) \end{aligned}$$

Die Verteilungen von Familien unabhängiger Zufallsvariablen lassen sich als Produktmaße charakterisieren:

9.9 Satz (Verteilungen von Familien unabhängiger Zufallsvariablen sind Produktmaße)
 Seien X_1, \dots, X_n Zufallsvariablen auf (Ω, \mathcal{A}, P) . Bezeichne \mathcal{X} die Abbildung $\mathcal{X} : \Omega \rightarrow \bar{\mathbb{R}}^n$, $\omega \mapsto (X_1(\omega), \dots, X_n(\omega))$. Dann gilt

$$X_1, \dots, X_n \text{ unabhängig} \iff P_{\mathcal{X}} = P_{X_1} \otimes \dots \otimes P_{X_n},$$

d.h. die gemeinsame Verteilung der X_i ist das Produkt der einzelnen Verteilungen.
 In diesem Fall gilt: Ist $f : \bar{\mathbb{R}}^n \rightarrow \bar{\mathbb{R}}$ nicht-negativ oder in $\mathcal{L}_{P_{\mathcal{X}}}^1$, so ist

$$E[f(X_1, \dots, X_n)] = \int_{\bar{\mathbb{R}}^n} f dP_{\mathcal{X}} = \int_{\bar{\mathbb{R}}^n} f d(P_{X_1} \otimes \dots \otimes P_{X_n}).$$

Beweis: X_1, \dots, X_n sind unabhängig genau dann, wenn für alle $A = A_1 \times \dots \times A_n \in \mathcal{Z}(\mathcal{A}_1, \dots, \mathcal{A}_n)$ gilt

$$P_{\mathcal{X}}(A) = P\left(\bigcap_{i=1}^n \{X_i \in A_i\}\right) \stackrel{!}{=} \prod_{i=1}^n P\{X_i \in A_i\} = \prod_{i=1}^n P_{X_i}(A_i) = (P_{X_1} \otimes \dots \otimes P_{X_n})(A)$$

Das wiederum ist äquivalent zu $P_{\mathcal{X}} = P_{X_1} \otimes \dots \otimes P_{X_n}$. □

9.10 Definition (Faltung von Wahrscheinlichkeitsmaßen auf \mathbb{R}) Bezeichne \mathcal{W} die Menge aller Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathcal{B})$ und sei $s : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x + y$. Für $\mu, \nu \in \mathcal{W}$ sei

$$\mu * \nu := (\mu \otimes \nu) \circ s^{-1} \in \mathcal{W}$$

die Faltung (oder das Faltungsprodukt) von μ und ν .

9.11 Satz (Eigenschaften der Faltung)

a) $\mu * \nu(A) = \int_{\mathbb{R}} \mu(A - x) d\nu(x) = \int_{\mathbb{R}} \nu(A - x) d\mu(x) = \nu * \mu(A).$

b) Ist $\mu = f\lambda$, so ist $\mu * \nu = h\lambda$ mit $h(x) := \int_{\mathbb{R}} f(x - y) d\nu(y)$. Ist darüberhinaus $\nu = g\lambda$, so ist $h(x) = f * g(x) := \int_{\mathbb{R}} f(x - y)g(y) dy$.

c) Sind X und Y unabhängige Zufallsvariablen, so hat $X + Y$ die Verteilung

$$P_{X+Y} = P_X * P_Y.$$

d) $*$ ist eine assoziative und kommutative Verknüpfung auf \mathcal{W} mit neutralem Element δ_0 .

Beweis: Übung.

Kapitel 10

Maße auf unendlichen Produktträumen

In diesem Kapitel zeigen wir zunächst, dass Wahrscheinlichkeitsmaße auf unendlichen Produktträumen durch Festlegung ihrer Werte auf Zylindermengen definiert werden können. Das versetzt uns in die Lage, die Ergebnisse aus Kapitel 9 auf einfache Weise auf den Fall unendlich vieler Faktoren zu übertragen.

Notation: Der allgemeinen Konvention folgend schreiben wir zur Vereinfachung der Notation \mathcal{B}^T statt $\mathcal{B}^{\otimes T}$.

10.1 Definition (Verträgliches System von Verteilungen) Bezeichne $\mathcal{E}(T)$ die Familie aller endlichen Teilmengen der Indexmenge T . Ein System $(\mu_I)_{I \in \mathcal{E}(T)}$ von Wahrscheinlichkeitsmaßen auf $(\mathbb{R}^I, \mathcal{B}^I)$ heißt verträglich (oder projektiv), falls

$$\mu_J = \mu_I \circ (\pi_J^I)^{-1} \quad \text{für alle } J \subseteq I \in \mathcal{E}(T).$$

10.2 Bemerkung (Marginalverteilungen) Sei ν ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^T, \mathcal{B}^T)$. Setze $\nu_I := \nu \circ (\pi_I^T)^{-1}$ für $I \in \mathcal{E}(T)$. Ist $J \subseteq I$, so folgt aus $\pi_J^T = \pi_J^I \circ \pi_I^T$, dass $(\nu_I)_{I \in \mathcal{E}(T)}$ verträglich ist. Die ν_I mit $I \in \mathcal{E}(T)$ heißen endlich-dimensionale Marginalverteilungen von ν .

Der folgende Satz ist die Umkehrung dieser Bemerkung:

10.3 Satz (Existenzsatz von Kolmogorov)

Sei $(\mu_I)_{I \in \mathcal{E}(T)}$ ein verträgliches System von Wahrscheinlichkeitsmaßen auf $(\mathbb{R}^I, \mathcal{B}^I)$, ($I \in \mathcal{E}(T)$). Dann gibt es genau ein Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}^T, \mathcal{B}^T)$ zu dem die μ_I die Randverteilungen sind, d.h. für das gilt

$$\mu_I = \mu \circ (\pi_I^T)^{-1} \quad \text{für alle } I \in \mathcal{E}(T). \quad (*)$$

Beweis: Wir betrachten die Algebra $\mathcal{A}_0 := \bigcup_{I \in \mathcal{E}(T)} \pi_I^{-1}(\mathcal{B}^I)$, die den Halbring \mathcal{Z} der Zylindermengen in $(\mathbb{R}^T, \mathcal{B}^T)$ enthält und daher die Produkt- σ -Algebra \mathcal{B}^T erzeugt. Jedes Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}^T, \mathcal{B}^T)$ ist wegen Satz 2.19 durch die Bedingung (*) eindeutig bestimmt, denn (*) legt die Werte von μ auf \mathcal{Z} fest. Also gibt es höchstens ein solches Maß.

Wir wenden uns der Existenz zu: Für $B = \pi_I^{-1}(A)$ mit $A \in \mathcal{B}^I$ und $I \in \mathcal{E}(T)$ setze

$$\mu(B) := \mu_I(A).$$

Wir zeigen, dass μ auf \mathcal{A}_0 dadurch eindeutig definiert ist: Hat B außerdem die Darstellung $B = \pi_{I'}^{-1}(A')$, so setze $J = I \cup I'$ und beachte

$$\pi_J^{-1} \left((\pi_I^J)^{-1}(A) \right) = \pi_I^{-1}(A) = B = \pi_I^{-1}(A') = \pi_J^{-1} \left((\pi_{I'}^J)^{-1}(A') \right).$$

Da π_J surjektiv ist, folgt daraus $(\pi_I^J)^{-1}(A) = (\pi_{I'}^J)^{-1}(A')$, so dass

$$\mu_I(A) \stackrel{\text{Verträglkt.}}{=} \mu_J \left((\pi_I^J)^{-1}(A) \right) = \mu_J \left((\pi_{I'}^J)^{-1}(A') \right) = \mu_{I'}(A'),$$

d.h. $\mu(B)$ ist eindeutig definiert.

Seien nun $B = \pi_I^{-1}(A)$, $B' = \pi_{I'}^{-1}(A')$ und $B \cap B' = \emptyset$. Wegen der vorhergehenden Überlegung kann man o.B.d.A. $I = I'$ annehmen, so dass insbesondere $A \cap A' = \emptyset$. Daher ist

$$\mu(B \cup B') = \mu(\pi_I^{-1}(A \cup A')) = \mu_I(A \cup A') = \mu_I(A) + \mu_I(A') = \mu(B) + \mu(B'),$$

d.h. μ ist *additiv* auf der Algebra \mathcal{A}_0 , die ja insbesondere selbst ein Halbring ist.

Um den Fortsetzungssatz 2.26 anwenden zu können, der die Fortsetzbarkeit von μ zu einem Wahrscheinlichkeitsmaß auf \mathcal{B}^T garantiert, müssen wir noch die σ -Subadditivität von μ auf \mathcal{A}_0 nachweisen. Seien also $B = \pi_I^{-1}(A) \in \mathcal{A}_0$ und $B_n = \pi_{I_n}^{-1}(A_n) \in \mathcal{A}_0$, $B \subseteq \bigcup_{n=1}^{\infty} B_n$. Setze $J_N := I \cup \bigcup_{n=1}^N I_n$ für $N \geq 1$ und $J \subseteq T$ eine abzählbar unendliche Indexmenge, die alle J_N enthält. Ist T abzählbar, so kann man einfach $J = T$ wählen, sonst nimmt man $J := \bigcup_{N=1}^{\infty} J_N$. Wir geben hier den Beweis nur für $J = T$, den nur von dere Notation her etwas aufwändigeren Beweis für allgemeines T findet man in Anhang C.

Zunächst ist

$$\pi_I^{-1}(A) = B \subseteq \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} \pi_{I_n}^{-1}(A_n).$$

Sei $\epsilon > 0$. Aus der Regularität der Wahrscheinlichkeitsmaße μ_{J_n} auf \mathbb{R}^{J_n} und μ_{J_n} auf \mathbb{R}^{J_n} (Bemerkung 2.30) folgt, dass es kompakte Teilmengen $K_n \subseteq (\pi_{I_n}^{J_n})^{-1}(A)$ und offene Mengen $G_n \supseteq A_n$ derart gibt, dass

$$\mu_{J_n}((\pi_{I_n}^{J_n})^{-1}A \setminus K_n) < 2^{-n}\epsilon \quad \text{und} \quad \mu_{I_n}(G_n \setminus A_n) < 2^{-n}\epsilon.$$

Für $m > 0$ sei $K^m := \bigcap_{n=1}^m (\pi_{J_n}^{J_m})^{-1}K_n \subseteq K_m \subseteq (\pi_I^{J_m})^{-1}(A)$. Dann ist K^m kompakt und

$$\begin{aligned} \mu_{J_m} \left((\pi_I^{J_m})^{-1}A \setminus K^m \right) &= \mu_{J_m} \left(\bigcup_{n=1}^m (\pi_I^{J_m})^{-1}A \setminus (\pi_{J_n}^{J_m})^{-1}K_n \right) \\ &\leq \sum_{n=1}^m \mu_{J_m} \left((\pi_I^{J_m})^{-1}A \setminus (\pi_{J_n}^{J_m})^{-1}K_n \right) \\ &= \sum_{n=1}^m \mu_{J_m} \left((\pi_{J_n}^{J_m})^{-1}((\pi_I^{J_n})^{-1}A \setminus K_n) \right) \\ &= \sum_{n=1}^m \mu_{J_n} \left((\pi_I^{J_n})^{-1}A \setminus K_n \right) < \epsilon \end{aligned}$$

und

$$(\pi_{J_m})^{-1}K^m = \bigcap_{n=1}^m (\pi_{J_n})^{-1}K_n \subseteq (\pi_I)^{-1}A \subseteq \bigcup_{n=1}^{\infty} (\pi_{I_n})^{-1}(A_n) \subseteq \bigcup_{n=1}^{\infty} (\pi_{I_n})^{-1}(G_n). \quad (*)$$

Am Ende des Beweises werden wir zeigen, dass die K^m eine gewisse Kompaktheitseigenschaft haben: Aus (*) folgt

$$\exists M > 0 : (\pi_{J_M})^{-1}K^M \subseteq \bigcup_{n=1}^M (\pi_{I_n})^{-1}(G_n) = (\pi_{J_M})^{-1} \left(\bigcup_{n=1}^M (\pi_{I_n}^{J_M})^{-1}G_n \right) \quad (**)$$

Da π_{J_M} surjektiv ist, folgt daraus

$$K^M \subseteq \bigcup_{n=1}^M (\pi_{I_n}^{J_M})^{-1}G_n.$$

Da μ_{J_M} ein Maß auf \mathcal{B}^{J_M} ist, folgt

$$\begin{aligned} \mu(B) &= \mu_{J_M} \left((\pi_I^{J_M})^{-1}A \right) \leq \mu_{J_M}(K^M) + \epsilon \leq \sum_{n=1}^M \mu_{J_M} \left((\pi_{I_n}^{J_M})^{-1}G_n \right) + \epsilon \\ &= \sum_{n=1}^M \mu_{I_n}(G_n) + \epsilon \leq \sum_{n=1}^M (\mu_{I_n}(A_n) + 2^{-n}\epsilon) + \epsilon \\ &\leq \sum_{n=1}^{\infty} \mu(B_n) + 2\epsilon. \end{aligned}$$

Im Limes $\epsilon \rightarrow 0$ erhält man die σ -Subadditivität von μ auf \mathcal{Z} .

Es bleibt zu zeigen, dass (**) aus (*) folgt. Angenommen (**) ist falsch. Dann gibt es zu jedem $m > 0$ einen Punkt

$$x_m \in (\pi_{J_m})^{-1}K^m \setminus \bigcup_{n=1}^m (\pi_{I_n})^{-1}(G_n) \subseteq \mathbb{R}^T. \quad (***)$$

Insbesondere ist für $m \geq M > 0$

$$\pi_{J_M}x_m = \pi_{J_M}^{J_m}(\pi_{J_m}x_m) \in \pi_{J_M}^{J_m}K^m \subseteq \bigcap_{n=1}^M \pi_{J_M}^{J_m} \left((\pi_{J_M}^{J_m})^{-1}(\pi_{J_n}^{J_m})^{-1}K_n \right) = K^M.$$

Da K^M kompakt ist, hat $(\pi_{J_M}x_m)_{m \geq M}$ eine in K^M konvergente Teilfolge, und durch Diagonalisierung erhält man eine Teilfolge $(x_{m_j})_{j > 0}$ derart, dass

$$y_M := \lim_{j \rightarrow \infty} \pi_{J_M}(x_{m_j}) \in K^M \subset \mathbb{R}^{J_M} \quad \text{für jedes } M > 0 \text{ existiert.}$$

Sei $M' > M$. Dann ist

$$\pi_{J_{M'}}^{J_{M'}}(y_{M'}) = \lim_{j \rightarrow \infty} \pi_{J_{M'}}^{J_{M'}} \pi_{J_{M'}}(x_{m_j}) = \lim_{j \rightarrow \infty} \pi_{J_M}(x_{m_j}) = y_M,$$

und es gibt einen Punkt $y \in \mathbb{R}^T$ derart, dass $\pi_{J_M}(y) = y_M \in K^M$ für alle $M > 0$. Es folgt aus (*), dass

$$\exists N > 0 : \pi_{I_N}^{J_N}(y_N) = \pi_{I_N}^{J_N}(\pi_{J_N}y) = \pi_{I_N}(y) \in G_N.$$

Da G_N offen ist und da $\pi_{I_N}^{J_N}(y_N) = \lim_{j \rightarrow \infty} \pi_{I_N}^{J_N}(\pi_{J_N} x_{m_j}) = \lim_{j \rightarrow \infty} \pi_{I_N} x_{m_j}$, ist $\pi_{I_N}(x_{m_j}) \in G_N$ für alle hinreichend großen j . Für $m_j \geq N$ steht das aber im Widerspruch zur Wahl von x_{m_j} in (***) . \square

10.4 Bemerkung (Satz von Kolmogorov für polnische Räume) Der Satz von Kolmogorov bleibt richtig, wenn $(\mathbb{R}, \mathcal{B})$ durch einen beliebigen *polnischen Raum* X mit Borel- σ -Algebra \mathcal{B} ersetzt wird. (X heißt polnisch, wenn X ein topologischer Raum mit abzählbarer Basis ist, dessen Topologie durch eine vollständige Metrik definiert wird.) Insbesondere ist auch $(\mathbb{R}, \mathcal{B})$ polnisch. Als Metrik kann $d(x, y) = |\arctan(x) - \arctan(y)|$ dienen, wobei $\arctan(\pm\infty) = \pm \frac{\pi}{2}$.

Nun suchen wir, wie eingangs angekündigt, den Anschluss an Kapitel 9. Zunächst betrachten wir eine Familie (μ_i) von Wahrscheinlichkeitsmaßen auf $(\mathbb{R}, \mathcal{B}^1)$. Zu $I \in \mathcal{E}(T)$ sei $\mu_I := \bigotimes_{i \in I} \mu_i$ ein Produktmaß wie in Bemerkung 9.8a. Die Verträglichkeit des Systems $(\mu_I)_{I \in \mathcal{E}(T)}$ folgt sofort aus dem Satz von Fubini. Daher erhalten wir als Korollar zum Satz von Kolmogorov:

10.5 Korollar (Existenz und Eindeutigkeit von Produktmaßen – unendlich viele Faktoren)

Sei $(\Omega, \mathcal{A}) = (\mathbb{R}^T, \mathcal{B}^T)$. Dann gibt es genau ein Wahrscheinlichkeitsmaß $\mu = \bigotimes_{i \in T} \mu_i$ auf (Ω, \mathcal{A}) derart, dass

$$\mu \left(\bigtimes_{i \in T} A_i \right) = \prod_{i \in T} \mu_i(A_i)$$

wenn $A_i \in \mathcal{B}^1$ ($i \in T$) und $A_i \neq \mathbb{R}$ für höchstens endlich viele i .

10.6 Korollar (Existenz von unabhängigen Prozessen) Ist $P = \bigotimes_{i \in T} \mu_i$ eine Produktwahrscheinlichkeit auf $\Omega = \mathbb{R}^T$ und sind $f_i : \mathbb{R} \rightarrow \mathbb{R}$ \mathcal{B}^1 -messbar, so sind die Zufallsvariablen $X_i := f_i \circ \pi_i : \Omega \rightarrow \mathbb{R}$ unabhängig und $P_{X_i} = \mu_i \circ f_i^{-1}$ ($i \in T$).

Beweis: Es ist $P_{X_i}(A) = P(\pi_i^{-1}(f_i^{-1}A)) = \mu_i(f_i^{-1}A)$. Zum Beweis der Unabhängigkeit kann man o.B.d.A. annehmen, dass T endlich ist (Satz 4.6a). Aber dann ist

$$\begin{aligned} P \left(\bigcap_{i \in T} \{X_i \in A_i\} \right) &= P \left(\bigcap_{i \in T} \pi_i^{-1}(f_i^{-1}A_i) \right) = P \left(\bigtimes_{i \in T} f_i^{-1}A_i \right) \\ &= \prod_{i \in T} \mu_i(f_i^{-1}A_i) = \prod_{i \in T} P\{X_i \in A_i\}. \end{aligned}$$

\square

10.7 Satz (Produktverteilungen und Unabhängigkeit)

Seien $(X_i)_{i \in T}$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Der Prozess $(X_i)_{i \in T}$ ist unabhängig genau dann, wenn

$$P_{\mathcal{X}} = \bigotimes_{i \in T} P_{X_i}$$

wo $\mathcal{X} : \Omega \rightarrow \mathbb{R}^T$, $\mathcal{X}(\omega) = (X_i(\omega))_{i \in T}$.

Beweis: $(X_i)_{i \in T}$ ist unabhängig genau dann, wenn $(X_i)_{i \in I}$ für alle $I \in \mathcal{E}(T)$ unabhängig ist (Satz 4.6a und Definition 4.8). Das ist äquivalent zu $P_{\pi_I \circ \mathcal{X}} = \prod_{i \in I} P_{X_i}$ für $I \in \mathcal{E}(T)$ (Satz 9.9), und da $P_{\pi_I \circ \mathcal{X}} = P_{\mathcal{X}} \circ \pi_I^{-1}$, ist das äquivalent zu $P_{\mathcal{X}} = \prod_{i \in T} P_{X_i}$. \square

Kapitel 11

Zerlegungssätze für Maße, Satz von Radon-Nikodym

Seien μ und ν zwei σ -endliche Maße auf dem messbaren Raum (Ω, \mathcal{A}) . Ziel dieses Kapitels ist es, zu charakterisieren, wann es eine Dichte h gibt, so dass $\nu = h\mu$.

Zunächst wiederholen wir ein paar einfache Tatsachen aus vorangegangenen Kapiteln. Sei (Ω, \mathcal{A}) ein messbarer Raum.

- ▷ Ist $h : \Omega \rightarrow [0, +\infty]$ \mathcal{A} -messbar, so wird durch $\nu(A) := \int_A h d\mu$ ($A \in \mathcal{A}$) ein Maß auf \mathcal{A} definiert. ν ist endlich genau dann, wenn $h \in \mathcal{L}_\mu^1$. (Korollar 6.3) Man sagt: ν hat Dichte h bzgl. μ , kurz:

$$\nu = h\mu \quad \text{und} \quad h = \frac{d\nu}{d\mu}.$$

- ▷ Ist $\nu = h\mu$, so gilt für messbare $f : \Omega \rightarrow [0, +\infty]$

$$\int f d\nu = \int fh d\mu \quad (*)$$

Für beliebige messbare $f : \Omega \rightarrow \mathbb{R}$ gilt daher: $f \in \mathcal{L}_\nu^1 \iff fh \in \mathcal{L}_\mu^1$, und für jedes $f \in \mathcal{L}_\nu^1$ gilt (*) ebenfalls. (Diese Aussagen wurden nicht explizit als Satz bewiesen, folgen aber leicht durch "algebraische Induktion", vergl. den Beweis zu Satz 5.19.

- ▷ Beispiele sind u.a.:

a) Normalverteilung auf \mathbb{R} : μ Lebesgue-Maß auf \mathbb{R} , $h(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$,
 $m \in \mathbb{R}$, $\sigma^2 > 0$.

b) Exponentialverteilung auf $[0, \infty)$: μ Lebesgue-Maß auf $[0, \infty)$, $h(x) = \alpha e^{-\alpha x}$, $\alpha > 0$.

In den uns interessierenden Fällen bestimmt ein Maß seine Dichte eindeutig (so es überhaupt eine hat).

11.1 Satz (Eindeutigkeit der Dichte)

Seien μ, ν σ -endlich. Hat ν sowohl Dichte h_1 als auch Dichte h_2 bzgl. μ , so ist $h_1 = h_2$ μ -f.s. Also ist die messbare Funktion $\frac{d\nu}{d\mu}$ nur μ -f.s. eindeutig definiert.

Beweis: Da μ und ν (und damit auch $\mu + \nu$) σ -endlich sind, gibt es $A_n \in \mathcal{A}$ mit $\Omega = \bigcup_{n=1}^{\infty} A_n$ und $\mu(A_n) + \nu(A_n) < \infty$. Es reicht zu zeigen, dass $h_1 = h_2$ μ -f.s. sicher auf jedem A_n . Sei daher $B_n := A_n \cap \{h_1 \geq h_2\}$. Dann ist $\int 1_{B_n} \cdot (h_1 - h_2) d\mu = \nu(B_n) - \mu(B_n) = 0$ und da $1_{B_n} \cdot (h_1 - h_2) \geq 0$, folgt $h_1 - h_2 = 0$ μ -f.s. auf B_n , siehe Satz 5.4a. Also ist $h_1 \leq h_2$ μ -f.s. auf A_n und durch Vertauschung der Rollen von h_1 und h_2 folgt $h_1 = h_2$ μ -f.s. auf A_n . \square

Ist $\nu = h\mu$, so gilt

$$A \in \mathcal{A}, \mu(A) = 0 \implies \nu(A) = 0$$

Dadurch motiviert definiert man

11.2 Definition (Absolut stetig) Sind μ, ν zwei Maße auf (Ω, \mathcal{A}) und gilt

$$A \in \mathcal{A}, \mu(A) = 0 \implies \nu(A) = 0,$$

so heißt ν absolut stetig bzgl. μ , kurz: $\nu \ll \mu$. Man sagt μ und ν sind äquivalent, kurz $\mu \approx \nu$, falls $\mu \ll \nu$ und $\nu \ll \mu$.

11.3 Bemerkung Aus dieser Definition folgt sofort: Hat ν Dichte bzgl. μ , so ist $\nu \ll \mu$. Wichtigstes Ergebnis dieses Kapitels wird sein, dass für σ -endliche Maße auch die Umkehrung davon gilt (Satz von Radon-Nikodym).

In vielen Fällen lässt sich Absolutstetigkeit durch ein ϵ - δ -Kriterium charakterisieren.

11.4 Satz (ϵ - δ -Kriterium)

Seien μ, ν Maße auf (Ω, \mathcal{A}) . Gilt

$$\forall \epsilon > 0 \exists \delta > 0 \forall A \in \mathcal{A} : \mu(A) < \delta \implies \nu(A) < \epsilon, \quad (**)$$

so ist $\nu \ll \mu$. Ist ν ein endliches Maß, so gilt auch die Umkehrung.

Beweis:

“ \implies ” Sei $A \in \mathcal{A}$ mit $\mu(A) = 0$. Wegen (**), ist $\nu(A) < \epsilon$ für jedes $\epsilon > 0$ und daher $\nu(A) = 0$.

“ \impliedby ” Sei ν endlich. Angenommen (**) ist falsch. Dann gibt es ein $\epsilon > 0$ und eine Folge von $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$) mit $\mu(A_n) < 2^{-n}$ aber $\nu(A_n) \geq \epsilon$ für alle n . Betrachte $A := \limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k$. Dann ist

$$\mu(A) \leq \lim_{n \rightarrow \infty} \mu \left(\bigcup_{k \geq n} A_k \right) \leq \lim_{n \rightarrow \infty} \sum_{k \geq n} \mu(A_k) \leq \lim_{n \rightarrow \infty} \sum_{k \geq n} 2^{-k} = 0$$

aber aus Satz 2.10 folgt

$$\nu(A) \stackrel{\nu \text{ endlich}}{=} \lim_{n \rightarrow \infty} \nu \left(\bigcup_{k \geq n} A_k \right) \geq \inf_n \nu(A_n) \geq \epsilon > 0$$

im Widerspruch zu $\nu \ll \mu$. \square

11.5 Beispiel Die Äquivalenz im vorhergehenden Satz gilt ohne die Endlichkeitsannahme an ν nicht. Beispiel: *Übung*

11.6 Definition (σ -additive Mengenfunktion, signiertes Maß) Sei (Ω, \mathcal{A}) ein messbarer Raum. Eine σ -additive Mengenfunktion $\varphi : \mathcal{A} \rightarrow \mathbb{R}$ heißt signiertes Maß.

11.7 Bemerkungen

- a) Ist φ ein signiertes Maß, so ist $\varphi(\emptyset) = 0$, denn aus der Additivität folgt $\varphi(\emptyset) = \varphi(\emptyset) + \varphi(\emptyset)$.
- b) Eine σ -additive Mengenfunktion ist in der Regel nicht σ -subadditiv.
- c) Da $\varphi(A) \in \mathbb{R}$, also $|\varphi(A)| < \infty$ für alle $A \in \mathcal{A}$, ist $\lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} \varphi(A_k) = \varphi(\bigcup_{k=1}^{\infty} A_k)$ konvergent, also $\lim_{n \rightarrow \infty} \sum_{k>n} \varphi(A_k) = 0$ für paarweise disjunkte A_n . In diesem Sinn ist auch die Konvergenz in Punkt ii) der obigen Definition zu verstehen. Aus Korollar 11.11 folgt dann allerdings später die absolute Konvergenz.

11.8 Beispiel Sind μ, ν endliche Maße auf (Ω, \mathcal{A}) , so ist $\varphi = \mu - \nu$ ein signiertes Maß. Wir werden zeigen, dass jedes signierte Maß so darstellbar ist.

11.9 Satz (Zerlegungssatz von Hahn)

Sei $\varphi : \mathcal{A} \rightarrow \mathbb{R}$ ein signiertes Maß. Dann gibt es disjunkte Mengen $\Omega^+, \Omega^- \in \mathcal{A}$, $\Omega^+ \cup \Omega^- = \Omega$, sodass $\varphi(E) \geq 0$ für alle $E \subseteq \Omega^+$, $E \in \mathcal{A}$, und $\varphi(E) \leq 0$ für alle $E \subseteq \Omega^-$, $E \in \mathcal{A}$.

Beweis: Sei $\alpha := \sup\{\varphi(A) : A \in \mathcal{A}\}$. Wir werden zeigen, dass es ein $\Omega^+ \in \mathcal{A}$ mit $\varphi(\Omega^+) = \alpha$ gibt. Dann ist $\alpha \in \mathbb{R}$ und für $E \subseteq \Omega^+$, $E \in \mathcal{A}$, gilt

$$\alpha \geq \varphi(\Omega^+ \setminus E) = \varphi(\Omega^+) - \varphi(E) = \alpha - \varphi(E),$$

also $\varphi(E) \geq 0$. Für $E \subseteq \Omega^- := \Omega \setminus \Omega^+$, $E \in \mathcal{A}$, gilt

$$\alpha \geq \varphi(\Omega^+ \cup E) = \varphi(\Omega^+) + \varphi(E) = \alpha + \varphi(E),$$

also $\varphi(E) \leq 0$.

Wir zeigen nun die Existenz von $\Omega^+ \in \mathcal{A}$ mit $\varphi(\Omega^+) = \alpha$. Wähle $A_n \in \mathcal{A}$ mit $\varphi(A_n) \rightarrow \alpha$ und setze $A = \bigcup_{n=1}^{\infty} A_n$. Da φ auch negative Werte annehmen kann, besteht das Problem darin, dass die A_n auch für große n noch "Inseln negativer Masse" enthalten können. So kann man nicht, wie im Fall eines Maßes schließen, dass $\varphi(A) \geq \lim_{n \rightarrow \infty} \varphi(A_n) = \alpha$. Deshalb müssen wir genauer in die A_n „hineinschauen“. *Beispiel:* $\Omega = \mathbb{Z} \setminus \{0\}$, $\varphi(\{z\}) = \frac{1}{z^3}$. Dann ist offensichtlich $\Omega^+ = \mathbb{N}$, $\Omega^- = -\mathbb{N}$ und $\alpha = \varphi(\mathbb{N}) = \sum_{k \in \mathbb{N}} \frac{1}{k^3}$. Betrachte nun $A_n := \mathbb{N} \cup \{-n\}$ für $n = 1, 2, \dots$. Es gilt $\varphi(A_n) = \alpha - \frac{1}{n^3} \rightarrow \alpha$, aber $A = \bigcup_{n \in \mathbb{N}} A_n = \mathbb{Z} = \Omega$.

Für jedes n betrachte die Mengenfamilie $\mathcal{P}_n := \{\bigcap_{k=1}^n A'_k : A'_k = A_k \text{ oder } A'_k = A \setminus A_k\}$. Diese Mengen bilden eine messbare Partition von A , d.h. $\mathcal{P}_n \subset \mathcal{A}$ und $A = \bigcup_{P \in \mathcal{P}_n} P$. (Einige der Mengen können auch leer sein.) Setze

$$C_n := \bigcup_{P \in \mathcal{P}_n, \varphi(P) > 0} P.$$

Da $A_n = A_n \cap A = \bigcup_{P' \in \mathcal{P}_{n-1}} A_n \cap P'$ ist, folgt

$$\varphi(A_n) = \sum_{P' \in \mathcal{P}_{n-1}} \varphi(A_n \cap P') = \sum_{P \in \mathcal{P}_n, P \subseteq A} \varphi(P) \leq \sum_{P \in \mathcal{P}_n, \varphi(P) > 0} \varphi(P) = \varphi(C_n).$$

Betrachte für $m \leq n$ die Mengen $E_m^n = C_m \cup \dots \cup C_n$. Da jedes $P \in \mathcal{P}_n$ in genau einem $Q \in \mathcal{P}_{n-1}$ enthalten ist, ist für $m < n$ die Menge $E_m^n \setminus E_m^{n-1}$ disjunkte Vereinigung von Mengen $P \in \mathcal{P}_n$ mit $\varphi(P) \geq 0$, also $\varphi(E_m^n \setminus E_m^{n-1}) \geq 0$. Für $E_m := \bigcup_{n \geq m} C_n$ gilt dann $E_m^n \nearrow E_m$ ($n \rightarrow \infty$) und

$$\begin{aligned} \varphi(A_m) &\leq \varphi(C_m) = \varphi(E_m^m) \leq \varphi(E_m^m) + \sum_{n=m+1}^{\infty} \varphi(E_m^n \setminus E_m^{n-1}) \\ &= \varphi\left(E_m^m \cup \bigcup_{n=m+1}^{\infty} (E_m^n \setminus E_m^{n-1})\right) = \varphi\left(\bigcup_{n \geq m} E_m^n\right) = \varphi(E_m). \end{aligned}$$

Setze $\Omega^+ = \bigcap_{m=1}^{\infty} E_m$, also $E_m \searrow \Omega^+$. Dann ist

$$\varphi(E_m) = \varphi\left(\Omega^+ \cup \bigcup_{n \geq m} (E_n \setminus E_{n+1})\right) = \varphi(\Omega^+) + \sum_{n \geq m} \varphi(E_n \setminus E_{n+1}) \xrightarrow{11.7c} \varphi(\Omega^+) \quad (m \rightarrow \infty).$$

Also:

$$\alpha = \lim_{m \rightarrow \infty} \varphi(A_m) \leq \lim_{m \rightarrow \infty} \varphi(E_m) = \varphi(\Omega^+),$$

was zu zeigen war. \square

Der Satz von Hahn führt direkt zur Zerlegung eines signierten Maßes in einen positiven und einen negativen Teil, die *singulär* zu einander sind.

11.10 Definition (Singularität von Maßen) Zwei Maße μ und ν auf \mathcal{A} sind singulär zueinander, kurz $\mu \perp \nu$, falls es ein $A \in \mathcal{A}$ mit $\mu(A) = 0 = \nu(\Omega \setminus A)$ gibt.

11.11 Korollar (Zerlegungssatz von Jordan) Sei φ ein signiertes Maß auf (Ω, \mathcal{A}) . Dann gibt es endliche Maße φ^+ und φ^- auf \mathcal{A} mit $\varphi = \varphi^+ - \varphi^-$ und $\varphi^+ \perp \varphi^-$.

Beweis: Sei $\Omega = \Omega^+ \cup \Omega^-$ die Zerlegung aus dem Hahn'schen Satz. Setze $\varphi^+(A) := \varphi(A \cap \Omega^+)$, $\varphi^-(A) := -\varphi(A \cap \Omega^-)$. \square

Zur Vorbereitung des wichtigsten Satzes dieses Kapitels, des Satzes von Radon-Nikodym, benötigen wir noch eine Konsequenz aus dem Hahnschen Zerlegungssatz.

