Maßtheorie für Statistiker

Mathias Trabs*

1. Juni 2015

Inhaltsverzeichnis

*Die Vorlesung und die vorliegende Gliederung beruht fundamental auf einem ausführlichen Skript von Uwe Küchler, welches demnächst in Buchform erscheinen soll. Ich bedanke mich dafür, dass er es mir freundlicher Weise zur Verfügung gestellt hat.

Motivation

Einer der größten Vorteile der Maßtheorie aus Sicht eines Statistikers ist, diskrete und kontinuierliche Wahrscheinlichkeitsverteilung in einem allgemeinen Rahmen gemeinsam behandeln zu können. Dabei werden Wahrscheinlichkeitsverteilungen durch (Wahrscheinlichkeits-)Maße beschrieben. Fundamental Größen wie Momente oder allgemein Erwartungswerte, können dann als (Lebesgue-)Integrale bezüglich dieser Maße aufgefasst werden. Hier kommt der zweite große Vorzug der Maßtheorie ins Spiel, nämlich eine sehr allgemeine und leistungsfähige Integrationstheorie.

1 Grundbegriffe

1.1 Begriffe der Mengenlehre

Unter einer *Menge* versteht man eine "Zusammenfassung bestimmter, wohlunterschiedener Objekte" (Georg Cantor). Dabei spielen keinerlei Beziehungen (Größe etc.) der Objekte untereinander eine Rolle. Die Objekte in der Mengen heißen *Elemente der Menge*.

Beispiel 1.1.

(i) $A := \{1, 2, 3, 4, 5, 6\}$ alle möglichen Ergebnisse eines Würfelwurfes,

(ii) $B := \{2k : k = 0, 1, 2, \dots\}$ Menge der gerade Zahlen,

(iii) $C :=$ {Berlin, Hamburg, Köln, München} Menge der deutschen Millionenstädte

Wir beginnen mit einigen Bezeichnungen:

- $x \in A$: x ist ein *Element* von A,
- $x \notin A : x$ ist kein Element von A,
- $A \subseteq B$: A ist eine *Teilmenge* von B, d.h. ist $x \in A$ so gilt auch $x \in B$,
- $A = B$: A und B sind gleich, d.h. $A \subseteq B$ und $B \subseteq A$,
- $\emptyset$ ist die *leere Menge*, d.h. sie enthält kein Element.
- A heißt *endliche Menge*, falls sie nur endlich viele Elemente enthält.

Definition 1.2. Seien A und B Mengen.

(i) $A \cup B := \{x : x \in A \text{ oder } x \in B\}$ heißt *Vereinigung.*

(ii) $A \cap B := \{x : x \in A \text{ und } x \in B\}$ heißt *Durchschnitt oder Schnittmenge.*

(iii) A und B heißen *disjunkt* falls $A \cap B = \emptyset$.

(iv) $A \setminus B := \{x \in A : x \notin B\}$ heißt *Differenz.*

(v) $A \Delta B := (A \setminus B) \cup (B \setminus A)$ heißt *symmetrische Differenz.*

Beispiel 1.3 (Zahlenmengen)**.**

(i) $\mathbb{N} := \{1, 2, 3, \dots\}$ ist die Menge der *natürlichen Zahlen*. Die Mengen der *ganzen Zahlen* ist $\mathbb{Z} := \mathbb{N} \cup \{0\} \cup (-\mathbb{N})$, wobei $-\mathbb{N} := \{-n : n \in \mathbb{N}\}$. Die Menge der *rationalen Zahlen* ist geben durch $\mathbb{Q} := \{\frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{Z} \setminus \{0\}\}$. Es gilt $\mathbb{Q} \subseteq \mathbb{R}$ für die Menge der *reellen Zahlen* $\mathbb{R}$.

(ii) $\{x \in \mathbb{R} : x > 0\} \subseteq \mathbb{R}$ (bzw. $\{x \in \mathbb{R} : x \geqslant 0\}$) ist die Teilmenge der *positiven* (bzw. *nichtnegativen*) Zahlen. Analog sind $\{x \in \mathbb{R} : x < 0\}$ (bzw. $\{x \in \mathbb{R} : x \leqslant 0\}$) die *negativen* (bzw. *nichtpositiven*) Zahlen.

Lemma 1.4. *Seien A, B und C beliebige Mengen. Dann gilt*

(i) Assoziativität:

$$A \cup (B \cup C) = (A \cup B) \cup C \quad \textit{und} \quad (A \cap B) \cap C = A \cap (B \cap C),$$

(ii) Distributivität:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad \textit{und} \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

(iii) Kommutativität:

$$A \cup B = B \cup A \quad \textit{und} \quad A \cap B = B \cap A.$$

Beweis. Bild. □

Aufgabe 1.5.

(i) Welche Relationen bestehen zwischen den Mengen $A = \{\text{Stuhl}, \text{Hocker}, \text{Tisch}\}$, $B = \{\text{Stuhl}, \text{Tisch}, \text{Stuhl}, \text{Hocker}\}$, $C = \{\text{Stuhl}, \text{Hocker}\}$ und $D = \{\text{Hocker}\}$? Geben Sie $B \cap C$ und $A \cup D$ an.

(ii) Es seinen A und B zwei Mengen. Wie kann man paarweise disjunkte Mengen A_1, A_2, A_3 wählen, dass $A_1 \cup A_2 \cup A_3 = A \cup B$ gilt?

Definition 1.6. Sei E eine Menge, dann heißt $\mathscr{P}(E) := \{A : A \subseteq E\}$ *Potenzmenge* von E. Für $A \subseteq E$ heißt $A^c := E \setminus A$ *Komplement* von A bzgl. E. Jede Menge $\mathcal{S} \subseteq \mathscr{P}(E)$ von Teilmengen von E nennen wir *Mengensystem* aus E.

Beispiel 1.7. $\mathcal{S} := \{(a,b) : -\infty \leqslant a \leqslant b \leqslant \infty\}$ ist das Mengensystem der *offenen Intervalle* aus $\mathbb{R}$.

Lemma 1.8. *Für alle Teilmengen $A, B \in \mathscr{P}(E)$ für eine Menge E gilt*

$$(A^c)^c = A, \quad A \cap A^c = \emptyset \quad \textit{und} \quad A \cup A^c = E.$$

Weiterhin gilt

$$(A \cup B)^c = A^c \cap B^c \quad \textit{und} \quad A \setminus B = A \cap B^c.$$

Beweis. Übung ☑. □

Definition 1.9. Sei $(A_i)_{i \in I}$ eine Familie von Mengen für eine Indexmenge I.

(i) $\bigcap_{i \in I} A_i := \{x : x \in A_i \text{ für alle } i \in I\}$ und

(ii) $\bigcup_{i \in I} A_i := \{x : x \in A_i \text{ für mindestens ein } i \in I\}$.

(iii) Mengen $A_i, i \in I$, heißen *paarweise disjunkt, falls $A_i \cap A_j = \emptyset$ für alle $i, j \in I$ mit $i \neq j$.*

(iv) Ein Mengensystem $\mathcal{Z} = \{Z_i : i \in I\}$ aus E heißt *Zerlegung* von E, falls $Z_i, i \in I$, paarweise disjunkt sind und $\bigcup_{i \in I} Z_i = E$.

Beispiel 1.10.

(i) $\{[0, 1/3), [1/3, 2/3), [2/3, 1)\}$ ist eine Zerlegung von $[0, 1)$.

(ii) $\{[k, k+1) : k \in \mathbb{Z}\}$ ist eine Zerlegung von $\mathbb{R}$.

Relevanz für Stochastik/Statistik: Mit einer Menge E werden wir alle möglichen Versuchsausgänge eines Zufallsexperimentes beschreiben, bspw. $E = \{1, 2, 3, 4, 5, 6\}$ für einen Würfelwurf. Welches Ergebnis tatsächlich eintritt ist ungewiss. Teilmengen von E beschreiben Ereignisse wie $A = \{2, 4, 6\} \subseteq E$ "es wurde eine gerade Zahl gewürfelt". Alle interessierenden Ereignisse fassen wir schließlich durch ein Mengensystem $\mathcal{S}$ zusammen. Für endlich viele Versuchsausgänge wählt man typischerweise die Potenzmenge $\mathscr{P}(E)$.

1.2 Abbildungen

Definition 1.11. Es seien E, F zwei (nichtleere) Mengen. Eine *Abbildung* (oder *Funktion*) X von E in F, geschrieben $X\colon E \to F$, ist eine Vorschrift, die jedem Element $y \in E$ ein eindeutig bestimmtes Element $X(y) \in F$ zuordnet, geschrieben $y \mapsto X(y)$. Die Menge E heißt *Definitionsbereich.* X heißt *injektiv*, falls für alle $y, z \in E$ gilt

$$X(y) = X(z) \quad \Rightarrow \quad y = z.$$

Die Menge $X(E) := \{X(y) : y \in E\} \subseteq F$ heißt *Werte-* oder *Bildbereich* von X. Gilt $X(E) = F$, so heißt X *surjektiv.* Ist X injektiv und surjektiv, heißt X *bijektiv.* Für jede Teilmenge $B \subseteq F$ ist das *Urbild* von B (vermittels X) definiert als

$$\{X \in B\} := X^{-1}(B) := \{u \in E : X(u) \in B\}.$$

Aufgabe 1.12.

(i) $E = [-1, 1], F = \mathbb{R}$ und $X\colon E \to F$ ist definiert durch $X(y) = y^2$. Ist X injektiv oder surjektiv? Wie können E bzw. F abgeändert werden, damit X surjektiv oder sogar bijektiv wird?

(ii) Man zeige, dass $X(y) := y_+ := \max\{0, y\}$ und $Y(y) := y_- := \max\{0, -y\}, y \in \mathbb{R}$, surjektive Abbildungen von $\mathbb{R}$ auf $[0, \infty)$ sind. Dasselbe gilt für $Z(y) := |y| := y_+ + y_-$. Sind diese Abbildungen bijektiv?

Lemma 1.13. *Seien E, F zwei nichtleere Mengen und $X\colon E \to F$ eine Abbildung. Die Urbildoperation X^{-1} ist mit den Mengenoperation vertauschbar, d.h. für $B, C \subseteq F$ gelten*

$$X^{-1}(B\cap C) = X^{-1}(B)\cap X^{-1}(C), \quad X^{-1}(B\cup C) = X^{-1}(B)\cup X^{-1}(C), \quad X^{-1}(F\setminus B) = E\setminus X^{-1}(B)$$

und für eine Familie von Teilmengenengen $(B_i)_{i\in I} \subseteq \mathscr{P}(F)$ mit Indexmenge I gilt

$$X^{-1}\Big(\bigcap_{i\in I} B_i\Big) = \bigcap_{i\in I} X^{-1}(B), \quad X^{-1}\Big(\bigcup_{i\in I} B_i\Big) = \bigcup_{i\in I} X^{-1}(B).$$

Beweis. Übung ☑. □

Definition 1.14. Ist $X\colon E \to F$ injektiv, existiert für jedes $z \in X(E)$ eine eindeutige Lösung $y \in E$ der Gleichung $X(y) = z$. Sie wird mit $X^{-1}(z)$ bezeichnet. Dadurch wird die zu X *inverse Abbildung* $X^{-1}\colon F \supseteq X(E) \to E$ definiert.

Lemma 1.15. *Ist $X\colon E \to F$ injektiv, so gilt*

$$X(X^{-1}(z)) = z \quad \textit{und} \quad X^{-1}(X(y)) = y \quad \textit{für alle } y \in E, z \in X(E).$$

Beweis. Übung ☑. □

Falls X eine injektive Abbildung ist, so gilt $X^{-1}(\{z\}) = \{X^{-1}(z)\}$, d.h. für einelementige Teilmengen $\{z\}$ von F stimmt die Urbildoperation mit der inversen Abbildung $X^{-1}(z)$ überein.