11.12 Lemma Seien μ, ν endliche Maße auf (Ω, \mathcal{A}) mit $\nu(\Omega) > 0$ und $\nu \ll \mu$. Dann gibt es ein $A \in \mathcal{A}$ mit $\mu(A) > 0$ und ein $N > 0$ derart, dass

$$\mu(E) \leq N \cdot \nu(E) \quad \text{für alle } E \subseteq A, E \in \mathcal{A}.$$

Beweis: Sei $\Omega = \Omega_n^+ \cup \Omega_n^-$ eine Hahn-Zerlegung zum signierten Maß $(\nu - \frac{1}{n}\mu)$. Setze $M := \bigcap_{n > 0} \Omega_n^-$. Da $M \subseteq \Omega_n^-$ für alle n , ist $(\nu - \frac{1}{n}\mu)(M) \leq 0$, d.h. $\nu(M) \leq \frac{1}{n}\mu(M)$ für alle n und daher $\nu(M) = 0$, also $\nu(\bigcup_{n > 0} \Omega_n^+) = \nu(\Omega \setminus M) = \nu(\Omega) > 0$. Es gibt also ein $N > 0$ mit $\nu(\Omega_N^+) > 0$, und da $\nu \ll \mu$, ist auch $\mu(\Omega_N^+) > 0$. Mit $A = \Omega_N^+$ ist dann $\mu(A) > 0$ und $(\nu - N^{-1}\mu)(E) \geq 0$ für alle $E \subseteq A, E \in \mathcal{A}$. \square

11.13 Satz (Satz von Radon-Nikodym)

Sei $(\Omega, \mathcal{A}, \mu)$ σ -endlich und sei ν ein weiteres σ -endliches Maß auf (Ω, \mathcal{A}) . Dann hat ν Dichte bzgl. μ genau dann, wenn $\nu \ll \mu$. Ist das der Fall, so ist $\frac{d\nu}{d\mu}$ \mathcal{A} -messbar und μ -f.s. endlich. $\frac{d\nu}{d\mu}$ heißt auch die Radon-Nikodym-Ableitung von ν nach μ .

Beweis: “ \Rightarrow ” ist Bemerkung 11.3.

“ \Leftarrow ”: Wir betrachten zunächst den Spezialfall, dass μ und ν endliche Maße sind. Unser erstes Ziel ist es, ein maximales Element in der Familie

$$\mathcal{G} := \left\{ g : \Omega \rightarrow [0, \infty] \mid g \text{ } \mathcal{A}\text{-messbar, } \int_A g \, d\mu \leq \nu(A) \forall A \in \mathcal{A} \right\} .$$

zu konstruieren. Von dem zeigen wir dann, dass es die gesuchte Dichte von ν ist.

Natürlich ist $g = 0 \in \mathcal{G}$, also $\mathcal{G} \neq \emptyset$. Wir zeigen:

$$g_1, g_2 \in \mathcal{G} \quad \Rightarrow \quad \max\{g_1, g_2\} \in \mathcal{G} . \quad (*)$$

Setze dazu $E = \{g_1 \geq g_2\}$. Dann ist für $A \in \mathcal{A}$

$$\int_A \max\{g_1, g_2\} \, d\mu = \int_{A \cap E} g_1 \, d\mu + \int_{A \setminus E} g_2 \, d\mu \leq \nu(A \cap E) + \nu(A \setminus E) = \nu(A) .$$

Sei nun $\gamma := \sup\{\int g \, d\mu : g \in \mathcal{G}\}$. Wir konstruieren ein $f \in \mathcal{G}$ mit $\int f \, d\mu = \gamma$. Wähle dazu $g_n \in \mathcal{G}$ mit $\int g_n \, d\mu \rightarrow \gamma$ und setze $f_n := \max\{g_1, \dots, g_n\}$. Wegen (*) ist $f_n \in \mathcal{G}$ für alle n , und es ist $0 \leq f_1 \leq f_2 \leq \dots$. Aus dem Satz von der monotonen Konvergenz folgt daher für $f = \sup_n f_n$:

$$\int_A f \, d\mu = \sup_n \int_A f_n \, d\mu \leq \nu(A) \quad \forall A \in \mathcal{A}$$

und

$$\int f \, d\mu = \sup_n \int f_n \, d\mu \geq \sup_n \int g_n \, d\mu \geq \gamma .$$

Also ist $f \in \mathcal{G}$ und $\int f \, d\mu = \gamma \leq \nu(\Omega) < \infty$.

Definiere nun $\hat{\nu}$ durch

$$\hat{\nu}(B) := \nu(B) - \int_B f \, d\mu \geq 0 .$$

$\hat{\nu}$ ist als Differenz zweier endlicher Maße σ -additiv, also wegen der Positivität selbst ein Maß. Außerdem ist $\hat{\nu} \ll \mu$, da ja $\nu \ll \mu$. Zu zeigen: $\hat{\nu} = 0$.

Wäre das nicht der Fall, dann gäbe es wegen des nachfolgenden Lemmas 11.12 eine Menge $A \in \mathcal{A}$ mit $\mu(A) > 0$ und ein $N > 0$ derart, dass $\mu(E) \leq N \cdot \hat{\nu}(E) \forall E \in \mathcal{A}, E \subseteq A$. Also wäre für alle $B \in \mathcal{A}$

$$\begin{aligned} \int_B (f + N^{-1}1_A) \, d\mu &= \int_B f \, d\mu + N^{-1}\mu(B \cap A) \leq \int_B f \, d\mu + \hat{\nu}(B \cap A) \\ &\leq \int_B f \, d\mu + \hat{\nu}(B) = \nu(B) , \end{aligned}$$

d.h. $(f + N^{-1}1_A) \in \mathcal{G}$, im Widerspruch zu $\int (f + N^{-1}1_A) \, d\mu = \gamma + N^{-1}\mu(A) > \gamma$. Daher ist $\hat{\nu} = 0$ und der Satz ist für endliche μ und ν bewiesen.

Um daraus die Aussage für σ -endliche μ und ν herzuleiten, benutzen wir die folgende Charakterisierung der σ -Endlichkeit (Übung!): Es gibt strikt positive $f \in \mathcal{L}_\mu^1$ und $g \in \mathcal{L}_\nu^1$. Sei $\tilde{\mu} = f\mu$, $\tilde{\nu} = g\nu$. $\tilde{\mu}$ und $\tilde{\nu}$ sind endliche Maße auf (Ω, \mathcal{A}) mit $\tilde{\mu} \ll \tilde{\nu}$. Aus dem ersten Teil des Beweises folgt:

$$g\nu = \tilde{\nu} = \frac{d\tilde{\nu}}{d\tilde{\mu}} \tilde{\mu} = \frac{d\tilde{\nu}}{d\tilde{\mu}} f\mu ,$$

also

$$\nu = \frac{d\tilde{\nu}}{d\tilde{\mu}} \frac{f}{g} \mu =: \frac{d\nu}{d\mu} \mu .$$

□

Kapitel 12

Bedingte Erwartung

Wir beginnen dieses Kapitel mit einer Diskussion des Begriffs der bedingten Wahrscheinlichkeit, die sich an die Vorstellungen der elementaren Stochastik anlehnt.

Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $A, B \in \mathcal{A}$. Die *bedingte Wahrscheinlichkeit von A gegeben B* ist $P(A|B) = \frac{P(A \cap B)}{P(B)}$. Ist Ω höchstens abzählbar und $P(\{\omega\}) > 0$ für jedes $\omega \in \Omega$, so kann man auf diese Weise die bedingten Wahrscheinlichkeiten $P(A|\omega) := P(A|\{\omega\})$ definieren. Ist aber $P(\{\omega\}) = 0$ für jedes $\omega \in \Omega$, so muss man dazu etwas weniger naiv vorgehen. Wir illustrieren eine solche Vorgehensweise zunächst im Rahmen der elementaren Stochastik.

12.1 Beispiel Sei $\beta = (B_i : i \in I)$ eine endliche oder abzählbare Partition von Ω in messbare Mengen positiven Maßes, $\mathcal{F} = \sigma(\beta)$. (\mathcal{F} ist die σ -Algebra aller Mengen, die sich als Vereinigung von Mengen B_i darstellen lassen.) Dann definieren wir für $A \in \mathcal{A}$

$$P[A|\mathcal{F}](\omega) := \sum_{i \in I} 1_{B_i}(\omega) \cdot P(A|B_i).$$

$P[A|\mathcal{F}]$ ist also eine \mathcal{F} -messbare Funktion. Sie nimmt auf B_i den Wert $P(A|B_i)$ an und zeichnet sich durch folgende Eigenschaft aus: Für alle $B \in \mathcal{F}$ ist

$$\begin{aligned} \int_B P[A|\mathcal{F}] dP &= \sum_{i \in I} P(A|B_i) \int_B 1_{B_i} dP = \sum_{i \in I} P(A|B_i) \underbrace{P(B \cap B_i)}_{=0 \text{ oder } =P(B_i)} \\ &= \sum_{i \in I, B_i \subseteq B} P(A|B_i) P(B_i) = \sum_{i \in I, B_i \subseteq B} P(A \cap B_i) = P(A \cap B) \\ &= \int_B 1_A dP = (1_A P)(B). \end{aligned}$$

Da $B \in \mathcal{F}$, folgt

$$\int_B P[A|\mathcal{F}] dP|_{\mathcal{F}} = (1_A P)|_{\mathcal{F}}(B).$$

Mit den Bezeichnungen des letzten Kapitels ist daher:

$$P[A|\mathcal{F}] = \frac{d((1_A P)|_{\mathcal{F}})}{d(P|_{\mathcal{F}})}$$

eine Radon-Nikodym-Ableitung über dem messbaren Raum (Ω, \mathcal{F}) . Damit ist $P[A|\mathcal{F}](\omega)$ zwar nur noch für P -f.a. ω eindeutig definiert, aber dafür kann man diese Vorgehensweise auf sehr allgemeine Wahrscheinlichkeitsräume übertragen. Das ist das Ziel dieses Kapitels.

Sei nun \mathcal{F} eine Teil- σ -Algebra von \mathcal{A} . Beachte, dass für messbares $f : \Omega \rightarrow \mathbb{R}$ gilt: $(fP)|_{\mathcal{F}} \ll P|_{\mathcal{F}}$.

12.2 Definition (Bedingte Wahrscheinlichkeit, bedingte Erwartung)

a) Sei $A \in \mathcal{A}$. Die \mathcal{F} -messbare Funktion

$$P[A|\mathcal{F}] := \frac{d((1_A P)|_{\mathcal{F}})}{d(P|_{\mathcal{F}})}$$

heißt bedingte Wahrscheinlichkeit von A gegeben \mathcal{F} .

b) Sei $0 \leq f \in \mathcal{L}_P^1$. Die \mathcal{F} -messbare Funktion

$$E[f|\mathcal{F}] := \frac{d((fP)|_{\mathcal{F}})}{d(P|_{\mathcal{F}})}$$

heißt bedingte Erwartung von f gegeben \mathcal{F} .

c) Sei $f = f^+ - f^- \in \mathcal{L}_P^1$. Setze $E[f|\mathcal{F}] := E[f^+|\mathcal{F}] - E[f^-|\mathcal{F}]$.

Wenn nötig schreibt man zur Verdeutlichung auch $E_P[f|\mathcal{F}]$.

12.3 Bemerkungen

a) $P[A|\mathcal{F}] = E[1_A|\mathcal{F}]$.

b) $P[\emptyset|\mathcal{F}] = 0$ und $P[\Omega|\mathcal{F}] = 1$.

c) $0 \leq E[f^\pm|\mathcal{F}]$ ist P -f.s. definiert und endlich, da $f^\pm P$ ein endliches Maß ist (Satz von Radon-Nikodym).

12.4 Lemma (Charakterisierung bedingter Erwartungen und bedingter Wahrscheinlichkeiten)

Seien $f, g \in \mathcal{L}_P^1$, g sei \mathcal{F} -messbar. Dann gilt:

a) Ist $g \cdot E[f|\mathcal{F}] \in \mathcal{L}_P^1$, so ist $\int g \cdot E[f|\mathcal{F}] dP = \int gf dP$. Äquivalent dazu: $E[g \cdot E[f|\mathcal{F}]] = E[gf]$. Insbesondere ist $\int_B E[f|\mathcal{F}] dP = \int_B f dP$ für alle $B \in \mathcal{F}$.

b) Es gilt $\int_B g dP \stackrel{(\geq)}{=} \int_B f dP$ für alle $B \in \mathcal{F}$ genau dann, wenn $g \stackrel{(\geq)}{=} E[f|\mathcal{F}]$.

c) Ist \mathcal{C} ein \cap -stabiler Erzeuger von \mathcal{F} , $\Omega \in \mathcal{C}$, und ist $\int_C g dP = \int_C f dP$ für alle $C \in \mathcal{C}$, so ist $g = E[f|\mathcal{F}]$.

Beachte: Alle Gleichungen und Ungleichungen, die bedingte Erwartungen (oder bedingte Wahrscheinlichkeiten) involvieren, können immer nur fast sicher verstanden werden, da Radon-Nikodym-Ableitungen nur fast sicher festgelegt sind! Wollen wir eine bedingte Erwartung oder Wahrscheinlichkeit als überall definiert verstehen, so wählen wir eine Version davon. Die ist dann messbar bzgl. der σ -Algebra, auf die bedingt wird.

Beweis:

a) Da g \mathcal{F} -messbar ist, ist

$$\int g E[f|\mathcal{F}] dP = \int g E[f|\mathcal{F}] dP|_{\mathcal{F}} = \int g d((fP)|_{\mathcal{F}}) = \int g d(fP) = \int gf dP .$$

b) Wir beweisen die Äquivalenz für “ \leq ”; für “ \geq ” verläuft der Beweis genau so.

“ \Leftarrow ”: Sei $B \in \mathcal{F}$. Dann ist $\int_B g dP \leq \int_B E[f|\mathcal{F}] dP \stackrel{a)}{=} \int_B f dP$.

“ \Rightarrow ”: Sei $h := E[f|\mathcal{F}] - g$. Dann ist h \mathcal{F} -messbar und

$$\int_{\{h < 0\}} h dP = \int_{\{h < 0\}} E[f|\mathcal{F}] dP - \int_{\{h < 0\}} g dP \stackrel{a)}{=} \int_{\{h < 0\}} f dP - \int_{\{h < 0\}} g dP \stackrel{n.V.}{\geq} 0 ,$$

also $P\{h < 0\} = 0$, d.h. $g \leq E[f|\mathcal{F}]$.

c) Man zeigt leicht: $\mathcal{D} := \{C \in \mathcal{F} : \int_C g dP = \int_C f dP\}$ ist ein Dynkin-System. Da $\mathcal{C} \subseteq \mathcal{D}$ ein \cap -stabiler Erzeuger für \mathcal{F} ist, folgt aus Satz 1.16, dass $\mathcal{F} = \sigma(\mathcal{C}) = \mathcal{D}(\mathcal{C}) \subseteq \mathcal{D}$. Aus Teil b) des Lemmas folgt nun $g = E[f|\mathcal{F}]$.

□

12.5 Beispiele a) $\mathcal{F} = \mathcal{A}$. Dann ist $E[f|\mathcal{F}] = f$, da $\frac{d((fP)|_{\mathcal{A}})}{dP|_{\mathcal{A}}} = \frac{d(fP)}{dP} = f$.

b) \mathcal{F} trivial (d.h. $P(A) = 0$ oder 1 für alle $A \in \mathcal{F}$). Dann ist $E[f|\mathcal{F}] = E[f]$, denn die konstante Funktion $E[f]$ ist trivialerweise \mathcal{F} -messbar, und für $B \in \mathcal{F}$ gilt:

$$\int_B E[f] dP = P(B) \cdot E[f] = \left\{ \begin{array}{ll} 0, & \text{falls } P(B) = 0 \\ \int f dP, & \text{falls } P(B) = 1 \end{array} \right\} = \int_B f dP .$$

c) Ist $Y \in \mathcal{L}_P^1$ unabhängig von \mathcal{F} , so ist $E[Y|\mathcal{F}] = E[Y]$, denn für alle $B \in \mathcal{F}$ ist

$$\int_B Y dP = \int 1_B Y dP = E[Y] P(B) = \int_B E[Y] dP$$

und daraus folgt $E[Y|\mathcal{F}] = E[Y]$ nach Lemma 12.4b. (Beachte, dass das vorhergehende Beispiel ein Spezialfall dieser Aussage ist.)

d) Sei $\mathcal{F} = \sigma(\beta)$, $\beta = (B_i | i \in I)$ eine höchstens abzählbare messbare Partition von Ω mit $P(B_i) > 0$. Dann ist

$$P[A|\mathcal{F}] = \frac{d((1_A P)|_{\mathcal{F}})}{dP|_{\mathcal{F}}} = \sum_{i \in I} 1_{B_i} P(A|B_i)$$

wie in Beispiel 12.1. Allgemeiner ist in diesem Fall (Übung!)

$$E[f|\mathcal{F}] = \sum_{i \in I} \frac{1}{P(B_i)} \int_{B_i} f dP \cdot 1_{B_i} .$$

- e) Wir betrachten die \mathbb{R}^2 -wertige ZV (X, Y) , wo wir X standard-normalverteilt wählen und Y in standard-normalverteilter Distanz um X . Genauer sei (X, Y) verteilt nach der Dichte

$$\frac{1}{2\pi} e^{-\frac{x^2}{2}} e^{-\frac{(y-x)^2}{2}}.$$

Sei unsere Vorinformation der Wert von X , also $\mathcal{F} = \sigma(X)$. Dann gilt $E[Y|\mathcal{F}] = X$, denn:

- 1) X ist \mathcal{F} -messbar und
- 2) für $F = X^{-1}(B) \in \mathcal{F}$ gilt:

$$\begin{aligned} \int_F Y dP &= \int (1_B \circ X) \cdot Y dP = \int_{\mathbb{R}^2} 1_B(x) \cdot y dP_{(X,Y)}(x, y) \\ &= \int_{\mathbb{R}^2} 1_B(x) \cdot y \frac{1}{2\pi} e^{-\frac{x^2}{2}} e^{-\frac{(y-x)^2}{2}} d\lambda^2(x, y) \\ &= \frac{1}{\sqrt{2\pi}} \int_B e^{-\frac{x^2}{2}} \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} y e^{-\frac{(y-x)^2}{2}} dy \right) dx = \frac{1}{\sqrt{2\pi}} \int_B e^{-\frac{x^2}{2}} \cdot x dx \\ &= \int 1_B(x) \cdot x dP_X(x) = \int (1_B \circ X) \cdot X dP = \int_F X dP. \end{aligned}$$

12.6 Satz (Elementare Eigenschaften der bedingten Erwartung)

Für $X, Y, X_n \in \mathcal{L}_P^1$ und $\alpha, \beta \in \mathbb{R}$ gilt:

- a) $E[\alpha X + \beta Y|\mathcal{F}] = \alpha E[X|\mathcal{F}] + \beta E[Y|\mathcal{F}]$.
- b) $X \leq Y$ P-f.s. $\Rightarrow E[X|\mathcal{F}] \leq E[Y|\mathcal{F}]$ (Insbesondere: $0 \leq P[A|\mathcal{F}] \leq 1$)
- c) X \mathcal{F} -messbar $\Rightarrow E[X|\mathcal{F}] = X$
- d) $|E[X|\mathcal{F}]| \leq E[|X||\mathcal{F}]$
- e) Ist $\lim_{n \rightarrow \infty} X_n = X$ P-f.s. und $|X_n| \leq Y$ P-f.s. für alle $n, Y \in \mathcal{L}_P^1$, so ist

$$\lim_{n \rightarrow \infty} E[X_n|\mathcal{F}] = E[X|\mathcal{F}] \quad \text{P-f.s.}$$

Beweis:

- a) Sei $\varphi := \alpha E[X|\mathcal{F}] + \beta E[Y|\mathcal{F}]$. φ ist \mathcal{F} -messbar und für $B \in \mathcal{F}$ ist

$$\begin{aligned} \int_B \varphi dP &= \alpha \int_B E[X|\mathcal{F}] dP + \beta \int_B E[Y|\mathcal{F}] dP = \alpha \int_B X dP + \beta \int_B Y dP \\ &= \int_B (\alpha X + \beta Y) dP \end{aligned}$$

Die Behauptung folgt aus Lemma 12.4b.

- b) Für jedes $B \in \mathcal{F}$ ist $\int_B E[X|\mathcal{F}] dP = \int_B X dP \leq \int_B Y dP$, also nach Lemma 12.4b: $E[X|\mathcal{F}] \leq E[Y|\mathcal{F}]$.
- c) Folgt aus Lemma 12.4b für $f = g = X$.

d) Folgt aus Aussage b), da $\pm X \leq |X|$.

e) Sei $Z_n := \sup_{k \geq n} |X_k - X|$. Dann gilt: $Z_n \searrow 0$ P -f.s. und $0 \leq Z_n \leq 2Y$, so dass $\int Z_n dP \rightarrow 0$ nach dem Satz von der majorisierten Konvergenz. Außerdem ist $|E[X_n|\mathcal{F}] - E[X|\mathcal{F}]| \leq E[Z_n|\mathcal{F}]$ wegen Aussage a) und d), und es fällt $E[Z_n|\mathcal{F}] \searrow Z \geq 0$ P -f.s. wegen Aussage b). Zu zeigen ist $Z = 0$ P -f.s., was wegen $Z \geq 0$ äquivalent ist zu $E[Z] = 0$. Das folgt aber aus $\int Z dP \leq \int E[Z_n|\mathcal{F}] dP = \int Z_n dP \rightarrow 0$.

□

12.7 Satz

Seien $X \in \mathcal{L}_P^1, Y$ \mathcal{F} -messbar und $XY \in \mathcal{L}_P^1$. Dann ist

$$E[XY|\mathcal{F}] = Y \cdot E[X|\mathcal{F}] \tag{*}$$

Beweis: $Y \cdot E[X|\mathcal{F}]$ ist \mathcal{F} -messbar und für $B \in \mathcal{F}$ ist

$$\int_B Y \cdot E[X|\mathcal{F}] dP = \int \underbrace{1_B Y}_{\mathcal{F}\text{-mb.}} \cdot E[X|\mathcal{F}] dP \stackrel{\text{S.12.4a}}{=} \int 1_B Y X dP = \int_B Y X dP$$

Die Behauptung folgt aus Lemma 12.4b.

□

12.8 Satz

Seien $\mathcal{F}_1 \subseteq \mathcal{F}_2$ Teil- σ -Algebren von \mathcal{A} , $X \in \mathcal{L}_P^1$. Dann ist

$$E[E[X|\mathcal{F}_2]|\mathcal{F}_1] = E[X|\mathcal{F}_1].$$

Insbesondere gilt (wenn \mathcal{F}_2 trivial ist): $E[E[X|\mathcal{F}_1]] = E[X]$.

Beweis: Die linke Seite ist \mathcal{F}_1 -messbar, und für $B \in \mathcal{F}_1 \subseteq \mathcal{F}_2$ gilt

$$\int_B E[E[X|\mathcal{F}_2]|\mathcal{F}_1] dP = \int_B E[X|\mathcal{F}_2] dP = \int_B X dP.$$

Die Behauptung folgt aus Lemma 12.4b.

□

12.9 Bemerkung Sind $A_k \in \mathcal{A}$ paarweise disjunkt ($k > 0$), so ist

$$\begin{aligned} P \left[\bigcup_{k=1}^{\infty} A_k \mid \mathcal{F} \right] &= E \left[\lim_{n \rightarrow \infty} \sum_{k=1}^n 1_{A_k} \mid \mathcal{F} \right] \stackrel{\text{S.12.6e}}{=} \lim_{n \rightarrow \infty} E \left[\sum_{k=1}^n 1_{A_k} \mid \mathcal{F} \right] \stackrel{\text{S.12.6a}}{=} \lim_{n \rightarrow \infty} \sum_{k=1}^n E [1_{A_k} \mid \mathcal{F}] \\ &= \sum_{k=1}^{\infty} P [A_k \mid \mathcal{F}] \end{aligned} \tag{*}$$

Da außerdem $P[\emptyset|\mathcal{F}] = 0$ und $P[\Omega|\mathcal{F}] = 1$ (Bemerkung 12.3b), liegt die Vermutung nahe, dass die Abbildung $A \mapsto P[A|\mathcal{F}](\omega)$ für jedes (oder für P -fast jedes) ω ein Wahrscheinlichkeitsmaß auf \mathcal{A} ist.

Im nächsten Abschnitt werden wir lernen, dass das in vielen (wichtigen) Fällen tatsächlich so ist. In voller Allgemeinheit ist diese Vermutung leider falsch, denn für jede Folge $(A_n)_{n>0}$

paarweise disjunkter Mengen in \mathcal{A} kann die σ -Additivität von $P[\cdot|\mathcal{F}](\omega)$ in (*) auf einer anderen Nullmenge von ω 's verletzt sein, und die (überabzählbare) Vereinigung solcher Ausnahmemengen muss keine Nullmenge sein. Ein Beispiel dafür, das auf der Existenz nicht Borel-messbarer Teilmengen von $[0, 1]$ beruht, wird in [3, Problem 33.13] vorgestellt.

In der folgenden Bemerkung werden bedingte Erwartungen geometrisch als Orthogonalprojektoren in einem Hilbertraum interpretiert.

12.10 Bemerkung (Bedingte Erwartungen als Orthogonalprojektionen) Sei wieder (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $\mathcal{F} \subseteq \mathcal{A}$ eine Teil- σ -Algebra. Mit $L^2_P(\mathcal{F})$ bezeichnen wir die Menge der P -Äquivalenzklassen \mathcal{F} -messbarer, quadratintegrierbarer Funktionen $f : \Omega \rightarrow \mathbb{R}$. Durch das Skalarprodukt $\langle f, g \rangle = \int fg \, dP$ wird $L^2_P(\mathcal{F})$ zu einem Hilbertraum, siehe Bemerkung 7.6 mit daraus abgeleiteter Norm $\|f\|_2 = \langle f, f \rangle^{\frac{1}{2}}$.

Da das Skalarprodukt auf $L^2_P(\mathcal{F})$ die Einschränkung des Skalarprodukts auf dem größeren Raum $L^2_P(\mathcal{A})$ ist, und da der $L^2_P(\mathcal{F})$ mit seiner Norm vollständig ist, ist er ein abgeschlossener linearer Teilraum des $L^2_P(\mathcal{A})$.

Definiere nun

$$\Phi : L^2_P(\mathcal{A}) \rightarrow L^2_P(\mathcal{A}), \quad \Phi(f) := E[f|\mathcal{F}].$$

Dann gilt

- ▷ Φ ist linear,
- ▷ $\Phi(f) \in L^2_P(\mathcal{F})$ für alle $f \in L^2_P(\mathcal{A})$ (Jensen'sche Ungleichung für bedingte Erwartungen angewandt auf die konvexe Funktion $t \mapsto t^2$) und
- ▷ $\Phi(\Phi f) = E[E[f|\mathcal{F}]|\mathcal{F}] = \Phi(f)$, also $\Phi^2 = \Phi$,
- ▷ $\langle \Phi f, f - \Phi f \rangle = \int \Phi f \cdot (f - \Phi f) \, dP = \int E[\Phi f \cdot (f - \Phi f)|\mathcal{F}] \, dP = \int \Phi f \cdot E[f - \Phi f|\mathcal{F}] \, dP = 0$, da $\Phi f = E[f|\mathcal{F}]$.

Also ist Φ die Orthogonalprojektion auf $L^2_P(\mathcal{F})$ und $f = E[f|\mathcal{F}] + (f - E[f|\mathcal{F}])$ die orthogonale Zerlegung von f bzgl. $L^2_P(\mathcal{F})$.

Zum Abschluss ist hier noch einmal zusammengefasst die **Definition 12.2** und die **Charakterisierung** bedingter Erwartungen aus Lemma 12.4b:

1. Sei $X \in \mathcal{L}^1_P$. Für $X \geq 0$ ist $E[X|\mathcal{F}] := \frac{d((X \cdot P)|_{\mathcal{F}})}{d(P|_{\mathcal{F}})}$.
2. Für allgemeine $X \in \mathcal{L}^1_P$ ist $E[X|\mathcal{F}] := E[X^+|\mathcal{F}] - E[X^-|\mathcal{F}]$.
3. Für $X \in \mathcal{L}^1_P$ ist die Zufallsvariable Y eine Version von $E[X|\mathcal{F}]$ genau dann, wenn $\int_B Y \, dP = \int_B X \, dP$ für alle Mengen B aus einem \cap -stabilen Erzeuger von \mathcal{F} , der auch Ω enthält.

Kapitel 13

Bedingte Wahrscheinlichkeitsverteilungen

13.1 Definition (Übergangskern, stochastischer Kern) Seien $(\Omega_i, \mathcal{A}_i), i = 1, 2$, messbare Räume. Eine Abbildung $K : \Omega_1 \times \mathcal{A}_2 \rightarrow [0, 1]$ heißt stochastischer Kern oder Übergangskern von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$, falls gilt:

- i) $A \mapsto K(\omega_1, A)$ ist ein Wahrscheinlichkeitsmaß auf $(\Omega_2, \mathcal{A}_2)$ für jedes $\omega_1 \in \Omega_1$, und
- ii) $\omega_1 \mapsto K(\omega_1, A)$ ist \mathcal{A}_1 - \mathcal{B} -messbar für jedes $A \in \mathcal{A}_2$.

Integration bzgl. eines Maßes $K(\omega_1, \cdot)$ wird als $\int f(\omega_2) K(\omega_1, d\omega_2)$ geschrieben.

13.2 Bemerkung Es reicht, ii) der Definition für alle Mengen A aus einem \cap -stabilen Erzeuger \mathcal{C} von \mathcal{A}_2 mit $\Omega_2 \in \mathcal{C}$ zu fordern, denn

$$\mathcal{D} := \{A \in \mathcal{A}_2 : \omega_1 \mapsto K(\omega_1, A) \text{ ist } \mathcal{A}_1\text{-}\mathcal{B}\text{-messbar}\}$$

ist ein Dynkin-System (Übung!) und $\mathcal{C} \subseteq \mathcal{D}$, also $\mathcal{D} = \mathcal{A}_2$ nach Satz 1.16.

13.3 Beispiele a) $\Omega_1 = \mathbb{Z}, \Omega_2 = \mathbb{Z}$. Gegeben sei $k : \mathbb{Z} \times \mathbb{Z} \rightarrow [0, 1]$ mit

$$\forall x \in \mathbb{Z} : \sum_{y \in \mathbb{Z}} k(x, y) = 1.$$

Für $x \in \mathbb{Z}$ und $A \subseteq \mathbb{Z}$ setze

$$K(x, A) = \sum_{y \in A} k(x, y).$$

Ein einfaches Beispiel ist $k(x, x-1) = k(x, x+1) = \frac{1}{2}, k(x, y) = 0$ sonst. Es beschreibt die Übergänge einer einfachen symmetrischen Irrfahrt auf \mathbb{Z} .

b) $\Omega_1 = \mathbb{R}, \Omega_2 = \mathbb{R}$ mit $\mathcal{A}_1 = \mathcal{B}^1 = \mathcal{A}_2$. Für $x \in \mathbb{R}$ und $A \in \mathcal{B}^1$ setze

$$K(x, A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{(y-x)^2}{2}} dy$$

oder allgemeiner für jedes $t > 0$:

$$K_t(x, A) = \frac{1}{\sqrt{2\pi t}} \int_A e^{-\frac{(y-x)^2}{2t}} dy$$

Letzteres beschreibt die Übergänge einer Brownschen Bewegung in einem Zeitraum der Länge t .

c) $K(x, A) = \mu(A)$ für ein gegebenes Wahrscheinlichkeitsmaß μ auf $(\Omega_2, \mathcal{A}_2)$. In diesem Fall hängt die Verteilung des „neuen“ Zustands ω_2 nicht vom „aktuellen“ Zustand ω_1 ab.

Ein Wahrscheinlichkeitsmaß P auf $(\Omega_1, \mathcal{A}_1)$ und ein stochastischer Kern K von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$ bestimmen eine Art „schiefes“ Produktmaß auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$:

13.4 Satz (Produktmaße mit Übergangskern)

Seien $(\Omega_i, \mathcal{A}_i)$ messbare Räume ($i = 1, 2$), sei P ein Wahrscheinlichkeitsmaß auf $(\Omega_1, \mathcal{A}_1)$ und K ein stochastischer Kern von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$. Dann gibt es genau ein Wahrscheinlichkeitsmaß $P \otimes K$ auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$, so dass für alle $A \times B \in \mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$ gilt

$$(P \otimes K)(A \times B) = \int_A K(\omega_1, B) dP(\omega_1) = \int_{\Omega_1} 1_A(\omega_1) K(\omega_1, B) dP(\omega_1). \quad (*)$$

Beweis: $\mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$ ist ein Halbring, der die σ -Algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ erzeugt (Definition 8.4). Also gibt es höchstens ein Wahrscheinlichkeitsmaß, das (*) erfüllt.

Wegen Satz 2.26 ist für die Existenz nur zu zeigen, dass durch (*) eine positive und σ -additive Mengenfunktion auf $\mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$ definiert wird. Die Positivität ist klar. Seien daher $A_n \times B_n \in \mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$ ($n \in \mathbb{N}$) so, dass auch $A \times B := \bigcup_{n=1}^{\infty} A_n \times B_n \in \mathcal{Z}(\mathcal{A}_1, \mathcal{A}_2)$. Dann ist für jedes $\omega_1 \in A$ und jedes $\omega_2 \in \Omega_2$:

$$1_B(\omega_2) = 1_{A \times B}(\omega_1, \omega_2) = \sum_{n=1}^{\infty} 1_{A_n}(\omega_1) \cdot 1_{B_n}(\omega_2),$$

also

$$K(\omega_1, B) = K\left(\omega_1, \bigcup_{n \in \mathbb{N}, \omega_1 \in A_n} B_n\right) = \sum_{n \in \mathbb{N}, \omega_1 \in A_n} K(\omega_1, B_n) = \sum_{n \in \mathbb{N}} 1_{A_n}(\omega_1) K(\omega_1, B_n),$$

und da $A_n \subseteq A$ für alle $n \in \mathbb{N}$, folgt mit Satz 6.1b

$$\begin{aligned} (P \otimes K)(A \times B) &= \int 1_A(\omega_1) \sum_{n \in \mathbb{N}} 1_{A_n}(\omega_1) K(\omega_1, B_n) dP(\omega_1) \\ &= \sum_{n \in \mathbb{N}} \int 1_{A_n}(\omega_1) K(\omega_1, B_n) dP(\omega_1) \\ &= \sum_{n \in \mathbb{N}} (P \otimes K)(A_n \times B_n). \end{aligned}$$

□

13.5 Bemerkung Für $A \in \mathcal{A}_1$ ist $P(A) = \int 1_A(\omega_1) K(\omega_1, \Omega_2) dP(\omega_1) = (P \otimes K)(A \times \Omega_2)$.