Abbildungen können wir insbesondere nutzen um die Größe von unendlichen Mengen (deren Kardinalität) zu vergleichen. Ein im Laufe der Vorlesung wichtiger Begriff ist folgender:

Definition 1.16. Eine Menge A heißt *abzählbar unendlich*, falls es eine Bijektion $f\colon \mathbb{N} \to A$ gibt, d.h. wir können A schreiben als $A = \{x_n : n \in \mathbb{N}\}$.

Beispiel 1.17. $\mathbb{Q}$ ist abzählbar unendlich.

1.3 Folgen

Definition 1.18. Es sei $B \subseteq \mathbb{R}$. Existiert ein $c \in \mathbb{R}$, so dass $x \leqslant c$ für alle $x \in B$, heißt c *obere Schranke* von B und B ist *nach oben beschränkt.* Ist c_s eine obere Schranke von B und gilt $c_s \leqslant c$ für alle oberen Schranken c von B, dann nennt man c_s *kleinste obere Schranke* von B und bezeichnet sie als *Supremum* von B, geschrieben $\sup B$. Gibt es keine obere Schranke von B, so setzen wir $\sup B := \infty$. Falls die Zahl $\sup B$ endlich ist und zu B gehört, nennt man sie auch *Maximum* und schreibt $\max B$.

Die Definitionen von *unterer Schranke* von B, *größter untere Schranke* von B, dem so genannten *Infimum* $\inf B$ und dem Minimum von B, $\min B$ erfolgen analog.

Mit $-B := \{-x : x \in B\}$ gilt für alle $B \subseteq \mathbb{R}$:

$$\sup(-B) = -\inf B \quad \text{und} \quad \inf(-B) = -\sup B.$$

Aufgabe 1.19.

(i) Die Menge aller oberen Schranken von $B \subseteq \mathbb{R}$ ist entweder die leere Menge oder ein abgeschlossenes Intervall der Form $[c, \infty)$ für ein $c \in \mathbb{R}$.

(ii) Ist $B \subseteq \mathbb{R}$ nach oben beschränkt, dann gilt $\overline{x} = \sup B$ genau dann, wenn

  (a) $\overline{x} \geqslant y$ für alle $y \in B$ und

  (b) Für jedes $x < \overline{x}$ gibt es mindestens ein $y \in B$ mit $x < y \leqslant \overline{x}$.

Definition 1.20. Eine Abbildung $X : \mathbb{N} \to E, n \mapsto x_n$ heißt Folge aus E und wird mit $(x_n : n \in \mathbb{N}) = (x_n)_{n\in\mathbb{N}}$ bezeichnet.

Man unterscheide zwischen einer Folge $(x_n)_{n\in\mathbb{N}}$ und der Menge $\{x_n : n \in \mathbb{N}\}$, da es im Gegensatz zu einer Folge in einer Menge nicht auf die Reihenfolgen der Elemente ankommt!

Definition 1.21. Eine Folge $(x_n)_{n\geqslant 1}$ reeller Zahlen heißt *konvergent*, falls ein $x \in \mathbb{R}$ existiert, so dass für jedes $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ existiert, so dass $|x - x_n| < \varepsilon$ für alle $n > n_0$. In diesem Fall heißt x *Grenzwert* von (x_n) wird mit $\lim_{n\to\infty} x_n$ bezeichnet. Eine Folge heißt *divergent* gegen ∞, falls es für jedes $c > 0$ ein $n_0 = n_0(c)$ gibt, so dass $x_n > c$ für alle $n > n_0$ (analog Divergenz gegen $-\infty$).

Beispiel 1.22. Die Mengen der reellen Zahlen $\mathbb{R}$ ist Menge aller Grenzwerte von Folgen rationaler Zahlen, d.h. für alle $x \in \mathbb{R}$ existiert eine Folge $x_n \in \mathbb{Q}, n \in \mathbb{N}$ mit $x = \lim_{n\to\infty} x_n$. Beispielsweise gilt für die Eulersche Zahl

$$e = \lim_{n\to\infty} \big(1 + \frac{1}{n}\big)^n.$$

1.4 Produktmengen

Definition 1.23. Es seien $n \geqslant 1$ und $E_1, E_2, \dots E_n$ nichtleere Mengen. Als *Produktmenge*

$$E_1 \times E_2 \times \cdots \times E_n = \prod_{k=1}^{n} E_k$$

bezeichnen wir die Menge aller n-Tupel $x = (x_1, \dots, x_n)$ mit $x_k \in E_k, k = 1, \dots, n$. Gilt $E_k = E$ für alle $k = 1, \dots, n$ und eine Menge E schreiben wir kurz $E^n = \prod_{k=1}^{n} E_k$.

Ist $I := \{k_1, k_2, \dots, k_l\} \subseteq \{1, \dots, n\}$, so wird durch

$$\pi_I(x) := (x_{k_1}, \dots, x_{k_l}) \quad \text{für } x = (x_1, \dots, x_n) \in \prod_{k=1}^{n} E_k$$

die zur Menge I gehörende *Projektion* $\pi_I \colon \prod_{k=1}^{n} E_k \to \prod_{j=1}^{l} E_{k_j}$ definiert. Im Fall $l = 1$ und $I = \{k\}$ nennt man $\pi_k = \pi_{\{k\}}$die zum Index k gehörende *Koordinatenprojektion.*

Beispiel 1.24. Für $n \geqslant 1$ ist die Produktmenge

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ mal}}$$

der Raum aller Vektoren der Dimension n. Für $a = (a_1, \ldots, a_n)$ und $b = (b_1, \ldots, b_n)$ mit $-\infty \leqslant a_k \leqslant b_k \leqslant \infty, k = 1, \ldots, n$ heißt

$$(a, b] := \prod_{k=1}^{n} (a_k, b_k]$$

ein (nach links) halboffener Quader in $\mathbb{R}^n$ (für $n = 2$ sprechen wir von Rechteck)

Beispiel 1.25. Betrachten wir einen 10-fachen Würfelwurf, so ist die Menge aller möglichen Ergebnisse gegeben durch

$$\Omega = \{1, \ldots, 6\}^{10}.$$

Wetten wir auf das Ereignis "beim fünften Wurf fällt eine fünf" entspricht dies Menge

$$\{\pi_k = 5\} := \pi_5^{-1}(\{5\}) = \{\omega \in \Omega : \pi_k(\omega) = 5\} = \{(\omega_1, \ldots, \omega_{10}) \in \Omega : \omega_5 = 5\}.$$

Interessieren wir uns für die Augensumme, führt dies auf die Abbildung

$$X : \Omega \to \{10, 11, \ldots, 60\}, \quad \{\omega_1, \ldots, \omega_{10}\} \mapsto \sum_{k=1}^{10} \omega_k.$$

Das Ereignis "Die Augensumme ist 55" kann dann als $\{X = 55\} = X^{-1}(\{55\})$ modelliert werden.

Aufgabe 1.26. Es seien $E = E_1 \times E_2$ die Produktmenge zweier nichtleerer Mengen E_1, E_2 sowie $\pi_1$und π_2 die zugehörigen Koordinatenabbildungen. Seien $B_1 \subseteq E_1$ und $B_2 \subseteq E_2$ zwei beliebige Teilmengen.

(i) Bestimmen Sie die Urbildmengen $\pi_j^{-1}(B_j), j = 1, 2$.

(ii) Zeigen Sie $B_1 \times B_2 = \pi_1^{-1}(B_1) \cap \pi_2^{-1}(B_2)$.

1.5 Ergänzungen*

Definition 1.27. Ist $(x_n)_{n \geqslant 1}$ eine Folge aus $\mathbb{R}$, so ist $(\overline{x}_n)_{n \geqslant 1}$ mit $\overline{x}_n := \sup_{m \geqslant n} x_m$ eine monoton nichtwachsende Folge reeller Zahlen, d.h. es gilt $\overline{x}_n \geqslant \overline{x}_{n+1}, n \in \mathbb{N}$. Der Wert

$$\limsup_{n \to \infty} x_n := \inf_n \overline{x}_n = \inf_n \sup_{m \geqslant n} x_m$$

heißt *Limes superior* der Folge $(x_n)_n$. Analog ist $(\underline{x}_n)_{n \geqslant 1}$ mit $\underline{x}_n := \inf_{m \geqslant n} x_m$ eine monoton nichtfallende Folge reeller Zahlen, d.h. es gilt $\overline{x}_n \leqslant \overline{x}_{n+1}, n \in \mathbb{N}$ und

$$\liminf_{n \to \infty} x_n := \sup_n \underline{x}_n = \sup_n \inf_{m \geqslant n} x_m$$

heißt *Limes inferior* der Folge $(x_n)_n$.

Aufgabe 1.28. Es sei $(x_n)_{n \geqslant 1}$ eine Folge reeller Zahlen. Zeigen Sie, dass $c = \limsup_{n \to \infty} x_n$ genau dann gilt, wenn

(i) Für jedes $\varepsilon > 0$ und jedes $n \in \mathbb{N}$, gibt es mindestens ein $m \geqslant n$, so dass gilt $x_m > c - \varepsilon$ und

(ii) Für jedes $\varepsilon > 0$ gibt es ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass $x_n \leqslant c + \varepsilon$ für alle $n \geqslant n_0$.

Lemma 1.29. *Für jede Folge* $(x_n)_{n\geqslant 1}$ *reeller Zahlen gilt*

$$-\infty \leqslant \liminf_{n\to\infty} x_n \leqslant \limsup_{n\to\infty} x_n \leqslant \infty. \tag{1.1}$$

Die Gleichheit in (1.1) gilt genau dann, wenn $(x_n)_{n\geqslant 1}$ *konvergiert(divergiert) und in diesem Fall gilt* $\lim_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.

Beweis. Übung ☑. □

Kommen wir schließlich zu Folgen von Mengen:

Definition 1.30. Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge von Mengen, definieren wir den *Limes superior* bzw. den *Limes inferior von* (A_n) *durch*

$$\limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad \textit{bzw.} \quad \liminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Lemma 1.31. *Sei* (A_n) *ein Folge aus Teilmengen einer Menge* E. *Es gilt*

$$\begin{aligned}
\liminf_{n\to\infty} A_n &= \big\{x \in E : \textit{es gibt ein } n_0(x), \textit{ so dass } x \in A_n \textit{ für alle } n \geqslant n_0(x)\big\} \\
&= \big\{x \in E : x \in A_n \textit{ für alle außer endlich vielen } n \geqslant 1\big\}, \\
\limsup_{n\to\infty} A_n &= \big\{x \in E : \textit{für alle } n \geqslant 1 \textit{ gibt es ein } n_1(x) \geqslant n, \textit{ so dass } x \in A_{n_1(x)}\big\} \\
&= \big\{x \in E : x \in A_n \textit{ für unendlich viele } n \geqslant 1\big\}.
\end{aligned}$$

Insbesondere gilt damit $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$.

Beweis. Übung ☑. □

2 Algebren und σ-Algebren

2.1 Algebren

Wie bereits erwähnt modellieren wir die Gesamtheit aller möglichen Ausgänge eines Zufallsexperimentes durch eine Menge E. Alle uns interessierenden Ereignisse fassen wir durch ein Mengensystem $\mathcal{F}$ aus E zusammen.

Beispiel 2.1. Ein Polizist misst die Geschwindigkeit vorbeifahrender Fahrzeuge in einer 30 Zone. Dann ist $E = \mathbb{R}_+$. Aufgrund des Bußgeldkatalogs interessiert sich der Polizist für die Ereignisse $A_1 = (30, \infty)$ (zu schnell gefahren) , $A_2 = [51, 61)$ (ein Punkt), $A_3 = [61, \infty)$ (zwei Punkte) aber auch für

$$A_1^c, \quad A_2 \cup A_3, \quad A_1 \cap A_3^c.$$

Das Beispiel illustriert, dass es wünschenswert ist, wenn das Mengensystem der Ereignisse abgeschlossen unter den Operationen $\cup, \cap, ^c$ ist. Gleichzeit wäre im Allgemeinen $\mathscr{P}(\mathbb{R}_+)$ zu groß.

Definition 2.2. Es sei E eine nichtleere Menge und $\mathcal{A} \subseteq \mathscr{P}(E)$ ein Mengensystem aus E. Mann nennt $\mathcal{A}$ eine *Algebra* aus E, wenn

(i) $E \in \mathcal{A}$,

(ii) für alle $A \in \mathcal{A}$ gilt $A^c \in \mathcal{A}$ sowie

(iii) für alle $A, B \in \mathcal{A}$ gilt $A \cup B \in \mathcal{A}$.

Lemma 2.3. *Für jede Algebra $\mathcal{A}$ aus einer Menge E und alle $A, B \in \mathcal{A}$ gilt :*

$$\emptyset \in \mathcal{A}, \quad A \cap B \in \mathcal{A}, \quad A \setminus B \in \mathcal{A}.$$

Ist $A_1, \ldots, A_n \in \mathcal{A}$ so gilt stets

$$\bigcup_{k=1}^{n} A_k \in \mathcal{A} \quad \textit{und} \quad \bigcap_{k=1}^{n} A_k \in \mathcal{A}.$$

Beweis. Übung ☑. □

Beispiel 2.4. Sei E eine beliebige nichtleere Menge.

(i) Für $A \subseteq E$ ist $\{\emptyset, A, A^c, E\}$ eine Algebra.

(ii) Das Mengensystem

$$\mathcal{A} := \{A \subseteq E : A \text{ oder } A^c \text{ besteht aus endlichen vielen Elementen}\}$$

ist die kleinste Algebra aus E, die alle einelementigen Mengen $\{x\}, x \in E$, umfasst.

(iii) Für $d \geqslant 1$ bildet das folgende Mengensystem eine Algebra aus $\mathbb{R}^d$:

$$\mathcal{A} := \Big\{ \bigcup_{k=1}^{n} (a_k, b_k] : (a_k, b_k] := \prod_{l=1}^{d} (a_{k,l}, b_{k,l}] \text{ mit } -\infty \leqslant a_{k,l} \leqslant b_{k,l} \leqslant \infty, k = 1, \ldots, n, l = 1, \ldots, d$$
$$\text{und } (a_k, b_k] \text{ sind paarweise disjunkt} \Big\}.$$

Aufgabe 2.5.

(i) Zeigen Sie, dass der Durchschnitt $\bigcap_{i \in I} \mathcal{A}_i$beliebig vieler Algebren $\mathcal{A}_i, i \in I$, wieder eine Algebra ist.

(ii) Es seien $\mathcal{A}_n, n \geqslant 1$, Algebren aus E mit der Eigenschaft $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ für alle $n \geqslant 1$. Weisen Sie nach, dass

$$\mathcal{A}_\infty := \bigcup_{n=1}^{n} \mathcal{A}_n = \big\{B \subseteq E : \text{es gibt ein } n \geqslant 1 \text{ mit } B \in \mathcal{A}_n\big\}$$

eine Algebra ist.

2.2 σ-Algebren

Definition 2.6. Es sei E eine nichtleere Menge und $\mathcal{A} \subseteq \mathscr{P}(E)$ ein Mengensystem aus E. Mann nennt $\mathcal{A}$ eine *σ-Algebra* aus E, falls $\mathcal{A}$ eine Algebra ist und zusätzlich gilt

$$\text{falls } A_n \in \mathcal{A}, n \geqslant 1, \text{ so ist } \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Eine σ-Algebra ist also eine Algebra, die zusätzlich abgeschlossen ist bezüglich der Vereinigung (und damit wegen $\bigcap A_n = (\bigcup A_n^c)^c$ auch bezüglich der Durchschnittsbildung) von Folgen aus E.

Beispiel 2.7. Die Menge $\mathcal{A} := \{\bigcup_{k=1}^{n} (a_k, b_k] : -\infty \leqslant a_k \leqslant b_k \leqslant \infty, k = 1, \ldots, n, \text{ und } b_k \leqslant a_{k+1} \text{ für } k = 1, \ldots, n-1\}$ ist eine Algebra, vgl. Beispiel 2.4, aber keine σ-Algebra, denn

$$\bigcup_{n \geqslant 1} (0, 1 - \tfrac{1}{n}] = (0, 1) \notin \mathcal{A}.$$

Aufgabe 2.8. Zeigen Sie:

(i) Ist E eine nichtleere Menge, dann ist die Potenzmenge $\mathscr{P}(E)$ eine σ-Algebra aus E.

(ii) Ist $\{Z_n : n \in \mathbb{N}\}$ eine Zerlegung der Menge E, so bildet

$$\Big\{ \bigcup_{i \in I} Z_i : I \text{ durchläuft alle Teilemengen von } \mathbb{N} \Big\}$$

eine σ-Algebra. Es ist die kleinste σ-Algebra, die alle $Z_i, i \in \mathbb{N}$, umfasst.

Lemma 2.9. *Der Durchschnitt*

$$\mathcal{A} := \bigcap_{i \in I} \mathcal{A}_i = \big\{ B \subseteq E : B \in \mathcal{A}_i \text{ für alle } i \in I \big\}$$

jeder Menge $\{\mathcal{A}_i : i \in I\}$ von σ-Algebren aus einer Menge E ist ebenfalls eine σ-Algebra aus E.

Beweis. Es gilt nach Voraussetzung $E \in \mathcal{A}_i$ für alle $i \in I$ und somit $E \in \mathcal{A}$. Sind $A_n \in \mathcal{A}, n \geqslant 1$, gilt $A_n \in \mathcal{A}_i, n \geqslant 1$ und $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_i$ für alle $i \in I$. Es folgt $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Analog zeigt, dass $A^c \in \mathcal{A}$ für alle $A \in \mathcal{A}$ gilt. □

Satz 2.10. *Ist $\mathcal{S}$ irgendein Mengensystem aus E, so gilt*

(i) Das Mengensystem

$$\sigma(\mathcal{S}) := \bigcap \big\{ \mathcal{A} : \mathcal{A} \text{ ist } \sigma\text{-Algebra aus } E \text{ mit } \mathcal{S} \subseteq \mathcal{A} \big\} \tag{2.1}$$

ist eine σ-Algebra aus E, die $\mathcal{S}$ umfasst, d.h. $\mathcal{S} \subseteq \sigma(\mathcal{S})$.

(ii) $\sigma(\mathcal{S})$ ist die kleinste σ-Algebra mit $\mathcal{S} \subseteq \sigma(\mathcal{S})$, d.h. ist $\mathcal{B}$ eine σ-Algebra aus E mit $\mathcal{S} \subseteq \mathcal{B}$, so gilt $\sigma(\mathcal{S}) \subseteq \mathcal{B}$.

Beweis. Aus dem vorherigen Lemma folgt, dass $\sigma(\mathcal{S})$ eine σ-Algebra ist und nach Definition gilt $\mathcal{S} \subseteq \sigma(\mathcal{S})$.

Angenommen $\mathcal{B}$ eine σ-Algebra aus E mit $\mathcal{S} \subseteq \mathcal{B}$, dann folgt aus der Definition $\sigma(\mathcal{S}) \subseteq \mathcal{B}$, da $\mathcal{B}$ eine der σ-Algebren bei der Durchschnittsbildung auf der rechten Seite von (2.1) ist. □

Definition 2.11. Für jedes beliebige Mengensystem $\mathcal{S}$ aus E heißt $\sigma(\mathcal{S})$ aus (2.1) die von $\mathcal{S}$ *erzeugte σ-Algebra.* $\mathcal{S}$ heißt *Erzeuger* der σ-Algebra $\sigma(\mathcal{S})$.

Lemma 2.12. *Für eine nichtleere Menge E gilt:*

(i) Ist $\mathcal{A}$ eine σ-Algebra aus E, so ist $\mathcal{A} = \sigma(\mathcal{A})$.

(ii) Sind $\mathcal{S}, \mathcal{S}'$ Mengensysteme aus E so folgt aus $\mathcal{S} \subseteq \mathcal{S}'$, dass $\sigma(\mathcal{S}) \subseteq \sigma(\mathcal{S}')$ gilt.

Beweis. (i) Wegen $\mathcal{A} \subseteq \sigma(\mathcal{A})$ bleibt nur zu zeigen, dass $\sigma(\mathcal{A}) \subseteq \mathcal{A}$ gilt. Dies folgt, aber aus Satz 2.10(ii).

(ii) Für jede σ-Algebra $\mathcal{B}$ mit $\mathcal{S}' \subseteq \mathcal{B}$ gilt auch $\mathcal{S} \subseteq \mathcal{B}$. Damit sind bei der Durchschnittsbildung in (2.1) für $\mathcal{S}$ mindestens die σ-Algebren beteiligt die auch bei $\mathcal{S}'$ vorkommen. □

Beispiel 2.13. Betrachten wir die beiden Mengensysteme

$$\mathcal{S}_d := \Big\{ (a,b] := \prod_{l=1}^{d} (a_l, b_l] : -\infty \leqslant a_l \leqslant b_l \leqslant \infty, l = 1, \dots, d \Big\} \quad \text{und}$$

$$\mathcal{S}'_d := \Big\{ (a,b) := \prod_{l=1}^{d} (a_l, b_l) : -\infty \leqslant a_l \leqslant b_l \leqslant \infty, l = 1, \dots, d \Big\}. \tag{2.2}$$

Dann gilt

$$\sigma(\mathcal{S}_d) = \sigma(\mathcal{S}'_d).$$

Ist nämlich $(a,b) := \prod_{l=1}^d (a_l, b_l) \in \mathcal{S}_d^{'}$, so gilt für

$$(a, b_n] := \prod_{l=1}^d (a_l, b_{n,l}] \in \mathcal{S}_d \quad \text{mit} \quad b_{n,l} := \begin{cases} b_k - \frac{1}{n}, & b_k < \infty, \\ n, & b_k = \infty, \end{cases}$$

die Beziehung $(a,b) = \bigcup_{n \geqslant 1} (a, b_n] \in \sigma(\mathcal{S}_d)$. Folglich $\mathcal{S}'_d \subseteq \sigma(\mathcal{S}_d)$ und

$$\sigma(\mathcal{S}'_d) \subseteq \sigma(\sigma(\mathcal{S}_d)) = \sigma(\mathcal{S}_d).$$

Analog erhält man auf der Grundlage von $(a_k, b_k] = \bigcap_{n \geqslant 1} (a_k, b_k + \frac{1}{n})$, falls $b_k < \infty$ und wegen $(a_k, \infty] = (a_k, \infty)$ falls $b_k = \infty$ die Beziehung $\mathcal{S}_d \subseteq \sigma(\mathcal{S}'_d)$ und damit $\sigma(\mathcal{S}_d) \subseteq \sigma(\mathcal{S}'_d)$.

Definition 2.14. Jedes Paar $(E, \mathcal{A})$, wobei E eine nichtleere Menge ist und $\mathcal{A}$ eine σ-Algebra aus E bildet, heißt *messbarer Raum*. Eine Teilmenge $A \subseteq E$ heißt *messbar,* falls $A \in \mathcal{A}$ gilt. Sind keine Verwechslungen möglich, so spricht man einfach von messbaren Mengen.

2.3 Borelmengen

Definition 2.15. Für $d \geqslant 1$ und $\mathcal{S}_d := \{\prod_{l=1}^d (a_l, b_l] : -\infty \leqslant a_l \leqslant b_l \leqslant \infty, l = 1, \dots, d\}$ heißt $\mathscr{B}_{\mathbb{R}^d} := \sigma(\mathcal{S}_d)$ die *Borelsche σ-Algebra* aus $\mathbb{R}^d$. Deren Elemente werden *Borelmengen* genannt.