13.6 Bemerkung Ein stochastischer Kern K erzeugt lineare Abbildungen auf Räumen von Funktionen und Maßen auf $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$, ähnlich wie eine reelle $d \times d$ -Matrix eine lineare Abbildung und auch ihre adjungierte erzeugt:

Bezeichne $\mathcal{M}(\Omega_i, \mathcal{A}_i)$ den Raum aller signierten Maße ν auf $(\Omega_i, \mathcal{A}_i)$ und $\text{mb}_b(\Omega_i, \mathcal{A}_i)$ den Raum aller beschränkten messbaren Funktionen $f : \Omega_i \rightarrow \mathbb{R}$. Dann haben wir die linearen Abbildungen

$$\begin{aligned} \mathcal{M}(\Omega_1, \mathcal{A}_1) &\rightarrow \mathcal{M}(\Omega_2, \mathcal{A}_2), \quad \nu \mapsto \nu K \text{ mit } (\nu K)(A_2) := \int_{\Omega_1} K(\omega_1, A_2) d\nu(\omega_1) \\ \text{mb}_b(\Omega_2, \mathcal{A}_2) &\rightarrow \text{mb}_b(\Omega_1, \mathcal{A}_1), \quad f \mapsto Kf \text{ mit } (Kf)(\omega_1) := \int_{\Omega_2} f(\omega_2) K(\omega_1, d\omega_2). \end{aligned}$$

(Man definiert sie zunächst für endliche Maße ν , dann gemäß $\nu = \nu^+ - \nu^-$ (Korollar 11.11) für beliebige signierte Maße.) Sie erfüllen folgende Dualitätsrelation:

$$\int_{\Omega_1} Kf d\nu = \int_{\Omega_2} f d(\nu K).$$

Für $f = 1_B$, $B \in \mathcal{A}_2$, gilt das, weil

$$\begin{aligned} \int_{\Omega_1} K 1_B d\nu &= \int_{\Omega_1} \left(\int_{\Omega_2} 1_B(\omega_2) K(\omega_1, d\omega_2) \right) d\nu(\omega_1) = \int_{\Omega_1} K(\omega_1, B) d\nu(\omega_1) \\ &= (\nu K)(B) = \int_{\Omega_2} 1_B d(\nu K). \end{aligned}$$

Für beliebige $f \in \text{mb}_b(\Omega_2, \mathcal{A}_2)$ wendet man nun algebraische Induktion an.

13.7 Bemerkung Seien P, K und $P \otimes K$ wie in Satz 13.4 und sei $\pi_1^{-1}(\mathcal{A}_1) \subseteq \mathcal{A}_1 \otimes \mathcal{A}_2$ die von der Projektion auf die erste Koordinate erzeugte σ -Algebra auf $\Omega_1 \times \Omega_2$. Dann ist für jedes $B \in \mathcal{A}_2$

$$(P \otimes K)[\Omega_1 \times B \mid \pi_1^{-1}(\mathcal{A}_1)](\omega_1, \omega_2) = K(\omega_1, B) \quad \text{für } P\text{-f.a. } \omega_1 \in \Omega_1 \text{ und beliebige } \omega_2 \in \Omega_2.$$

Das gilt, weil $(\omega_1, \omega_2) \mapsto K(\omega_1, B)$ bzgl. $\pi_1^{-1}(\mathcal{A}_1)$ messbar ist und weil für jedes $A \times \Omega_2 \in \pi_1^{-1}(\mathcal{A}_1)$ gilt:

$$\begin{aligned} \int_{A \times \Omega_2} 1_{\Omega_1 \times B} d(P \otimes K)(\omega_1, \omega_2) &= (P \otimes K)(A \times B) = \int_A K(\omega_1, B) dP(\omega_1) \\ &= \int_{\Omega_1} 1_A(\omega_1) K(\omega_1, B) dP(\omega_1) \\ &\stackrel{(*)}{=} \int_{\Omega_1 \times \Omega_2} 1_A(\omega_1) K(\omega_1, B) d(P \otimes K)(\omega_1, \omega_2) \\ &= \int_{A \times \Omega_2} K(\omega_1, B) d(P \otimes K)(\omega_1, \omega_2), \end{aligned}$$

wobei die Gleichung (*) ein Spezialfall (mit $f(\omega_1) = 1_A(\omega_1) K(\omega_1, B)$) der für beliebige $f \in \text{mb}_b(\Omega_1, \mathcal{A}_1)$ gültigen Identität

$$\int_{\Omega_1} f(\omega_1) dP(\omega_1) = \int_{\Omega_1 \times \Omega_2} f(\omega_1) dQ(\omega_1, \omega_2)$$

ist: Wegen Bemerkung 13.5 ist $(P \otimes K) \circ \pi_1^{-1} = P$, also

$$\int_{\Omega_1} f(\omega_1) dP(\omega_1) = \int_{\Omega_1 \times \Omega_2} f(\pi_1(\omega_1, \omega_2)) d(P \otimes K)(\omega_1, \omega_2) = \int_{\Omega_1 \times \Omega_2} f(\omega_1) d(P \otimes K)(\omega_1, \omega_2)$$

13.8 Definition (Reguläre bedingte Wahrscheinlichkeit, bedingte Verteilung) Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $\mathcal{F} \subset \mathcal{A}$ eine Teil- σ -Algebra.

a) Ein stochastischer Kern $K_{\mathcal{F}}$ von (Ω, \mathcal{F}) nach (Ω, \mathcal{A}) heißt eine reguläre bedingte Wahrscheinlichkeit bzgl. \mathcal{F} , falls gilt:

Für jedes $A \in \mathcal{A}$ ist $\omega \mapsto K_{\mathcal{F}}(\omega, A)$ eine Version von $P[A \mid \mathcal{F}]$.

b) Sei X eine reellwertige Zufallsvariable auf (Ω, \mathcal{A}, P) . Ein stochastischer Kern $P_{X|\mathcal{F}}$ von (Ω, \mathcal{F}) nach $(\mathbb{R}, \mathcal{B}^1)$ heißt bedingte Verteilung von X bzgl. \mathcal{F} , falls gilt:

Für jedes $B \in \mathcal{B}^1$ ist $\omega \mapsto P_{X|\mathcal{F}}(\omega, B)$ eine Version von $P[\{X \in B\} \mid \mathcal{F}]$.

Ist $\mathcal{F} = \sigma(Y)$ für eine weitere Zufallsvariable Y auf (Ω, \mathcal{A}, P) , so schreibt man auch $P_{X|Y}(\omega, B)$ oder, mit Argumenten $B \in \mathcal{B}^1$ und $y \in \mathbb{R}$,

$$P_{X|Y=y}(B) = P_{X|Y}(\omega, B) \quad \text{wobei } y = Y(\omega).$$

Da $P_{X|Y}$ bzgl. $\sigma(Y)$ messbar ist, hängt das tatsächlich nur von $y = Y(\omega)$ ab.

(Genauer: Da $\omega \mapsto P_{X|Y}(\omega, B)$ bzgl. $\sigma(Y)$ -messbar ist, gibt es für jedes $B \in \mathcal{B}^1$ eine $\mathcal{B}^1 - \mathcal{B}^1$ -messbare Funktion $h_B : \mathbb{R} \rightarrow \mathbb{R}$ derart, dass $P_{X|Y}(\omega, B) = h_B(Y(\omega))$ für alle $\omega \in \Omega$ (Übung!). Man setzt dann $P_{X|Y=y}(B) := h_B(y)$.)

13.9 Bemerkung Ist $K_{\mathcal{F}}$ wie in Definition 13.8a) und ist X eine Zufallsvariable wie in Teil b) der Definition, so ist $(\omega, A) \mapsto K_{\mathcal{F}}(\omega, X^{-1}(A))$ eine bedingte Verteilung von X .

13.10 Satz (Existenz bedingter Verteilungen)

Sei X eine Zufallsvariable auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) und \mathcal{F} eine Teil- σ -Algebra von \mathcal{A} . Dann hat X eine bedingte Verteilung $P_{X|\mathcal{F}}$.

Beweis: Für $r \in \mathbb{Q}$ sei $F(\cdot, r)$ eine Version von $P[\{X \leq r\} \mid \mathcal{F}] = E[1_{\{X \leq r\}} \mid \mathcal{F}]$, insbesondere ist $\omega \mapsto F(\omega, r)$ \mathcal{F} -messbar. Für $r \leq s$ ist $1_{\{X \leq r\}} \leq 1_{\{X \leq s\}}$, also gibt es wegen Lemma 12.4b eine P -Nullmenge $A_{r,s}$ mit

$$F(\omega, r) \leq F(\omega, s) \quad \forall \omega \notin A_{r,s}.$$

Wegen Satz 12.6e gibt es außerdem P -Nullmengen B_r ($r \in \mathbb{Q}$) und C derart, dass

$$\lim_{n \rightarrow \infty} F(\omega, r + \frac{1}{n}) = F(\omega, r) \quad \forall \omega \notin B_r \quad \text{und} \\ \lim_{n \rightarrow \infty} F(\omega, -n) = 0, \quad \lim_{n \rightarrow \infty} F(\omega, n) = 1 \quad \forall \omega \notin C.$$

$N := \bigcup_{r,s \in \mathbb{Q}} A_{r,s} \cup \bigcup_{r \in \mathbb{Q}} B_r \cup C$ ist eine P -Nullmenge. Da die $F(\cdot, r)$ alle \mathcal{F} -messbar sind, gehören auch all diese Ausnahmemengen zu \mathcal{F} . Insbesondere $N \in \mathcal{F}$. Für $\omega \in \Omega \setminus N$ und $z \in \mathbb{R}$ ist

$$F(\omega, z) := \inf\{F(\omega, r) : r \in \mathbb{Q}, r \geq z\}$$

wohl definiert. Man prüft leicht nach, dass $F(\omega, \cdot)$ eine Verteilungsfunktion ist (Übung!). Für $\omega \in N$ setze $F(\omega, z) := F_0(z)$, wo F_0 eine beliebige fest gewählte Verteilungsfunktion ist. Dann ist $z \mapsto F(\omega, z)$ für jedes $\omega \in \Omega$ eine Verteilungsfunktion. Das durch sie bestimmte Wahrscheinlichkeitsmaß auf \mathcal{B}^1 werde mit $P_{X|\mathcal{F}}(\omega, \cdot)$ bezeichnet.

Für $B = (-\infty, r]$, $r \in \mathbb{Q}$, ist

$$P_{X|\mathcal{F}}(\omega, B) = F(\omega, r) = P[\{X \leq r\}|\mathcal{F}](\omega) \cdot 1_{\Omega \setminus N}(\omega) + F_0(r) \cdot 1_N(\omega)$$

als Funktion von ω \mathcal{F} -messbar. Wegen Bemerkung 13.2 ist dann $\omega \mapsto K_{\mathcal{F}}(\omega, B)$ für alle $B \in \mathcal{B}^1$ \mathcal{F} -messbar, d.h. $P_{X|\mathcal{F}}$ ist ein stochastischer Kern.

Betrachte nun ein $F \in \mathcal{F}$. Für $B = (-\infty, r]$, $r \in \mathbb{Q}$, ist

$$\int_F P_{X|\mathcal{F}}(\omega, B) dP(\omega) = \int_F P[\{X \leq r\}|\mathcal{F}] dP = P(X^{-1}(B) \cap F).$$

Da beide Seiten dieser Gleichung (als Funktion von B) ein endliches Maß auf \mathcal{B}^1 beschreiben, folgt aus dem Eindeutigkeitsatz 2.19, dass diese Gleichheit für alle $B \in \mathcal{B}^1$ gilt. Aus Lemma 12.4c folgt dann, dass $P_{X|\mathcal{F}}(\omega, B) = P[X^{-1}(B)|\mathcal{F}](\omega) = P[\{X \in B\}|\mathcal{F}](\omega)$ P -f.s. für jedes $B \in \mathcal{B}^1$. \square

13.11 Satz (Bedingte Erwartungen durch bedingte Wahrscheinlichkeitsverteilungen)

Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $\mathcal{F} \subseteq \mathcal{A}$ eine Teil- σ -Algebra mit bedingten Wahrscheinlichkeitsverteilungen $K_{\mathcal{F}}(\omega, \cdot)$. Dann ist für jedes $f \in \mathcal{L}_P^1$

$$E_P[f|\mathcal{F}](\omega) = \int_{\Omega} f(\omega') K_{\mathcal{F}}(\omega, d\omega') = E_{K_{\mathcal{F}}(\omega, \cdot)}[f].$$

Insbesondere ist für jede Zufallsvariable X und $\psi \in \mathcal{L}_{P_X}^1$:

$$E_P[\psi(X)|\mathcal{F}](\omega) = \int_{\mathbb{R}} \psi(x) P_{X|\mathcal{F}}(\omega, dx) = E_{P_{X|\mathcal{F}}(\omega, \cdot)}[\psi].$$

Beweis: Wir zeigen diese Identität zunächst für $f = 1_A$, $A \in \mathcal{A}$. Die rechte Seite ist dann gleich $K_{\mathcal{F}}(\omega, A)$, die linke gleich $P[A|\mathcal{F}](\omega)$, d.h. beide stimmen P -fast sicher überein.

Durch algebraische Induktion folgt die Gleichheit für $f \in \mathcal{L}_P^1$. Dabei wird zur Behandlung von $E_P[f|\mathcal{F}]$ mehrfach Satz 12.6 heran gezogen. \square

13.12 Beispiel (Erwartungswert eines Verzweigungsprozesses) Sei $(Z_n)_{n \geq 0}$ ein Verzweigungsprozess, d.h.

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_i^{(n)},$$

wobei $Z_0 = 1$ und die Zufallsvariablen $X_i^{(n)}$ ($n, i \in \mathbb{N}$) unabhängig und identisch verteilt sind. Die $X_i^{(n)}$ haben Werte in \mathbb{N}_0 , und es gelte $E[X_i^{(n)}] < \infty$. Aus der Definition folgt, dass Z_n

eine Funktion der $X_i^{(k)}$ für $k < n$ und $i \in \mathbb{N}$ ist, also $\sigma(Z_n) \subseteq \sigma\left(\bigcup_{k < n, i \in \mathbb{N}} \sigma(X_i^{(k)})\right)$ und

daher Z_n unabhängig von den $X_i^{(n)}$ ist.

Wir schreiben abkürzend $E[Z_{n+1}|Z_n]$ für $E[Z_{n+1}|\sigma(Z_n)]$. Dann ist

$$\begin{aligned}
 E[Z_{n+1}] &= E \left[\sum_{m=0}^{\infty} Z_{n+1} \cdot 1_{\{Z_n=m\}} \right] \stackrel{6.1b}{=} \sum_{m=0}^{\infty} E [Z_{n+1} \cdot 1_{\{Z_n=m\}}] \\
 &\stackrel{12.8}{=} \sum_{m=0}^{\infty} E [E[Z_{n+1} \cdot 1_{\{Z_n=m\}}|Z_n]] = \sum_{m=0}^{\infty} E \left[E \left[\sum_{i=1}^m X_i^{(n)} \cdot 1_{\{Z_n=m\}} \mid Z_n \right] \right] \\
 &\stackrel{12.7}{=} \sum_{m=0}^{\infty} E \left[\sum_{i=1}^m E [X_i^{(n)}|Z_n] \cdot 1_{\{Z_n=m\}} \right] \stackrel{12.5c}{=} \sum_{m=0}^{\infty} E \left[\sum_{i=1}^m E [X_i^{(n)}] \cdot 1_{\{Z_n=m\}} \right] \\
 &= E[X_1^{(0)}] \cdot \sum_{m=0}^{\infty} m P(\{Z_n = m\}) = E[X_1^{(0)}] \cdot E[Z_n],
 \end{aligned}$$

also induktiv

$$E[Z_{n+1}] = \left(E[X_1^{(0)}] \right)^{n+1} \cdot E[Z_0] = \left(E[X_1^{(0)}] \right)^{m+1} .$$

Kapitel 14

Stationäre Prozesse

Alle Definitionen und Ergebnisse dieses Kapitels gelten für die Indexmengen \mathbb{N} , \mathbb{N}_0 und \mathbb{Z} . Deshalb benutzen wir \mathbb{I} als Symbol für eine dieser Indexmengen.

14.1 Definition und Bemerkung (Stochastischer Prozess) Eine Familie $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ von Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) heißt (stochastischer) Prozess. \mathcal{X} kann als messbare Abbildung $\mathcal{X} : \Omega \rightarrow \mathbb{R}^{\mathbb{I}}$ aufgefasst werden, denn $\pi_n \circ \mathcal{X} = X_n$ ist für jedes $n \in \mathbb{I}$ messbar. $P_{\mathcal{X}} := P \circ \mathcal{X}^{-1}$ ist die Verteilung des Prozesses $(X_n)_{n \in \mathbb{I}}$ auf $\mathbb{R}^{\mathbb{I}}$.

14.2 Definition und Bemerkung (Stationarität)

a) Es bezeichne

$$S : \mathbb{R}^{\mathbb{I}} \rightarrow \mathbb{R}^{\mathbb{I}}, \quad (Sx)_n = x_{n+1}$$

die Links-Verschiebung (Links-shift) auf $\mathbb{R}^{\mathbb{I}}$. (Beachte: Da $\pi_n \circ S = \pi_{n+1}$ für alle $n \in \mathbb{I}$, ist S nach Satz 8.8 messbar.)

b) Der Prozess $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ heißt stationär, falls gilt: $P_{\mathcal{X}} \circ S^{-1} = P_{\mathcal{X}}$.

14.3 Satz (Charakterisierung der Stationarität)

Der Prozess $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ ist stationär genau dann, wenn für jede endliche Menge $J \subseteq \mathbb{I}$, jedes $n \in \mathbb{I}$ und jede Familie $(B_j)_{j \in J}$ aus \mathcal{B} gilt:

$$P\{X_j \in B_j \forall j \in J\} = P\{X_{j+n} \in B_j \forall j \in J\} \quad (*)$$

Beweis: Seien $(B_j)_{j \in J}$ gegeben. Setze $B_j = \mathbb{R}$ für $j \in \mathbb{I} \setminus J$ und betrachte den Zylinder $B = \times_{j \in \mathbb{I}} B_j$. Dann ist (*) äquivalent zu $P_{\mathcal{X}}(B) = P_{\mathcal{X}}(S^{-n}B)$, d.h. (*) gilt für alle solchen Mengen genau dann, wenn die Wahrscheinlichkeitsmaße $P_{\mathcal{X}} \circ S^{-n}$ und $P_{\mathcal{X}}$ für alle $n \in \mathbb{I}$ auf \mathcal{Z} übereinstimmen. Da \mathcal{Z} ein Semiring und $\sigma(\mathcal{Z}) = \mathcal{B}^{\mathbb{I}}$ ist (Satz 8.3), ist das gleichbedeutend mit $P_{\mathcal{X}} \circ S^{-n} = P_{\mathcal{X}}$ für alle $n \in \mathbb{I}$ (Satz 2.19), und das ist gleichbedeutend mit $P_{\mathcal{X}} \circ S^{-1} = P_{\mathcal{X}}$, denn daraus folgt $P_{\mathcal{X}} \circ S^{-2} = (P_{\mathcal{X}} \circ S^{-1}) \circ S^{-1} = P_{\mathcal{X}} \circ S^{-1} = P_{\mathcal{X}}$ usw. \square

14.4 Korollar Ist $(X_n)_{n \in \mathbb{I}}$ unabhängig identisch verteilt, so ist $(X_n)_{n \in \mathbb{I}}$ stationär.

Beweis:

$$\begin{aligned} P\{X_j \in B_j \forall j \in J\} &= \prod_{j \in J} P_{X_j}(B_j) = \prod_{j \in J} P_{X_{j+n}}(B_j) \\ &= P\{X_{j+n} \in B_j \forall j \in J\} \end{aligned}$$

□

14.5 Bemerkung (Identisch verteilt versus stationär) Ist $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ stationär, so ist für alle $n, k \in \mathbb{I}$

$$P_{X_{n+k}} = P_{\mathcal{X}} \circ \pi_{n+k}^{-1} = P_{\mathcal{X}} \circ (\pi_n \circ S^k)^{-1} = P_{\mathcal{X}} \circ S^{-k} \circ \pi_n^{-1} = P_{\mathcal{X}} \circ \pi_n^{-1} = P_{X_n},$$

d.h. der Prozess ist identisch verteilt. Die Umkehrung gilt aber nicht: Aus identischer Verteilung folgt nicht die Stationarität (Übung!).

14.6 Bemerkung (Kanonisches Modell eines Prozesses) Sei $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ ein stochastischer Prozess auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Betrachte den stochastischen Prozess $\Pi = (\pi_n)_{n \in \mathbb{I}}$ auf $(\mathbb{R}^{\mathbb{I}}, \mathcal{B}^{\mathbb{I}}, P_{\mathcal{X}})$. $\Pi : \mathbb{R}^{\mathbb{I}} \rightarrow \mathbb{R}^{\mathbb{I}}$ ist die Identität auf $\mathbb{R}^{\mathbb{I}}$. Daher ist $P_{\Pi} = P_{\mathcal{X}} \circ \Pi^{-1} = P_{\mathcal{X}}$, d.h. beide Prozesse haben dieselbe Verteilung, sind also von einem Beobachter nicht zu unterscheiden. Π heißt das kanonische Modell für \mathcal{X} . Es gilt: $\Pi = (\pi_0 \circ S^n)_{n \in \mathbb{I}}$.

14.7 Definition (Ergodischer Prozess) Ein stationärer Prozess $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ heißt ergodisch, falls für alle $A \in \mathcal{B}^{\mathbb{I}}$ gilt

$$S^{-1}(A) = A \implies P_{\mathcal{X}}(A) = 0 \text{ oder } 1,$$

oder anders ausgedrückt, falls $\mathcal{I}(S) := \{A \in \mathcal{B}^{\mathbb{I}} : S^{-1}A = A\}$, die σ -Algebra der shift-invarianten messbaren Mengen, für $P_{\mathcal{X}}$ trivial ist.

14.8 Satz (u.i.v. Prozesse sind ergodisch)

Ist $(X_n)_{n \in \mathbb{I}}$ unabhängig identisch verteilt, so ist $(X_n)_{n \in \mathbb{I}}$ ergodisch.

Beweis: Sei $A \in \mathcal{B}^{\mathbb{I}}$, $S^{-1}(A) = A$. Sei $\epsilon > 0$ beliebig. Wegen des Approximationssatzes 2.29 (in Verbindung mit Satz 8.3) gibt es eine endliche Vereinigung A' von Zylindern aus \mathcal{Z} derart, dass $P_{\mathcal{X}}(A \Delta A') < \epsilon$. Da \mathcal{X} stationär ist (Korollar 14.4), folgt für jedes $n \in \mathbb{I}$

$$P_{\mathcal{X}}(S^{-n}A \Delta S^{-n}A') = P_{\mathcal{X}}(S^{-n}(A \Delta A')) = P_{\mathcal{X}}(A \Delta A') < \epsilon,$$

$$P_{\mathcal{X}}\left(\underbrace{(A \cap S^{-n}A) \Delta (A' \cap S^{-n}A')}_{\subseteq (A \Delta A') \cup S^{-n}(A \Delta A')}\right) \leq P_{\mathcal{X}}(A \Delta A') + P_{\mathcal{X}}(S^{-n}(A \Delta A')) \leq 2\epsilon.$$

Da A' eine endliche Vereinigung von Zylindern ist, gibt es eine endliche Teilmenge $J \subseteq \mathbb{I}$ und ein $B' \in \mathcal{B}^J$ derart, dass $A' = \pi_J^{-1}B'$. Sei nun $n \in \mathbb{I}$ so, dass $(J+n) \cap J = \emptyset$. Dann ist $S^{-n}A' = S^{-n}(\pi_J^{-1}B') = (\pi_{J+n} \circ S^n)^{-1}(B') = \pi_{J+n}^{-1}B'$ und daher

$$P_{\mathcal{X}}(A' \cap S^{-n}A') = P_{\mathcal{X}}(\pi_J^{-1}B' \cap \pi_{J+n}^{-1}B') = P_{\mathcal{X}}(\pi_J^{-1}B') \cdot P_{\mathcal{X}}(\pi_{J+n}^{-1}B')$$

$$= P_{\mathcal{X}}(A') \cdot P_{\mathcal{X}}(S^{-n}A')$$

Zusammen folgt

$$P_{\mathcal{X}}(A) - P_{\mathcal{X}}(A)^2 = P_{\mathcal{X}}(A \cap S^{-n}A) - P_{\mathcal{X}}(A) \cdot P_{\mathcal{X}}(S^{-n}A)$$

$$\leq P_{\mathcal{X}}(A' \cap S^{-n}A') - P_{\mathcal{X}}(A') \cdot P_{\mathcal{X}}(S^{-n}A') + 4\epsilon = 4\epsilon$$

Da $\epsilon > 0$ beliebig war, folgt $P_{\mathcal{X}}(A) - P_{\mathcal{X}}(A)^2 = 0$, d.h. $P_{\mathcal{X}}(A) = 0$ oder 1 .

□

14.9 Satz (Stationäre Prozesse aus maßerhaltenden Transformationen)

Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $T : \Omega \rightarrow \Omega$ messbar und maßerhaltend, d.h. $P \circ T^{-1} = P$. Sei $f : \Omega \rightarrow \mathbb{R}$ messbar.

- a) $\mathcal{X} := (f \circ T^n)_{n \in \mathbb{N}}$ ist ein stationärer Prozess.
 b) Ist (T, P) ergodisch, d.h. ist $\mathcal{I}(T) := \{A \in \mathcal{A} : T^{-1}A = A\}$ für P trivial, so ist auch der Prozess \mathcal{X} ergodisch,

Beweis: Für jede endliche Menge $J \subseteq \mathbb{I}$, jedes $n \in \mathbb{I}$ und jede Familie $(B_j)_{j \in J}$ aus \mathcal{B} gilt:

$$\begin{aligned} P\{X_j \in B_j \forall j \in J\} &= P\left(\bigcap_{j \in J} T^{-j}(f^{-1}B_j)\right) = (P \circ T^{-n})\left(\bigcap_{j \in J} T^{-j}(f^{-1}B_j)\right) \\ &= P\left(\bigcap_{j \in J} T^{-(j+n)}(f^{-1}B_j)\right) = P\{X_{j+n} \in B_j \forall j \in J\} \end{aligned}$$

Also ist der Prozess stationär. Außerdem ist

$$S \circ \mathcal{X} = (f \circ T^{n+1})_{n \in \mathbb{N}} = \mathcal{X} \circ T.$$

Daraus folgt die Ergodizität von \mathcal{X} , denn für $B \in \mathcal{I}(S)$ gilt:

$$T^{-1}(\mathcal{X}^{-1}(B)) = (\mathcal{X} \circ T)^{-1}(B) = (S \circ \mathcal{X})^{-1}(B) = \mathcal{X}^{-1}(S^{-1}B) = \mathcal{X}^{-1}(B),$$

also $P_{\mathcal{X}}(B) = P(\mathcal{X}^{-1}(B)) \in \{0, 1\}$. □

14.10 Satz (Funktionale stationärer Prozesse)

Sei $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ ein stationärer Prozess, $F : \mathbb{R}^{\mathbb{I}} \rightarrow \mathbb{R}$ messbar. Für $n \in \mathbb{I}$ definiere $Y_n := F \circ S^n \circ \mathcal{X} : \Omega \rightarrow \mathbb{R}$. Dann ist $\mathcal{Y} := (Y_n)_{n \in \mathbb{I}}$ ein stationärer Prozess, und falls \mathcal{X} ergodisch ist, ist auch \mathcal{Y} ergodisch.

Beweis: Da Stationarität und Ergodizität Eigenschaften der Verteilung eines Prozesses sind, können wir annehmen, dass $\mathcal{X} = \Pi$ ein kanonischer Prozess auf $(\mathbb{R}^{\mathbb{I}}, \mathcal{B}^{\mathbb{I}}, P_{\mathcal{X}})$ ist. Die Ergodizität von \mathcal{X} ist gleichbedeutend zur Ergodizität der maßtreuen Transformation S auf $(\mathbb{R}^{\mathbb{I}}, \mathcal{B}^{\mathbb{I}}, P_{\mathcal{X}})$. Dann ist $Y_n = F \circ S^n$, und die Behauptung folgt aus Satz 14.9.

Alternativ geben ich hier noch einen Beweis, der ohne den Übergang zum kanonischen Modell des Prozesses und ohne Satz 14.9 auskommt: Betrachte $\hat{F} : \mathbb{R}^{\mathbb{I}} \rightarrow \mathbb{R}^{\mathbb{I}}$, $\hat{F}(\underline{x}) := (F(S^n \underline{x}))_{n \in \mathbb{I}}$. Dann ist $\mathcal{Y} = \hat{F} \circ \mathcal{X}$ und

$$(\hat{F} \circ S)(\underline{x}) = (F(S^{n+1} \underline{x}))_{n \in \mathbb{I}} = S((F(S^n \underline{x}))_{n \in \mathbb{I}}) = (S \circ \hat{F})(\underline{x}),$$

so dass

$$P_{\mathcal{Y}} \circ S^{-1} = P \circ \mathcal{X}^{-1} \circ \hat{F}^{-1} \circ S^{-1} = P_{\mathcal{X}} \circ S^{-1} \circ \hat{F}^{-1} = P_{\mathcal{X}} \circ \hat{F}^{-1} = P \circ \mathcal{X}^{-1} \circ \hat{F}^{-1} = P_{\mathcal{Y}}.$$

Sei nun \mathcal{X} ergodisch und $A \in \mathcal{I}(S)$. Dann ist $S^{-1}(\hat{F}^{-1}A) = \hat{F}^{-1}(S^{-1}A) = \hat{F}^{-1}(A)$, so dass

$$P_{\mathcal{Y}}(A) = P_{\mathcal{X}}(\hat{F}^{-1}(A)) \in \{0, 1\}$$

und damit auch \mathcal{Y} ergodisch ist. □

14.11 Beispiel (Moving average Prozesse) Seien $X_n \in \mathcal{L}_P^1$ u.i.v. ($n \in \mathbb{Z}$) und seien $a_n \in \mathbb{R}$ mit $\sum_{n \in \mathbb{Z}} |a_n| < \infty$. Betrachte $F(\underline{x}) := \sum_{k \in \mathbb{Z}} a_k x_k$. Man zeigt leicht, dass die Summe $F(\mathcal{X})$ fast sicher konvergiert und dass $F(\mathcal{X}) \in \mathcal{L}_P^1$ (Übung). Ist $a_n = 0$ außer für endlich viele n , so ist das trivial! Setze wieder $Y_n := F \circ S^n \circ \mathcal{X} = \sum_{k \in \mathbb{Z}} a_k X_{k+n}$. Dann ist $\mathcal{Y} = (Y_n)_{n \in \mathbb{Z}}$ ein stationärer ergodischer Prozess.

14.12 Beispiel (Irrationale Rotation) Sei $\Omega = \{z \in \mathbb{C} : |z| = 1\}$, \mathcal{B} die Borel- σ -Algebra auf Ω . Sei $\alpha \in \mathbb{R}$. Betrachte die Rotation $R_\alpha : \Omega \rightarrow \Omega$, $R_\alpha(z) = e^{2\pi i \alpha} z$, um den Winkel $2\pi\alpha$ auf Ω . Da \mathcal{B} von den Kreissegmenten (= Intervallen) in Ω erzeugt wird und da sich die Länge von Kreissegmenten unter Rotation nicht ändert, gilt für das normierte 1-dimensionale Lebesgue-Maß m auf Ω : $m \circ R_\alpha^{-1} = m$. Ist $f : \Omega \rightarrow \mathbb{R}$ messbar, so ist $(f \circ R_\alpha^n)_{n \in \mathbb{Z}}$ ein stationärer Prozess auf (Ω, \mathcal{B}, m) , siehe Satz 14.9.

Ist α irrational, so ist dieser Prozess sogar ergodisch. Nach Satz 14.9 reicht es dazu zu zeigen, dass $m(A) \in \{0, 1\}$ für jedes $A \in \mathcal{I}(R_\alpha)$. Ein eleganter Beweis dafür benutzt Fourier-Reihen: Es gibt $a_n \in \mathbb{C}$ mit $\sum_{n \in \mathbb{Z}} |a_n|^2 < \infty$, so dass $1_A(z) = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n z}$ (mit Konvergenz im \mathcal{L}_m^2 -Sinn). Ist nun $A \in \mathcal{I}(R_\alpha)$, also $1_A = 1_A \circ R_\alpha$, so gilt

$$\sum_{n \in \mathbb{Z}} a_n e^{2\pi i n z} = 1_A(z) = 1_A(e^{2\pi i \alpha} z) = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n \alpha} e^{2\pi i n z},$$

woraus wegen der Eindeutigkeit der Fourier-Entwicklung folgt, dass $a_n = a_n e^{2\pi i n \alpha}$ für alle $n \in \mathbb{Z}$, also $a_n = 0$ oder $e^{2\pi i n \alpha} = 1$ für alle $n \in \mathbb{Z}$ ist. Da $\alpha \notin \mathbb{Q}$, kann die zweite Möglichkeit nur für $n = 0$ auftreten, so dass $a_n = 0$ für alle $n \in \mathbb{Z} \setminus \{0\}$. Daher ist $1_A = a_0$ P -f.s., also $1_A = 1$ P -f.s. oder $1_A = 0$ P -f.s.

Kapitel 15

Der Ergodensatz - ein starkes Gesetz der großen Zahl für stationäre Prozesse

15.1 Satz (Birkhoffscher Ergodensatz für stationäre Prozesse)

Sei $\mathcal{X} = (X_n)_{n \in \mathbb{I}}$ ein stationärer Prozess auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit $X_n \in \mathcal{L}^1_P$. Dann existiert

$$\bar{X} := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} X_k$$

P -f.s. und in L^1_P , und für die Zufallsvariable \bar{X} gilt:

- $E|\bar{X}| \leq E|X_0| < \infty$
- Ist \mathcal{X} ergodisch, so ist $\bar{X} = EX_0$ konstant P -f.s. Insbesondere gilt das für unabhängige identisch verteilte X_n .
- Ist \mathcal{X} ein kanonischer Prozess auf $(\mathbb{R}^{\mathbb{I}}, \mathcal{B}^{\mathbb{I}}, P)$, so ist $\bar{X} = E[X_0 | \mathcal{I}(S)]$.

15.2 Korollar (Starkes Gesetz der großen Zahl für u.i.v. Prozesse) Sind $X_1, X_2, \dots \in \mathcal{L}^1_P$ unabhängig identisch verteilt, so ist $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = EX_1$ P -f.s. und in L^1_P .

Satz 15.1 folgt aus

15.3 Satz (Birkhoffscher Ergodensatz für maßerhaltende Transformationen)

Sei (X, \mathcal{F}, μ) ein Wahrscheinlichkeitsraum, $T : X \rightarrow X$ messbar und maßerhaltend (d.h. $\mu \circ T^{-1} = \mu$). Sei $f \in \mathcal{L}^1_\mu$. Dann existiert

$$\bar{f}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$$

μ -f.s. und in \mathcal{L}^1_μ , und \bar{f} ist eine Version von $E[f | \mathcal{I}(T)]$.