Beispiel 2.16. Jede einelementige Menge $\{x\}, x \in \mathbb{R}^d$ ist eine Borelmenge, da $\{x\} = \bigcap_n (x - \frac{1}{n}, x]$. Damit ist auch jede Menge $B \subseteq \mathbb{R}^d$ mit endlich vielen oder abzählbar vielen Elementen eine Borelmenge (insbesondere $\mathbb{Q}^d$).

Definition 2.17. Eine Menge $U \subseteq \mathbb{R}^d$ heißt *offen*, wenn man für jedes $x \in U$ einen Quader $Q \in \mathcal{S}'_d$, vgl. (2.2), finden kann mit $x \in Q \subseteq U$. Eine Menge $V \subseteq \mathbb{R}^d$ heißt *abgeschlossen*, falls ihr Komplement $V^c = \mathbb{R}^d \setminus V$ offen ist.

Aufgabe 2.18. Weisen Sie nach, dass $\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ offen ist.

Aus der Definition folgt sofort, dass die Vereinigung beliebig vieler offener Mengen in $\mathbb{R}^d$ ebenfalls offen ist. Geht man zu Komplementen über, ergibt sich, dass der Durchschnitt beliebig vieler abgeschlossener Mengen abgeschlossen ist.

Satz 2.19. *Es sei $d \geqslant 1$.*

(i) *Im $\mathbb{R}^d$ sind alle Quader der Form $(a,b) = \prod_{k=1}^d (a_k, b_k)$ aus $\mathcal{S}'_d$ offen und alle Quader der Form $[a,b] = \prod_{k=1}^d [a_k, b_k]$ mit $-\infty \leqslant a \leqslant b \leqslant \infty, k = 1, \dots, d$, sind abgeschlossen.*

(ii) *Jede offene und jede abgeschlossene Menge im $\mathbb{R}^d$ ist eine Borelmenge.*

(iii) *Bezeichnen $\mathcal{U}_d$ bzw. $\mathcal{V}_d$ die Systeme aller offenen bzw. abgeschlossenen Mengen im $\mathbb{R}^d$, so gilt*

$$\sigma(\mathcal{U}_d) = \sigma(\mathcal{V}_d) = \mathscr{B}_{\mathbb{R}^d}.$$

Beweis. *(i)* Für alle $x = (x_1, \dots, x_d) \in (a,b)$ ist (a,b) selbst ein Quader aus $\mathcal{S}_d^{'}$, der x enthält, so dass (a,b) offen ist.

Im Fall $d = 1$ ist $[a,b]^c = (-\infty, a) \cup (b, \infty)$ offen. Für $d \geqslant 2$ gilt

$$[a,b]^c = \bigcup_{k=1}^d \pi_k^{-1}\big((-\infty, a_k) \cup (b_k, \infty)\big).$$

Wegen

$$\pi_k^{-1}((-\infty, a_k) \cup (b_k, \infty)) = (-\infty, \infty) \times \cdots \times (-\infty, \infty) \times (-\infty, a_k) \times (-\infty, \infty) \times \cdots \times (-\infty, \infty)$$
$$\cup\, (-\infty, \infty) \times \cdots \times (-\infty, \infty) \times (b_k, \infty) \times (-\infty, \infty) \times \cdots \times (-\infty, \infty)$$

ist $[a, b]^c$ eine Vereinigung offener Mengen, also offen.

(ii) Sei U eine offene Menge. Zu jeden $x \in U$ existiert ein offener Quader $Q = (a, b) \in \mathcal{S}'_d$ mit $x \in Q \subseteq U$. Da jede reelle Zahl der Grenzwert einer Folge von rationalen Zahlen ist existiert ein Quader $Q' = (a', b')$ mit $a', b' \in \mathbb{Q}^d$, so dass $x \in Q' \subseteq Q \subseteq U$. Da $\mathbb{Q}^d$ eine abzählbare Menge ist, folgt

$$U \subseteq \bigcup \big\{(a, b) : a, b \in \mathbb{Q}^d \text{ mit } (a, b) \subseteq U\big\} \subseteq U.$$

Daher ist U die abzählbare Vereinigung von offenen Mengen und somit $U \in \mathscr{B}_{\mathbb{R}^d}$. Folglich liegen auch die Komplemente aller offenen Mengen, also die abgeschlossenen Mengen, in $\mathscr{B}_{\mathbb{R}^d}$.

(iii) Aus $\mathcal{S}'_d \subseteq \mathcal{U}_d \subseteq \mathscr{B}_{\mathbb{R}^d}$ folgt $\mathscr{B}_{\mathbb{R}^d} = \sigma(\mathcal{S}'_d) \subseteq \sigma(\mathcal{U}_d) \subseteq \sigma(\mathscr{B}_{\mathbb{R}^d}) = \mathscr{B}_{\mathbb{R}^d}$. Beachte schließlich $\sigma(\mathcal{U}_d) = \sigma(\mathcal{V}_d)$. □

Bemerkung 2.20. Die Borelsche σ-Algebra ist also die kleinste σ-Algebra aus $\mathbb{R}^d$, die alle offenen (bzw. alle abgeschlossenen) Teilmengen von $\mathbb{R}^d$ umfasst. Es sei erwähnt, dass es Borelmengen gibt, die nicht durch abzählbare Vereinigungen oder Durchschnitte von Intervallen dargestellt werden könne.

Wir werden sehen, dass die messbaren Räume $(\mathbb{R}^d, \mathscr{B}_{\mathbb{R}^d})$ sehr gut geeignete Objekte für die Maßtheorie im $\mathbb{R}^d$ (und somit für die Stochastik) sind. Dagegen umfasst die Potenzmenge $\mathscr{P}(\mathbb{R}^d)$ zu viele Elemente, um eine fruchtbare Maßtheorie für den messbaren Raum $(\mathbb{R}^d, \mathscr{P}(\mathbb{R}^d))$ zu entwickeln.

2.4 σ-Algebren und Abbildungen

Lemma 2.21. *Es sei X eine Abbildung von einer nichtleeren Mengen E in einen messbaren Raum $(F, \mathcal{F})$.*

(i) Dann ist das Mengensystem

$$\mathcal{A}^X := X^{-1}(\mathcal{F}) := \{X^{-1}(A) : A \in \mathcal{F}\}$$

eine σ-Algebra aus E. $\mathcal{A}^X$ wird als die von X erzeugte σ-Algebra *in E bezeichnet.*

(ii) Für jedes Mengensystem $\mathcal{S}$ in F gilt

$$\sigma(X^{-1}(\mathcal{S})) = X^{-1}(\sigma(\mathcal{S})).$$

Beweis. (i) Es gilt $E \in \mathcal{A}^X$, wegen $X^{-1}(F) = E$. Die übrigen Eigenschaften folgen aus der Operationstreue der Urbildabbildung.

(ii) Aus $X^{-1}(\mathcal{S}) \subseteq X^{-1}(\sigma(\mathcal{S}))$ folgt $\sigma(X^{-1}(\mathcal{S})) \subseteq X^{-1}(\sigma(\mathcal{S}))$. Für die umgekehrte Inklusion, setzen wir $\mathcal{B} := \{C \subseteq F : X^{-1}(C) \in \sigma(X^{-1}(\mathcal{S}))\}$. Das Mengensystem $\mathcal{B}$ ist eine σ-Algebra in F und nach Definition von $\mathcal{B}$ gilt $\mathcal{S} \subseteq \mathcal{B}$. Also ist $\sigma(\mathcal{S}) \subseteq \mathcal{B}$ und somit gilt $X^{-1}(\sigma(\mathcal{S})) \subseteq X^{-1}(\mathcal{B}) \subseteq \sigma(X^{-1}(\mathcal{S}))$. □

Aufgabe 2.22. Wir setzen $(F, \mathcal{F}) = (\mathbb{R}, \mathscr{B}_{\mathbb{R}})$. Bestimmen Sie die von X erzeugten σ-Algebren $X^{-1}(\mathscr{B}_{\mathbb{R}})$ für folgenden Abbildungen:

(i) Betrachten Sie für eine nichtleere Menge E und $A \subseteq E$ die Abbildung $X = \mathbb{1}_A$ gegeben durch die *Indikatorfunktion*

$$\mathbb{1}_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

(ii) Sei $E = \mathbb{R}$ und $X(y) = e^y, y \in E$.

3 Maße

Maße sind Abbildungen von einer σ-Algebra nach $[0,\infty]$, die intuitiv jeder Menge aus der σ-Algebra ein "Größe" zuordnen, bspw. Flächen oder Volumina. Es ist intuitiv, dass diese Abbildungen nicht-negativ und additiv sein sollten. In der Stochastik dienen Maße dazu den Ereignissen Wahrscheinlichkeiten zuzuordnen, was den Wertebereich von Wahrscheinlichkeitsmaßen in natürlich Weise auf $[0,1]$ beschränkt.

3.1 Definition und Eigenschaften

Definition 3.1. Es seien E eine nichtleere Menge und $\mathcal{A}$ eine Algebra aus E. Jede Abbildung $\mu\colon \mathcal{A} \to [0,\infty]$ heißt *Mengenfunktion.* μ heißt *additiv*, falls $\mu(\emptyset) = 0$ und falls für jedes $n \in \mathbb{N}$ und für jede Folge paarweise disjunkter Mengen $A_1, \dots, A_n \in \mathcal{A}$ gilt

$$\mu\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} \mu(A_k).$$

Die Mengenfunktion μ heißt *σ-additiv*, falls sie additiv ist und falls für jede abzählbare Folge $(A_k)_{k\geqslant 1}$ paarweise disjunkter Mengen aus $\mathcal{A}$ mit $\bigcup_{k\geqslant 1} A_k \in \mathcal{A}$ gilt

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mu(A_k).$$

Satz 3.2. *Für jede additive Mengenfunktion μ auf einer Algebra $\mathcal{A}$ aus einer Menge E gelten folgende Aussagen:*

*(i) Für alle $A, B \in \mathcal{A}$ mit $A \subseteq B$ und $\mu(A) < \infty$ gilt $\mu(B \setminus A) = \mu(B) - \mu(A)$, insbesondere $\mu(A) \leqslant \mu(B)$ (*Monotonie*).*

(ii) Für alle $A, B \in \mathcal{A}$ mit $\mu(A \cap B) < \infty$ gilt $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

(iii) Die beiden folgenden Eigenschaften sind äquivalent:

(a) Die additive Mengenfunktion μ ist σ-additiv.

(b) Für jede Folge $(A_n)_{n\geqslant 1} \subseteq \mathcal{A}$ mit $A_n \subseteq A_{n+1}, \mu(A_n) < \infty$ für $n \geqslant 1$ und $\bigcup_{n\geqslant 1} A_n \in \mathcal{A}$, gilt

$$\lim_{n\to\infty} \mu(A_n) = \mu\Big(\bigcup_{n\geqslant 1} A_n\Big) \quad \textit{(Stetigkeit von unten).}$$

(iv) Die beiden folgenden Eigenschaften sind äquivalent:

(a) Für jede Folge $(A_n)_{n\geqslant 1} \subseteq \mathcal{A}$ mit $A_n \supseteq A_{n+1}, n \geqslant 1$, und $\mu(A_1) < \infty$ und $\bigcap_{n\geqslant 1} A_n \in \mathcal{A}$, gilt

$$\lim_{n\to\infty} \mu(A_n) = \mu\Big(\bigcap_{n\geqslant 1} A_n\Big) \quad \textit{(Stetigkeit von oben).}$$

*(b) Für jede Folge $(A_n)_{n\geqslant 1} \subseteq \mathcal{A}$ mit $A_n \supseteq A_{n+1}, n \geqslant 1$, und $\mu(A_1) < \infty$ und $\bigcap_{n\geqslant 1} A_n = \emptyset$ gilt $\lim_{n\to\infty} \mu(A_n) = 0$ (*Stetigkeit in der leeren Menge*).*

(v) Die Eigenschaften in (iii) implizieren (iv). Im Fall $\mu(E) < \infty$ sind die Eigenschaften in (iii) und (iv) äquivalent.