Beweis von Satz 15.1 aus Satz 15.3: Es reicht, statt \mathcal{X} den kanonischen Prozess Π auf $(X, \mathcal{F}, \mu) := (\mathbb{R}^{\mathbb{I}}, \mathcal{B}^{\mathbb{I}}, P_{\mathcal{X}})$ zu betrachten. Mit $T = S$ und $f = \pi_0 = X_0$ ist dann $X_k = \pi_k = \pi_0 \circ S^k = f \circ S^k$, also

$$\bar{X} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} X_k = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ S^k = \bar{f} = E[X_0 | \mathcal{I}(S)] \quad P\text{-f.s. und in } \mathcal{L}_{\mu}^1.$$

- a) $E|\bar{X}| = E|E[X_0 | \mathcal{I}(S)]| \leq E[E[|X_0| | \mathcal{I}(S)]] = E|X_0|$
 b) Ist \mathcal{X} ergodisch, so ist $\mathcal{I}(S)$ trivial, also $\bar{X} = E[X_0 | \mathcal{I}(S)] = EX_0$. \square

Beweis von Satz 15.3: Sei \bar{f} eine Version von $E[f | \mathcal{I}(T)]$, insbesondere also $\mathcal{I}(T)$ -messbar. Dann ist $\bar{f} \circ T = \bar{f}$, denn für alle $\alpha \in \mathbb{R}$ gilt:

$$\{\bar{f} \circ T = \alpha\} = T^{-1}\{\bar{f} = \alpha\} = \{\bar{f} = \alpha\}.$$

Sei nun $S_n f := \sum_{k=0}^{n-1} f \circ T^k$ ($n \geq 0$). Wir werden zeigen, dass

$$F := \limsup_{n \rightarrow \infty} \frac{1}{n} S_n f \leq \bar{f} \quad \mu\text{-f.s.} \quad (*)$$

Durch Anwendung des gleichen Arguments auf die Funktion $-f$ folgt dann auch

$$\limsup_{n \rightarrow \infty} \frac{1}{n} S_n(-f) \leq -\bar{f} \quad \mu\text{-f.s.}$$

Beide Ungleichungen zusammen ergeben die Behauptung $\lim_{n \rightarrow \infty} \frac{1}{n} S_n f = \bar{f}$ μ -f.s.

Beachte zunächst, dass $F \circ T = F$:

$$F \circ T = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f \circ T^k = \limsup_{n \rightarrow \infty} \frac{n+1}{n} \left(\frac{1}{n+1} \sum_{k=0}^n f \circ T^k - \frac{1}{n+1} f \right) = F.$$

Insbesondere ist $\{F > \bar{f} + \epsilon\} \in \mathcal{I}(T)$ für jedes $\epsilon > 0$. Setze $g_{\epsilon} := (f - \bar{f} - \epsilon)1_{\{F > \bar{f} + \epsilon\}}$. Dann ist

$$G_{\epsilon} := \limsup_{n \rightarrow \infty} \frac{1}{n} S_n g_{\epsilon} = (F - \bar{f} - \epsilon) \cdot 1_{\{F > \bar{f} + \epsilon\}}$$

T -invariant. Also ist $\{G_{\epsilon} > 0\}$ eine T -invariante Menge,

Wir werden zeigen, dass

$$\mu\{G_{\epsilon} > 0\} = 0 \quad \text{für jedes } \epsilon > 0. \quad (**)$$

Wegen $\{F > \bar{f} + \epsilon\} = \{G_{\epsilon} > 0\}$ folgt daraus $\mu\{F > \bar{f} + \frac{1}{k}\} = 0$ für jedes $k \in \mathbb{N}$, so dass

$$\mu\{F > \bar{f}\} = \mu\left(\bigcup_{k=1}^{\infty} \{F > \bar{f} + \frac{1}{k}\}\right) \leq \sum_{k=1}^{\infty} \mu\{F > \bar{f} + \frac{1}{k}\} = 0,$$

also $F \leq \bar{f}$ μ -f.s.

Nun zum Beweis von (**): Wir bemerken für später, dass

$$g_{\epsilon} = (f - \bar{f} - \epsilon)1_{\{F > \bar{f} + \epsilon\}} = (f - \bar{f} - \epsilon)1_{\{F > \bar{f} + \epsilon\}} \cdot 1_{\{G_{\epsilon} > 0\}} = g_{\epsilon}1_{\{G_{\epsilon} > 0\}}, \quad (***)$$

da ja $\{F > \bar{f} + \epsilon\} = \{G_\epsilon > 0\}$. Schreibe nun zur Vereinfachung der Notation $g := g_\epsilon$. Da $|g| \leq |f| + |\bar{f}| + \epsilon$ und da $f, \bar{f} = E[f|\mathcal{I}(T)] \in \mathcal{L}_\mu^1$, ist $g \in \mathcal{L}_\mu^1$. Es folgt, dass $\int |S_k g| d\mu \leq \sum_{i=0}^{k-1} |g| \circ T^i d\mu = \sum_{i=0}^{k-1} |g| d\mu < \infty$, also auch $S_k g \in \mathcal{L}_\mu^1$ für alle $k \in \mathbb{N}$.

Für $n > 0$ sei $M_n := \max\{0, S_1 g, \dots, S_n g\}$. Dann ist $0 \leq M_n \leq |S_1 g| + \dots + |S_n g|$, insbesondere $M_n \in \mathcal{L}_\mu^1$. Außerdem ist $0 \leq M_1 \leq M_2 \leq M_3 \leq \dots$, und da $S_k g(Tx) = S_{k+1} g(x) - g(x)$ für alle $k \geq 0$, folgt für jedes $x \in \{M_n > 0\}$ Hopfs Maximalungleichung

$$\begin{aligned} M_n(x) - M_n(Tx) &= \max\{0, S_1 g(x), \dots, S_n g(x)\} - \max\{0, S_1 g(Tx), \dots, S_n g(Tx)\} \\ &= \max\{S_1 g(x), \dots, S_n g(x)\} - \max\{S_1 g(x), S_2 g(x), \dots, S_{n+1} g(x)\} + g(x) \\ &\leq g(x). \end{aligned}$$

Da $M_n \geq 0$ für alle n , folgt daraus (beachte, dass $M_n \in \mathcal{L}_\mu^1$!)

$$\int_{\{M_n > 0\}} g d\mu \geq \int_{\{M_n > 0\}} M_n d\mu - \int_{\{M_n > 0\}} M_n \circ T d\mu \geq \int M_n d\mu - \int M_n \circ T d\mu = 0.$$

Nun gilt aber $\{M_n > 0\} \nearrow \bigcup_{n=1}^\infty \{S_n g_\epsilon > 0\} \supseteq \{G_\epsilon > 0\}$, so dass wegen (***)

$$g_\epsilon 1_{\{M_n > 0\}} = g_\epsilon 1_{\{G_\epsilon > 0\}} 1_{\{M_n > 0\}} \nearrow g_\epsilon 1_{\{G_\epsilon > 0\}} = g_\epsilon \quad \text{mit } n \rightarrow \infty.$$

Aus Korollar 6.10 zum Satz von der majorisierten Konvergenz folgt dann

$$\int_{\{G_\epsilon > 0\}} g_\epsilon d\mu = \lim_{n \rightarrow \infty} \int_{\{M_n > 0\}} g_\epsilon d\mu \geq 0.$$

Da $\{F > \bar{f} + \epsilon\} = \{G_\epsilon > 0\}$, ist $g_\epsilon = (f - \bar{f} - \epsilon) 1_{\{G_\epsilon > 0\}}$, so dass

$$0 \leq \int_{\{G_\epsilon > 0\}} g_\epsilon d\mu = \int_{\{G_\epsilon > 0\}} f d\mu - \int_{\{G_\epsilon > 0\}} \bar{f} d\mu - \epsilon \cdot \mu\{G_\epsilon > 0\} = -\epsilon \cdot \mu\{G_\epsilon > 0\},$$

wobei wir benutzt haben, dass $\bar{f} = E[f|\mathcal{I}(T)]$ und dass $\{G_\epsilon > 0\} \in \mathcal{I}(T)$. Also ist $\mu\{G_\epsilon > 0\} = 0$ für jedes $\epsilon > 0$.

Zu zeigen bleibt noch die \mathcal{L}_μ^1 -Konvergenz von $(n^{-1} S_n f)_{n \in \mathbb{N}}$ gegen \bar{f} . Wegen Satz 6.18 muss dazu nur die gleichgradige Integrierbarkeit der Familie $(n^{-1} S_n f : n \in \mathbb{N})$ gezeigt werden, und die folgt wegen Lemma 6.17b aus der gleichgradigen Integrierbarkeit der Familie $(f \circ T^k : k \in \mathbb{N})$. Das wiederum zeigt man so: Sei $\epsilon > 0$. Nach Lemma 6.17d und Satz 6.16iii gibt es ein $\delta > 0$, so dass $\int_A f d\mu < \epsilon$ für alle $A \in \mathcal{A}$ mit $\mu(A) < \delta$. Für solche A und jedes $k \in \mathbb{N}$ ist dann auch $\mu(T^{-k} A) = \mu(A) < \delta$, so dass $\int_A f \circ T^k d\mu = \int_{T^{-k} A} f d\mu < \epsilon$, woraus, ebenfalls nach Satz 6.16iii die gleichgradige Integrierbarkeit der Familie $(f \circ T^k : k \in \mathbb{N})$ folgt. \square

Kapitel 16

Verteilungskonvergenz

In diesem Kapitel führen wir ein Konvergenzkonzept ein, das es uns gestattet, das asymptotische Verhalten von normierten Summen $n^{-1/2} \sum_{i=1}^n X_i$ für unabhängig identisch verteilte X_i zu untersuchen.

16.1 Definition (Schwache Konvergenz von Verteilungsfunktionen) Seien F, F_n Verteilungsfunktionen auf \mathbb{R} , siehe Definition 4.1. Wir sagen (F_n) konvergiert schwach gegen F (symbolisch: $F_n \implies F$), falls $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ für alle Stetigkeitsstellen x von F . (Beachte, dass F als monotone Funktion höchstens abzählbar viele Unstetigkeitsstellen hat.)

16.2 Beispiel (Satz von de Moivre-Laplace) Seien X_1, X_2, \dots unabhängig identisch verteilte Bernoulli-Zufallsvariablen zum Parameter p ,

$$F_n(x) := P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - p) \leq x\right), \quad F(x) := \frac{1}{\sqrt{2\pi p(1-p)}} \int_{-\infty}^x e^{-\frac{t^2}{2p(1-p)}} dt.$$

Dann beschreibt " $F_n \implies F$ " gerade die Aussage des Satzes von de Moivre-Laplace.

Auch das schwache Gesetz der großen Zahl lässt sich so formulieren. Setze dazu

$$\tilde{F}_n(x) := P\left(\frac{1}{n} \sum_{i=1}^n X_i \leq x\right), \quad \tilde{F}(x) := \begin{cases} 1 & \text{falls } x \geq p \\ 0 & \text{falls } x < p. \end{cases}$$

Dann besagt das schwache GGZ gerade $\lim_{n \rightarrow \infty} \tilde{F}_n(x) = \tilde{F}(x)$ für alle $x \neq p$, d.h. $F_n \implies F$.

Da Verteilungsfunktionen eineindeutig Wahrscheinlichkeitsmaßen auf \mathbb{R} entsprechen, soll versucht werden, die schwache Konvergenz durch Maße zu beschreiben. Dazu verlassen wir den messbaren Raum $(\mathbb{R}, \mathcal{B})$ und betrachten eine etwas allgemeinere Situation:

Generalvoraussetzung: (E, d) ist ein metrischer Raum mit Borel- σ -Algebra \mathcal{E} .

16.3 Definition (Schwache Konvergenz von Wahrscheinlichkeitsmaßen auf (E, \mathcal{E}))

Eine Folge μ_1, μ_2, \dots von Wahrscheinlichkeitsmaßen auf (E, \mathcal{E}) konvergiert schwach gegen das Wahrscheinlichkeitsmaß μ (symbolisch $\mu_n \implies \mu$), falls

$$\lim_{n \rightarrow \infty} \int_E f d\mu_n = \int_E f d\mu \quad \forall f \in C_b(E; \mathbb{R}),$$

d.h. für alle stetigen beschränkten $f : E \rightarrow \mathbb{R}$.

16.4 Bemerkung Schwache Limiten sind eindeutig bestimmt, denn gilt sowohl $\mu_n \implies \mu$ als auch $\mu_n \implies \nu$, so ist $\int_E f d\mu = \int_E f d\nu$ für alle $f \in C_b(E; \mathbb{R})$. Um $\mu = \nu$ zu zeigen reicht es, $\mu(A) = \nu(A)$ für alle abgeschlossenen $A \subseteq E$ zu zeigen (\cap -stabiler Erzeuger der Borel- σ -Algebra). Für $A \subseteq E$ sei dazu

$$f_{A,\epsilon} : E \rightarrow \mathbb{R}, \quad f_{A,\epsilon}(x) = \begin{cases} 1 - \frac{d(x,A)}{\epsilon} & \text{falls } d(x,A) \leq \epsilon \\ 0 & \text{falls } d(x,A) > \epsilon. \end{cases}$$

Es ist $1_{\bar{A}} \leq f_{A,\epsilon} \leq 1$, $\lim_{\epsilon \rightarrow 0} f_{A,\epsilon}(x) = 1_{\bar{A}}(x)$ für alle x und $f_{A,\epsilon}$ ist Lipschitz-stetig (mit Lipschitz-Konstante ϵ^{-1}). Also folgt für abgeschlossenes A aus dem Satz von der majorisierten Konvergenz

$$\mu(A) = \lim_{k \rightarrow \infty} \int f_{A, \frac{1}{k}} d\mu = \lim_{k \rightarrow \infty} \int f_{A, \frac{1}{k}} d\nu = \nu(A).$$

Der folgende Satz wird auch als *Portemanteau-Theorem* bezeichnet.

16.5 Satz (Charakterisierungen der schwachen Konvergenz von W'maßen)

Seien μ_n, μ Wahrscheinlichkeitsmaße auf (E, \mathcal{E}) . Folgende Aussagen sind äquivalent:

- (i) $\mu_n \implies \mu$.
- (ii) $\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A)$ für alle $A \in \mathcal{E}$ mit $\mu(\partial A) = 0$.
- (iii) $\lim_{n \rightarrow \infty} \int_E f d\mu_n = \int_E f d\mu$ für alle beschränkten Lipschitz-stetigen $f : E \rightarrow \mathbb{R}$.

Dabei ist $\partial A = \bar{A} \setminus \overset{\circ}{A}$ der Rand von A .

Beweis: (i) \implies (iii) trivial.

(iii) \implies (ii) Sei $A \in \mathcal{E}$ mit $\mu(\partial A) = 0$. Mit $f_{A,\epsilon}$ wie in der vorangehenden Bemerkung gilt:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mu_n(A) &\leq \liminf_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \int f_{A, \frac{1}{k}} d\mu_n = \liminf_{k \rightarrow \infty} \int f_{A, \frac{1}{k}} d\mu = \mu(\bar{A}) \\ &\leq \mu(A) + \mu(\partial A) = \mu(A). \end{aligned}$$

(Dabei wurde für die erste Ungleichung nur die Monotonie des Integrals benutzt.) Angewandt auf A^c liefert diese Ungleichung

$$\liminf_{n \rightarrow \infty} \mu_n(A) = 1 - \limsup_{n \rightarrow \infty} \mu_n(A^c) \geq 1 - \mu(A^c) = \mu(A),$$

und zusammen folgt daraus (ii).

(ii) \implies (i) Sei $f \in C_b(E, \mathbb{R})$. Sei $T_k := \{t \in \mathbb{R} : \mu(f^{-1}\{t\}) \geq 1/k\}$. Dann ist

$$\#T_k \leq \sum_{t \in T_k} k \cdot \mu(f^{-1}\{t\}) \leq k,$$

so dass $T := \{t \in \mathbb{R} : \mu(f^{-1}\{t\}) > 0\} = \bigcup_{k \in \mathbb{N}} T_k$ höchstens abzählbar ist. Zu $\epsilon > 0$ gibt es daher $t_0 \leq -\|f\|_\infty < t_1 < \dots < t_{k-1} < \|f\|_\infty \leq t_k$ derart, dass

$$t_i \notin T \quad \text{und} \quad |t_{i+1} - t_i| \leq \epsilon \quad \text{für alle } i.$$

Da $t_0 \leq f(x) \leq t_k$ für alle $x \in E$, ist

$$E = \bigcup_{i=1}^k \underbrace{\{x : t_{i-1} \leq f(x) < t_i\}}_{=: E_i},$$

wobei $\mu(\partial E_i) \leq \mu(f^{-1}\{t_{i-1}\}) + \mu(f^{-1}\{t_i\}) = 0$. Daher ist

$$\begin{aligned} \int f d\mu_n &\leq \sum_{i=1}^k \mu_n\{t_{i-1} \leq f < t_i\} \cdot t_i \\ &\leq \sum_{i=1}^k \mu_n(E_i) \cdot t_{i-1} + \sum_{i=1}^k \mu_n(E_i) \cdot \epsilon \\ &\leq \sum_{i=1}^k \mu(E_i) \cdot t_{i-1} + \sum_{i=1}^k (\mu_n(E_i) - \mu(E_i)) \cdot t_{i-1} + \epsilon \\ &\leq \int f d\mu + \epsilon + \sum_{i=1}^k \underbrace{(\mu_n(E_i) - \mu(E_i))}_{\rightarrow 0 \text{ mit } n \rightarrow \infty \text{ für jedes } i} \cdot t_{i-1}, \end{aligned}$$

also

$$\limsup_{n \rightarrow \infty} \int f d\mu_n \leq \int f d\mu$$

und durch Betrachtung von $(-f)$ erhält man

$$\liminf_{n \rightarrow \infty} \int f d\mu_n \geq \int f d\mu$$

□

16.6 Beispiele Sei $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$.

a) Sei $\mu_n = h_n \lambda$, $\mu = h \lambda$ und gelte $\lim_{n \rightarrow \infty} \|h_n - h\|_1 = 0$. Dann gilt für jedes $A \in \mathcal{B}$:

$$|\mu_n(A) - \mu(A)| \leq \int_A |h_n - h| d\lambda \leq \|h_n - h\|_1 \rightarrow 0 \text{ mit } n \rightarrow \infty.$$

b) Seien $\mu_n = \delta_{c_n}$, $\mu = \delta_c$, wo $c, c_n \in \mathbb{R}$ und $\lim_{n \rightarrow \infty} c_n = c$. Dann gilt für stetiges f :

$$|\int f d\mu_n - \int f d\mu| = |f(c_n) - f(c)| \rightarrow 0 \text{ mit } n \rightarrow \infty.$$

c) Seien $\mu_n = \frac{1}{n} \sum_{k=1}^n \delta_{k/n}$, $\mu = 1_{[0,1]} \lambda$. Dann gilt für Lipschitz-stetiges f :

$$\left| \int f d\mu_n - \int f d\mu \right| \leq \sum_{k=1}^n \int_{(k-1)/n}^{k/n} |f(k/n) - f(x)| dx \leq \frac{1}{n} \text{Lip}(f) \rightarrow 0 \text{ mit } n \rightarrow \infty.$$

16.7 Definition (Verteilungskonvergenz) Seien X, X_n Zufallsvariablen. Die Folge $(X_n)_{n>0}$ konvergiert in Verteilung (oder auch schwach) gegen X , kurz $X_n \Longrightarrow X$, falls $P_{X_n} \Longrightarrow P_X$.

16.8 Satz (Charakterisierungen der Verteilungskonvergenz)

Seien X, X_n Zufallsvariablen mit Verteilungsfunktionen F bzw. F_n . Äquivalent sind

(i) $X_n \implies X$

(ii) $F_n \implies F$

(iii) $\lim_{n \rightarrow \infty} E[f(X_n)] = E[f(X)]$ für alle $f \in C_b(\mathbb{R}; \mathbb{R})$.

Beweis: (i) \implies (ii) Sei F bei x stetig, d.h. $P_X\{x\} = 0$. Aus $X_n \implies X$ und Satz 16.5 folgt

$$F_n(x) = P_{X_n}(-\infty, x] \rightarrow P_X(-\infty, x] = F(x).$$

(ii) \implies (i) Zu zeigen ist $\lim_{n \rightarrow \infty} \int f dP_{X_n} = \int f dP$ für alle Lipschitz-stetigen $f : \mathbb{R} \rightarrow \mathbb{R}$, siehe Satz 16.5. Sei $\epsilon > 0$. Wir können $t_0 < t_1 < \dots < t_k$ in \mathbb{R} so wählen, dass

$$F(t_0) = P(-\infty, t_0] < \epsilon, \quad 1 - F(t_k) = P(t_k, +\infty) < \epsilon,$$

$$t_i - t_{i-1} < \epsilon \quad \forall i = 1, \dots, k, \text{ und } F \text{ an allen Stellen } t_0, \dots, t_k \text{ stetig ist.}$$

Insbesondere ist $\lim_{n \rightarrow \infty} F_n(t_i) = F(t_i) \quad \forall i = 0, \dots, k$, so dass

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 \quad \forall i = 0, \dots, k : |F_n(t_i) - F(t_i)| < \frac{\epsilon}{k+1}.$$

Setze

$$\bar{f}_i := \max_{t_{i-1} \leq x \leq t_i} f(x), \quad \underline{f}_i := \min_{t_{i-1} \leq x \leq t_i} f(x).$$

Dann ist für $n \geq n_0$:

$$\begin{aligned} & \int f dP_{X_n} \\ & \leq \|f\|_\infty \cdot (F_n(t_0) + 1 - F_n(t_k)) + \sum_{i=1}^k \bar{f}_i \cdot (F_n(t_i) - F_n(t_{i-1})) \\ & \leq \|f\|_\infty \cdot \left(F(t_0) + 1 - F(t_k) + 2 \frac{\epsilon}{k+1} \right) + \sum_{i=1}^k \bar{f}_i \cdot (F(t_i) - F(t_{i-1})) + \sum_{i=1}^k 2 \frac{\epsilon}{k+1} \|f\|_\infty \\ & \leq \|f\|_\infty \cdot 5\epsilon + \sum_{i=1}^k (\underline{f}_i + \epsilon \cdot \text{Lip}(f)) (F(t_i) - F(t_{i-1})) \\ & \leq 5\epsilon \|f\|_\infty + \int_{t_0}^{t_k} f dP_X + \epsilon \text{Lip}(f) \\ & \leq \int f dP_X + \epsilon \cdot (5\|f\|_\infty + 2\|f\|_\infty + \text{Lip}(f)) \end{aligned}$$

und es folgt $\limsup_{n \rightarrow \infty} \int f dP_{X_n} \leq \int f dP_X$. Durch Betrachtung von $(-f)$ an Stelle von f folgt $\liminf_{n \rightarrow \infty} \int f dP_{X_n} \geq \int f dP_X$ und damit die Behauptung.

(i) \Leftrightarrow (iii) Ist nur eine Frage der Schreibweise: $E[f(X)] = \int f dP_X$. □

16.9 Übung Seien X, X_n Zufallsvariablen, $X_n \implies X$. Dann gilt:

- a) $E|X| \leq \liminf_{n \rightarrow \infty} E|X_n|$.
- b) Sind die X_n gleichgradig integrierbar, so ist $EX = \lim_{n \rightarrow \infty} EX_n$.
- c) Aussage b) bleibt ohne die Annahme der gleichgradigen Integrierbarkeit nicht richtig:

Zum Begriff der schwachen Konvergenz gehört auch ein Kompaktheitskonzept. Das werden wir im Rest dieses Kapitels untersuchen.

16.10 Definition (Folgenkompaktheit von Wahrscheinlichkeitsmaßen) Eine Familie \mathcal{M} von Wahrscheinlichkeitsmaßen auf (E, \mathcal{E}) heißt relativ folgenkompakt, falls jede Folge in \mathcal{M} eine schwach konvergente Teilfolge hat.

16.11 Definition (Straffheit) Eine Familie \mathcal{M} von Wahrscheinlichkeitsmaßen auf (E, \mathcal{E}) heißt straff, falls für jedes $\epsilon > 0$ eine kompakte Menge $K \subseteq E$ existiert, so dass $\sup_{\mu \in \mathcal{M}} \mu(E \setminus K) \leq \epsilon$.

Beide Definitionen übertragen sich auf Verteilungsfunktionen, wenn wir sie mit den zugehörigen Wahrscheinlichkeitsmaßen identifizieren.

16.12 Satz (Prohorov: relativ folgenkompakt \iff straff)

Sei (E, d) ein vollständiger, separabler metrischer Raum. Eine Familie \mathcal{M} von Wahrscheinlichkeitsmaßen auf (E, \mathcal{E}) ist genau dann relativ folgenkompakt, wenn sie straff ist.

Beweis: Wir beweisen diesen Satz nur im Fall $E = \mathbb{R}$.

“ \implies ” Angenommen, \mathcal{M} ist nicht straff. Dann gibt es ein $\epsilon > 0$ derart, dass

$$\forall n \in \mathbb{N} \exists \mu_n \in \mathcal{M} : \mu_n(M_n) > \epsilon, \quad \text{wo } M_n := \mathbb{R} \setminus [-n, n].$$

Wähle eine Teilfolge $\mu_{n_j} \implies \mu$ mit $j \rightarrow \infty$. Zu jedem n gibt es ein $\eta_n \in [0, 1)$ mit $\mu\{-(n + \eta_n), n + \eta_n\} = 0$. Also folgt für alle $k \in \mathbb{N}$ aus Satz 16.5

$$\begin{aligned} \mu(M_k) &\geq \mu(\mathbb{R} \setminus [-(k + \eta_k), k + \eta_k]) = \lim_{j \rightarrow \infty} \mu_{n_j}(\mathbb{R} \setminus [-(k + \eta_k), k + \eta_k]) \\ &\geq \limsup_{j \rightarrow \infty} \mu_{n_j}(M_{n_j}) \geq \epsilon, \end{aligned}$$

und wir erhalten den Widerspruch $0 < \epsilon \leq \liminf_{k \rightarrow \infty} \mu(M_k) = 0$.

“ \impliedby ” Zum Beweis der Gegenrichtung benötigen wir den folgenden Satz:

16.13 Satz (Helly)

Ist $(F_n)_{n>0}$ eine Folge von Verteilungsfunktionen, so gibt es eine Teilfolge $(F_{n_i})_{i>0}$ und eine nichtfallende, rechtsseitig stetige Funktion F derart, dass $\lim_{i \rightarrow \infty} F_{n_i}(x) = F(x)$ an allen Stetigkeitsstellen x von F .

Es folgt sofort, dass $0 \leq F \leq 1$ und dass $F \nearrow$, aber es ist nicht notwendig

$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \text{und} \quad \lim_{x \rightarrow +\infty} F(x) = 1. \tag{*}$$

Wir kehren zum Beweis von Satz 16.12 zurück. Sei $(\mu_n)_{n>0}$ eine Folge in \mathcal{M} , $(F_n)_{n>0}$ die Folge der zugehörigen Verteilungsfunktionen. Seien F_{n_i} und F wie im Satz von Helly. Ist F eine Verteilungsfunktion, so folgt $F_{n_i} \implies F$, also $\mu_{n_i} \implies \mu$ für das durch F bestimmte Wahrscheinlichkeitsmaß μ auf $E = \mathbb{R}$. Zu zeigen ist also, dass F unter der Straffheitsannahme tatsächlich eine Verteilungsfunktion ist, d.h. (*).

Sei $\epsilon > 0$. Da \mathcal{M} straff ist, folgt

$$\begin{aligned} & \exists K \text{ kompakt } \forall i > 0 : \mu_{n_i}(\mathbb{R} \setminus K) \leq \epsilon \\ \implies & \exists M \in \mathbb{N} \forall i > 0 : \mu_{n_i}(\mathbb{R} \setminus [-M, M]) \leq \epsilon \\ \implies & \exists M \in \mathbb{N} \forall i > 0 : \forall x \leq -M : F_{n_i}(x) \leq \epsilon \text{ und } \forall y \geq M : 1 - F_{n_i}(y) \leq \epsilon \end{aligned}$$

Da $\epsilon > 0$ beliebig ist, folgt (*). □

Beweis von Satz 16.13: Da $0 \leq F_n(x) \leq 1$ für alle $n > 0$ und $x \in \mathbb{R}$, gibt es zu jedem x und jeder Teilfolge von $(F_n(x))$ eine Teilfolge dieser Teilfolge, die konvergiert. Durch einen Diagonalschluss erhält man:

Es gibt eine Teilfolge $(F_{n_i})_{i>0}$ derart, dass $G(q) := \lim_{i \rightarrow \infty} F_{n_i}(q)$ existiert $\forall q \in \mathbb{Q}$.

Setze $F(x) := \inf\{G(q) : q > x\}$. Dann gilt:

- i) F_{n_i} nichtfallend $\implies G$ nichtfallend $\implies F$ nichtfallend.
- ii) F ist rechtsseitig stetig, denn: Zu $x \in \mathbb{R}$ und $\epsilon > 0$ existiert ein $q > x$ in \mathbb{Q} mit $G(q) < F(x) + \epsilon$. Ist dann $x \leq y < q$, so folgt $F(y) \leq G(q) < F(x) + \epsilon$.
- iii) $\lim_{i \rightarrow \infty} F_{n_i}(x) = F(x)$ an allen Stetigkeitsstellen x von F . Denn: Sei F stetig bei x , $\epsilon > 0$. Wähle $y < x$ derart, dass $F(x) < F(y) + \epsilon$ und wähle $q, r \in \mathbb{Q}$ mit $y < q < x < r$ und $G(r) < F(x) + \epsilon$. Dann ist

$$F(x) - \epsilon \leq F(y) \leq G(q) \leq G(r) < F(x) + \epsilon .$$

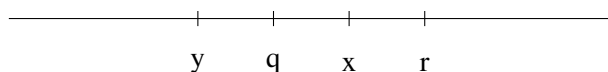
Da

$$G(q) = \lim_{i \rightarrow \infty} F_{n_i}(q) \leq \liminf_{i \rightarrow \infty} F_{n_i}(x) \leq \limsup_{i \rightarrow \infty} F_{n_i}(x) \leq \lim_{i \rightarrow \infty} F_{n_i}(r) = G(r) ,$$

folgt

$$F(x) - \epsilon \leq \liminf_{i \rightarrow \infty} F_{n_i}(x) \leq \limsup_{i \rightarrow \infty} F_{n_i}(x) \leq F(x) + \epsilon ,$$

also die behauptete Stetigkeit von F bei x .



$$F(x) - \epsilon < F(y) \leq G(y), \quad G(r) < F(x) + \epsilon$$

□

Kapitel 17

Charakteristische Funktionen und der Zentrale Grenzwertsatz

Ein wichtiges Hilfsmittel zur Untersuchung der Verteilungskonvergenz reellwertiger Zufallsvariablen sind die *charakteristischen Funktionen*. Die Grundidee ist, dass man jedem Wahrscheinlichkeitsmaß eine auf ganz \mathbb{R} definierte \mathbb{C} -wertige charakteristische Funktion derart eindeutig zuordnet, dass eine Folge von Wahrscheinlichkeitsmaßen genau dann schwach konvergiert, wenn die Folge ihrer charakteristischen Funktionen punktweise konvergiert.

17.1 Definition (Charakteristische Funktion) a) Sei μ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B})$. Die Fourier-Transformierte $\varphi_\mu : \mathbb{R} \rightarrow \mathbb{C}$ von μ ,

$$\varphi_\mu(t) := \int e^{itx} d\mu(x) \left[= \int \cos(tx) d\mu(x) + i \int \sin(tx) d\mu(x) \right]$$

wird auch als charakteristische Funktion von μ bezeichnet.

b) Ist X eine Zufallsvariable mit $P\{|X| < \infty\} = 1$, so heißt $\varphi_X := \varphi_{P_X}$ die charakteristische Funktion von X . Es ist

$$\varphi_X(t) = \int e^{itx} dP_X(x) = E[e^{itX}] .$$

17.2 Lemma Sei φ eine charakteristische Funktion.

- $\varphi(0) = 1$ und $|\varphi(t)| \leq 1$ für alle $t \in \mathbb{R}$.
- $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ ist gleichmäßig stetig.
- Ist $P\{|X| < \infty\} = 1$, so gilt $\varphi_{aX+b}(t) = e^{itb} \cdot \varphi_X(at)$ für alle $a, b \in \mathbb{R}$.

Beweis:

- trivial.
- Für alle $t, h \in \mathbb{R}$ ist

$$|\varphi(t+h) - \varphi(t)| = \left| \int e^{itx} (e^{ihx} - 1) d\mu(x) \right| \leq \int |e^{ihx} - 1| d\mu(x)$$

unabhängig von t , und da $|e^{ihx}| = 1$ für alle h und $\lim_{h \rightarrow 0} e^{ihx} = 1$, folgt aus dem Satz von der majorisierten Konvergenz $\lim_{h \rightarrow 0} \int |e^{ihx} - 1| d\mu(x) = 0$.

$$c) E[e^{it(aX+b)}] = e^{itb} \cdot E[e^{iatX}].$$

□

17.3 Lemma Sind X_1, \dots, X_n unabhängige Zufallsvariablen mit $P\{|X_i| < \infty\} = 1$, so ist

$$\varphi_{X_1+\dots+X_n} = \varphi_{X_1} \cdot \dots \cdot \varphi_{X_n}.$$

Beweis: Wir zeigen das Lemma für $n = 2$. Der Rest folgt durch Induktion.

$$E[e^{it(X_1+X_2)}] = E[e^{itX_1} \cdot e^{itX_2}] \stackrel{\text{Unabhängigkeit}}{=} E[e^{itX_1}] \cdot E[e^{itX_2}]$$

Dabei haben wir benutzt, dass ganz allgemein für unabhängige \mathbb{C} -wertige beschränkte Zufallsvariablen X, Y gilt $E[XY] = EX \cdot EY$, denn

$$\begin{aligned} E[XY] &= E[\Re X \cdot \Re Y - \Im X \cdot \Im Y + i(\Re X \cdot \Im Y + \Im X \cdot \Re Y)] \\ &= E[\Re X] \cdot E[\Re Y] + E[\Im X] \cdot E[\Im Y] + i(E[\Re X] \cdot E[\Im Y] + E[\Im X] \cdot E[\Re Y]) \\ &= (E[\Re X] + iE[\Im X]) \cdot (E[\Re Y] + iE[\Im Y]) \\ &= EX \cdot EY. \end{aligned}$$

□

Zum Rechnen mit charakteristischen Funktionen benötigen wir ein paar analytische Abschätzungen.

17.4 Lemma Für $n \geq 0$ und $x \in \mathbb{R}$ ist

$$e^{ix} = \sum_{k=0}^n \frac{(ix)^k}{k!} + \underbrace{\frac{i^{n+1}}{n!} \int_0^x (x-s)^n e^{is} ds}_{=: I_n} \quad \text{mit} \quad |I_n| \leq \min \left\{ \frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right\}, \quad (*)$$

insbesondere

$$|e^{ix} - 1| \leq \min\{|x|, 2\} \quad \text{und} \quad \left| e^{ix} - \left(1 + ix - \frac{x^2}{2}\right) \right| \leq \min \left\{ \frac{|x|^3}{6}, x^2 \right\}.$$

Beweis: Durch partielle Integration erhält man für $n \geq 1$:

$$\int_0^x (x-s)^{n-1} e^{is} ds = \frac{x^n}{n} + \frac{i}{n} \int_0^x (x-s)^n e^{is} ds,$$

also

$$I_{n-1} = \frac{i^n}{(n-1)!} \int_0^x (x-s)^{n-1} e^{is} ds = \frac{(ix)^n}{n!} + I_n. \quad (**)$$

Da außerdem $e^{ix} = 1 + (e^{ix} - 1) = 1 + i \int_0^x e^{is} ds = 1 + I_0$, folgt (*) aus (**) durch vollständige Induktion.

Außerdem ist

$$|I_n| \leq \frac{1}{n!} \int_0^{|x|} s^n ds = \frac{|x|^{n+1}}{(n+1)!},$$

woraus zusammen mit (**) schließlich folgt

$$|I_n| \leq |I_{n-1}| + \frac{|x|^n}{n!} \leq 2 \frac{|x|^n}{n!}.$$

□

17.5 Satz (Charakteristische Funktionen und Momente)

 Sei $\varphi(t)$ die charakteristische Funktion der Zufallsvariablen X . Ist $E[X^2] < \infty$, so ist

$$\varphi(t) = 1 + itE[X] - \frac{t^2}{2}E[X^2] + o(t^2) \quad \text{im Limes } t \rightarrow 0.$$

Beweis: Aus Lemma 17.4 folgt

$$\begin{aligned} \left| \varphi(t) - \left(1 + itE[X] - \frac{t^2}{2}E[X^2] \right) \right| &\leq E \left| e^{itX} - \left(1 + i(tX) - \frac{1}{2}(tX)^2 \right) \right| \\ &\leq E \left[\min \left\{ \frac{t^3}{6}|X|^3, t^2|X|^2 \right\} \right] \\ &\leq \frac{t^3}{6} E \left[|X|^3 \cdot 1_{\{|X| \leq t^{-\frac{1}{4}}\}} \right] + t^2 E \left[X^2 \cdot 1_{\{|X| \geq t^{-\frac{1}{4}}\}} \right] \\ &\leq \frac{1}{6}t^{\frac{9}{4}} + t^2 \cdot o(1) \quad \text{im Limes } t \rightarrow 0. \end{aligned}$$

(Für den letzten Schritt siehe Bemerkung 6.15.)

□

Nun wollen wir sehen, wie charakteristische Funktionen uns bei der Untersuchung von zentrierten und standardisierten Summen unabhängiger Zufallsvariablen helfen können.

17.6 Satz

 Seien $X_1, X_2, \dots \in \mathcal{L}_P^2$ unabhängige identisch verteilte Zufallsvariablen, $m := E[X_i]$ und $\sigma^2 := V(X_i)$. Sei

$$S_n^* := \frac{1}{\sqrt{n\sigma^2}} \sum_{k=1}^n (X_k - m)$$

 die zentrierte und standardisierte Partialsumme von X_1, \dots, X_n . Dann ist

$$\lim_{n \rightarrow \infty} \varphi_{S_n^*}(t) = e^{-\frac{t^2}{2}} = \varphi_{\mathcal{N}(0,1)}(t).$$

 Beweis: Sei $\varphi(t)$ die charakteristische Funktion der Zufallsvariablen $X_k - m$, also nach Satz 17.5

$$\varphi(t) = 1 - \frac{\sigma^2}{2}t^2 + t^2 \beta(t) \quad \text{wo } \lim_{t \rightarrow 0} \beta(t) = 0.$$

Wegen Lemma 17.3 und Lemma 17.2c) ist dann

$$\varphi_{S_n^*}(t) = \left(\varphi \left(\frac{t}{\sqrt{n\sigma^2}} \right) \right)^n = \left(1 - \frac{t^2}{2n} \right)^n - \left[\left(1 - \frac{t^2}{2n} \right)^n - \left(\varphi \left(\frac{t}{\sqrt{n\sigma^2}} \right) \right)^n \right]$$

 und im Limes $n \rightarrow \infty$ gilt für alle $t \in \mathbb{R}$

$$\left(1 - \frac{t^2}{2n} \right)^n \rightarrow e^{-\frac{t^2}{2}} \quad \text{und}$$

$$\left| \left(1 - \frac{t^2}{2n} \right)^n - \left(\varphi \left(\frac{t}{\sqrt{n\sigma^2}} \right) \right)^n \right| \leq n \cdot \left| 1 - \frac{t^2}{2n} - \varphi \left(\frac{t}{\sqrt{n\sigma^2}} \right) \right| \leq n \frac{t^2}{n\sigma^2} \beta \left(\frac{t}{\sqrt{n\sigma^2}} \right) \rightarrow 0,$$

da $|u^n - v^n| \leq |u - v| \cdot n \cdot \max\{|u|, |v|\}^{n-1}$ für beliebige $u, v \in \mathbb{C}$.

Es bleibt zu zeigen, dass $e^{-t^2/2} = \varphi_{\mathcal{N}(0,1)}(t)$. Wir zeigten dazu, dass $\varphi_{\mathcal{N}(0,1)}(t)$ und $e^{-t^2/2}$ Differentialgleichung $\varphi'(t) = -t \varphi(t)$ lösen mit gemeinsamem Anfangswert $\varphi(0) = 1$: $(e^{-t^2/2})' = -te^{-t^2/2}$ und

$$\begin{aligned} \sqrt{2\pi} \cdot \frac{d}{dt} \varphi_{\mathcal{N}(0,1)}(t) &= \frac{d}{dt} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx \\ \text{Satz 6.12:} \quad &= \int_{-\infty}^{\infty} ix e^{itx} e^{-\frac{x^2}{2}} dx = -i \int_{-\infty}^{\infty} e^{itx} \left(e^{-\frac{x^2}{2}} \right)' dx \\ &= -i \left[e^{itx} e^{-\frac{x^2}{2}} \right]_{-\infty}^{\infty} + i \int_{-\infty}^{\infty} it e^{itx} e^{-\frac{x^2}{2}} dx \\ &= 0 - \sqrt{2\pi} \cdot t \varphi_{\mathcal{N}(0,1)}(t). \end{aligned}$$

□

Im Rest dieses Kapitels werden wir hauptsächlich zeigen, dass daraus folgt

$$P_{S_n^*} \implies \mathcal{N}(0, 1) \quad (n \rightarrow \infty),$$

d.h. wir werden den *Zentralen Grenzwertsatz* beweisen. Vorher geben wir noch einige charakteristische Funktionen an.

17.7 Beispiel Hier sind einige Beispiele charakteristischer Funktionen:

Verteilung	Formel	Charakteristische Funktion
Normalverteilung	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx \quad (-\infty < x < \infty)$	$e^{-\frac{\sigma^2 t^2}{2}}$
Gleichverteilung	$1_{[0,1]}(x) dx$	$\frac{e^{it} - 1}{it}$
Exponentialverteilung	$\lambda e^{-\lambda x} dx \quad (0 \leq x < \infty)$	$\frac{\lambda}{\lambda - it}$
Binomialverteilung	$\binom{n}{k} p^k (1-p)^{n-k} \delta_k \quad (k = 0, 1, \dots, n)$	$(1 + p(e^{it} - 1))^n$
Poisson-Verteilung	$e^{-\alpha} \frac{\alpha^k}{k!} \delta_k \quad (k = 0, 1, 2, \dots)$	$e^{\alpha(e^{it} - 1)}$

Für die Normalverteilung wurde das in Satz 17.6 gezeigt, für die übrigen Verteilungen ist die Bestimmung der charakteristischen Funktionen eine einfache *Übung*.

Der folgende Satz zeigt, wie man aus einer charakteristischen Funktion auf eindeutige Weise die zu Grunde liegende Verteilung zurückgewinnen kann.

17.8 Satz (Umkehrformel und Eindeutigkeitsatz)

Seien μ und ν Verteilungen auf \mathbb{R} .

a) Ist $\mu(\{a\}) = \mu(\{b\}) = 0$, so ist

$$\mu([a, b]) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_\mu(t) dt.$$

b) Ist $\varphi_\mu = \varphi_\nu$, so ist auch $\mu = \nu$.

Bevor wir diesen Satz beweisen, halten wir zunächst zwei einfache Korollare fest:

17.9 Korollar Ist $\varphi_\mu(t) = e^{-\frac{t^2}{2}}$, so ist $\mu = \mathcal{N}(0, 1)$.

17.10 Übung μ symmetrisch $\iff \forall t \in \mathbb{R} : \varphi_\mu(t) \in \mathbb{R}$.

Den Beweis von Satz 17.8 bereiten wir durch das folgende Lemma vor.

17.11 Lemma Für $T > 0$ sei $S(T) := \frac{2}{\pi} \int_0^T \frac{\sin t}{t} dt$. Dann existiert $S(\infty) := \lim_{T \rightarrow \infty} S(T)$, es ist $0 < S(\infty) \leq \|S\|_\infty < \infty$, und für $\alpha \in \mathbb{R}$ ist

$$\frac{2}{\pi} \int_0^T \frac{\sin(\alpha t)}{t} dt = \text{sign}(\alpha) \cdot S(|\alpha|T). \quad (*)$$

(Achtung: Das heißt nicht, dass $\frac{\sin(\alpha t)}{t}$ über $(0, \infty)$ integrierbar ist! Vergleiche die alternierende harmonische Reihe!)

Beweis: Beachte, dass für $k = 1, 2, \dots$

$$\begin{aligned} \frac{2}{(2k+1)\pi} &< \int_{2k\pi}^{(2k+1)\pi} \frac{\sin t}{t} dt \leq \frac{2}{2k\pi} \\ \frac{-2}{(2k+1)\pi} &< \int_{(2k+1)\pi}^{2(k+1)\pi} \frac{\sin t}{t} dt \leq \frac{-2}{2(k+1)\pi}. \end{aligned}$$

Addiert man diese beiden Zeilen, so folgt

$$0 < \int_{2k\pi}^{2(k+1)\pi} \frac{\sin t}{t} dt \leq \frac{1}{k(k+1)\pi}$$

so dass $S(T)$ aus dem gleichen Grund wie die alternierende harmonische Reihe konvergiert. Ebenfalls folgt sofort, dass $0 < S(\infty) \leq \|S\|_\infty < \infty$, und (*) folgt durch Substitution, indem man der Reihe nach die Fälle $\alpha = 0$, $\alpha > 0$ und $\alpha = -1$ betrachtet. \square

Beweis von Satz 17.8:

a) Beachte zunächst, dass wegen Lemma 17.4

$$\left| \frac{e^{-ita} - e^{-itb}}{it} \right| = \left| \frac{e^{-itb}(e^{it(b-a)} - 1)}{t} \right| \leq \left| \frac{t(b-a)}{t} \right| = |b-a|. \quad (**)$$

Also ist der Integrand im folgenden Integral beschränkt, und man kann den Satz von Fubini anwenden:

$$\begin{aligned}
 I_{\mu}^{a,b}(T) &:= \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\mu}(t) dt \\
 &= \frac{1}{2\pi} \int_{-T}^T \int_{\mathbb{R}} \frac{e^{-ita} - e^{-itb}}{it} e^{itx} d\mu(x) dt \\
 &= \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{-T}^T \frac{e^{it(x-a)} - e^{it(x-b)}}{it} dt \right] d\mu(x) \\
 &= \frac{1}{2\pi} \int_{\mathbb{R}} \left[\underbrace{\frac{1}{i} \int_{-T}^T \frac{\cos(t(x-a))}{t} - \frac{\cos(t(x-b))}{t} dt}_{=0, \text{ da der Integrand ungerade in } t \text{ ist}} \right. \\
 &\quad \left. + \underbrace{\int_{-T}^T \frac{\sin(t(x-a))}{t} - \frac{\sin(t(x-b))}{t} dt}_{\text{Integrand gerade}} \right] d\mu(x) \\
 &= \frac{1}{2} \cdot \int_{\mathbb{R}} \left[\frac{2}{\pi} \int_0^T \frac{\sin(t(x-a))}{t} - \frac{\sin(t(x-b))}{t} \right] d\mu(x) \\
 &= \frac{1}{2} \int_{\mathbb{R}} \left[\text{sign}(x-a) \cdot S(|x-a|T) - \text{sign}(x-b) \cdot S(|x-b|T) \right] d\mu(x) \\
 &\xrightarrow{T \rightarrow \infty} \frac{1}{2} S(\infty) \cdot \int_{\mathbb{R}} (\text{sign}(x-a) - \text{sign}(x-b)) d\mu(x)
 \end{aligned}$$

nach dem Satz von der majorisierten Konvergenz, da $S(|x-a|T), S(|x-b|T) \rightarrow S(\infty)$ μ -f.s. im Limes $T \rightarrow \infty$ und $\|S\|_{\infty} < \infty$. Da $\text{sign}(x-a) - \text{sign}(x-b) = 1_{\{a\}}(x) + 2 \cdot 1_{(a,b)}(x) + 1_{\{b\}}(x)$ und da $\mu(\{a\}) = \mu(\{b\}) = 0$, folgt

$$\lim_{T \rightarrow \infty} I_{\mu}^{a,b}(T) = \frac{S(\infty)}{2} (\mu(\{a\}) + 2\mu((a,b)) + \mu(\{b\})) = S(\infty) \cdot \mu([a,b]) \quad (***)$$

Um den Beweis von Teil a) zu beenden, bleibt zu zeigen, dass $S(\infty) = 1$. Das geschieht in Lemma C.1 des Anhangs. Aussage b) können wir ohne die Kenntnis des Werts von $S(\infty)$ beweisen.

- b) Seien μ, ν Verteilungen auf \mathbb{R} mit $\varphi_{\mu} = \varphi_{\nu}$. Sei $N := \{x \in \mathbb{R} : \mu\{x\} > 0 \text{ oder } \nu\{x\} > 0\}$. N ist höchstens abzählbar, und für jedes Intervall (a, b) gibt es $a_n, b_n \in \mathbb{R} \setminus N$ mit $[a_n, b_n] \nearrow (a, b)$. Daher ist

$$\begin{aligned}
 \mu((a, b)) &= \sup_n \mu([a_n, b_n]) \stackrel{(***)}{=} \lim_{T \rightarrow \infty} \frac{I_{\mu}^{a_n, b_n}(T)}{S(\infty)} \\
 &= \lim_{T \rightarrow \infty} \frac{I_{\nu}^{a_n, b_n}(T)}{S(\infty)} \stackrel{(***)}{=} \sup_n \nu([a_n, b_n]) = \nu((a, b)),
 \end{aligned}$$

also $\mu = \nu$.

□

Um aus $\lim_{n \rightarrow \infty} \varphi_{S_n^*} = \varphi_{\mathcal{N}(0,1)}$ die Konvergenz $P_{S_n^*} \Rightarrow \mathcal{N}(0, 1)$, d.h. den *Zentralen Grenzwertsatz*, schließen zu können, benötigen wir noch den folgenden Satz.

17.12 Satz (Stetigkeitssatz)

Seien μ, μ_n Wahrscheinlichkeitsmaße auf \mathbb{R} . Dann gilt:

$$\mu_n \implies \mu \quad \text{genau dann, wenn} \quad \lim_{n \rightarrow \infty} \varphi_{\mu_n}(t) = \varphi_\mu(t) \quad \text{für alle } t \in \mathbb{R}.$$

Daraus folgt nun sofort unter Berücksichtigung von Satz 17.6 der Zentrale Grenzwertsatz.

17.13 Satz (Zentraler Grenzwertsatz)

Seien $X_1, X_2, \dots \in \mathcal{L}_P^2$ unabhängige identisch verteilte Zufallsvariablen, $m := E[X_i]$ und $\sigma^2 := V(X_i)$. Sei

$$S_n^* := \frac{1}{\sqrt{n\sigma^2}} \sum_{k=1}^n (X_k - m)$$

die zentrierte und standardisierte Partialsumme von X_1, \dots, X_n . Dann konvergiert

$$P_{S_n^*} \implies \mathcal{N}(0, 1).$$

Beweis von Satz 17.12: Konvergiere $\mu_n \implies \mu$. Da $\Re e^{itx} = \cos(tx)$ und $\Im e^{itx} = \sin(tx)$ für jedes $t \in \mathbb{R}$ beschränkte stetige Funktionen in $x \in \mathbb{R}$ sind, folgt $\lim_{n \rightarrow \infty} \varphi_{\mu_n}(t) = \varphi_\mu(t)$ aus der Definition der schwachen Konvergenz.

Der Beweis der Umkehrung beruht auf den beiden folgenden Lemmata.

17.14 Lemma Existiert $g(t) := \lim_{n \rightarrow \infty} \varphi_{\mu_n}(t)$ für alle $t \in \mathbb{R}$ und ist g stetig bei 0, so ist $(\mu_n)_n$ straff.

17.15 Lemma Existiert $g(t) := \lim_{n \rightarrow \infty} \varphi_{\mu_n}(t)$ für alle $t \in \mathbb{R}$ und ist $(\mu_n)_n$ straff, so ist $g(t)$ charakteristische Funktion eines Wahrscheinlichkeitsmaßes ν auf \mathbb{R} und $\mu_n \implies \nu$.

Gelte nun $\lim_{n \rightarrow \infty} \varphi_{\mu_n}(t) = \varphi_\mu(t)$ für alle $t \in \mathbb{R}$ in Satz 17.12. Da φ_μ stetig ist (Lemma 17.2b), folgt aus Lemma 17.14, dass $(\mu_n)_n$ straff ist, und aus Lemma 17.15, dass es ein Wahrscheinlichkeitsmaß ν gibt mit $\mu_n \implies \nu$ und $\varphi_\nu(t) = \lim_{n \rightarrow \infty} \varphi_{\mu_n}(t) = \varphi_\mu(t)$ für alle $t \in \mathbb{R}$. Aus dem Eindeutigkeitsatz 17.8 folgt schließlich $\nu = \mu$ und daher $\mu_n \implies \mu$. \square

Beweis von Lemma 17.14: Aus dem Satz von Fubini folgt für $u > 0$

$$\begin{aligned} \frac{1}{u} \int_{-u}^u (1 - \varphi_{\mu_n}(t)) dt &= \int_{-\infty}^{\infty} \left[\frac{1}{u} \int_{-u}^u (1 - e^{itx}) dt \right] d\mu_n(x) \\ &= \int_{-\infty}^{\infty} \left[\frac{1}{u} \int_{-u}^u (1 - \cos(tx)) dt \right] d\mu_n(x) - i \int_{-\infty}^{\infty} \underbrace{\left[\frac{1}{u} \int_{-u}^u \sin(tx) dt \right]}_{=0} d\mu_n(x) \\ &= \int_{-\infty}^{\infty} \left[\frac{1}{u} \left(t - \frac{1}{x} \sin(tx) \right) \right]_{t=-u}^u d\mu_n(x) = \int_{-\infty}^{\infty} \left(2 - 2 \frac{\sin(ux)}{ux} \right) d\mu_n(x) \quad (*) \\ &\geq 2 \cdot \int_{\{|x| > \frac{2}{u}\}} \underbrace{\left(1 - \frac{\sin(ux)}{ux} \right)}_{\geq \frac{1}{2}} d\mu_n(x) \geq \mu_n \left(\mathbb{R} \setminus \left[-\frac{2}{u}, \frac{2}{u} \right] \right). \end{aligned}$$

Sei nun $\epsilon > 0$ beliebig. Da $g(0) = \lim_{n \rightarrow \infty} \varphi_{\mu_n}(0) = 1$ und da $g(t)$ bei $t = 0$ stetig ist, gibt es zu $\epsilon > 0$ ein $u > 0$ mit

$$\left| \frac{1}{u} \int_{-u}^u (1 - g(t)) dt \right| < \epsilon .$$

Daher folgt aus (*) und dem Satz von der majorisierten Konvergenz, dass

$$\limsup_{n \rightarrow \infty} \mu_n \left(\mathbb{R} \setminus \left[-\frac{2}{u}, \frac{2}{u} \right] \right) \leq \limsup_{n \rightarrow \infty} \frac{1}{u} \int_{-u}^u (1 - \varphi_{\mu_n}(t)) dt \leq \left| \frac{1}{u} \int_{-u}^u (1 - g(t)) dt \right| < \epsilon .$$

Daraus folgt leicht, dass $(\mu_n)_n$ straff ist (*Übung!*). □

Beweis von Lemma 17.15: Da $(\mu_n)_n$ straff ist, hat jede Teilfolge $(\mu_{n_i})_i$ von $(\mu_n)_n$ eine Teilfolge $(\mu_{n_i(j)})_j$, die schwach gegen ein ν konvergiert. Aus der schon bewiesenen “ \Rightarrow ”-Richtung von Satz 17.12 folgt für jedes solche Limes-Maß ν

$$\varphi_\nu(t) = \lim_{j \rightarrow \infty} \varphi_{\mu_{n_i(j)}}(t) = \lim_{n \rightarrow \infty} \varphi_{\mu_n}(t) = g(t) .$$

Aus dem Eindeutigkeitssatz 17.8 folgt dann, dass alle Teilfolgen $(\mu_{n_i(j)})_j$ gegen das gleiche Maß ν konvergieren, also $\mu_n \Rightarrow \nu$. □

17.16 Übung Sei λ_n die Gleichverteilung auf $[-n, n]$. Zeigen Sie, dass $(\lambda_n)_{n \in \mathbb{N}}$ nicht schwach konvergent, ja nicht einmal straff ist.

17.17 Übung Sei $\alpha > 0$ und seien X_1, \dots, X_n unabhängig identisch verteilt mit Bernoulli-Verteilung zum Parameter $p = \frac{\alpha}{n}$, $S_n := X_1 + \dots + X_n$. Sei $P(\alpha)$ die Poisson-Verteilung zum Parameter α , siehe Beispiel 17.7. Zeigen Sie: $P_{S_n} \Rightarrow P(\alpha)$.

Kapitel 18

Martingale, Submartingale, Stoppzeiten

Allen Definitionen und Sätzen dieses Kapitels liegt ein Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) zu Grunde.

18.1 Definition (Filtration, adaptierter Prozess)

- Eine Familie $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ von Teil- σ -Algebren von \mathcal{A} heißt eine Filtration, falls $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ für alle $n \in \mathbb{N}_0$.
- Eine Familie $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ heißt ein adaptierter Prozess, falls X_n für jedes $n \in \mathbb{N}_0$ eine \mathcal{F}_n -messbare Zufallsvariable ist.

18.2 Definition (Martingal, Submartingal) Ein adaptierter Prozess \mathcal{X} heißt ein Submartingal, falls für alle $n \in \mathbb{N}_0$ gilt

- $E|X_n| < \infty$ und
- $E[X_{n+1} | \mathcal{F}_n] \geq X_n$.

Ein Prozess \mathcal{X} ist ein Martingal, falls in Eigenschaft ii) sogar Gleichheit gilt.

Wir sagen auch, dass $(X_n)_{n \in \mathbb{N}_0}$ ein (Sub-) Martingal ist, falls $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ mit $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$ eines ist.

18.3 Bemerkungen

- Die Summe endlich vieler (Sub-) Martingale bzgl. einer gegebenen Filtration ist ein (Sub-) Martingal.
- Ist \mathcal{X} ein (Sub-)Martingal, so gilt $E[X_{n+k} | \mathcal{F}_n] \underset{(\geq)}{=} X_n$ für alle $n, k \in \mathbb{N}_0$. Das folgt sofort induktiv aus Eigenschaft 18.2ii.

18.4 Beispiele

- A) Seien $\xi_0, \xi_1, \xi_2, \dots$ unabhängige Zufallsvariablen, $E|\xi_i| < \infty$ für alle $i \geq 0$, $\mathcal{F}_n = \sigma(\xi_0, \dots, \xi_n)$.

a) Sei $E\xi_i = 0$ (bzw. ≥ 0) für alle i , $X_n = \xi_0 + \dots + \xi_n$. Dann ist $(X_n, \mathcal{F}_n)_{n \geq 0}$ ein Martingal (bzw. Submartingal), denn für $n \geq 0$ ist

$$E|X_n| \leq E|\xi_0| + \dots + E|\xi_n| < \infty,$$

$$E[X_{n+1} | \mathcal{F}_n] = E[X_n | \mathcal{F}_n] + E[\xi_{n+1} | \mathcal{F}_n] \stackrel{S.12.6c, B.12.5c}{=} X_n + E\xi_{n+1} = (\geq) X_n.$$

b) Sei $E\xi_i = 1$ für alle i , $X_n = \xi_0 \cdot \xi_1 \cdot \dots \cdot \xi_n$. Dann ist $(X_n, \mathcal{F}_n)_{n \geq 0}$ ein Martingal, denn für $n \geq 0$ ist

$$E|X_n| = E|\xi_0| \cdot \dots \cdot E|\xi_n| < \infty,$$

$$E[X_{n+1} | \mathcal{F}_n] = E[X_n \cdot \xi_{n+1} | \mathcal{F}_n] \stackrel{S.12.7}{=} X_n \cdot E[\xi_{n+1} | \mathcal{F}_n] \stackrel{B.12.5c}{=} X_n \cdot E\xi_{n+1} = X_n.$$

B) Sei $\xi \in \mathcal{L}_P^1$, $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration, $X_n := E[\xi | \mathcal{F}_n]$. Dann ist $(X_n, \mathcal{F}_n)_{n \geq 0}$ ein Martingal, denn für $n \geq 0$ ist

$$E|X_n| \stackrel{S.12.6d}{\leq} E[E[|\xi| | \mathcal{F}_n]] = E|\xi| < \infty,$$

$$E[X_{n+1} | \mathcal{F}_n] = E[E[\xi | \mathcal{F}_{n+1}] | \mathcal{F}_n] \stackrel{S.12.8}{=} E[\xi | \mathcal{F}_n] = X_n.$$

18.5 Satz (Martingaleigenschaften von Partialsummenprozessen)

Seien $(\xi_n)_{n \in \mathbb{N}}$ u.i.v. integrierbare Zufallsvariablen mit Erwartungswert 0 und $S_n := \sum_{k=0}^n \xi_k$. Dann sind die folgenden Prozesse Martingale:

i) $(S_n)_{n \in \mathbb{N}_0}$ selbst.

ii) $(S_n^2 - nV(\xi_1))_{n \in \mathbb{N}_0}$, falls $\xi_1 \in \mathcal{L}_P^2$.

iii) $(M_n^\alpha)_{n \in \mathbb{N}_0}$, wobei $M_n^\alpha := \exp(\alpha S_n - n\psi(\alpha))$ für jedes $\alpha \in \mathbb{R}$, für das $\psi(\alpha) := \log E[e^{\alpha \xi_1}] < \infty$.

Beweis: Übung! □

18.6 Lemma (Konvexe Transformation von Submartingalen) Sei $I \subseteq \mathbb{R}$ ein (nicht notwendig endliches) Intervall, $\varphi : I \rightarrow \mathbb{R}$ konvex. Sei $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Submartingal mit Werten in I , und seien die $\varphi(X_n) \in \mathcal{L}_P^1$.

Ist a) φ nicht fallend, oder b) \mathcal{X} ein Martingal, so ist $\varphi(\mathcal{X}) = (\varphi(X_n), \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Submartingal.

Beweis: Mit X_n ist auch $\varphi(X_n)$ bzgl. \mathcal{F}_n messbar. Außerdem folgt aus der bedingten Jensenschen Ungleichung

$$E[\varphi(X_{n+1}) | \mathcal{F}_n] \geq \varphi(E[X_{n+1} | \mathcal{F}_n]) \begin{cases} \geq \varphi(X_n) & \text{im Fall a)} \\ = \varphi(X_n) & \text{im Fall b).} \end{cases}$$

Die bedingte Jensensche Ungleichung kann ähnlich wie die gewöhnliche bewiesen werden. Wenn für die bedingte Erwartung bzgl. einer σ -Algebra \mathcal{F}_n eine reguläre bedingte Wahrscheinlichkeitsverteilung $K_{\mathcal{F}_n}(\omega, \cdot)$ existiert, so folgt sie direkt aus der üblichen Jensenschen Ungleichung:

$$\begin{aligned} E[\varphi(X_{n+1}) | \mathcal{F}_n](\omega) &\stackrel{S.13.11}{=} E_{K_{\mathcal{F}_n}(\omega, \cdot)}[\varphi(X_{n+1})] \stackrel{\text{Jensen}}{\geq} \varphi\left(E_{K_{\mathcal{F}_n}(\omega, \cdot)}[X_{n+1}]\right) \\ &= \varphi(E[X_{n+1} | \mathcal{F}_n]). \end{aligned}$$

□

18.7 Beispiel Ist $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Martingal, so sind $|\mathcal{X}|^p = (|X_n|^p, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ für $p \geq 1$, $\mathcal{X}^+ = (X_n^+, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ und $\mathcal{X}^- = (X_n^-, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ Submartingale. Das folgt aus dem vorhergehenden Lemma für $\varphi(x) = |x|^p, x^+, x^-$. (Für \mathcal{X}^+ bleibt die Aussage sogar richtig, wenn \mathcal{X} nur ein Submartingal ist.)

Die Bedeutung von Martingalen in der Stochastik liegt zu einem großen Teil darin, dass man aus einem gegebenen Martingal auf vielfältige Weise neue Martingale und Submartingale gewinnen kann. Lemma 18.6 war ein erstes einfaches Beispiel dafür. Für die Theorie und die Anwendungen von Martingalen von zentraler Bedeutung ist die *Transformation durch Stoppzeiten*.

18.8 Definition (Stoppzeit, Abbruchregel) Sei $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ eine Filtration. Eine Abbildung $\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$ heißt Stoppzeit (bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$), falls

$$\{\tau \leq n\} := \{\omega \in \Omega : \tau(\omega) \leq n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}_0.$$

18.9 Bemerkung Es gilt

$$\{\tau = n\} \in \mathcal{F}_n \quad \text{und} \quad \{\tau < n\}, \{\tau \geq n\} \in \mathcal{F}_{n-1} \quad \text{für alle } n,$$

denn

$$\begin{aligned} \{\tau = n\} &= \{\tau \leq n\} \setminus \{\tau \leq n-1\}, \\ \{\tau < n\} &= \{\tau \leq n-1\} \quad \text{und} \quad \{\tau \geq n\} = \Omega \setminus \{\tau < n\}. \end{aligned}$$

18.10 Beispiel Ist $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ adaptiert und $B \subseteq \bar{\mathbb{R}}$ messbar, so ist

$$\tau_B := \inf\{n \geq 0 : X_n \in B\} \quad (\tau_B = +\infty, \text{ falls } \forall n \geq 0 : X_n \notin B)$$

eine Stoppzeit, die Zeit des ersten Eintritts in B , denn $\{\tau_B \leq n\} = \bigcup_{k=0}^n \{X_k \in B\} \in \mathcal{F}_n$.

Es gilt sogar mehr: Sei τ eine beliebige Stoppzeit bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Dann ist auch

$$\tau^B := \inf\{n > \tau : X_n \in B\}$$

eine Stoppzeit, denn

$$\{\tau^B \leq n\} = \bigcup_{\substack{k \in \mathbb{N}_0 \\ \tau < k \leq n}} \{X_k \in B\} = \bigcup_{j \in \mathbb{N}_0} \bigcup_{\substack{k \in \mathbb{N}_0 \\ j < k \leq n}} \underbrace{\{\tau = j\}}_{\in \mathcal{F}_j \subseteq \mathcal{F}_n} \cap \underbrace{\{X_k \in B\}}_{\in \mathcal{F}_k \subseteq \mathcal{F}_n} \in \mathcal{F}_n.$$

18.11 Lemma Sind τ_1, τ_2 Stoppzeiten, so auch $\tau_1 \wedge \tau_2 := \min\{\tau_1, \tau_2\}$.

Beweis: $\{\tau_1 \wedge \tau_2 \leq n\} = \{\tau_1 \leq n\} \cup \{\tau_2 \leq n\} \in \mathcal{F}_n$. □

18.12 Definition Ist τ eine Stoppzeit bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$, so heißt

$$\mathcal{F}_\tau := \{A \in \mathcal{A} : A \cap \{\tau \leq n\} \in \mathcal{F}_n \text{ für alle } n \in \mathbb{N}_0\}$$

die σ -Algebra der τ -Vergangenheit.

18.13 Lemma a) \mathcal{F}_τ ist tatsächlich eine σ -Algebra.

b) X_τ ist messbar bzgl. \mathcal{F}_τ .

c) Sind σ und τ Stoppzeiten bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ und ist $\sigma \leq \tau$, so ist $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.

Beweis:

a) Übung.

b) Für $B \in \mathcal{B}^1$ und $n \in \mathbb{N}_0$ ist

$$\{X_\tau \in B\} \cap \{\tau \leq n\} = \bigcup_{k=0}^n \{\tau = k\} \cap \{X_k \in B\} \in \bigcup_{k=0}^n \mathcal{F}_k = \mathcal{F}_n,$$

also $\{X_\tau \in B\} \in \mathcal{F}_\tau$.

c) Sei $A \in \mathcal{F}_\sigma$, $n \in \mathbb{N}_0$. Dann ist

$$A \cap \{\tau \leq n\} = (A \cap \{\sigma \leq n\}) \cap \{\tau \leq n\} \in \mathcal{F}_n,$$

also $A \in \mathcal{F}_\tau$.

□

18.14 Satz (Diskretes stochastisches Integral)

Sei $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein (Sub-)Martingal.

a) Seien H_n beschränkte, nicht negative \mathcal{F}_{n-1} -messbare Zufallsvariablen ($n = 1, 2, \dots$). Setze

$$Y_0 = X_0, Y_{n+1} = Y_n + H_{n+1}(X_{n+1} - X_n), \quad \text{also } Y_n = X_0 + \sum_{k=1}^n H_k(X_k - X_{k-1}).$$

Dann ist $(Y_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein (Sub-)Martingal. Y_n ist das diskrete stochastische Integral von H bzgl. X und wird mit $(H \cdot X)_n$ bezeichnet. Ist \mathcal{X} ein Martingal, so kann auf die Nichtnegativitätsannahme an die H_n verzichtet werden.

b) Sei τ eine Stoppzeit bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Dann ist der gestoppte Prozess $(X_{\tau \wedge n}, \mathcal{F}_{\tau \wedge n})_{n \in \mathbb{N}_0}$ ein (Sub-) Martingal.