(vi) Ist μ σ-additiv, so gilt für jede Folge $(A_n) \subseteq \mathcal{A}$ mit $\bigcup_{n\geqslant 1} A_n \in \mathcal{A}$ die Ungleichung

$$\mu\Big(\bigcup_{n\geqslant 1} A_n\Big) \leqslant \sum_{n=1}^{\infty} \mu(A_n) \quad \textit{(Subadditivität).}$$

Beweis. Siehe Aussage 3.2 aus Küchler [3]. □

Definition 3.3. Es sei $(E, \mathcal{A})$ ein messbarer Raum und μ eine Mengenfunktion auf $\mathcal{A}$.

(i) Ist μ σ-additiv, so nennt man μ ein *Maß* und das Triplet $(E, \mathcal{A}, \mu)$ ein *Maßraum*.

(a) Gilt $\mu(E) < \infty$, dann heißt μ *endliches Maß*. Gilt $\mu(E) = 1$, heißt μ *normiert*.

(b) Gibt es eine Folge $(E_n)_{n\geqslant 1} \subseteq \mathcal{A}$ mit $\bigcup_{n\geqslant 1} E_n = E$ und $\mu(E_n) < \infty$ für alle $n \geqslant 1$, dann nennt man μ *σ-endliches Maß*.

Normierte Maße heißen auch *Wahrscheinlichkeitsmaße* und ein Maßraum mit einem normiert Maß heißt *Wahrscheinlichkeitsraum*.

Beispiel 3.4. Es sei $\lambda > 0$, $E := \mathbb{N} \cup \{0\}, \mathcal{A} = \mathscr{P}(E)$. Durch

$$\mu(B) = \sum_{k\in B} \frac{\lambda^k}{k!}, \quad B \subseteq E,$$

ist ein endliches Maß auf $(E, \mathcal{A})$ mit $\mu(E) = e^{\lambda}$ gegeben. Mit der Normierung $\mu^*(A) := e^{-\lambda}\mu(A), A \subseteq E$, wird $(E, \mathcal{A}, \mu^*)$ zu einem Wahrscheinlichkeitsraum.

Aufgabe 3.5. Es sei $(E, \mathcal{A}, \mu)$ ein Maßraum und $C \in \mathcal{A}$. Wir definieren

$$\mu|_C(B) := \mu(B \cap C), B \in \mathcal{A}, \quad \text{und} \quad \mathcal{A}_C := \{A \in \mathcal{A} : A \subseteq C\}.$$

Zeigen Sie, dass $(C, \mathcal{A}_C, \mu|_C)$ ebenfalls ein Maßraum ist.

Definition 3.6. Eine messbare Menge A aus einem Maßraum $(E, \mathcal{A}, \mu)$ heißt *μ-Nullmenge*, falls $\mu(A) = 0$.

Aufgabe 3.7. Ist $(A_n, n \geqslant 1)$ eine Folge von μ-Nullmengen aus einem Maßraum $(E, \mathcal{A}, \mu)$, dann ist auch $\bigcup_{n\geqslant 1} A_n$ eine μ-Nullmenge.

3.2 Konstruktion von Maßen

Da man im Allgemeinen nicht einmal alle Elemente der σ-Algebra eines messbaren Raumes $(E, \mathcal{A})$ kennt (z.B. Borelmengen), erscheint es hoffnungslos $\mu(A)$ für alle $A \in \mathcal{A}$ angeben zu können. Wie gelangt man dann aber zum konkreten Maß? Wir beginnen mit diskreten Maßen.

Definition 3.8. Es sei $M \subseteq \mathbb{N}$, E eine beliebige Menge und $Y := \{y_n : n \in M\} \subseteq E$ eine endliche oder abzählbar unendliche Teilmenge von E. Weiterhin sei $(q_n, n \in M)$ eine Folge positiver reeller Zahlen. Durch

$$\mu(B) := \sum_{n:y_n\in B} q_n, \quad B \subseteq E,$$

ist auf der σ-Algebra $\mathscr{P}(E)$ ein Maß μ definiert. Damit gilt $\mu(\{y_n\}) = q_n, n \in M$, und $\mu(E\backslash Y) = 0$. Das Maß μ heißt *diskretes Maß* und Y heißt *Träger* von μ.

Gilt $q_n = 1, n \geqslant 1$, dann heißt μ *Zählmaß* auf Y, da

$$\mu(B) = \text{Anzahl der Elemente } y \in Y \text{, die auch in } B \text{ liegen.}$$

Diskrete Maße sind also durch $(y_n, q_n)_{n\in M}$ eindeutig definiert und leicht zu handhaben.

Beispiel 3.9. Gilt $\sum_{n\in M} q_n = 1$ ist μ ein Wahrscheinlichkeitsmaß und wird folglich *diskretes Wahrscheinlichkeitsmaß (Wahrscheinlichkeitsverteilung)* genannt. Ist $Y = \{y_1, \ldots, y_n\}$ endlich und $q_k = 1/n$, so heißt das dadurch gegebene Wahrscheinlichkeitsmaß *gleichmäßige Verteilung* auf $\{y_1, \ldots, y_n\}$. Beispiele sind der faire Münz- und Würfelwurf oder Lotto “6 aus 49”.

Aufgabe 3.10. Sei μ ein diskretes Maß auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ mit Träger $\mathbb{N}$. Bestimmen Sie die μ-Nullmengen.

Statt Maße wie bisher explizit auf gegeben σ-Algebren zu definieren, werden Maße im Allgemeinen nur auf kleineren Mengensystemen vorgeben und müssen dann im Sinne der folgenden Definition fortgesetzt werden.

Definition 3.11. Ist ν ein σ-additive Mengenfunktion auf einem System $\mathcal{S} \subseteq \mathscr{P}(E)$ und gibt es ein Maß μ auf einer σ-Algebra $\mathcal{A}$ auf E mit $\mathcal{S} \subseteq \mathcal{A}$ und $\nu(A) = \mu(A)$ für alle $A \in \mathcal{S}$, so heißt μ *Fortsetzung* von ν von $\mathcal{S}$ auf $\mathcal{A}$.

Es folgt einer der zentralen Sätze der Maßtheorie:

Satz 3.12 (Fortsetzungssatz von Caratheodory)**.** *Es sei $E \neq \emptyset$ eine Menge und μ eine σ-additive σ-endliche Mengenfunktion auf einer Algebra $\mathcal{A}^0 \subseteq \mathscr{P}(E)$. Dann gibt es ein eindeutig bestimmtes σ-endliches Maß $\widetilde{\mu}$ auf $(E, \sigma(\mathcal{A}^0))$ für das gilt*

$$\widetilde{\mu}(A) = \mu(A) \textit{ für alle } A \in \mathcal{A}^0.$$

Beweisskizze. Für alle $B \subseteq E$ definieren wir

$$\mu^*(B) := \inf\Big\{ \sum_{n=1}^{\infty} \mu(B_n) : B_n \in \mathcal{A}^0, n \geqslant 1, B \subseteq \bigcup_{n=1}^{\infty} B_n \Big\}$$

mit $\inf \emptyset := 0$. Die Mengenfunktion $\mu^* \colon \mathscr{P}(E) \to [0, \infty]$ heißt das von μ erzeugte *äußere Maß*. Weiterhin sei

$$\mathcal{A}_{\mu^*} := \big\{ B \in \mathscr{P}(E) : \mu^*(C) = \mu^*(C \cap B) + \mu^*(C \cap B^c) \text{ für alle } C \subseteq E \text{ mit } \mu^*(C) < \infty \big\}.$$

Beweis. Man kann nun zeigen, das für das Mengensystem $\mathcal{A}_{\mu^*}$ und das äußere Maß μ^* gilt:

(i) $\mathcal{A}_{\mu^*}$ ist eine σ-Algebra,

(ii) $\mathcal{A}^0 \subseteq \mathcal{A}_{\mu^*}$ und folglich $\sigma(\mathcal{A}^0) \subseteq \mathcal{A}_{\mu^*}$,

(iii) die Einschränkung $\mu^*|_{\mathcal{A}_{\mu^*}}$ ist ein Maß auf $(E, \mathcal{A}_{\mu^*})$,

(iv) $\mu^*(B) = \mu(B)$ für alle $B \in \mathcal{A}^0$, d.h. μ^* ist eine Fortsetzung von μ auf die σ-Algebra $\mathcal{A}_{\mu^*}$,

(v) μ^* ist die einzige σ-endliche Fortsetzung von μ auf $\sigma(\mathcal{A}^0)$.

□

Als Anwendung des Fortsetzungssatzes wollen wir endliche Maße auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ durch ihre Verteilungsfunktionen charakterisieren.

□

Definition 3.13. Für jedes endliche Maß μ auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ ist die *Verteilungsfunktion* $F \colon \mathbb{R} \to \mathbb{R}_+$ von μ definiert durch

$$F(x) := \mu((-\infty, x]), \quad x \in \mathbb{R}.$$

Es gilt also insbesondere $\mu((a, b]) = F(b) - F(a)$.

Beispiel 3.14. Sei μ ein diskretes Maß auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ gegeben durch $(y_n, q_n)_{n \geqslant 1}$. Dann ist seine Verteilungsfunktion gegeben durch

$$F(x) = \mu((-\infty, x]) = \sum_{k : y_k \leqslant x} q_k.$$

(konstant auf $[y_k, y_{k+1})$ mit Sprunghöhen q_k bei y_k).

Lemma 3.15. *Sei F die Verteilungsfunktion eines endlichen Maßes μ auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$. Dann gilt:*

(i) *F ist monoton nichtfallend: Für alle $x, y \in \mathbb{R}$ gilt $x \leqslant y \Rightarrow F(x) \leqslant F(y)$.*

(a) *$F(-\infty) := \lim_{x\downarrow -\infty} F(x) = 0$ und $F(\infty) := \lim_{x\uparrow\infty} F(x) = \mu(\mathbb{R}) < \infty$.*

(b) *F ist von rechts stetig: $F(x) = \lim_{y\downarrow x} F(y)$ für alle $x \in \mathbb{R}$.*

(c) *Für alle $x \in \mathbb{R}$ gilt $F(x) - \lim_{y\uparrow x} F(y) = \mu(\{x\})$.*

Beweis. Übung ☑. □

Definition 3.16. Erfüllt eine Funktion $F\colon \mathbb{R} \to \mathbb{R}_+$ die Eigenschaften (i) bis (iii) aus Lemma 3.15, nennt man sie ebenfalls *Verteilungsfunktion.*

Sein nun $F\colon \mathbb{R} \to \mathbb{R}_+$ eine Verteilungsfunktion. Auf der Algebra

$$\mathcal{A}^0 := \Big\{ \bigcup_{k=1}^n (a_k, b_k] : -\infty \leqslant a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \ldots \leqslant b_n \leqslant \infty, n \in \mathbb{N} \Big\}$$

betrachten wir die Mengenfunktion

$$\mu_F(A) := \sum_{k=1}^n \big(F(b_k) - F(a_k)\big), \quad A = \bigcup_{k=1}^n (a_k, b_k] \in \mathcal{A}^0. \tag{3.1}$$

Lemma 3.17. *Die durch (3.1) definierte Mengenfunktion μ_F ist σ-additiv.*

Beweis. Wir zeigen hier nur die Additivität und verweisen für die σ-Additivität auf Küchler [3, Aussage 3.10]. Für $A = \bigcup_{k=1}^n (a_k, b_k], A' = \bigcup_{k=1}^m (a'_k, b'_k]$ disjunkte Mengen aus $\mathcal{A}^0$ mit $n, m \in \mathbb{N}$ seinen $(c_k)_{k=1,\ldots,n+m}$ bzw. $(d_k)_{k=1,\ldots,n+}$. die monoton wachsend geordneten Folgen aus $\{a_1, \ldots, a_n, a'_1, \ldots, a'_m\}$ bzw. $\{b_1, \ldots, b_k, b'_1, \ldots, b'_m\}$. Dann gilt

$$A \cup B = \bigcup_{k=1}^{n+m} (c_k, d_k]$$

und somit

$$\begin{aligned} \mu_F(A \cup B) &= \sum_{k=1}^{n+m} \big(F(d_k) - F(c_k)\big) \\ &= \sum_{k=1}^{n} \big(F(b_k) - F(a_k)\big) - \sum_{k=1}^{m} \big(F(b'_k) - F(a'_k)\big) = \mu_F(A) + \mu_F(B). \end{aligned}$$ □

Korollar 3.18. *Die σ-additive Mengenfunktion μ_F von (3.1) auf $\mathcal{A}^0$ besitzt eine eindeutig bestimmte Fortsetzung auf die σ-Algebra der Borelmengen $\mathscr{B}_\mathbb{R} = \sigma(\mathcal{A}^0)$. Damit gibt es für jede Verteilungsfunktion $F\colon \mathbb{R} \to \mathbb{R}_+$ genau ein endliches Maß μ auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$.*

Beispiel 3.19. Für $a, b \in \mathbb{R}$ mit $a < b$ ist

$$L(y) := \begin{cases} 0, & \text{falls } y \leqslant a, \\ y - a, & \text{falls } y \in (a, b), \\ b - a, & \text{falls } y \geqslant b \end{cases} \tag{3.2}$$

die Verteilungsfunktion eines Maßes $\mu_L =: \lambda_{(a,b]}$ auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$. Für jedes Intervall $(c, d] \subseteq (a, b]$ gilt $\lambda_{(a,b]}((c, d]) = d - c$ und jede zu $[a, b]$ disjunkte Borelmenge besitzt das $\lambda_{(a,b]}$-Maß Null. Die Wahrscheinlichkeitsverteilung gegeben durch das normierte Maß $\mu^* := \frac{1}{b-a}\lambda_{(a,b]}$ heißt *gleichmäßige Verteilung* auf $[a, b]$.

Aufgabe 3.20. Zeigen Sie, dass $\lambda_{(a,b]}(\{x\}) = 0$ für alle $x \in (a,b]$ gilt. Ist jede abzählbare Teilmenge von $(a,b]$ eine $\lambda_{(a,b]}$-Nullmenge?

Aufgabe 3.21. Es seien F eine Verteilungsfunktion auf $\mathbb{R}$, definiert durch

$$F(x) = \frac{1}{4}\mathbb{1}_{[0,\infty)}(x) + \frac{1}{2}\mathbb{1}_{[1,\infty)}(x) + \frac{1}{4}\mathbb{1}_{[2,\infty)}(x), \quad x \in \mathbb{R},$$

und μ_F das durch F bestimmte Maß auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$. Berechnen Sie das μ_F-Maß folgender Mengen:

$$[1,\infty), \quad (-\frac{2}{3}, \frac{3}{2}), \quad [0,2), \quad (\frac{2}{3}, \frac{5}{2}), \quad (3,\infty), \quad [0,1] \cup [2,5).$$

3.3 Das Lebesguemaß

Für $a < b$ sei $\lambda_{(a,b]}$ das Maß welches durch die Verteilungsfunktion (3.2) bestimmt wird. Mit $I_k := (k, k+1], k \in \mathbb{Z}$, gilt

$$\sum_{k=-n}^{n} \lambda_{I_k} = \lambda_{(-n,n+1]}, \quad n \geqslant 1.$$

Satz 3.22. *Durch*

$$\lambda(B) := \sup\left\{\lambda_{(-n,n+1]}(B) : n \geqslant 1\right\}, \quad B \in \mathscr{B}_\mathbb{R},$$

ist ein σ-endliches Maß λ auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ definiert für das $\lambda([a,b]) = b-a$ für alle $a,b \in \mathbb{R}$ mit $a \leqslant b$ gilt.

Beweis. Seien $B_n \subseteq \mathbb{R}, n \geqslant 1$, paarweise disjunkte Borelmengen und $B := \bigcup_{n\geqslant 1} B_n$. Da $(\lambda_{(-n,n+1]}(B))_{n\geqslant 1}$ eine nichtfallende Folge bildet, gilt

$$\lambda(B) = \lim_{n\to\infty} \lambda_{(-n,n+1]}(B) = \lim_{n\to\infty} \sum_{k=-n}^{n} \lambda_{I_k}(B) = \sum_{k\in\mathbb{Z}} \lambda_{I_k}(B).$$

Die σ-Additivität der Maße λ_{I_k} ergibt

$$\sum_{k\in\mathbb{Z}} \lambda_{I_k}(B) = \sum_{k\in\mathbb{Z}} \sum_{n=1}^{\infty} \lambda_{I_k}(B_n).$$

Da alle Summanden nichtnegativ sind, können wir die beiden Summen vertauschen (großer Umordnungssatz) und erhalten

$$\sum_{k\in\mathbb{Z}} \sum_{n=1}^{\infty} \lambda_{I_k}(B_n) = \sum_{n=1}^{\infty} \sum_{k\in\mathbb{Z}} \lambda_{I_k}(B_n) = \sum_{n=1}^{\infty} \lambda(B_n).$$

Damit ist die Mengenfunktion λ σ-additiv. Da λ auf der σ-Algebra $\mathscr{B}_\mathbb{R}$ definiert ist, handelt es sich um ein Maß. Wegen $\lambda((-n,n]) < \infty$ und $\bigcup_{n\geqslant 1}(-n,n] = \mathbb{R}$, ist λ auch σ-endlich.

Für alle $a \leqslant b$ gibt es ein hinreichend großes $N \in \mathbb{N}$ mit $[a,b] \subseteq (-N,N]$. Für diese N gilt

$$\lambda([a,b]) = \sup\left\{\lambda_{(-n,n+1]}([a,b])\right\} = \lambda_{(-N-1,N+2]}([a,b]) = L(b) - \lim_{y\uparrow a} L(a) = b-a. \qquad \square$$

Definition 3.23. Das Maß λ aus Satz 3.22 heißt *Lebesguemaß*.

Für das Lebesguemaß gilt die Translationsinvarianz $\lambda(B + \{x\}) = \lambda(B)$ für alle $B \in \mathscr{B}_\mathbb{R}$, wobei $B + \{x\} := \{y + x : y \in B\}$.

Definition 3.24. Ein Maßraum $(E, \mathcal{A}, \mu)$ heißt *vollständig*, falls jede Teilmenge einer μ-Nullmenge aus $\mathcal{A}$ ebenfalls zu $\mathcal{A}$ gehört (und damit auch eine μ-Nullmenge ist).

Beispiel 3.25. Der im Beweis des Fortsetzungssatzes konstruierte Maßraum $(E, \mathcal{A}_{\mu^*}, \mu^*|_{\mathcal{A}_{\mu^*}})$ ist vollständig.

Bemerkung 3.26 (Vervollständigung). Ist $(E, \mathcal{A}, \mu)$ ein Maßraum und $\mathcal{N}$ die Menge aller Teilmengen von μ-Nullmengen aus $\mathcal{A}$, dann ist

$$\overline{\mathcal{A}} := \{B \cup N : B \in \mathcal{A}, N \in \mathcal{N}\}$$

eine σ-Algebra. $(E, \overline{\mathcal{A}})$ wird *Vervollständigung* von $(E, \mathcal{A})$ bezüglich μ genannt. Durch

$$\overline{\mu}(B \cup N) := \mu(B), \quad B \in \mathcal{A}, N \in \mathcal{N},$$

wird auf $\overline{\mathcal{A}}$ mein Maß $\overline{\mu}$ definiert. $\overline{\mu}$ setzt μ von $\mathcal{A}$ auf $\overline{\mathcal{A}}$ fort und $(E, \overline{\mathcal{A}}, \overline{\mu})$ ist ein vollständiger Maßraum.

Definition 3.27. Ist λ das Lebesguemaß auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$, so bezeichnet man die Vervollständigung von$(\mathbb{R}, \mathscr{B}_\mathbb{R})$ bezüglich λ mit $(\mathbb{R}, \mathscr{L}_\mathbb{R})$ und $\mathscr{L}_\mathbb{R}$ heißt σ-Algebra der *Lebesguemessbaren Mengen*. Das vervollständigte Maß $\overline{\lambda}$ auf $\mathscr{L}_\mathbb{R}$ heißt ebenfalls *Lebesguemaß*.

4 Messbare Abbildungen

4.1 Definition und Eigenschaften

Es seien $(E, \mathcal{A})$ und $(F, \mathcal{B})$ zwei messbare Räume und $X\colon E \to F$ eine Abbildung.

Definition 4.1. Die Abbildung X heißt $(\mathcal{A}, \mathcal{B})$-*messbar* (wenn keine Verwechslungen möglich sind, auch kürzer $\mathcal{A}$-messbar, oder einfach messbar), falls $X^{-1}(\mathcal{B}) \subseteq \mathcal{A}$, d.h.

$$X^{-1}(B) \in \mathcal{A} \quad \text{für alle } B \in \mathcal{B}. \tag{4.1}$$

Ist $(F, \mathcal{B}) = (\mathbb{R}^d, \mathscr{B}_{\mathbb{R}^d})$, so heißt jede $(\mathcal{A}, \mathscr{B}_{\mathbb{R}^d})$-messbare Abbildung von E nach $\mathbb{R}^d$ auch *Borel-messbar*.

Beispiel 4.2. Gegeben sei der messbare Raum $(E, \mathcal{A})$.

(i) Sind $\alpha \in \mathbb{R} \setminus \{0\}, A \subseteq E$, ist die Funktion $X(y) = \alpha \mathbb{1}_A(y), y \in E$, $\mathcal{A}$-messbar genau dann, wenn $A \in \mathcal{A}$. Insbesondere ist jede konstante Funktion messbar.

(ii) Gilt $A_1, \ldots, A_n \in \mathcal{A}$ und $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, so ist

$$X(y) = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{A_k}(y), \quad y \in E, \tag{4.2}$$

$\mathcal{A}$-messbar. Funktionen von der Gestalt (4.2) nennen wir *einfache Funktionen*. Man prüft leicht nach, das für einfache Funktionen X und Y auch $\alpha X + \beta Y$ eine einfache Funktion ist für alle $\alpha, \beta \in \mathbb{R}$.