Beweis:

a) Die Y_n sind offensichtlich adaptiert, und da die H_n beschränkt sind, sind die Y_n integrierbar. Schließlich ist

$$E[Y_{n+1} | \mathcal{F}_n] = Y_n + H_{n+1}(E[X_{n+1} | \mathcal{F}_n] - X_n) \stackrel{H_{n+1} \geq 0}{\geq} Y_n$$

mit Gleichheit, falls \mathcal{X} ein Martingal ist.

b) Setze $H_n := 1_{\{\tau \geq n\}}$. Dann ist $H_n = 1 - 1_{\{\tau \leq n-1\}}$ \mathcal{F}_{n-1} -messbar und $(H \cdot X)_n = X_{\tau \wedge n}$, denn $(H \cdot X)_0 = X_0 = X_{\tau \wedge 0}$, und induktiv folgt

$$\begin{aligned} (H \cdot X)_{n+1} &= X_{\tau \wedge n} + 1_{\{\tau \geq n+1\}}(X_{n+1} - X_n) = \begin{cases} X_\tau & \text{falls } \tau \leq n \\ X_{n+1} & \text{falls } \tau \geq n+1 \end{cases} \\ &= X_{\tau \wedge (n+1)}. \end{aligned}$$

Beachte nun, dass die $\tau \wedge n$ wegen Lemma 18.11 Stoppzeiten sind und dass $(X_{\tau \wedge n}, \mathcal{F}_{\tau \wedge n})_{n \in \mathbb{N}_0}$ wegen Lemma 18.13b) und c) ein adaptierter Prozess ist. Es bleibt die (Sub-)Martingaleigenschaft zu zeigen: Da $\mathcal{F}_{\tau \wedge n} \subseteq \mathcal{F}_n$ wegen Lemma 18.13c), ist

$$\begin{aligned} E[X_{\tau \wedge (n+1)} | \mathcal{F}_{\tau \wedge n}] &= E[E[X_{\tau \wedge (n+1)} | \mathcal{F}_n] | \mathcal{F}_{\tau \wedge n}] \\ &= E[E[(H \cdot X)_{n+1} | \mathcal{F}_n] | \mathcal{F}_{\tau \wedge n}] \\ &\stackrel{(\geq)}{=} E[(H \cdot X)_n | \mathcal{F}_{\tau \wedge n}] \\ &= E[\mathcal{F}_{\tau \wedge n} | \mathcal{F}_{\tau \wedge n}] \\ &= \mathcal{F}_{\tau \wedge n}, \end{aligned}$$

wobei für die letzte Gleichung Lemma 18.13a) benutzt wurde. □

18.15 Satz (Optionales Stoppen, einfachste Version)

Sei $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein (Sub-)Martingal und seien $\sigma \leq \tau$ beschränkte Stoppzeiten bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Dann ist $X_\sigma \stackrel{(\leq)}{=} E[X_\tau | \mathcal{F}_\sigma]$.

Beweis: Sei $\sigma \leq \tau \leq M < \infty$. Dann sind

$$|X_\sigma|, |X_\tau| \leq \max\{|X_1|, \dots, |X_M|\} \in \mathcal{L}_P^1.$$

Setze $H_n^\sigma = 1_{\{\sigma \geq n\}}$, $H_n^\tau = 1_{\{\tau \geq n\}}$. Es ist $0 \leq H_n^\tau - H_n^\sigma = 1_{\{\sigma < n \leq \tau\}} \leq 1$, also

$$((H^\tau - H^\sigma) \cdot X)_n - X_0 = (H^\tau \cdot X)_n - (H^\sigma \cdot X)_n = X_{\tau \wedge n} - X_{\sigma \wedge n} = X_\tau - X_\sigma$$

falls $n \geq M$. (Es ist $(H^\tau \cdot X)_n = X_{\tau \wedge n}$ wie im Beweis von Satz 18.14b.) Da $((H^\tau - H^\sigma) \cdot X)$ nach Satz 18.14 ein (Sub-)Martingal ist, erhält man

$$\begin{aligned} E[X_\tau] - E[X_\sigma] &= E[((H^\tau - H^\sigma) \cdot X)_n] - E[X_0] \stackrel{(\geq)}{=} E[((H^\tau - H^\sigma) \cdot X)_0] - E[X_0] \\ &= E[X_0] - E[X_0] = 0. \end{aligned} \tag{*}$$

Betrachte nun ein beliebiges $B \in \mathcal{F}_\sigma$ und setze $\tilde{\sigma} := \sigma 1_B + M 1_{B^c}$, $\tilde{\tau} := \tau 1_B + M 1_{B^c}$. Dann sind $\tilde{\sigma} \leq \tilde{\tau} \leq M$ Stoppzeiten, denn

$$\{\tilde{\sigma} \leq n\} = \begin{cases} B \cap \{\sigma \leq n\} & (n < M) \\ \Omega & (n \geq M) \end{cases} \quad \text{gehört zu } \mathcal{F}_n.$$

Wendet man (*) auf $\tilde{\sigma}$ und $\tilde{\tau}$ an, so erhält man

$$E[X_\sigma 1_B + X_M 1_{B^c}] = E[X_{\tilde{\sigma}}] \stackrel{(\leq)}{=} E[X_{\tilde{\tau}}] = E[X_\tau 1_B + X_M 1_{B^c}],$$

so dass $E[X_\sigma 1_B] \stackrel{(\leq)}{=} E[X_\tau 1_B]$. Da das für alle $B \in \mathcal{F}_\sigma$ gilt, folgt $X_\sigma \stackrel{(\leq)}{=} E[X_\tau | \mathcal{F}_\sigma]$, siehe Lemma 12.4b. □

Eine erste Anwendung dieses Satzes ist der Beweis der folgenden *Maximalungleichung*.

18.16 Satz (Maximalungleichung)

Sei $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Submartingal. Dann gilt für jedes $N \in \mathbb{N}_0$ und jedes $\lambda > 0$

a)
$$P \left\{ \max_{0 \leq n \leq N} X_n \geq \lambda \right\} \leq \lambda^{-1} E|X_N|.$$

b) Ist $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ sogar ein Martingal, so gilt auch

$$P \left\{ \min_{0 \leq n \leq N} X_n \leq -\lambda \right\} \leq \lambda^{-1} E|X_N|.$$

Die Wahrscheinlichkeit, dass der Pfad $(X_n)_{n \in \mathbb{N}_0}$ das Intervall $[-\lambda, \lambda]$ bis zur Zeit N verlässt, kann also allein aufgrund der Martingaleigenschaft und der Kenntnis von $E|X_N|$ abgeschätzt werden.

Beweis: a) Sei

$$\tau = \begin{cases} \min \{n \leq N : X_n \geq \lambda\} & \text{falls diese Menge nicht leer ist,} \\ N & \text{sonst.} \end{cases}$$

$\tau \leq N$ ist eine Stoppzeit (siehe Beispiel 18.10). Daher folgt aus Satz 18.15, dass

$$\begin{aligned} E[X_N] &\geq E[X_\tau] = E[X_\tau 1_{\{\max_n X_n \geq \lambda\}}] + E[X_\tau 1_{\{\max_n X_n < \lambda\}}] \\ &\geq \lambda P \left\{ \max_n X_n \geq \lambda \right\} + E[X_N 1_{\{\max_n X_n < \lambda\}}]. \end{aligned}$$

Daher ist

$$P \left\{ \max_n X_n \geq \lambda \right\} \leq \lambda^{-1} E[X_N 1_{\{\max_n X_n \geq \lambda\}}] \leq E|X_N|.$$

b) Ist $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Martingal, so folgt b) direkt aus a), denn dann ist auch $(-X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Martingal, so dass

$$P \left\{ \min_n X_n \leq -\lambda \right\} = P \left\{ \max_n (-X_n) \geq \lambda \right\} \leq \lambda^{-1} E| -X_N | = \lambda^{-1} E|X_N|.$$

□

Kapitel 19

Der Martingal-Konvergenzsatz

Das Hauptergebnis dieses Kapitels ist Doob's Konvergenzsatz für (Sub-)Martingale mit Indexmenge \mathbb{N}_0 .

19.1 Satz (Martingal-Konvergenzsatz)

Sei $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Submartingal und gelte

$$\sup_{n \in \mathbb{N}_0} E[X_n^+] < +\infty.$$

Dann existiert $X_\infty := \lim_{n \rightarrow \infty} X_n$ fast sicher und $E|X_\infty| < \infty$, also insbesondere $|X_\infty| < \infty$ f.s. Genauer: $X_\infty := \lim_{n \rightarrow \infty} X_n$, wo immer der Limes existiert und $X_\infty := 0$ sonst. Damit ist X_∞ messbar bzgl. $\mathcal{F}_\infty := (\bigcup_{n \in \mathbb{N}_0} \mathcal{F}_n)$.

Man beachte, dass zwar $X_\infty \in \mathcal{L}_P^1$, aber trotzdem die Konvergenz $X_n \rightarrow X_\infty$ nur fast sicher und nicht in \mathcal{L}_P^1 gilt. Die damit verbundenen Schwierigkeiten bei der Interpretation des Martingal-Konvergenzsatzes soll das folgende Beispiel illustrieren.

19.2 Beispiel Seien ξ_1, ξ_2, \dots unabhängige Zufallsvariablen, die die Werte $\xi_n = 2 + \frac{2\alpha}{n^2}$ und $\xi_n = 0$ je mit Wahrscheinlichkeit $\frac{1}{2}$ annehmen, wobei $\alpha \geq 0$. Sei $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ und $X_n = \prod_{k=1}^n \xi_k$. X_n kann als Guthaben eines Spielers zur Zeit n interpretiert werden, der bei einem fairen Münzwurfspiel alles setzt und diesen Einsatz entweder verliert oder ihn etwas mehr als verdoppelt. Dann ist $(X_n, \mathcal{F}_n)_{n \geq 1}$ ein Submartingal. In der Tat ist $E[X_{n+1} | \mathcal{F}_n] = X_n \cdot (1 + \frac{\alpha}{(n+1)^2}) \geq X_n$ mit strikter Ungleichung, falls $\alpha > 0$ und $X_n > 0$. Ein Spieler, der im Fall $\alpha > 0$ immer nur einen Schritt in die Zukunft schaut, mag versucht sein, endlos zu spielen, da das erwartete Guthaben $E[X_{n+1} | \mathcal{F}_n]$ nach dem nächsten Spiel immer echt größer als sein gegenwärtiges Guthaben ist (solange er überhaupt noch ein Guthaben hat). Trotzdem ist natürlich $X_\infty = \lim_{n \rightarrow \infty} X_n = 0$ f.s., d.h. der Spieler verliert mit dieser Strategie fast sicher sein gesamtes Guthaben. Beachte auch, dass

$$1 \leq E[X_n^+] = E[X_n] = \prod_{k=1}^n (1 + \frac{\alpha}{k^2}) \leq \prod_{k=1}^{\infty} (1 + \frac{\alpha}{k^2}) < \infty$$

für alle n ist, so dass die Voraussetzungen des Martingal-Konvergenzsatzes erfüllt sind.

Im Fall $\alpha = 0$ liegt ein Martingal vor, und man hat es mit einer fairen Spielstrategie zu tun, die den Martingalen ihren Namen gegeben haben mag. Im Provenzalischen heißt diese Strategie

gie beim Roulettespiel “jouga a la martegalo”, wie man im Buch *Wahrscheinlichkeitstheorie* von Bauer auf Seite 144 nachlesen kann, wo auch noch weitere Hinweise zu finden sind.

Den Beweis des Martingal-Konvergenzatzes bereiten wir mit einer Bemerkung und einem Lemma vor.

19.3 Bemerkung (Absteigende Überquerungen (“downcrossings”)) Sei $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Submartingal. Für $a, b \in \mathbb{R}$ mit $a < b$ definieren wir induktiv Zufallsvariablen

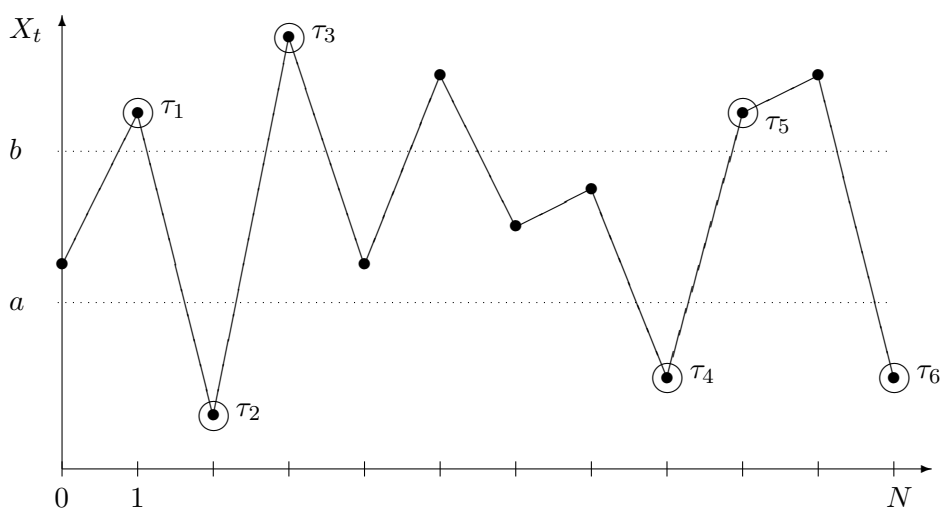
$$\begin{aligned} \hat{\tau}_1 &= \inf \{k \in \mathbb{N}_0 : X_k > b\}, \\ \hat{\tau}_2 &= \inf \{k \in \mathbb{N}_0 : k > \hat{\tau}_1 : X_k < a\}, \\ &\vdots \\ \hat{\tau}_{2n+1} &= \inf \{k \in \mathbb{N}_0 : k > \hat{\tau}_{2n} : X_k > b\}, \\ \hat{\tau}_{2n+2} &= \inf \{k \in \mathbb{N}_0 : k > \hat{\tau}_{2n+1} : X_k < a\}, \\ &\vdots \end{aligned}$$

wobei $\inf \emptyset = \infty$ gesetzt wird. Wegen Beispiel 18.10 sind die $\hat{\tau}_n$ Stoppzeiten.

Für festes $N \in \mathbb{N}_0$ setze nun $\tau_n := \hat{\tau}_n \wedge N$. Die τ_n sind dann wegen Lemma 18.11 auch Stoppzeiten. Setze

$$D(N, [a, b]) := \sup \{n : \tau_{2n} < N\}.$$

$D(N, [a, b])$ ist die Anzahl der Male, die der durch $(1, X_1), \dots, (N-1, X_{N-1})$ definierte zufällige Polygonzug das Intervall $[a, b]$ “von oben nach unten” überspringt (“downcrossings”).



In diesem Beispiel ist $D(N, [a, b]) = 2$ (nicht =3!).

Daher ist

$$D(N, [a, b]) = \max_{n=1, \dots, N} (n \cdot 1_{\{\tau_{2n} < N\}})$$

messbar, so dass auch

$$D([a, b]) := \sup \{D(N, [a, b]) : N \in \mathbb{N}_0\}$$

messbar ist.

19.4 Lemma (“Downcrossing”-Ungleichung) Sei $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Submartingal. Dann gilt für jedes Intervall $[a, b]$

$$(b - a) \cdot E[D([a, b])] \leq \sup_{n \in \mathbb{N}_0} E[(X_n - b)^+].$$

Beweis: Es reicht offensichtlich, diese Ungleichung für alle $E[D(N, [a, b])]$ zu zeigen. Seien die τ_k wie in Bemerkung 19.3. Setze $A_k := \{\tau_k < N\}$. Dann ist $A_{k+1} \subseteq A_k$ und $A_k = \emptyset$ für $k \geq N$. Außerdem ist $A_k \in \mathcal{F}_{\tau_k}$, denn für alle $n < N$ ist $A_k \cap \{\tau_k = n\} = \{\tau_k = n\} \in \mathcal{F}_n$ und für $n \geq N$ ist $A_k \cap \{\tau_k = n\} = \emptyset \in \mathcal{F}_n$.

Da $X_{\tau_{2n-1}} > b$ auf A_{2n-1} und $X_{\tau_{2n}} < a$ auf A_{2n} , können wir folgendermaßen abschätzen:

$$\begin{aligned} 0 &\leq \int_{A_{2n-1}} (X_{\tau_{2n-1}} - b) dP \stackrel{\text{S.18.15}}{\leq} \int_{A_{2n-1}} E[X_{\tau_{2n}} - b | \mathcal{F}_{\tau_{2n-1}}] dP \\ &= \int_{A_{2n-1}} (X_{\tau_{2n}} - b) dP \leq (a - b) P(A_{2n}) + \int_{A_{2n-1} \setminus A_{2n}} (X_{\tau_{2n}} - b) dP \end{aligned}$$

Da $\tau_{2n} = N$ auf A_{2n}^c , folgt

$$(b - a) P(A_{2n}) \leq \int_{A_{2n-1} \setminus A_{2n}} (X_N - b) dP$$

und daher

$$\begin{aligned} E[D(\mathcal{X}, N, [a, b])] &= \sum_{n=1}^{\infty} P(\underbrace{D(\mathcal{X}, N, [a, b]) \geq n}_{\Leftrightarrow \tau_{2n} < N}) = \sum_{n=1}^{\infty} P(A_{2n}) \\ &\leq \frac{1}{b - a} \sum_{n=1}^{\infty} \int_{A_{2n-1} \setminus A_{2n}} (X_N - b) dP \\ &\leq \frac{1}{b - a} \int (X_N - b)^+ dP, \end{aligned}$$

da die Mengen $A_{2n-1} \setminus A_{2n}$ paarweise disjunkt sind. □

Beweis von Satz 19.1: Es ist

$$\{X_n \text{ konvergiert nicht}\} = \left\{ \liminf_{n \rightarrow \infty} X_n < \limsup_{n \rightarrow \infty} X_n \right\} = \bigcup_{\substack{a, b \in \mathbb{Q} \\ a < b}} \left\{ \liminf_{n \rightarrow \infty} X_n < a < b < \limsup_{n \rightarrow \infty} X_n \right\}$$

und

$$\left\{ \liminf_{n \rightarrow \infty} X_n < a < b < \limsup_{n \rightarrow \infty} X_n \right\} \subseteq \{D(\mathcal{X}, [a, b]) = \infty\}.$$

Dabei ist $P(\{D(\mathcal{X}, [a, b]) = \infty\}) = 0$, da

$$E[D(\mathcal{X}, [a, b])] \stackrel{\text{L.19.4}}{\leq} \frac{1}{b - a} \sup_n E[(X_n - a)^+] \leq \frac{1}{b - a} \left(|a| + \sup_n E[X_n^+] \right) < \infty.$$

Also existiert $X_\infty = \lim_{n \rightarrow \infty} X_n$ f.s., und aus dem Lemma von Fatou folgt

$$E|X_\infty| \leq \liminf_{n \rightarrow \infty} E|X_n| \leq \sup_n E[\underbrace{2X_n^+ - X_n}_{=|X_n|}] \leq 2 \sup_n E[X_n^+] - E[X_0] < \infty,$$

denn $E[X_n] \geq E[X_0]$, weil \mathcal{X} ein Submartingal ist. □

19.5 Korollar a) Ein nach oben beschränktes Submartingal $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ konvergiert fast sicher.

b) Ein nach unten beschränktes Martingal $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ konvergiert fast sicher.

Beweis: a) folgt direkt aus Satz 19.1, und für b) beachte, dass $(-X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein nach oben beschränktes Martingal ist. \square

Als erste Anwendung betrachten wir das Martingal $E[\xi | \mathcal{F}_n]$ aus Beispiel 18.4B.

19.6 Satz (Konvergenz bedingter Erwartungen)

Sei $\xi \in \mathcal{L}_P^1$, $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ eine Filtration, $\mathcal{F}_\infty = \sigma(\bigcup_{n \in \mathbb{N}_0} \mathcal{F}_n)$. Dann konvergiert $E[\xi | \mathcal{F}_n] \rightarrow E[\xi | \mathcal{F}_\infty]$ fast sicher und in \mathcal{L}_P^1 .

Zum Beweis benötigen wir folgendes Lemma.

19.7 Lemma (Gleichgradige Integrierbarkeit bedingter Erwartungen) Sei $\xi \in \mathcal{L}_P^1$. Dann ist die Familie

$$\{X_{\mathcal{G}} := E[\xi | \mathcal{G}] : \mathcal{G} \subseteq \mathcal{F} \text{ Teil-}\sigma\text{-Algebra}\}$$

gleichgradig integrierbar.

Beweis: Sei \mathcal{G} eine Teil- σ -Algebra von \mathcal{F} und $a > 0$.

$$\begin{aligned} \int_{\{|X_{\mathcal{G}}| \geq a^2\}} |X_{\mathcal{G}}| dP &\stackrel{S.12.6d}{\leq} \int_{\{|X_{\mathcal{G}}| \geq a^2\}} E[|\xi| | \mathcal{G}] dP = \int_{\{|X_{\mathcal{G}}| \geq a^2\}} |\xi| dP \\ &= \int_{\{|X_{\mathcal{G}}| \geq a^2\} \cap \{|\xi| \leq a\}} |\xi| dP + \int_{\{|X_{\mathcal{G}}| \geq a^2\} \cap \{|\xi| > a\}} |\xi| dP \\ &\leq \underbrace{a \cdot P\{|X_{\mathcal{G}}| \geq a^2\}}_{\leq \frac{a}{a^2} E|X_{\mathcal{G}}|} + \int_{\{|\xi| > a\}} |\xi| dP \\ &\leq \frac{1}{a} E|\xi| + \int_{\{|\xi| > a\}} |\xi| dP \\ &\rightarrow 0 \quad (a \rightarrow \infty) \end{aligned}$$

gleichmäßig in \mathcal{G} . \square

Beweis von Satz 19.6: Sei $X_n := E[\xi | \mathcal{F}_n]$. Nach Beispiel 18.4B ist $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Martingal, und wegen des vorhergehenden Lemmas sind die X_n gleichgradig integrierbar. Insbesondere ist auch $\sup_n E[X_n^+] < \infty$. Daher kann man den Martingal-Konvergenzatz 19.1 anwenden, so dass $X_\infty := \lim_{n \rightarrow \infty} X_n$ fast sicher existiert. Setze $X_\infty = 0$, wo dieser Limes nicht existiert. Dann ist X_∞ bzgl. \mathcal{F}_∞ messbar. Die \mathcal{L}_P^1 -Konvergenz folgt aus Satz 6.21. Bleibt zu zeigen: $X_\infty = E[\xi | \mathcal{F}_\infty]$. Für $A \in \mathcal{F}_n$ und $m \geq n \geq 0$ ist

$$\int_A X_m dP = \int_A E[\xi | \mathcal{F}_m] dP = \int_A \xi dP = \int_A E[\xi | \mathcal{F}_\infty] dP,$$

also

$$\int_A X_\infty dP = \lim_{m \rightarrow \infty} \int_A X_m dP = \int_A E[\xi | \mathcal{F}_\infty] dP,$$

und da $\mathcal{F}_\infty = \sigma(\bigcup_{n \in \mathbb{N}_0} \mathcal{F}_n)$, folgt $X_\infty = E[\xi | \mathcal{F}_\infty]$ aus Lemma 12.4c. \square

Der Martingal-Konvergenzatz garantiert, dass der fast sichere Grenzwert $X_\infty = \lim_{n \rightarrow \infty} X_n$ integrierbar ist. Trotzdem liegt, anders als im gerade bewiesenen Satz, im Allgemeinen keine \mathcal{L}^1 -Konvergenz vor, wie schon Beispiel 19.2 zeigte. Die Frage, wann Submartingale in \mathcal{L}^1 konvergieren, klärt der folgende Satz. Seine Aussage iii) zeigt, dass in Satz 19.6 schon der allgemeine Fall eines \mathcal{L}^1 -konvergenten Martingals behandelt worden ist.

19.8 Satz (\mathcal{L}^1 -Konvergenz und gleichgradige Integrierbarkeit)

a) Sei $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein Martingal und gelte $\sup_{n \in \mathbb{N}_0} E[X_n^+] < +\infty$. Seien X_∞ und \mathcal{F}_∞ wie in Satz 19.1. Dann sind folgende Aussagen äquivalent:

- i) $(X_n)_{n \in \mathbb{N}_0}$ ist gleichgradig integrierbar.
- ii) $\lim_{n \rightarrow \infty} E|X_n - X_\infty| = 0$.
- iii) $\bar{\mathcal{X}} := (X_n, \mathcal{F}_n)_{0 \leq n \leq \infty}$ ist ein Martingal, d.h. $X_n = E[X_\infty | \mathcal{F}_n]$ für alle $n \in \mathbb{N}$.

b) Ist \mathcal{X} nur ein Submartingal, so gilt i) \Leftrightarrow ii) \Rightarrow “ $\bar{\mathcal{X}}$ ist ein Submartingal”.

Beweis: i) \Leftrightarrow ii): Folgt aus Satz 6.21.

ii) \Rightarrow iii): Aus Satz 19.1 folgt, dass $E|X_\infty| < \infty$ und dass X_∞ bzgl. \mathcal{F}_∞ messbar ist. Außerdem ist für jedes $k \in \mathbb{N}_0$

$$E \left[(E[X_\infty | \mathcal{F}_k] - X_k)^- \right] = E \left[\lim_{n \rightarrow \infty} (E[X_n | \mathcal{F}_k] - X_k)^- \right] \stackrel{\text{Fatou}}{\leq} \liminf_{n \rightarrow \infty} E \left[(E[X_n | \mathcal{F}_k] - X_k)^- \right]$$

und $(E[X_n | \mathcal{F}_k] - X_k)^- = 0$ ($n \geq k$) gilt sogar für Submartingale. Also ist $E[X_\infty | \mathcal{F}_k] \geq X_k$, und für Martingale folgt die umgekehrte Ungleichung durch Betrachtung des Submartingals $-\mathcal{X}$.

iii) \Rightarrow ii): Da $X_\infty \in \mathcal{L}_P^1$ und $X_n = E[X_\infty | \mathcal{F}_n]$ nach Voraussetzung, folgt das aus Satz 19.6. \square

19.9 Korollar Jedes gleichgradig integrierbare Submartingal $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ konvergiert fast sicher und in \mathcal{L}_P^1 .

Dieser Satz hilft uns, einen erweiterten Satz über optionales Stoppen zu beweisen. Anders als in Satz 18.15 müssen die Stopzeiten nicht mehr beschränkt sein.

19.10 Satz (Optionales Stoppen)

Seien $\sigma \leq \tau$ fast sicher endliche Stopzeiten bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ und sei $\mathcal{X} = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$ ein gleichgradig integrierbares Submartingal. Dann ist $X_\tau \in \mathcal{L}_P^1$ und $X_\sigma \leq E[X_\tau | \mathcal{F}_\sigma]$ mit Gleichheit, falls \mathcal{X} ein Martingal ist.

19.11 Bemerkung Ohne die Annahme der gleichgradigen Integrierbarkeit wäre dieser Satz falsch. Betrachte z.B. das Martingal $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}_0}$, wo $X_n := \prod_{i=1}^n \xi_i$, $\mathcal{F}_n := \sigma(\xi_1, \dots, \xi_n)$ mit unabhängigen, identisch verteilten ξ_i , $P\{\xi_i = 0\} = P\{\xi_i = 2\} = \frac{1}{2}$. Dann ist $\tau := \inf\{n > 0 : X_n = 0\} \geq 1$ eine fast sicher endliche Stopzeit, aber $E[X_\tau] = 0 \neq 1 = E[X_1]$.

Beweis von Satz 19.10: Wir beweisen den Satz hier nur für Martingale.

Da \mathcal{X} gleichgradig integrierbar ist, ist $X_n = E[X_\infty | \mathcal{F}_n]$ (Satz 19.8). Sei $Y_n := E[X_\infty | \mathcal{F}_{\tau \wedge n}]$. Dann ist auch $(Y_n)_{n \in \mathbb{N}_0}$ gleichgradig integrierbar (Lemma 19.7) und $Y_n = X_{\tau \wedge n}$, denn für

$A \in \mathcal{F}_{\tau \wedge n}$ ist

$$\int_A X_{\tau \wedge n} dP = \sum_{k=0}^n \int_{\underbrace{A \cap \{\tau \wedge n = k\}}_{\in \mathcal{F}_k}} X_k dP = \sum_{k=0}^n \int_{A \cap \{\tau \wedge n = k\}} X_\infty dP = \int_A X_\infty dP .$$

Da $X_{\tau \wedge n} \rightarrow X_\tau$ f.s. und da $X_{\tau \wedge n} = Y_n$ f.s. und in \mathcal{L}^1 konvergiert (Satz 19.6), konvergiert auch $X_{\tau \wedge n} \rightarrow X_\tau$ in \mathcal{L}^1 .

Sei nun $A \in \mathcal{F}_\sigma$ und $A_m := A \cap \{\sigma \leq m\}$. Dann ist $A_m \in \mathcal{F}_{\sigma \wedge m}$, denn für $n \geq 0$ ist

$$A_m \cap \{\sigma \wedge m \leq n\} = \begin{cases} A \cap \{\sigma \leq m\} \in \mathcal{F}_m \subseteq \mathcal{F}_n, & \text{falls } n \geq m \\ A \cap \{\sigma \leq m\} \cap \{\sigma \leq n\} = A \cap \{\sigma \leq n\} \in \mathcal{F}_n, & \text{falls } n < m. \end{cases}$$

Es folgt

$$\int_{A_m} X_\tau dP = \lim_{n \rightarrow \infty} \int_{A_m} X_{\tau \wedge n} dP \stackrel{\text{S.18.15}}{=} \lim_{n \rightarrow \infty} \int_{A_m} X_{\sigma \wedge n} dP = \int_{A_m} X_\sigma dP$$

und da $\lim_{m \rightarrow \infty} P(A \setminus A_m) = 0$, folgt $\int_A X_\sigma dP = \int_A X_\tau dP$, also $X_\sigma = E[X_\tau | \mathcal{F}_\sigma]$. \square

Eine andere Verallgemeinerung des optionalen Stoppens mit beschränkten Stoppzeiten ist die *Wald'sche Gleichung*.

19.12 Satz (Wald'sche Gleichung)

Seien ξ_1, ξ_2, \dots unabhängige identisch verteilte integrierbare Zufallsvariablen, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ und τ eine Stoppzeit zur Filtration $(\mathcal{F}_n)_{n \geq 0}$. Ist $E[\tau] < \infty$, so folgt

$$E[\xi_1 + \dots + \xi_\tau] = E[\tau] \cdot E[\xi_1] .$$

Beweis: Sei $X_n := \xi_1 + \dots + \xi_n - nE[\xi_1]$. Dann ist $(X_n, \mathcal{F}_n)_{n \geq 0}$ ein Martingal (Beispiel 18.4A). Aus Satz 18.15 folgt $EX_{\tau \wedge n} = EX_1 = 0$, also

$$E[\xi_1 + \dots + \xi_{\tau \wedge n}] = E[\tau \wedge n] \cdot E[\xi_1] \tag{*}$$

Sind die $\xi_i \geq 0$, so folgt aus dem Satz von der monotonen Konvergenz

$$E[\xi_1 + \dots + \xi_\tau] = \sup_n E[\xi_1 + \dots + \xi_{\tau \wedge n}] = \sup_n E[\tau \wedge n] \cdot E[\xi_1] = E[\tau] \cdot E[\xi_1] < \infty .$$

Für allgemeine ξ_i folgt insbesondere, dass $E[|\xi_1| + \dots + |\xi_\tau|] = E[\tau] \cdot E[|\xi_1|] < \infty$, also $|\xi_1 + \dots + \xi_{\tau \wedge n}| \leq |\xi_1| + \dots + |\xi_\tau| \in \mathcal{L}_P^1$ für alle n , so dass aus (*) und dem Satz von der majorisierten Konvergenz folgt

$$E[\xi_1 + \dots + \xi_\tau] = \lim_{n \rightarrow \infty} E[\xi_1 + \dots + \xi_{\tau \wedge n}] = \lim_{n \rightarrow \infty} E[\tau \wedge n] \cdot E[\xi_1] = E[\tau] \cdot E[\xi_1] .$$

\square

Anhang A

Maßtheorie

A.1 Beispiel (Intervalle im \mathbb{R}^k)

$\Omega = \mathbb{R}^k$, $\mathcal{S} = \mathcal{I}^k := \{(a, b] : -\infty < a \leq b < +\infty\}$ mit $(a, b] := \{x \in \mathbb{R}^k : a < x \leq b\}$. Dabei sei $\underline{x} \leq \underline{y}$ für $\underline{x}, \underline{y} \in \mathbb{R}^k$, falls $x_i \leq y_i$ für alle $i = 1, \dots, k$ und es sei $\underline{x} < \underline{y}$, falls $\underline{x} \leq \underline{y}$ und $\underline{x} \neq \underline{y}$. Die weiteren Ordnungssymbole werden analog benutzt. $(a, b] \setminus (c, d]$ kann als disjunkte Vereinigung von höchstens $2k$ k -dimensionalen Intervallen geschrieben werden.

A.2 Lemma Ist $(\mathcal{A}_i)_{i \in I}$ eine Familie von (σ) -Algebren, I eine beliebige Indexmenge, so ist $\bigcap_{i \in I} \mathcal{A}_i$ eine (σ) -Algebra.

Beweis:

$$A1) \forall i \in I : \Omega \in \mathcal{A}_i \implies \Omega \in \bigcap_{i \in I} \mathcal{A}_i$$

A3 σ)

$$\begin{aligned} \forall n \in \mathbb{N} : A_n \in \bigcap_{i \in I} \mathcal{A}_i &\implies \forall n \forall i : A_n \in \mathcal{A}_i \xrightarrow{A3\sigma} \forall i : \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_i \\ &\implies \bigcup_{n=1}^{\infty} A_n \in \bigcap_{i \in I} \mathcal{A}_i \end{aligned}$$

A2), A3) ähnlich.

□

A.3 Lemma Ein \cap -stabiles Dynkin-System \mathcal{D} ist eine σ -Algebra.

Beweis:

$$A1) D1 \quad \checkmark$$

$$A2) \Omega \in \mathcal{D} \implies A^c = \Omega \setminus A \in \mathcal{D} \text{ für } A \in \mathcal{D}.$$

$$A3) A, B \in \mathcal{D} \implies A \cup B = A \cup (B \setminus (A \cap B)) \in \mathcal{D}.$$

$$A3\sigma) A_n \in \mathcal{D} (n \geq 1) \xrightarrow{A3} B_n := \bigcup_{i=1}^n A_i \in \mathcal{D} (n \geq 0) \implies \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (B_n \setminus B_{n-1}) \in \mathcal{D} \text{ wegen D3.}$$

□

A.4 Satz (Dynkin-System / σ -Algebra)

Sei $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$ \cap -stabil. Dann ist $\sigma(\mathcal{C}) = \mathcal{D}(\mathcal{C})$.

Beweis:

“ \supseteq ”: $\sigma(\mathcal{C})$ ist ein Dynkin-System und $\mathcal{C} \subseteq \sigma(\mathcal{C})$. Also $\mathcal{D}(\mathcal{C}) \subseteq \mathcal{D}(\sigma(\mathcal{C})) = \sigma(\mathcal{C})$.

“ \subseteq ”: Da $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{D}(\mathcal{C}))$, reicht es zu zeigen, dass $\mathcal{D}(\mathcal{C})$ eine σ -Algebra ist, denn dann ist $\sigma(\mathcal{D}(\mathcal{C})) = \mathcal{D}(\mathcal{C})$. Also ist zu zeigen, dass $\mathcal{D}(\mathcal{C})$ \cap -stabil ist (Lemma 1.15): Für $C \in \mathcal{D}(\mathcal{C})$ sei

$$\mathcal{D}_C := \{A \in \mathcal{D}(\mathcal{C}) : A \cap C \in \mathcal{D}(\mathcal{C})\} \quad (\text{Beachte: } A \in \mathcal{D}_C \iff C \in \mathcal{D}_A)$$

Wir zeigen zunächst, dass \mathcal{D}_C ein Dynkin-System ist:

D1) $\Omega \cap C = C \in \mathcal{D}(\mathcal{C})$, also $\Omega \in \mathcal{D}_C$.

D2) $A, B \in \mathcal{D}_C, A \subseteq B \implies (B \setminus A) \cap C = (B \cap C) \setminus (A \cap C) \in \mathcal{D}(\mathcal{C})$, also $B \setminus A \in \mathcal{D}_C$.

D3) $A_n \in \mathcal{D}_C$ ($n \geq 1$) paarweise disjunkt $\implies (\bigcup_{n=1}^{\infty} A_n) \cap C = \bigcup_{n=1}^{\infty} (A_n \cap C) \in \mathcal{D}(\mathcal{C})$, also $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}_C$.

Die \cap -Stabilität folgt nun aus folgenden drei Überlegungen:

▷ Ist $G \in \mathcal{C}$, so ist $C \in \mathcal{D}_G$ für alle $C \in \mathcal{C}$, also $\mathcal{C} \subseteq \mathcal{D}_G$ und damit $\mathcal{D}(\mathcal{C}) \subseteq \mathcal{D}(\mathcal{D}_G) = \mathcal{D}_G$.

▷ Sei $B \in \mathcal{D}(\mathcal{C})$ und $G \in \mathcal{C}$. Dann ist $B \in \mathcal{D}_G$ und damit $G \in \mathcal{D}_B$ für alle $G \in \mathcal{C}$. Es folgt: $\mathcal{C} \subseteq \mathcal{D}_B$, also $\mathcal{D}(\mathcal{C}) \subseteq \mathcal{D}_B$.

▷ Sind nun $A, B \in \mathcal{D}(\mathcal{C})$, so folgt $A \in \mathcal{D}_B$, d.h. $A \cap B \in \mathcal{D}(\mathcal{C})$.

□

A.5 Satz (Lebesgue-Stieltjes Inhalt)

Sei $\mathcal{B} = \sigma(\mathcal{I}^1)$ die Borel- σ -Algebra auf \mathbb{R} und sei $F : \mathbb{R} \rightarrow \mathbb{R}$ monoton wachsend und rechtsseitig stetig (d.h. $\lim_{t \downarrow x} F(t) = F(x)$ für alle x). Dann wird durch

$$\lambda_F((a, b]) := F(b) - F(a)$$

eine additive und σ -subadditive Mengenfunktion auf dem Halbring \mathcal{I}^1 definiert.

Beweis: Additivität: Seien $A_i = (a_i, b_i] \in \mathcal{I}$ ($i = 1, \dots, n$) mit $A_i \cap A_j = \emptyset$ für $i \neq j$. O.B.d.A. sei $a_1 \leq b_1 \leq a_2 \leq b_2 \leq \dots \leq a_n \leq b_n$. Da $\bigcup_{i=1}^n A_i \in \mathcal{I}$, folgt $b_1 = a_2, \dots, b_{n-1} = a_n$, also $\bigcup_{i=1}^n A_i = (a_1, b_n]$. Daher

$$\lambda_F\left(\bigcup_{i=1}^n A_i\right) = F(b_n) - F(a_1) = \sum_{i=1}^n (F(b_i) - F(a_i)) = \sum_{i=1}^n \lambda_F(A_i).$$

Die endliche Subadditivität folgt nun aus Lemma A.6.

σ -Subadditivität: Seien nun $(a, b], (a_n, b_n] \in \mathcal{I}$ ($n \geq 1$) und $(a, b] \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n]$. O.B.d.A. $a < b$. Sei $\epsilon > 0$. Da F rechtsseitig stetig ist, gibt es $\delta, \delta_n > 0$ derart, dass

$F(a + \delta) < F(a) + \epsilon$ und $F(b_n + \delta_n) \leq F(b_n) + \epsilon 2^{-n}$ für alle $n \geq 1$. Es folgt

$$\begin{aligned}
 [a + \delta, b] &\subseteq \bigcup_{n=1}^{\infty} (a_n, b_n + \delta_n) \\
 \implies \exists m > 0 : [a + \delta, b] &\subseteq \bigcup_{n=1}^m (a_n, b_n + \delta_n) \quad (\text{Kompaktheit}) \\
 \implies F(b) - F(a + \delta) &= \lambda_F((a + \delta, b]) \leq \sum_{n=1}^m \lambda_F((a_n, b_n + \delta_n]) = \sum_{n=1}^m (F(b_n + \delta_n) - F(a_n)) \\
 &\quad (\text{endliche Subadditivität}) \\
 \implies F(b) - F(a) - \epsilon &\leq \sum_{n=1}^{\infty} (F(b_n) - F(a_n)) + \epsilon
 \end{aligned}$$

Da $\epsilon > 0$ beliebig war, folgt $\lambda_F((a, b]) \leq \sum_{n=1}^{\infty} \lambda_F((a_n, b_n])$. □

A.6 Lemma Sei \mathcal{S} ein Halbring, $\mu : \mathcal{S} \rightarrow [0, +\infty]$ additiv. Dann gilt:

- i) $A, B \in \mathcal{S}, A \subseteq B \implies \mu(A) \leq \mu(B)$ (Monotonie)
- ii) μ ist endlich subadditiv.

Beweis:

- i) $B \setminus A = \bigcup_{i=1}^n A_i$ für paarweise disjunkte $A_i \in \mathcal{S}$. Daher:

$$\mu(B) = \mu(A \cup A_1 \cup \dots \cup A_n) = \mu(A) + \mu(A_1) + \dots + \mu(A_n) \geq \mu(A)$$

- ii) Betrachte zunächst den Fall $A = \bigcup_{i=1}^n A_i$. Für jedes A_i gibt es $C_i^1, \dots, C_i^{k_i} \in \mathcal{S}$, so dass $A \setminus A_i = C_i^1 \cup \dots \cup C_i^{k_i}$. Setze $C_i^0 := A_i$. Dann ist $A = \bigcup_{j=0}^{k_i} C_i^j$ ($i = 1, \dots, n$). Setze $\mathcal{C} := \{C_1^{j_1} \cap \dots \cap C_n^{j_n} : 0 \leq j_i \leq k_i \text{ (} i = 1, \dots, n)\}$ und $\mathcal{C}_i := \{S \in \mathcal{C} : S \subseteq A_i\}$. Dann ist

$$A = \bigcup_{S \in \mathcal{C}} S, \quad A_i = \bigcup_{S \in \mathcal{C}_i} S \quad (i = 1, \dots, n)$$

und

$$\mathcal{C} = \bigcup_{i=1}^n \mathcal{C}_i \subseteq \mathcal{S}.$$

Da außerdem $\mu(S) \geq 0$ für jedes $S \in \mathcal{S}$, folgt aus der endlichen Additivität von μ auf \mathcal{S}

$$\mu(A) = \sum_{S \in \mathcal{C}} \mu(S) \leq \sum_{i=1}^n \sum_{S \in \mathcal{C}_i} \mu(S) = \sum_{i=1}^n \mu(A_i).$$

Ist nun A eine echte Teilmenge von $\bigcup_{i=1}^n A_i$, so setze $\tilde{A}_i = A \cap A_i$. Dann sind die $\tilde{A}_i \in \mathcal{S}$ und $A = \bigcup_{i=1}^n \tilde{A}_i$. Also: $\mu(A) \leq \sum_{i=1}^n \mu(\tilde{A}_i)$, und aus Teil i) folgt $\mu(\tilde{A}_i) \leq \mu(A_i)$ □

A.7 Satz (Konstruktion äußerer Maße)

Sei $\mathcal{C} \subseteq \mathfrak{P}(\Omega)$, $\emptyset \in \mathcal{C}$, und sei $\rho : \mathcal{C} \rightarrow [0, +\infty]$ eine Mengenfunktion mit $\rho(\emptyset) = 0$. (Z.B. kann \mathcal{C} ein Semiring sein.) Für $A \subseteq \Omega$ setze

$$\rho^*(A) := \inf \left\{ \sum_{n=1}^{\infty} \rho(C_n) : A \subseteq \bigcup_{n=1}^{\infty} C_n, C_n \in \mathcal{C} \forall n \right\}.$$

(Falls keine abzählbare \mathcal{C} -Überdeckung von A existiert, heißt das $\rho^*(A) = +\infty$.)

Dann ist ρ^* ein äußeres Maß.

Beweis: $\rho^*(\emptyset) = 0$ ist klar. Wir zeigen die σ -Subadditivität von ρ^* : Sei $A \subseteq \bigcup_{n=1}^{\infty} A_n$. Ist $\rho^*(A_n) = \infty$ für ein n , so ist $\rho^*(A) \leq \infty = \sum_{n=1}^{\infty} \rho^*(A_n)$. Also können wir annehmen, dass $\rho^*(A_n) < \infty$ für alle n .

Sei $\epsilon > 0$. Zu jedem A_n gibt es eine Überdeckung $A_n \subseteq \bigcup_{k=1}^{\infty} C_n^k$ mit $C_n^k \in \mathcal{C}$ und

$$\sum_{k=1}^{\infty} \rho(C_n^k) \leq \rho^*(A_n) + \epsilon 2^{-n}.$$

Es folgt:

$$\begin{aligned} \bigcup_{n=1}^{\infty} A_n &\subseteq \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} C_n^k \quad \text{und} \\ \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho(C_n^k) &\leq \sum_{n=1}^{\infty} \rho^*(A_n) + \epsilon, \end{aligned}$$

d.h. $\rho^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \rho^*(A_n) + \epsilon$. Lasse nun $\epsilon \rightarrow 0$ gehen. □

A.8 Satz (Carathéodory)

Ist μ^* ein äußeres Maß auf Ω , so ist $\mathcal{M}(\mu^*)$ eine σ -Algebra, und die Einschränkung von μ^* auf $\mathcal{M}(\mu^*)$ ist ein Maß.

Der Satz ist eine sofortige Konsequenz der folgenden drei Lemmata:

A.9 Lemma Ist μ^* ein äußeres Maß, so ist $\mathcal{M}(\mu^*)$ eine Algebra.

Beweis:

A1) $\Omega \in \mathcal{M}(\mu^*)$, denn $\mu^*(E) + \mu^*(\emptyset) = \mu^*(E)$ für alle $E \subseteq \Omega$.

A2) $A \in \mathcal{M}(\mu^*) \iff A^c \in \mathcal{M}(\mu^*)$ ✓

A3') Wir zeigen, dass mit $A, B \in \mathcal{M}(\mu^*)$ auch $A \cap B \in \mathcal{M}(\mu^*)$: Sei $E \subseteq \Omega$. Dann ist

$$\begin{aligned} &\mu^*((A \cap B) \cap E) + \mu^*((A \cap B)^c \cap E) \\ &= \mu^*(A \cap B \cap E) + \mu^*((A^c \cap B \cap E) \cup (A^c \cap B^c \cap E) \cup (A \cap B^c \cap E)) \\ \stackrel{\text{Subadditivität}}{\leq} &\mu^*(A \cap (B \cap E)) + \mu^*(A^c \cap (B \cap E)) + \mu^*(A^c \cap (B^c \cap E)) + \mu^*(A \cap (B^c \cap E)) \\ &\stackrel{A \in \mathcal{M}(\mu^*)}{=} \mu^*(B \cap E) + \mu^*(B^c \cap E) \\ &\stackrel{B \in \mathcal{M}(\mu^*)}{=} \mu^*(E) \end{aligned}$$

□

A.10 Lemma Ein äußeres Maß μ^* ist σ -additiv auf $\mathcal{M}(\mu^*)$.

Beweis: Einfache Additivität: Seien $A, B \in \mathcal{M}(\mu^*)$, $A \cap B = \emptyset$. Dann ist

$$\begin{aligned} \mu^*(A \cup B) &= \mu^*(A \cap (A \cup B)) + \mu^*(A^c \cap (A \cup B)) \quad \text{da } A \in \mathcal{M}(\mu^*) \\ &= \mu^*(A) + \mu^*(B) \end{aligned}$$

Die endliche Additivität folgt daraus durch Induktion, da $\mathcal{M}(\mu^*)$ eine Algebra ist (Lemma A.9).

σ -Additivität: Seien $A, A_n \in \mathcal{M}(\mu^*)$, $A = \bigcup_{n=1}^{\infty} A_n$. Für $m > 0$ ist

$$\sum_{n=1}^m \mu^*(A_n) \stackrel{\text{Additivität}}{=} \mu^*\left(\bigcup_{n=1}^m A_n\right) \stackrel{\text{Bem. 2.21}}{\leq} \mu^*(A)$$

Es folgt $\mu^*(A) \geq \sum_{n=1}^{\infty} \mu^*(A_n)$, und die umgekehrte Ungleichung folgt aus der σ -Subadditivität des äußeren Maßes μ^* . □

A.11 Lemma Ist μ^* ein äußeres Maß, so ist $\mathcal{M}(\mu^*)$ eine σ -Algebra.

Beweis: Da $\mathcal{M}(\mu^*)$ eine Algebra und damit \cap -stabil ist (Lemma A.9), folgt aus Satz A.4 dass $\mathcal{D}(\mathcal{M}(\mu^*)) = \sigma(\mathcal{M}(\mu^*))$. Es reicht daher zu zeigen, dass $\mathcal{M}(\mu^*)$ ein Dynkin-System ist, denn dann folgt $\mathcal{M}(\mu^*) = \mathcal{D}(\mathcal{M}(\mu^*))$:

D1) und D2) folgen direkt aus der Algebra-Eigenschaft von $\mathcal{M}(\mu^*)$.

Für D3) seien $A_n \in \mathcal{M}(\mu^*)$ paarweise disjunkt, $A = \bigcup_{n=1}^{\infty} A_n$. Zu zeigen ist:

$$\mu^*(A \cap E) + \mu^*(A^c \cap E) \leq \mu^*(E) \quad \text{für alle } E \subseteq \Omega. \quad (*)$$

Unter Ausnutzung von Lemma A.10 zeigen wir zunächst durch Induktion nach n :

$$\mu^*\left(E \cap \bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \mu^*(E \cap A_k)$$

$n = 1$: ✓

$n \Rightarrow n + 1$:

$$\begin{aligned} \mu^*\left(E \cap \bigcup_{k=1}^{n+1} A_k\right) &= \mu^*\left(\left(E \cap \bigcup_{k=1}^{n+1} A_k\right) \cap \bigcup_{k=1}^n A_k\right) + \mu^*\left(\left(E \cap \bigcup_{k=1}^{n+1} A_k\right) \cap \left(\bigcup_{k=1}^n A_k\right)^c\right) \\ &= \mu^*\left(E \cap \bigcup_{k=1}^n A_k\right) + \mu^*(E \cap A_{n+1}) \\ &= \sum_{k=1}^{n+1} \mu^*(E \cap A_k) \end{aligned}$$

Daraus und aus der Monotonie von μ^* (Bemerkung 2.21) folgt für alle n

$$\mu^*(E) = \mu^*\left(E \cap \bigcup_{k=1}^n A_k\right) + \mu^*\left(E \setminus \bigcup_{k=1}^n A_k\right) \geq \sum_{k=1}^n \mu^*(E \cap A_k) + \mu^*(E \cap A^c)$$

Für $n \rightarrow \infty$ folgt mit der σ -Subadditivität von μ^* :

$$\mu^*(E) \geq \sum_{k=1}^{\infty} \mu^*(E \cap A_k) + \mu^*(E \setminus A) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c),$$

also (*). □

A.12 Satz (Fortsetzungssatz für Maße, klassische Form)

Jeder σ -stetige Inhalt μ auf einer Algebra \mathcal{A} lässt sich zu einem Maß auf $\sigma(\mathcal{A})$ fortsetzen.

Ist μ σ -endlich auf \mathcal{A} , so ist diese Fortsetzung eindeutig.

Da jede Algebra insbesondere ein Halbring ist, ist dieser Satz enthalten in der folgenden, etwas technischeren, aber flexibler anwendbaren Aussage:

A.13 Satz (Fortsetzungssatz für Maße, flexible Form)

Sei \mathcal{S} ein Halbring. $\mu : \mathcal{S} \rightarrow [0, +\infty]$ sei additiv, σ -subadditiv und erfülle $\mu(\emptyset) = 0$. Dann lässt sich μ zu einem Maß auf $\sigma(\mathcal{S})$ fortsetzen. Ist μ σ -endlich auf \mathcal{S} , so ist diese Fortsetzung eindeutig.

Beweis: Die Eindeutigkeit wurde schon in Satz 2.19 bewiesen. Zum Beweis der Existenz wird ein äußeres Maß μ^* wie in Satz A.7 definiert:

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(S_n) : A \subseteq \bigcup_{n=1}^{\infty} S_n, S_n \in \mathcal{S} \forall n \right\}$$

Ist $A \in \mathcal{S}$, so folgt aus der σ -Subadditivität von μ auf \mathcal{S} : $\mu(A) \leq \sum_{n=1}^{\infty} \mu(S_n)$ wann immer $A \subseteq \bigcup_{n=1}^{\infty} S_n$ mit $S_n \in \mathcal{S}$. Daher ist $\mu(A) \leq \mu^*(A)$, und da $A \subseteq A \cup \emptyset \cup \emptyset \cup \dots$, ist auch $\mu^*(A) \leq \mu(A) + \mu(\emptyset) + \mu(\emptyset) + \dots = \mu(A)$. Also ist $\mu^*(A) = \mu(A)$ für $A \in \mathcal{S}$.

Bleibt zu zeigen, dass $\mathcal{S} \subseteq \mathcal{M}(\mu^*)$. Denn dann ist auch $\sigma(\mathcal{S}) \subseteq \mathcal{M}(\mu^*)$, und da nach Satz A.8 μ^* ein Maß auf $\mathcal{M}(\mu^*)$ ist, ist $\mu^*|_{\sigma(\mathcal{S})}$ ein Maß auf $\sigma(\mathcal{S})$, das auf \mathcal{S} mit μ übereinstimmt.

Seien also $A \in \mathcal{S}$ und $E \subseteq \Omega$ beliebig. Zu zeigen:

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \mu^*(E). \tag{**}$$

Sei o.B.d.A. $\mu^*(E) < \infty$. Dann gibt es zu $\epsilon > 0$ Mengen $E_n \in \mathcal{S}$ mit

$$E \subseteq \bigcup_{n=1}^{\infty} E_n \quad \text{und} \quad \sum_{n=1}^{\infty} \mu(E_n) < \mu^*(E) + \epsilon.$$

Da \mathcal{S} ein Halbring ist, gilt für alle $n \in \mathbb{N}$

$$B_n := E_n \cap A \in \mathcal{S},$$

$$E_n \setminus A = E_n \setminus B_n = \bigcup_{k=1}^{m_n} C_n^k \quad \text{für geeignete } C_n^k \in \mathcal{S}.$$

Also:

$$E \cap A \subseteq \bigcup_{n=1}^{\infty} B_n, \quad E \setminus A \subseteq \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{m_n} C_n^k$$

$$E_n = B_n \cup \bigcup_{k=1}^{m_n} C_n^k,$$

so dass aus der Definition von μ^* und aus der endlichen Additivität von μ auf \mathcal{S} folgt:

$$\begin{aligned} \mu^*(E \cap A) + \mu^*(E \cap A^c) &\leq \sum_{n=1}^{\infty} \mu(B_n) + \sum_{n=1}^{\infty} \sum_{k=1}^{m_n} \mu(C_n^k) \\ &= \sum_{n=1}^{\infty} \mu(E_n) < \mu^*(E) + \epsilon. \end{aligned}$$

Mit $\epsilon \rightarrow 0$ folgt daraus (**). □

A.14 Satz (Approximationssatz für Maße)

Seien \mathcal{S} und $\mu : \mathcal{S} \rightarrow [0, \infty]$ wie in Satz 2.26, μ σ -endlich.

- a) Zu $A \in \mathcal{M}(\mu^*)$ und $\epsilon > 0$ gibt es eine Folge $S_1, S_2, \dots \in \mathcal{S}$ paarweise disjunkter Mengen mit $A \subseteq \bigcup_{n=1}^{\infty} S_n$ und $\mu(\bigcup_{n=1}^{\infty} S_n \setminus A) < \epsilon$.
- b) Zu $A \in \mathcal{M}(\mu^*)$ mit $\mu(A) < \infty$ und $\epsilon > 0$ gibt es eine endliche Folge $S_1, \dots, S_n \in \mathcal{S}$ paarweise disjunkter Mengen mit $\mu(A \triangle \bigcup_{k=1}^n S_k) < \epsilon$.
- c) Zu $A \in \mathcal{M}(\mu^*)$ gibt es $A_-, A_+ \in \sigma(\mathcal{S})$ mit $A_- \subseteq A \subseteq A_+$ und $\mu(A_+ \setminus A_-) = 0$.

Beweis: Sei $A \in \mathcal{M}(\mu^*)$.

- a) Sei $\epsilon > 0$. Da μ σ -endlich auf \mathcal{S} ist, gibt es Mengen $C_n \in \mathcal{S}$ derart, dass $C_n \nearrow \Omega$ und $\mu(C_n) = \mu^*(C_n) < \infty$ für alle n . Nach Konstruktion von μ^* gibt es zu jedem n Mengen $B_n^k \in \mathcal{S}$ mit

$$C_n \cap A \subseteq \bigcup_{k=1}^{\infty} B_n^k \quad \text{und} \quad \sum_{k=1}^{\infty} \mu(B_n^k) < \mu^*(C_n \cap A) + \epsilon 2^{-n} = \mu(C_n \cap A) + \epsilon 2^{-n}. \quad (*)$$

Wir nummerieren die Mengen B_n^k zu einer Folge $(B_\ell)_\ell$ um und konstruieren daraus eine spezielle Überdeckung von A durch paarweise disjunkte Mengen aus \mathcal{S} : Sei $R_1 = B_1$, $R_2 = B_2 \setminus B_1 = B_2 \setminus (R_1 \cap B_2)$, und für allgemeine ℓ sei

$$R_{\ell+1} = B_{\ell+1} \setminus \left(\bigcup_{k=1}^{\ell} R_k \cap B_{\ell+1} \right).$$

Die R_ℓ sind paarweise disjunkt, und man überlegt sich leicht (Induktion, Übung!), dass jedes R_ℓ endliche disjunkte Vereinigung von Mengen aus \mathcal{S} ist. Daher gibt es $S_1, S_2, \dots \in \mathcal{S}$ derart, dass

$$A = \bigcup_{n=1}^{\infty} (C_n \cap A) \subseteq \bigcup_{\ell=1}^{\infty} B_\ell = \bigcup_{\ell=1}^{\infty} R_\ell = \bigcup_{j=1}^{\infty} S_j$$

und

$$\begin{aligned}
 \mu\left(\bigcup_{j=1}^{\infty} S_j \setminus A\right) &= \mu\left(\bigcup_{\ell=1}^{\infty} B_{\ell} \setminus A\right) \leq \mu\left(\bigcup_{n=1}^{\infty} \left(\bigcup_{k=1}^{\infty} B_n^k \setminus (C_n \cap A)\right)\right) \\
 &\leq \sum_{n=1}^{\infty} \mu\left(\bigcup_{k=1}^{\infty} B_n^k \setminus (C_n \cap A)\right) = \sum_{n=1}^{\infty} \left(\mu\left(\bigcup_{k=1}^{\infty} B_n^k\right) - \mu(C_n \cap A)\right) \\
 &\leq \sum_{n=1}^{\infty} \left(\sum_{k=1}^{\infty} \mu(B_n^k) - \mu(C_n \cap A)\right) \\
 &\leq \sum_{n=1}^{\infty} 2^{-n} \epsilon \\
 &= \epsilon,
 \end{aligned}$$

wobei (*) für die letzte Ungleichung benutzt wurde.

- b) Da $\mu(A) < \infty$, ist auch $\mu\left(\bigcup_{j=1}^{\infty} S_j\right) < \mu(A) + \epsilon < \infty$, und es gibt ein $m > 0$ derart, dass $\sum_{j=m+1}^{\infty} \mu(S_j) < \epsilon$. Also:

$$\mu\left(A \Delta \bigcup_{j=1}^m S_j\right) \leq \mu\left(\bigcup_{j=m+1}^{\infty} S_j\right) + \mu\left(\bigcup_{j=1}^m S_j \setminus A\right) < 2\epsilon,$$

d.h. Aussage b).

- c) Wegen a) gibt es zu $A \in \mathcal{M}(\mu^*)$ eine Folge von Mengen $A_n \in \sigma(\mathcal{S})$ mit $A \subseteq A_n$ und $\mu(A_n \setminus A) < \frac{1}{n}$. Für $A_+ := \bigcap_{n=1}^{\infty} A_n$ ist dann $A \subseteq A_+$ und $\mu(A_+ \setminus A) \leq \inf_n \mu^*(A_n \setminus A) = 0$. Das gleiche Argument angewandt auf A^c liefert $A_- := ((A^c)_+)^c$, denn $((A^c)_+)^c \subseteq A$ und $\mu(A \setminus ((A^c)_+)^c) = \mu(A \cap (A^c)_+) = \mu((A^c)_+ \setminus A^c) = 0$.

□

Anhang B

Integrationstheorie

B.1 Satz (Messbarkeit auf Erzeugern)

a) Seien (Ω, \mathcal{A}) , (Ω', \mathcal{A}') messbare Räume, $\mathcal{M}' \subseteq \mathcal{A}'$, und sei $T : \Omega \rightarrow \Omega'$. Dann ist

$$\sigma(T^{-1}\mathcal{M}') = T^{-1}\sigma(\mathcal{M}').$$

b) Ist außerdem $\sigma(\mathcal{M}') = \mathcal{A}'$ und ist $T^{-1}A' \in \mathcal{A}$ für alle $A' \in \mathcal{M}'$, so ist T messbar.

Beweis:

a) Da $T^{-1}\mathcal{M}' \subseteq T^{-1}\sigma(\mathcal{M}')$ und da $T^{-1}\sigma(\mathcal{M}')$ eine σ -Algebra ist (Lemma 3.3), ist auch $\sigma(T^{-1}\mathcal{M}') \subseteq T^{-1}\sigma(\mathcal{M}')$. Für die Umkehrung betrachte $\mathcal{A}'_0 := \{A' \in \sigma(\mathcal{M}') : T^{-1}A' \in \sigma(T^{-1}\mathcal{M}')\}$. \mathcal{A}'_0 ist eine σ -Algebra (Lemma 3.3) und $\mathcal{M}' \subseteq \mathcal{A}'_0$. Also ist $\sigma(\mathcal{M}') \subseteq \mathcal{A}'_0$ und daher $T^{-1}\sigma(\mathcal{M}') \subseteq T^{-1}\mathcal{A}'_0 = \sigma(T^{-1}\mathcal{M}')$.

b) Ist $\sigma(\mathcal{M}') = \mathcal{A}'$, so folgt aus Teil a):

$$T^{-1}\mathcal{A}' = T^{-1}\sigma(\mathcal{M}') = \sigma(T^{-1}\mathcal{M}') \subseteq \sigma(\mathcal{A}) = \mathcal{A}. \quad \square$$

B.2 Satz

Sei f_1, f_2, \dots eine Folge messbarer $\bar{\mathbb{R}}$ -wertiger Funktionen auf (Ω, \mathcal{A}) .

a) Die Funktionen $\sup_n f_n$, $\inf_n f_n$, $\limsup_{n \rightarrow \infty} f_n$ und $\liminf_{n \rightarrow \infty} f_n$ sind messbar.

b) $E := \{\omega : \lim_{n \rightarrow \infty} f_n(\omega) \text{ existiert in } \bar{\mathbb{R}}\} \in \mathcal{A}$ und $1_E \cdot \lim_{n \rightarrow \infty} f_n$ ist messbar.

Beweis:

a)

$$\left\{ \sup_n f_n \leq x \right\} = \bigcap_n \{f_n \leq x\} \in \mathcal{A} \quad \text{für alle } x \in \bar{\mathbb{R}},$$

$$\left\{ \inf_n f_n \geq x \right\} = \bigcap_n \{f_n \geq x\} \in \mathcal{A} \quad \text{für alle } x \in \bar{\mathbb{R}},$$

$$\limsup_{n \rightarrow \infty} f_n = \inf_n \sup_{k \geq n} f_k, \quad \liminf_{n \rightarrow \infty} f_n = \sup_n \inf_{k \geq n} f_k$$

b) $\{\lim_{n \rightarrow \infty} f_n \text{ existiert in } \bar{\mathbb{R}}\} \in \mathcal{A}$, da die Menge folgendermaßen geschrieben werden kann:

$$\begin{aligned} & \left\{ \liminf_{n \rightarrow \infty} f_n = +\infty \right\} \cup \left\{ \limsup_{n \rightarrow \infty} f_n = -\infty \right\} \\ & \cup \left(\left\{ \liminf_{n \rightarrow \infty} f_n > -\infty \right\} \cap \left\{ \limsup_{n \rightarrow \infty} f_n < +\infty \right\} \cap \left\{ \limsup_{n \rightarrow \infty} f_n - \liminf_{n \rightarrow \infty} f_n \leq 0 \right\} \right). \end{aligned}$$

Dann ist $1_E \cdot \lim_{n \rightarrow \infty} f_n = 1_E \cdot \limsup_{n \rightarrow \infty} f_n$ messbar.

□

B.3 Satz

Sei (Ω, \mathcal{A}) ein messbarer Raum und $f : \Omega \rightarrow [0, +\infty]$ messbar.

- a) Es gibt eine Folge von \mathcal{A} -Elementarfunktionen f_n derart, dass $f_n \nearrow f$.
- b) Es gibt eine Folge messbarer Mengen A_n und $\alpha_n \geq 0$ derart, dass $f = \sum_{n=1}^{\infty} \alpha_n 1_{A_n}$.

Beweis:

- a) Wähle

$$f_n(\omega) = \begin{cases} k2^{-n} & \text{falls } k2^{-n} \leq f(\omega) < (k+1)2^{-n}, 0 \leq k < n2^n, \\ n & \text{falls } n \leq f(\omega) \leq +\infty \end{cases}$$

$$= \sum_{k=0}^{n2^n-1} k \cdot 2^{-n} \cdot 1_{\{k2^{-n} \leq f < (k+1)2^{-n}\}} + n \cdot 1_{\{f \geq n\}}.$$

- b) Für jedes n ist $f_{n+1} - f_n$ eine endliche Linearkombination von Indikatorfunktionen. Da $f_{n+1} - f_n \geq 0$, zeigt man leicht, dass alle Koeffizienten nichtnegativ sind. Also ist $f = f_1 + \sum_{n=1}^{\infty} (f_{n+1} - f_n)$ eine Darstellung der gesuchten Form.

□

B.4 Satz

Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$ messbare Räume, $T : \Omega \rightarrow \Omega'$ messbar, und sei μ ein Maß auf (Ω, \mathcal{A}) . Definiere

$$\mu \circ T^{-1} : \mathcal{A}' \rightarrow [0, +\infty] \quad \text{durch} \quad \mu \circ T^{-1}(A') = \mu(T^{-1}A').$$

Dann ist $\mu \circ T^{-1}$ ein Maß auf \mathcal{A}' (das Bildmaß von μ unter T). Es ist ein Wahrscheinlichkeitsmaß genau dann, wenn μ eines ist.

Beweis: Elementar, denn $\mu \circ T^{-1}$ ist eine positive Mengenfunktion und

$$\mu \circ T^{-1}(\emptyset) = \mu(T^{-1}(\emptyset)) = \mu(\emptyset) = 0$$

und

$$\begin{aligned} \mu \circ T^{-1} \left(\bigcup_{n=1}^{\infty} A_n \right) &= \mu \left(T^{-1} \left(\bigcup_{n=1}^{\infty} A_n \right) \right) = \mu \left(\bigcup_{n=1}^{\infty} T^{-1}(A_n) \right) = \sum_{n=1}^{\infty} \mu(T^{-1}(A_n)) \\ &= \sum_{n=1}^{\infty} \mu \circ T^{-1}(A_n) \end{aligned}$$

□

B.5 Lemma

Seien f, g, f_n nichtnegative, messbare $\bar{\mathbb{R}}$ -wertige Funktionen.

- a) Ist $f = \sum_{i=1}^n \alpha_i 1_{A_i}$, $\alpha_i \in [0, +\infty]$, $A_i \in \mathcal{A}$, so ist $\int f d\mu = \sum_{i=1}^n \alpha_i \mu(A_i)$. Insbesondere ist $\int 1_A d\mu = \mu(A)$.
- b) Aus $0 \leq f \leq g$ folgt $0 \leq \int f d\mu \leq \int g d\mu$.

c) Aus $0 \leq f_n \nearrow f$ folgt $0 \leq \int f_n d\mu \nearrow \int f d\mu$.

d) Sind $\alpha, \beta \in \bar{\mathbb{R}}$, $\alpha, \beta \geq 0$, so gilt

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu .$$

Beweis:

a) Aus der Definition des Integrals folgt sofort, dass $\int f d\mu \geq \sum_{i=1}^n \alpha_i \mu(A_i)$.