Lemma 4.3. *Es seien E eine Menge, $(F, \mathcal{B})$ ein messbarer Raum und $X\colon E \to F$ eine Abbildung. Dann ist $\sigma(X) := X^{-1}(\mathcal{B})$ die kleinste σ-Algebra $\mathcal{A}$ aus E, so dass X eine $(\mathcal{A}, \mathcal{B})$-messbare Abbildung ist. Wir nennen $\sigma(X)$ die von X* erzeugte σ-Algebra*.*

Beweis. Dass $\sigma(X)$ eine σ-Algebra ist, hatten wir bereits in Lemma 2.21 gezeigt. Nach Konstruktion von $\sigma(X)$ ist X $(\sigma(X), \mathcal{B})$-messbar. Ist andererseits $\mathcal{B}$ eine σ-Algebra aus E so, dass X $(\mathcal{A}, \mathcal{B})$-messbar ist, folgt $\sigma(X) = X^{-1}(\mathcal{B}) \subseteq \mathcal{A}$. □

Im Allgemeinen ist Messbarkeit schwer nachzuprüfen. Oft hilft folgendes Kriterium:

Satz 4.4. *Es sei $\mathcal{S}$ ein Erzeuger von $\mathcal{B}$, d.h. $\mathcal{B} = \sigma(\mathcal{S})$. Dann ist X $(\mathcal{A}, \mathcal{B})$-messbar genau dann, wenn*

$$X^{-1}(\mathcal{S}) \subseteq \mathcal{A}. \tag{4.3}$$

Beweis. Aus (4.1) folgt offensichtlich (4.3). Andererseits folgt aus Lemma 2.21, dass

$$X^{-1}(\mathcal{B}) = X^{-1}(\sigma(\mathcal{S})) = \sigma(X^{-1}(\mathcal{S})) \subseteq \sigma(\mathcal{A}) = \mathcal{A}. \qquad \square$$

Korollar 4.5. *Es sei $(E, \mathcal{A})$ ein messbarer Raum und $X, X_n, n \geqslant 1$, Abbildungen von E nach $\mathbb{R}$.*

(i) $X \colon E \to \mathbb{R}$ ist genau dann Borel-messbar, wenn für alle $x \in \mathbb{R}$ gilt

$$\{y \in E : X(y) \leqslant x\} \in \mathcal{A}.$$

(ii) Sind $X_n \colon E \to \mathbb{R}, n \geqslant 1$, Borel-messbare, so sind es auch

$$\sup_{n \geqslant 1} X_n, \quad \inf_{n \geqslant 1} X_n, \quad \limsup_{n \to \infty} X_n, \quad \liminf_{n \to \infty} X_n$$

(hierbei folgt die Bildung von $\sup X_n, \inf X_n$ etc. punktweise, d.h. $(\sup_n X_n)(y) := \sup_n X_n(y)$).

(iii) Sind $X_n \colon E \to \mathbb{R}$ Borel-messbar und konvergiert X_n punktweise gegen X, dann ist X Borel-messbar.

Beweisskizze. (i) $\sigma(\{(-\infty, a] : a \in \mathbb{R}\}) = \mathscr{B}_{\mathbb{R}}$, da $(a, b] = (-\infty, b] \setminus (-\infty, a]$ für alle $a, b \in \mathbb{R}$.
(ii) $\{\sup_n X_n \leqslant x\} = \bigcap_n \{X_n \leqslant x\}$ und $\{\inf_n X_n \leqslant x\} = \bigcap_m \{\inf_n X_n < x + \frac{1}{m}\} = \bigcap_m \bigcup_n \{X_n < x + \frac{1}{m}\}$.
(iii) Es gilt $X = \limsup_{n \to \infty} X_n = \liminf_{n \to \infty} X_n$. $\square$

Bemerkung 4.6. Man kann auch zeigen, dass jede stetige Funktion $X \colon \mathbb{R} \to \mathbb{R}$ Borel-messbar ist. Die umgekehrte Implikation gilt nicht, da die *Dirichletsche Funktion* $X(y) = \mathbb{1}_{\mathbb{Q}}(y), y \in \mathbb{R}$, messbar, aber nirgendwo stetig ist.

Die nächsten Aussagen ermöglichen es uns, aus messbaren Funktionen weitere zu bilden.

Lemma 4.7. *Es seien $X \colon (E, \mathcal{A}) \to (F, \mathcal{B})$ und $Y \colon (F, \mathcal{B}) \to (G, \mathcal{C})$ messbare Abbildungen. Dann ist die zusammengesetzte Abbildung*

$$Z \colon E \to G, \quad y \mapsto Y(X(y))$$

(kurz $Z = Y \circ X$) eine $(\mathcal{A}, \mathcal{C})$-messbare Abbildung.

Beweis. Übung ☑ . $\square$

Lemma 4.8.

(i) Sind $X_1, X_2, \ldots, X_n$ reellwertige Borel-messbare Funktionen auf $(E, \mathcal{A})$ und ist $h \colon \mathbb{R}^n \to \mathbb{R}^d$ Borel-messbar, so ist $h(X_1, \ldots, X_n)$ eine $\mathcal{A}$-messbare Funktion.

(ii) Sind X, Y reellwertig und Borel-messbar, so sind es auch $X + Y, X \cdot Y, \max(X, Y), \min(X, Y)$ und, falls $Y \neq 0$, $\frac{X}{Y}$.

Beweis. Küchler [3, Aussage 4.9]. $\square$

Der folgende Satz zeigt, dass man jede messbare Funktion (zerlegbar in zwei nichtnegative Funktionen $X = X_+ - X_-$) durch einfache Funktionen beliebig genau angenähert (approximiert) werden kann. Dieses Resultat wird zentral für die Konstruktion von Integralen sein.

Satz 4.9 (Approximationssatz). *Für jede nichtnegative Borel-messbare Funktion $X\colon (E,\mathcal{A}) \to (\mathbb{R}_+, \mathscr{B}_{\mathbb{R}_+})$ gibt es eine nichtfallende Folge $(X_n)_{n\geqslant 1}$ einfacher nichtnegativer Borel-messbarer Funktionen, die punktweise von unten gegen X konvergiert:*

$$0 \leqslant X_n(y) \leqslant X_{n+1}(y) \leqslant X(y), \quad n \geqslant 1, y \in E$$

und

$$\lim_{n\to\infty} X_n(y) = X(y), \quad y \in E.$$

Beweis. Wir setzen

$$X_n(y) := \begin{cases} k \cdot 2^{-n}, & \text{falls } X(y) \in [k2^{-n}, (k+1)2^{-n}] \text{ und } k = 0, \dots, n2^n - 1, \\ n, & \text{falls } X(y) \geqslant n. \end{cases} \tag{4.4}$$

Die X_n sind einfache, insbesondere Borel-messbare, Funktionen mit $X_n(y) \leqslant X_{n+1}(y) \leqslant X(y)$ und

$$|X_n(y) - X(y)| \mathbb{1}_{\{X \leqslant n\}}(y) \leqslant 2^{-n}$$

für alle $y \in E$. Also gilt $\lim_{n\to\infty} X_n(y) = X(y)$ für alle $y \in E$. □

Aufgabe 4.10. Bestimmen Sie für die Funktion $X\colon \mathbb{R} \to \mathbb{R}, y \mapsto y^2$ und $n \in \mathbb{N}$ die in (4.4) konstruierte einfache Funktion und stellen Sie diese graphisch dar.

4.2 Produkt-σ-Algebren

Produktmengen und Produkt-σ-Algebren dienen in der Wahrscheinlichkeitstheorie zur Modellierung mehrerer gleichzeitig oder nacheinander ausgeführter zufälliger Versuche.

Für zwei messbare Räume $(E_i, \mathcal{A}_i), i = 1, 2$, ist die Produktmenge gegeben als

$$E := E_1 \times E_2 = \{(y_1, y_2)^\top : y_1 \in E_1, y_2 \in E_2\}$$

und die Koordinatenprojektionen

$$\pi_i \colon E \to E_i, \quad \pi_i(y) = y_i \quad \text{für} \quad y = (y_1, y_2) \in E, i = 1, 2.$$

Definition 4.11. Die kleinste σ-Algebra $\mathcal{A}$ auf E, bezüglich der beide π_i jeweils $(\mathcal{A}, \mathcal{A}_i)$-messbar sind, heißt die *Produkt-σ-Algebra* von $\mathcal{A}_1$ und $\mathcal{A}_2$. Sie wird mit mit

$$\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$$

bezeichnet.

Satz 4.12. *$\mathcal{S}_1$ und $\mathcal{S}_2$ seien Erzeuger von $\mathcal{A}_1$ bzw. $\mathcal{A}_2$. Dann wird $\mathcal{A}_1 \otimes \mathcal{A}_2$ vom System $\mathcal{S}_1 \times \mathcal{S}_2$ erzeugt, d.h.*

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{S}_1 \times \mathcal{S}_2) = \sigma\big(\{S_1 \times S_2 : S_i \in \mathcal{S}_i, i = 1, 2\}\big).$$

Beweis. Küchler [3, Aussage 4.13]. □

Auf analoge Weise definiert man für gegeben messbare Räume $(E_i, \mathcal{A}_i), i = 1, \dots, n$ (oder sogar größere Indexmengen), die Produkt-σ-Algebra

$$\mathcal{A} = \bigotimes_{i=1}^{n} \mathcal{A}_i$$

als die kleinste σ-Algebra auf dem Produktraum $E = \prod_{i=1}^n E_i$ für die Koordinatenprojektionen $\pi_i \colon (E, \mathcal{A}) \to (E_i, \mathcal{A}_i), (y_1, \dots, y_n) \mapsto y_i$ für alle $i = 1, \dots, n$ messbar sind. Gilt $(E_i, \mathcal{A}_i) = (E_1, \mathcal{A}_1)$ für alle $i = 1, \dots, n$, so schreibt man kurz $\bigotimes_{i=1}^n \mathcal{A}_i = \mathcal{A}_1^{\otimes n}$.

Bemerkung 4.13. Im Fall $(E_i, \mathcal{A}_i) = (\mathbb{R}, \mathscr{B}_{\mathbb{R}}), i = 1, \dots, d$, erhalten wir $\mathscr{B}_{\mathbb{R}}^{\otimes d} = \mathscr{B}_{\mathbb{R}^d}$.

4.3 Induzierte Maße

Satz 4.14. *Es seien $(E, \mathcal{A}, \mu)$ ein Maßraum und $X : E \to F$ eine $(\mathcal{A}, \mathcal{B})$-messbare Abbildung in einen messbaren Raum $(F, \mathcal{B})$. Durch*

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

ist auf $\mathcal{B}$ ein Maß μ^x definiert, das als von X induzierte Maß *oder als* Bildmaß *von X bezeichnet wird. Ist μ endlich, so ist auch μ^X endlich.*

Beweis. Es gilt $\mu^X(\emptyset) = 0$ und für jede Folge $(B_k)_{k\geqslant 1} \subseteq \mathcal{F}$ paarweiser diskjunkter Mengen

$$\mu^X\Big(\bigcup_{k\geqslant 1} B_k\Big) = \mu\Big(X^{-1}\big(\bigcup_{k\geqslant 1} B_k\big)\Big) = \mu\Big(\bigcup_{k\geqslant 1} X^{-1}(B_k)\Big) = \sum_{k\geqslant 1} \mu(X^{-1}(B_k)) = \sum_{k\geqslant 1} \mu^X(B_k). \qquad \square$$

Damit $X^{-1}(B) \in \mathcal{A}$ und somit obiger Ausdruck für alle $B \in \mathcal{B}$ wohldefiniert ist, benötigen wir also genau die Messbarkeit von X.

Beispiel 4.15. Es sei $(F, \mathcal{B}) = (\mathbb{R}, \mathscr{B}_{\mathbb{R}})$. Ist der Wertebereich von X diskret, d.h. $\{X(y) : y \in E\} = \{x_k, k \geqslant 1\}$ für reelle Zahlen x_k, so ist μ^x ein diskretes Maß mit

$$\mu^X(\{x_k\}) = \mu\big(\{y \in E : X(y) = x_k\}\big) = \mu\big(\{X = x_k\}\big), \quad k \geqslant 1.$$

Messbare Abbildung X auf Wahrscheinlichkeitsräumen heißen auch Zufallsvariablen und das Bildmaß μ^X wird auch Wahrscheinlichkeitsverteilung der Zufallsvariable X genannt.

Aufgabe 4.16. Wir betrachten das Modell $\big(E, \mathscr{P}(E), \mu\big)$ mit $E = \{1, 2, 3, 4, 5, 6\}^2$ und Wahrscheinlichkeitsmaß $\mu(\{(k_1, k_2)\}) = \frac{1}{36}, (k_1, k_2) \in E$, des zweifachen Würfelwurfes. Wir interessieren uns für die Augensumme $X : E \to F, (k_1, k_2) \mapsto k_1 + k_2$, mit $F = \{2, 3, \dots, 12\}$ und σ-Algebra $\mathcal{B} = \mathscr{P}(F)$ auf F.