Wir können nun annehmen, dass $A_1 \cup \dots \cup A_n = \Omega$ (siehe Bemerkung 3.11). Sei $(B_j)_{j=1, \dots, m}$ eine beliebige endliche messbare Partition von Ω und seien $\beta_1, \dots, \beta_m \geq 0$ so, dass $\sum_{j=1}^m \beta_j 1_{B_j} \leq f = \sum_{i=1}^n \alpha_i 1_{A_i}$. Ist $A_i \cap B_j \neq \emptyset$, so wertet man diese Ungleichung an einer Stelle $\omega \in A_i \cap B_j$ aus und erhält $\beta_j \leq \alpha_i$. Es folgt:

$$\sum_j \beta_j \mu(B_j) = \sum_{i,j} \beta_j \mu(A_i \cap B_j) \leq \sum_{i,j} \alpha_i \mu(A_i \cap B_j) = \sum_i \alpha_i \mu(A_i) ,$$

und da das für jede Wahl von $\sum_{j=1}^m \beta_j 1_{B_j} \leq f$ richtig ist, folgt aus der Definition des Integrals, dass $\int f d\mu \leq \sum_i \alpha_i \mu(A_i)$.

b) folgt sofort aus der Definition des Integrals.

c) Wegen b) gilt

$$0 \leq \int f_n d\mu \nearrow \sup_n \int f_n d\mu \leq \int f d\mu .$$

Es bleibt zu zeigen: $\int f d\mu \leq \sup_n \int f_n d\mu$. Sei also $f \geq \sum_{i=1}^k \alpha_i 1_{A_i} \in \mathcal{E}(\Omega, \mathcal{A})$. Zu zeigen ist

$$\sum_{i=1}^k \alpha_i \mu(A_i) \leq \sup_n \int f_n d\mu . \quad (*)$$

Sei $0 < \epsilon < 1$. Setze $u_i := (1 - \epsilon)\alpha_i$ falls $\alpha_i < \infty$ und $u_i := \epsilon^{-1}$ falls $\alpha_i = \infty$. Betrachte $\omega \in A_i$.

▷ Ist $\alpha_i = 0$, so ist $u_i = 0 \leq f_n(\omega)$ für alle n .

▷ Ist $\alpha_i > 0$, so ist $u_i < \alpha_i \leq f(\omega) = \sup_n f_n(\omega)$, also $u_i \leq f_n(\omega)$ für hinreichend große $n \geq n_0(\omega)$.

Setze nun

$$A_{i,n} = \{\omega \in A_i : f_n(\omega) \geq u_i\} .$$

Dann strebt $A_{i,n} \nearrow A_i$ mit $n \rightarrow \infty$ für $i = 1, \dots, k$, so dass $\mu(A_{i,n}) \nearrow \mu(A_i)$ (Satz 2.10).

Beachte nun, dass $f_n \geq \sum_{i=1}^k u_i \cdot 1_{A_{i,n}}$. Wegen b) und a) folgt:

$$\int f_n d\mu \geq \sum_{i=1}^k u_i \mu(A_{i,n}) \nearrow \sum_{i=1}^k u_i \mu(A_i) \quad \text{für } n \rightarrow \infty ,$$

d.h.

$$\sup_n \int f_n d\mu \geq \sum_{i=1}^k u_i \mu(A_i) = (1 - \epsilon) \sum_{\substack{i \\ \alpha_i < \infty}} \alpha_i \mu(A_i) + \epsilon^{-1} \sum_{\substack{i \\ \alpha_i = \infty}} \mu(A_i) .$$

Die Behauptung (*) folgt nun im Limes $\epsilon \rightarrow 0$.

d) Seien zunächst $f = \sum_i \alpha_i 1_{A_i}$ und $g = \sum_j \beta_j 1_{B_j}$ Elementarfunktionen. Dann ist

$$\alpha f + \beta g = \sum_{i,j} (\alpha \alpha_i + \beta \beta_j) 1_{A_i \cap B_j}$$

auch Elementarfunktion, und aus a) folgt

$$\begin{aligned} \int (\alpha f + \beta g) d\mu &= \sum_{i,j} (\alpha \alpha_i + \beta \beta_j) \mu(A_i \cap B_j) \\ &= \alpha \sum_i \left(\alpha_i \sum_j \mu(A_i \cap B_j) \right) + \beta \sum_j \left(\beta_j \sum_i \mu(A_i \cap B_j) \right) \\ &= \alpha \sum_i \alpha_i \mu(A_i) + \beta \sum_j \beta_j \mu(B_j) \\ &= \alpha \int f d\mu + \beta \int g d\mu. \end{aligned}$$

Beachte, dass diese Rechnung richtig bleibt, falls wir für α, β den Wert $+\infty$ zulassen!

Für beliebige nichtnegative, messbare f, g gibt es nach Satz 3.12 Folgen $(f_n)_{n>0}$ und $(g_n)_{n>0}$ \mathcal{A} -messbarer Elementarfunktionen mit $0 \leq f_n \nearrow f$ und $0 \leq g_n \nearrow g$. Daraus folgt $0 \leq \alpha f_n + \beta g_n \nearrow \alpha f + \beta g$, so dass wegen c)

$$\begin{aligned} \int (\alpha f + \beta g) d\mu &= \sup_n \int (\alpha f_n + \beta g_n) d\mu = \sup_n \left(\alpha \int f_n d\mu + \beta \int g_n d\mu \right) \\ &= \alpha \int f d\mu + \beta \int g d\mu. \end{aligned}$$

□

B.6 Satz

Seien $f, g : \Omega \rightarrow \bar{\mathbb{R}}$ messbar, $f \geq 0$.

- a) $f(\omega) = 0$ f.s. $\iff \int f d\mu = 0$
- b) $\int f d\mu < \infty \implies f(\omega) < \infty$ f.s.
- c) $f(\omega) \leq g(\omega)$ f.s. $\implies \int f d\mu \leq \int g d\mu$

Beweis:

a) " \implies " Sei $N = \{f > 0\}$, Also $\mu(N) = 0$. Setze $N_n := \{0 < f \leq n\} \subseteq N$. Dann gilt $0 \leq f \cdot 1_{N_n} + n \cdot 1_{N \setminus N_n} \nearrow f$, so dass

$$\begin{aligned} 0 &\stackrel{\text{Lemma B.5b)}}{\leq} \int f d\mu \stackrel{\text{Lemma B.5c)}}{=} \sup_n \int (f 1_{N_n} + n 1_{N \setminus N_n}) d\mu \stackrel{\text{Lemma B.5b)}}{\leq} \sup_n \int n 1_N d\mu \\ &\stackrel{\text{Lemma B.5a)}}{\leq} \sup_n n \cdot \mu(N) = 0 \end{aligned}$$

" \impliedby " Für $n \in \mathbb{N}$ sei $A_n := \{f > \frac{1}{n}\}$. Dann ist

$$0 = \int f d\mu \stackrel{\text{Lemma B.5b)}}{\geq} \int \frac{1}{n} \cdot 1_{A_n} d\mu \stackrel{\text{Lemma B.5a)}}{=} \frac{1}{n} \mu(A_n),$$

also $\mu(A_n) = 0$. Da $A_n \nearrow \{f > 0\}$, ist $\mu\{f > 0\} = \sup_n \mu(A_n) = 0$ (Satz 2.10), also $f(\omega) = 0$ f.s.

b) Es ist für alle $n \in \mathbb{N}$

$$n \cdot \mu\{f \geq n\} \stackrel{\text{Lemma B.5a)}}{=} \int n \cdot 1_{\{f \geq n\}} d\mu \stackrel{\text{Lemma B.5b)}}{\leq} \int f d\mu < \infty,$$

also $\mu\{f = \infty\} \leq \mu\{f \geq n\} \rightarrow 0$ mit $n \rightarrow \infty$, d.h. $f < \infty$ f.s.

c) Sei $N = \{f > g\}$, also $\mu(N) = 0$. Es folgt

$$\int f d\mu \stackrel{\text{Lemma B.5d)}}{=} \int \underbrace{f \cdot 1_N}_{=0 \text{ f.s.}} d\mu + \int f \cdot 1_{N^c} d\mu \stackrel{\text{Teil a) und Lemma B.5b)}}{\leq} \int g \cdot 1_{N^c} d\mu \stackrel{\text{Lemma B.5b)}}{\leq} \int g d\mu$$

□

B.7 Satz (Monotonie Linearität und Positivität des Integrals)

Seien $f, g \in \mathcal{L}_\mu^1$.

a) $f \leq g$ f.s. $\implies \int f d\mu \leq \int g d\mu$

b) Für alle $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g \in \mathcal{L}_\mu^1$ und

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu \quad (*)$$

c) $|\int f d\mu| \leq \int |f| d\mu$ (Insbesondere: $f \geq 0 \implies \int f d\mu \geq 0$)

Beweis:

a) Für $f, g \geq 0$ benutze Satz B.6c). Für beliebige f, g beachte, dass aus $f \leq g$ f.s. folgt: $f^+ \leq g^+$ f.s. und $g^- \leq f^-$ f.s., also

$$\int f^+ d\mu \leq \int g^+ d\mu \quad \text{und} \quad \int g^- d\mu \leq \int f^- d\mu,$$

und daher $\int f d\mu \leq \int g d\mu$.

b) Da $|\alpha f + \beta g| \leq |\alpha| \cdot |f| + |\beta| \cdot |g|$, folgt aus Lemma B.5b) und d), dass $(\alpha f + \beta g) \in \mathcal{L}_\mu^1$. Um (*) zu zeigen, beweisen wir folgende drei Aussagen:

i) $\int (f + g) d\mu = \int f d\mu + \int g d\mu$

ii) $\int \alpha f d\mu = \alpha \int f d\mu$ für $\alpha \geq 0$

iii) $\int (-f) d\mu = -\int f d\mu$.

i) Es gilt

$$(f + g)^+ - (f + g)^- = f + g = f^+ - f^- + g^+ - g^-,$$

also

$$(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+.$$

Diese Funktionen sind ≥ 0 , so dass nach Lemma B.5d)

$$\int (f + g)^+ d\mu + \int f^- d\mu + \int g^- d\mu = \int (f + g)^- d\mu + \int f^+ d\mu + \int g^+ d\mu,$$

also

$$\begin{aligned} \int (f + g) d\mu &= \int (f + g)^+ d\mu - \int (f + g)^- d\mu \\ &= \int f^+ d\mu - \int f^- d\mu + \int g^+ d\mu - \int g^- d\mu \\ &= \int f d\mu + \int g d\mu . \end{aligned}$$

ii) Sei $\alpha \geq 0$. Da $(\alpha f)^\pm = \alpha f^\pm$, ist

$$\int \alpha f d\mu = \int \alpha f^+ d\mu - \int \alpha f^- d\mu = \alpha \int f^+ d\mu - \alpha \int f^- d\mu = \alpha \int f d\mu$$

iii)

$$\int (-f) d\mu = \int (-f)^+ d\mu - \int (-f)^- d\mu = \int f^- d\mu - \int f^+ d\mu = - \int f d\mu$$

c) Da $f \leq |f|$ und $-f \leq |f|$, ist $\int f d\mu \leq \int |f| d\mu$ und $-\int f d\mu = \int (-f) d\mu \leq \int |f| d\mu$ nach Teil a) und b).

□

B.8 Satz (Monotone Konvergenz)

Seien $f, f_n : \Omega \rightarrow \overline{\mathbb{R}}$ messbar.

- a) Aus $0 \leq f_n \nearrow f$ f.s. folgt $\int f_n d\mu \nearrow \int f d\mu$.
- b) Sind die $f_n \geq 0$, so gilt $\int \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int f_n d\mu$.
- c) Sind die $f_n \in \mathcal{L}_\mu^1$, gilt $f_n \nearrow f$ f.s. und ist $\sup_n \int f_n d\mu < \infty$, so ist $f \in \mathcal{L}_\mu^1$ und $\int f_n d\mu \nearrow \int f d\mu$. (Diese Aussage heißt auch Satz von Beppo Levi.)

Beweis:

- a) Das ist gerade Lemma B.5 c).
- b) Das folgt aus a), da $0 \leq \sum_{n=1}^m \int f_n d\mu = \int \sum_{n=1}^m f_n d\mu \nearrow \int \sum_{n=1}^{\infty} f_n d\mu$.
- c) $f_1 \leq f_n \implies 0 \leq f_n - f_1 \nearrow f - f_1$, und aus Teil a) folgt

$$\int (f - f_1) d\mu = \sup_n \int (f_n - f_1) d\mu = \sup_n \int f_n d\mu - \int f_1 d\mu < \infty \quad (*)$$

da das Supremum endlich ist und $f_1 \in \mathcal{L}_\mu^1$. Daher $f - f_1 \in \mathcal{L}_\mu^1$, und da $f_1 \in \mathcal{L}_\mu^1$, ist auch $f \in \mathcal{L}_\mu^1$. Es folgt $\int (f - f_1) d\mu = \int f d\mu - \int f_1 d\mu$ und schließlich mit (*) auch $\sup_n \int f_n d\mu = \int f d\mu$.

□

B.9 Satz (Lemma von Fatou)

Seien $0 \leq f_n : \Omega \rightarrow \overline{\mathbb{R}}$ messbar. Dann ist

$$\int \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu .$$

Beweis: Sei $g_k := \inf_{n \geq k} f_n$ (punktweise!), so dass $0 \leq g_k \nearrow \liminf f_n$. Wegen der Monotonie des Integrals ist $\int g_k d\mu \leq \int f_n d\mu$ für alle $n \geq k$, also $\int g_k d\mu \leq \inf_{n \geq k} \int f_n d\mu$. Aus dem Satz von der monotonen Konvergenz folgt nun

$$\int \liminf_{n \rightarrow \infty} f_n d\mu = \sup_k \int g_k d\mu \leq \sup_k \inf_{n \geq k} \int f_n d\mu = \liminf_{n \rightarrow \infty} \int f_n d\mu$$

□

B.10 Satz (Majorisierte Konvergenz, auch Konvergenzsatz von Lebesgue)

Seien $f, f_n, g : \Omega \rightarrow \bar{\mathbb{R}}$ messbar, $|f_n| \leq g$ f.s. und $g \in \mathcal{L}_\mu^1$. Existiert $f = \lim_{n \rightarrow \infty} f_n$ f.s., so ist $f \in \mathcal{L}_\mu^1$ und

$$\lim_{n \rightarrow \infty} \int |f - f_n| d\mu = 0, \quad \text{insbesondere} \quad \int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Beweis: Da $|f_n| \leq g$ f.s., ist auch $|f| = \lim_{n \rightarrow \infty} |f_n| \leq g$ f.s., also $f \in \mathcal{L}_\mu^1$ wegen der Monotonie des Integrals. Sei $h_n := |f - f_n|$. Dann gilt

$$0 \leq h_n \leq |f| + |f_n| \leq 2g, \quad \text{also} \quad 0 \leq 2g - h_n \leq 2g \quad \text{für alle } n$$

und $h_n \rightarrow 0$ f.s., so dass aus dem Lemma von Fatou folgt

$$\begin{aligned} \int 2g d\mu &= \int \liminf_{n \rightarrow \infty} (2g - h_n) d\mu \leq \liminf_{n \rightarrow \infty} \int (2g - h_n) d\mu \\ &= \int 2g d\mu - \limsup_{n \rightarrow \infty} \int h_n d\mu \leq \int 2g d\mu, \end{aligned}$$

also $\limsup_{n \rightarrow \infty} \int h_n d\mu = 0$. Daher

$$\left| \int f d\mu - \int f_n d\mu \right| \stackrel{\text{Satz B.7}}{\leq} \int |f - f_n| d\mu = \int h_n d\mu \rightarrow 0.$$

□

B.11 Korollar (Riemann- und Lebesgue-Integral)) Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist f auch integrierbar bzgl. der Einschränkung des Lebesgue-Maßes auf $[a, b]$, und die beiden Integrale stimmen überein, d.h.

$$\mathbb{R}\text{-} \int_a^b f(x) dx = \int_{[a,b]} f d\lambda.$$

Beweis: Zu Riemann-integrierbarem f gibt es Treppenfunktionen (Unter- und Oberfunktionen) $u_1 \leq u_2 \leq \dots \leq f \leq \dots \leq v_2 \leq v_1$ auf $[a, b]$, so dass

$$\begin{aligned} \int_{[a,b]} u_n d\lambda &= \int_a^b u_n(x) dx \nearrow \mathbb{R}\text{-} \int_a^b f(x) dx \\ \int_{[a,b]} v_n d\lambda &= \int_a^b v_n(x) dx \searrow \mathbb{R}\text{-} \int_a^b f(x) dx. \end{aligned}$$

Sei $u_\infty = \sup_n u_n, v_\infty = \inf_n v_n$. Dann ist $u_\infty \leq f \leq v_\infty$ und da $|u_n|, |v_n| \leq \max\{|u_1|, |v_1|\} < \infty$, folgt aus Bemerkung 6.9, dass $\int_{[a,b]} u_\infty d\lambda = \int_{[a,b]} v_\infty d\lambda = \mathbb{R}\text{-} \int_a^b f(x) dx$. Da $u_\infty \leq v_\infty$, folgt $u_\infty = v_\infty$ λ -f.s., also auch $f = u_\infty$ λ -f.s. f stimmt also λ -f.s. mit einer beschränkten messbaren Funktion überein, und ist damit selbst Lebesgue-integrierbar (Vollständigkeit des Lebesgue-Maßes!). Insbesondere ist dann auch $\int_{[a,b]} f d\lambda = \int_{[a,b]} u_\infty d\lambda = \mathbb{R}\text{-} \int_a^b f(x) dx$. □

B.12 Satz (Differentiation parameterabhängiger Integrale)

Sei $G \subseteq \mathbb{R}$ offen, $f : G \times \Omega \rightarrow \mathbb{R}$ sei so, dass $\omega \mapsto f(t, \omega)$ für jedes $t \in G$ μ -integrierbar und $t \mapsto f(t, \omega)$ für jedes $\omega \in \Omega$ differenzierbar ist. Gibt es ein $\Psi \in \mathcal{L}^1_\mu$ derart, dass $|\frac{d}{dt} f(t, \omega)| \leq \Psi(\omega)$ für alle $(t, \omega) \in G \times \Omega$, so existiert $\frac{d}{dt} \int f(t, \omega) d\mu(\omega)$ für jedes $t \in G$ und es ist

$$\frac{d}{dt} \int f(t, \omega) d\mu(\omega) = \int \frac{d}{dt} f(t, \omega) d\mu(\omega).$$

(Entsprechendes gilt auch für \mathbb{C} -wertige f .)

Beweis: Sei $(h_n)_n$ eine Nullfolge reeller Zahlen, $h_n \neq 0$ für alle n . Bezeichne

$$g(t, \omega) := \frac{d}{dt} f(t, \omega), \quad g_n(t, \omega) := \frac{1}{h_n} (f(t + h_n, \omega) - f(t, \omega)).$$

Zu zeigen ist

$$\lim_{n \rightarrow \infty} \frac{1}{h_n} \left(\int f(t + h_n, \omega) d\mu(\omega) - \int f(t, \omega) d\mu(\omega) \right) = \lim_{n \rightarrow \infty} \int g_n(t, \omega) d\mu(\omega) \stackrel{!}{=} \int g(t, \omega) d\mu(\omega)$$

für alle t . Nach Voraussetzung ist $\lim_{n \rightarrow \infty} g_n(t, \cdot) = g(t, \cdot)$ für alle t und $|g_n| \leq \Psi$ (Mittelwertsatz der Differentialrechnung). Daher folgt die Behauptung aus Satz B.10. \square

B.13 Bemerkung (Jensen impliziert Hölder) Die Hölder-Ungleichung kann direkt aus der Jensen-Ungleichung gefolgert werden: Seien $f \in \mathcal{L}^p_\mu, g \in \mathcal{L}^q_\mu$. Setze

$$\tilde{f} := \left(\frac{|f|}{\|f\|_p} \right)^p, \quad \tilde{g} := \left(\frac{|g|}{\|g\|_q} \right)^q.$$

Dann ist $\int \tilde{f} d\mu = \int \tilde{g} d\mu = 1$. Insbesondere ist $P := \tilde{f} \mu$ ein Wahrscheinlichkeitsmaß (Korollar 6.3). Es folgt

$$\int \frac{|fg|}{\|f\|_p \|g\|_q} d\mu = \int \tilde{f}^{\frac{1}{p}} \tilde{g}^{\frac{1}{q}} d\mu = \int \tilde{f} \left(\frac{\tilde{g}}{\tilde{f}} \right)^{\frac{1}{q}} d\mu = E_P \left[\left(\frac{\tilde{g}}{\tilde{f}} \right)^{\frac{1}{q}} \right]$$

da $\frac{1}{p} + \frac{1}{q} = 1$. (Beachte auch, dass der Quotient P -f.s. wohldefiniert ist!) Da $x \mapsto x^q$ konvex auf $[0, \infty)$ ist, folgt aus der Jensen'schen Ungleichung

$$\left(\int \frac{|fg|}{\|f\|_p \|g\|_q} d\mu \right)^q = \left(E_P \left[\left(\frac{\tilde{g}}{\tilde{f}} \right)^{\frac{1}{q}} \right] \right)^q \leq E_P \left[\frac{\tilde{g}}{\tilde{f}} \right] = \int \tilde{g} d\mu = 1,$$

also die Hölder-Ungleichung.

B.14 Satz

$\mathcal{Z} = \mathcal{Z}(\mathcal{A}_i, i \in I)$ ist ein Semiring.

Beweis:

SR1) Setze $S = \{i_0\}, A_{i_0} = \emptyset$. Dann ist $\bigcap_{i \in I} A_i = \emptyset$.

SR2) Seien $(\bigcap_{i \in I} A_i), (\bigcap_{i \in I} B_i) \in \mathcal{Z}$. Dann ist

$$\bigcap_{i \in I} A_i \cap \bigcap_{i \in I} B_i = \bigcap_{i \in I} (A_i \cap B_i) \quad \text{mit } A_i \cap B_i \in \mathcal{A}_i \ (i \in I)$$

und $A_i \cap B_i = \Omega_i$ falls $A_i = B_i = \Omega_i$, also für alle i bis auf endlich viele.

SR3) Seien $(\prod_{i \in I} A_i), (\prod_{i \in I} B_i) \in \mathcal{Z}$, $(\prod_{i \in I} A_i) \subseteq (\prod_{i \in I} B_i)$. Dann ist $A_i \subseteq B_i$ für alle $i \in I$. Ist $S \subseteq I$ endlich und $A_i = \Omega_i \forall i \in S$, so ist auch $B_i = \Omega_i \forall i \in S$. Daher ist

$$\prod_{i \in I} B_i = \bigcup_{j \in \{0,1\}^S} \prod_{i \in I} C_i^j$$

mit $C_i^j = \Omega_i$ ($i \notin S$), $C_i^j = A_i$ ($i \in S, j_i = 0$), $C_i^j = B_i \setminus A_i$ ($i \in S, j_i = 1$). Es folgt, dass

$$\prod_{i \in I} B_i \setminus \prod_{i \in I} A_i = \prod_{i \in I} B_i \setminus \prod_{i \in I} C_i^{(0, \dots, 0)}$$

endliche disjunkte Vereinigung von Mengen in \mathcal{Z} ist.

□

Anhang C

Diverse Ergänzungen

Beweis des Satzes von Kolmogorov (Satz 10.3)

Wir schreiben $\mathcal{E} := \mathcal{E}(T)$ und betrachten den Semiring \mathcal{Z} der Zylindermengen in $(\mathbb{R}^T, \mathcal{B}^T)$. Jedes Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}^T, \mathcal{B}^T)$ ist wegen Satz 2.19 durch die Bedingung (*) eindeutig bestimmt, denn (*) legt die Werte von μ auf dem Halbring der Zylindermengen fest. Also kann es höchstens ein solches Maß geben.

Wir wenden uns der Existenz zu: Für $B = \pi_I^{-1}(A)$ setze

$$\mu(B) := \mu_I(A).$$

Wir zeigen, dass μ auf \mathcal{Z} dadurch eindeutig definiert ist: Hat B außerdem die Darstellung $B = \pi_{I'}^{-1}(A')$, so setze $J = I \cup I'$ und beachte

$$\pi_J^{-1} \left((\pi_I^J)^{-1}(A) \right) = \pi_I^{-1}(A) = B = \pi_I^{-1}(A') = \pi_J^{-1} \left((\pi_{I'}^J)^{-1}(A') \right).$$

Da π_J surjektiv ist, folgt daraus $(\pi_I^J)^{-1}(A) = (\pi_{I'}^J)^{-1}(A')$, so dass

$$\mu_I(A) \stackrel{\text{Verträglkt.}}{=} \mu_J \left((\pi_I^J)^{-1}(A) \right) = \mu_J \left((\pi_{I'}^J)^{-1}(A') \right) = \mu_{I'}(A'),$$

d.h. $\mu(B)$ ist eindeutig definiert.

Seien nun $B = \pi_I^{-1}(A)$, $B' = \pi_{I'}^{-1}(A')$ und $B \cap B' = \emptyset$. Wegen der vorhergehenden Überlegung kann man o.B.d.A. $I = I'$ annehmen, so dass insbesondere $A \cap A' = \emptyset$. Daher ist

$$\mu(B \cup B') = \mu(\pi_I^{-1}(A \cup A')) = \mu_I(A \cup A') = \mu_I(A) + \mu_I(A') = \mu(B) + \mu(B'),$$

d.h. μ ist *additiv* auf der Algebra $\bigcup_{I \in \mathcal{E}(T)} \pi_I^{-1}(\bigotimes_{i \in I} \mathcal{A}_i)$, insbesondere also auf \mathcal{Z} .

Um den Fortsetzungssatz 2.26 anwenden zu können, der die Fortsetzbarkeit von μ zu einem Wahrscheinlichkeitsmaß auf \mathcal{B}^T garantiert, müssen wir noch die σ -Subadditivität von μ auf \mathcal{Z} nachweisen. Seien also $B = \pi_I^{-1}(A) \in \mathcal{Z}$ und $B_n = \pi_{I_n}^{-1}(A_n) \in \mathcal{Z}$, $B \subseteq \bigcup_{n=1}^{\infty} B_n$. Setze $J_N := I \cup \bigcup_{n=1}^N I_n$ für $N \geq 1$ und $J \subseteq T$ eine abzählbar unendliche Indexmenge, die $I \cup \bigcup_{n=1}^{\infty} I_n$ enthält. Ist T abzählbar, so kann man einfach $J = T$ wählen, sonst nimmt man $J := I \cup \bigcup_{n=1}^{\infty} I_n$. Wir geben hier den Beweis nur für $J = T$, den nur unwesentlich schwierigeren Beweis für allgemeines T findet man in

Dann ist

$$\pi_J^{-1} \left((\pi_I^J)^{-1}(A) \right) = \pi_I^{-1}(A) = B \subseteq \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} \pi_{I_n}^{-1}(A_n) = \pi_J^{-1} \left(\bigcup_{n=1}^{\infty} (\pi_{I_n}^J)^{-1}(A_n) \right),$$

und da $\pi_J : \mathbb{R}^T \rightarrow \mathbb{R}^J$ surjektiv ist, folgt

$$(\pi_I^J)^{-1}A \subseteq \bigcup_{n=1}^{\infty} (\pi_{I_n}^J)^{-1}(A_n).$$

Sei $\epsilon > 0$. Aus der Regularität der Wahrscheinlichkeitsmaße μ_{I_n} auf \mathbb{R}^{J_n} und μ_{J_n} auf \mathbb{R}^{J_n} (Bemerkung 2.30) folgt, dass es kompakte Teilmengen $K_n \subseteq (\pi_{I_n}^{J_n})^{-1}(A)$ und offene Mengen $G_n \supseteq A_n$ derart gibt, dass

$$\mu_{J_n}((\pi_{I_n}^{J_n})^{-1}A \setminus K_n) < 2^{-n}\epsilon \quad \text{und} \quad \mu_{I_n}(G_n \setminus A_n) < 2^{-n}\epsilon.$$

Für $m > 0$ sei $K^m := \bigcap_{n=1}^m (\pi_{J_n}^{J_m})^{-1}K_n \subseteq K^m \subseteq (\pi_I^{J_m})^{-1}(A)$. Dann ist K^m kompakt und

$$(\pi_{J_m}^J)^{-1}K^m = \bigcap_{n=1}^m (\pi_{J_n}^J)^{-1}K_n \subseteq (\pi_I^J)^{-1}A \subseteq \bigcup_{n=1}^{\infty} (\pi_{I_n}^J)^{-1}(A_n) \subseteq \bigcup_{n=1}^{\infty} (\pi_{I_n}^J)^{-1}(G_n) \quad (*)$$

und

$$\begin{aligned} \mu_{J_m} \left((\pi_I^{J_m})^{-1}A \setminus K^m \right) &= \mu_{J_m} \left(\bigcup_{n=1}^m (\pi_{I_n}^{J_m})^{-1}A \setminus (\pi_{J_n}^{J_m})^{-1}K_n \right) \\ &\leq \sum_{n=1}^m \mu_{J_m} \left((\pi_{I_n}^{J_m})^{-1}A \setminus (\pi_{J_n}^{J_m})^{-1}K_n \right) \\ &= \sum_{n=1}^m \mu_{J_m} \left((\pi_{J_n}^{J_m})^{-1}((\pi_{I_n}^{J_n})^{-1}A \setminus K_n) \right) \\ &= \sum_{n=1}^m \mu_{J_n} \left((\pi_{I_n}^{J_n})^{-1}A \setminus K_n \right) < \epsilon. \end{aligned}$$

Am Ende des Beweises werden wir zeigen, dass die K^m eine gewisse Kompaktheitseigenschaft haben: Aus (*) folgt

$$\exists M > 0 : (\pi_{J_M}^J)^{-1}K^M \subseteq \bigcup_{n=1}^M (\pi_{I_n}^J)^{-1}(G_n) = (\pi_{J_M}^J)^{-1} \left(\bigcup_{n=1}^M (\pi_{I_n}^{J_M})^{-1}G_n \right) \quad (**)$$

Da $\pi_{J_M}^J$ surjektiv ist, folgt daraus

$$K^M \subseteq \bigcup_{n=1}^M (\pi_{I_n}^{J_M})^{-1}G_n.$$

Da μ_{J_M} ein Maß auf \mathcal{B}^{J_M} ist, folgt

$$\begin{aligned} \mu(B) &= \mu_{J_M} \left((\pi_I^{J_M})^{-1}A \right) \leq \mu_{J_M}(K^M) + \epsilon \leq \sum_{n=1}^M \mu_{J_M} \left((\pi_{I_n}^{J_M})^{-1}G_n \right) + \epsilon \\ &= \sum_{n=1}^M \mu_{I_n}(G_n) + \epsilon \leq \sum_{n=1}^M (\mu_{I_n}(A_n) + 2^{-n}\epsilon) + \epsilon \\ &\leq \sum_{n=1}^{\infty} \mu(B_n) + 2\epsilon. \end{aligned}$$

Im Limes $\epsilon \rightarrow 0$ erhält man die σ -Subadditivität von μ auf \mathcal{Z} .

Es bleibt zu zeigen, dass (***) aus (*) folgt. Angenommen (***) ist falsch. Dann gibt es zu jedem $m > 0$ einen Punkt

$$x_m \in (\pi_{J_m}^J)^{-1} K^m \setminus \bigcup_{n=1}^m (\pi_{I_n}^J)^{-1} (G_n) \subseteq \mathbb{R}^J. \quad (***)$$

Insbesondere ist für $m \geq M > 0$

$$\pi_{J_M}^J x_m = \pi_{J_M}^{J_m} (\pi_{J_m}^J x_m) \in \pi_{J_M}^{J_m} K^m \subseteq \bigcap_{n=1}^M \pi_{J_M}^{J_n} \left((\pi_{J_M}^{J_n})^{-1} (\pi_{J_n}^J)^{-1} K_n \right) = K^M.$$

Da K^M kompakt ist, hat $(\pi_{J_M}^J x_m)_{m \geq M}$ eine in K^M konvergente Teilfolge, und durch Diagonalisierung erhält man eine Teilfolge $(x_{m_j})_{j > 0}$ derart, dass

$$y_M := \lim_{j \rightarrow \infty} \pi_{J_M}^J (x_{m_j}) \in K^M \subset \mathbb{R}^{J_M} \quad \text{für jedes } M > 0 \text{ existiert.}$$

Sei $M' > M$. Dann ist

$$\pi_{J_{M'}}^{J_{M'}} (y_{M'}) = \lim_{j \rightarrow \infty} \pi_{J_{M'}}^{J_{M'}} \pi_{J_{M'}}^J (x_{m_j}) = \lim_{j \rightarrow \infty} \pi_{J_M}^J (x_{m_j}) = y_M,$$

und da $J = \bigcup_{M=1}^{\infty} J_M$, gibt es einen Punkt $y \in \mathbb{R}^J$ derart, dass $\pi_{J_M}^J (y) = y_M \in K^M$ für alle $M > 0$. Es folgt aus (*), dass

$$\exists N > 0 : \pi_{I_N}^{J_N} (y_N) = \pi_{I_N}^{J_N} (\pi_{J_N}^J y) = \pi_{I_N}^J (y) \in G_N.$$

Da G_N offen ist und da $\pi_{I_N}^{J_N} (y_N) = \lim_{j \rightarrow \infty} \pi_{I_N}^{J_N} (\pi_{J_N}^J x_{m_j}) = \lim_{j \rightarrow \infty} \pi_{I_N}^J x_{m_j}$, ist $\pi_{I_N}^J (x_{m_j}) \in G_N$ für alle hinreichend großen j . Für $m_j \geq N$ steht das aber im Widerspruch zur Wahl von x_{m_j} in (***) .

C.1 Lemma a) Sei μ eine Verteilung auf \mathbb{R} . Ist $\int_{-\infty}^{\infty} |\varphi_{\mu}(t)| dt < \infty$, so hat μ die Dichte $f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_{\mu}(t) dt$ zum Lebesgue-Maß.

b) $S(\infty) = 1$.

Beweis:

a) Sei $\int_{-\infty}^{\infty} |\varphi_{\mu}(t)| dt < \infty$. Dann ist wegen (***) im Beweis von Satz 17.8

$$\left| \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\mu}(t) \right| \leq |b - a| \cdot |\varphi_{\mu}(t)| \in \mathcal{L}_{\text{Leb}}^1$$

und aus dem Satz von der majorisierten Konvergenz und dem Satz von Fubini folgt unter Beachtung von (***) aus dem Beweis von Satz 17.8

$$\begin{aligned} \mu([a, b]) &= \frac{1}{S(\infty)} \cdot \lim_{T \rightarrow \infty} I_{\mu}^{a,b}(T) = \frac{1}{2\pi S(\infty)} \int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\mu}(t) dt \\ &= \frac{1}{2\pi S(\infty)} \int_{-\infty}^{\infty} \left[\int_a^b e^{-itx} dx \right] \varphi_{\mu}(t) dt \\ &= \frac{1}{2\pi S(\infty)} \int_a^b \left[\int_{-\infty}^{\infty} e^{-itx} \varphi_{\mu}(t) dt \right] dx \end{aligned}$$

d.h. μ hat die behauptete Dichte, falls $S(\infty) = 1$.

b) Wir zeigen nun, dass $S(\infty) = 1$. Sei dazu $\mu = \mathcal{N}(0, 1)$, also $\varphi_\mu(t) = e^{-\frac{t^2}{2}} \in \mathcal{L}_{\text{Leb}}^1$. Daher ist nach dem soeben gezeigten

$$S(\infty) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} e^{-itx} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \cdot \varphi_\mu(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}},$$

also $S(\infty) = 1$.

□