(i) Geben Sie die Elemente der σ-Algebra $\sigma(X) = X^{-1}(\mathcal{F})$ an.

(ii) Bestimmen Sie der Wahrscheinlichkeitsverteilung μ^X.

Aufgabe 4.17. Zeigen Sie: Wenn $(E, \mathcal{A}, \mu)$ ein diskreter Maßraum ist und wenn $X : (E, \mathcal{A}) \to (F, \mathcal{B})$ eine messbare Abbildung auf einen messbaren Raum $(F, \mathcal{B})$ ist, so ist auch μ^X ein diskretes Maß.

5 Integrale

Um Erwartungswerte, Varianzen, Kovarianzen, etc. von Zufallsgrößen zu berechnen benötigen wir einen Integralbegriff auf Wahrscheinlichkeitsräumen $(E, \mathcal{A}, \mu)$. Diesen einzuführen und dessen Eigenschaften kennen zu lernen ist Ziel dieses Kapitels. Ist $X \in E$ eine zufällige Größe, verteilt gemäß eines Wahrscheinlichkeitsmaßes μ auf einem messbaren Raum $(E, \mathcal{A})$, so ist ihr Erwartungswert gegeben durch $\mathbb{E}[X] := \int_E x\mu(\mathrm{d}x)$.

5.1 Konstruktion

Wir betrachten zunächst einen Maßraum $(E, \mathcal{A}, \mu)$ mit endlichem Maß μ. Wir beginnen mit einfachen reellwertigen Funktionen

$$X(y) = \sum_{i=1}^{m} a_i \mathbb{1}_{A_i}(y), \quad y \in E, \tag{5.1}$$

für $m \in \mathbb{N}, a_1, \dots, a_m \in \mathbb{R}$ und $A_1, \dots, A_m \in \mathcal{A}$.

Definition 5.1. Als *Integral* der einfachen Funktion X aus (5.1) über E bezüglich μ bezeichnet man die Zahl

$$\int_E X \mathrm{d}\mu := \int_E X(y)\mu(\mathrm{d}y) := \sum_{i=1}^m a_i \mu(A_i).$$

Bemerkung 5.2.

(i) Das Integral hängt nicht von der Wahl der Darstellung (5.1) ab.

(ii) Für einfache Funktionen X, Y mit $X(y) \leqslant Y(y), y \in E$, gilt

$$\int_E X \mathrm{d}\mu \leqslant \int_E Y \mathrm{d}\mu \quad \text{(Montonie des Integrals).}$$

(iii) Für einfache Funktionen X, Y und $a, b \in \mathbb{R}$ gilt

$$\int_E (aX + bY)\mathrm{d}\mu = a\int_E X \mathrm{d}\mu + b\int_E Y \mathrm{d}\mu \quad \text{(Linearität des Integrals).}$$

Ist nun allgemeiner $X\colon E \to \mathbb{R}_+$ eine nichtnegative Borel-messbare Abbildung, dann besagt der (Approximations-)Satz 4.9, dass es eine Folge einfacher nichtnegativer Funktionen $(X_n)_{n\geqslant 1}$ gibt mit $X_n \leqslant X_{n+1}$ und $\lim_{n\to\infty} X_n = X$ (beides punktweise). Aus der Monotonie folgt, dass

$$\Big(\int_E X_n \mathrm{d}\mu\Big)_{n\geqslant 1}$$

eine nichtfallende Folge (aus $[0, \infty]$) ist, die damit einen Grenzwert besitzt (ggf. ∞).

Definition 5.3. Als *Integral* der nichtnegativen Borel-messbaren Abbildung $X\colon E \to \mathbb{R}_+$ über E bezüglich μ bezeichnen wir

$$\int_E X \mathrm{d}\mu := \int_E X(y)\mu(\mathrm{d}y) := \lim_{n\to\infty}\int_E X_n(y)\mu(\mathrm{d}y).$$

Die Funktion X heißt dabei *Integrand*.

Für $\mathbb{R}$-wertige Borel-messbare Abbildungen $X\colon E \to \mathbb{R}$ verwenden wir die Zerlegung

$$X(y) = X_+(y) - X_-(y), \quad X_+(y) = \max\{0, X(y)\}, X_-(y) = -\min\{0, X(y)\}, \quad y \in E.$$

Definition 5.4. Eine Borel-messbare Abbildung $X\colon E \to \mathbb{R}$ heißt *μ-integrierbar*, wenn $\int_E X_+ \mathrm{d}\mu < \infty$ oder $\int_E X_- \mathrm{d}\mu < \infty$ gilt (oder beides endlich ist). In diesem Fall ist das *(Lebesgue)-Integral* definiert als

$$\int_E X \mathrm{d}\mu := \begin{cases} \int_E X_+ \mathrm{d}\mu - \int_E X_- \mathrm{d}\mu, & \text{falls } \int_E X_+ \mathrm{d}\mu < \infty \text{ und } \int_E X_- \mathrm{d}\mu < \infty, \\ +\infty, & \text{falls } \int_E X_+ \mathrm{d}\mu = \infty \text{ und } \int_E X_- \mathrm{d}\mu < \infty, \\ -\infty, & \text{falls } \int_E X_+ \mathrm{d}\mu < \infty \text{ und } \int_E X_- \mathrm{d}\mu = \infty. \end{cases}$$

Gilt $\int_E X_+ \mathrm{d}\mu = \int_E X_- \mathrm{d}\mu = \infty$, so nennen wir X *nicht intergierbar* bezüglich μ.

Ohne Beweis bemerken wir, dass der Wert $\int_E X \mathrm{d}\mu$ nicht von der Folge der approximierenden einfachen Funktionen abhängt.

Definition 5.5. Ist X eine μ-integrierbare Funktion und ist $A \in \mathcal{A}$, so ist offensichtlich auch $X \cdot \mathbb{1}_A$ μ-integrierbar und wir schreiben

$$\int_A X \mathrm{d}\mu := \int_E X \mathbb{1}_A \mathrm{d}\mu.$$

A heißt *Integrationsbereich* über den X integriert wird.

Wir benötigen schließlich noch den Integralbegriff für den Fall, dass μ ein σ-finites Maß ist. In diesem Fall sei $(B_n)_{n\geqslant 1}$ eine Zerlegung von E in messbare Mengen B_n mit $\mu(B_n) < \infty$. Für nichtnegative messbare Funktionen $X\colon E \to \mathbb{R}_+$ definieren wir

$$\int_E X \mathrm{d}\mu := \sum_{n=1}^{\infty} \int_{B_n} X \mathrm{d}\mu$$

und erweitern diesen Integralbegriff auf beliebige messbare Funktionen X wie oben durch Zerlegung $X = X_+ + X_-$.

Definition 5.6. Ist $(E, \mathcal{A}, \mu)$ ein Maßraum mit σ-endlichem Maß μ, bezeichnen wir mit $\mathcal{L}^1 = \mathcal{L}^1(E, \mathcal{A}, \mu)$ die Menge aller μ-integrierbaren Borel-messbaren Funktionen $X\colon E \to \mathbb{R}$ mit $\int X \mathrm{d}\mu < \infty$.

Bemerkung 5.7. Das Lebesgueintegral beruht auf der Annäherung von messbaren Abbildungen mittels einfacher Funktionen. Diese Approximation durch einfach Funktionen beruht auf einer Zerlegung des Wertebereichs, vgl. Satz 4.9. Im Gegensatz dazu wird das in der Schule gelehrte *Riemannintegral* von Funktionen $X\colon \mathbb{R} \to \mathbb{R}$ durch eine Zerlegung des Definitionsbereiches konstruiert.

Es stellt sich heraus, dass jede Riemannintegrierbare (z.B. stetige) Funktion auf einem Intervall $[a, b]$ auch Lebesgueintegrierbar bzgl. des Lebesguemaßes $\lambda_{[a,b]}$ ist und beide Integrale übereinstimmen. Da das Lebesgueintegral auch die Integration bezüglich sehr beliebiger σ-endlicher Maße zulässt, ist es sehr viel allgemeiner als das Riemannintegral.

5.2 Eigenschaften

In diesem Abschnitt sei $(E, \mathcal{A}, \mu)$ stets ein Maßraum mit σ-endlichem Maß. Viele grundlegende Eigenschaften folgen direkt aus der Definition des Integrals.

Satz 5.8. *Für $X, Y \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ gelten folgende Aussagen:*

(i) Für alle $a, b \in \mathbb{R}$ ist $aX + bY \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ und es gilt

$$\int_E (aX + bY)\mathrm{d}\mu = a \int_E X \mathrm{d}\mu + b \int_E Y \mathrm{d}\mu.$$

(ii) Wenn $X \leqslant Y$, so ist $\int X \mathrm{d}\mu \leqslant \int Y \mathrm{d}\mu$. Insbesondere gilt $\int_E Y \mathrm{d}\mu \geqslant 0$, falls $Y \geqslant 0$.

(iii) Es gilt

$$\Big| \int_E X(y)\mu(\mathrm{d}y) \Big| \leqslant \int_E |X(y)|\mu(\mathrm{d}y) < \infty.$$

(iv) Wenn $\mu(\{y \in E : X(y) \neq Y(y)\}) = 0$, so ist $\int_E X \mathrm{d}\mu = \int_E Y \mathrm{d}\mu$.

(v) Aus $X \geqslant 0$ und $\int_E X \mathrm{d}\mu = 0$ folgt $\mu(\{X \neq 0\}) = \mu(\{y \in E : X(y) \neq 0\}) = 0$.

Beweis. Küchler [3, Satz 5.5]. □

Bemerkung 5.9. Gilt eine Aussage auf einer Menge $A \in \mathcal{A}$ mit $\mu(A^c) = 0$, so sagen wir, die Aussage gilt μ-fast überall. (v) kann man also schreiben als: Aus $X \geqslant 0$ und $\int_E X \mathrm{d}\mu = 0$ folgt $X = 0$ μ-fast überall.

Die folgende wichtige Eigenschaft besagt, dass man zur Berechnung von $\int (h \circ X)\mathrm{d}\mu$ nur das Bildmaß μ^X kennen muss, nicht die Abbildung X noch das Maß μ selbst. Dies ist in der Statistik von fundamentaler Bedeutung, da wir häufig den zugrundeliegenden Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ nicht kennen, sondern nur abgeleitete Größen, die so genannten Zufallsvariablen, $X\colon \Omega \to F$ beobachten und daher (unter Verwendung von Stichproben) nur die Wahrscheinlichkeitsverteilung $\mathbb{P}^X$ zugänglich ist.

Satz 5.10 (Substitutionsregel)**.** *Es sei $X\colon E \to F$ eine $(\mathcal{A}, \mathcal{B})$-messbare Funktion in einen messbaren Raum $(F, \mathcal{B})$, $h\colon F \to \mathbb{R}$ eine Borel-messbare Funktion und μ^x das Bildmaß von μ vermittels der Abbildung X. Dann gilt*

(i) $h \circ X = h(X) \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ genau dann, wenn $h \in \mathcal{L}^1(F, \mathcal{B}, \mu^X)$ und

(ii) falls $h \geqslant 0$ oder $h \in\in \mathcal{L}^1(F, \mathcal{B}, \mu^X)$, dann

$$\int_E h(X(y))\mu(\mathrm{d}y) = \int_F h(x)\mu^X(\mathrm{d}x). \tag{5.2}$$

Beweis. Schritt 1: Die Gleichung (5.2) gilt für $h(x) = \mathbb{1}_B(x), B \in \mathcal{B}, x \in F$, nach Definition des Bildmaßes μ^x:

$$\begin{aligned}\int_E h(x)\mu^X(\mathrm{d}x) &= \mu^X(B) = \mu(X^{-1}(b))\\ &= \int_E \mathbb{1}_{X^{-1}(B)}(y)\mu(\mathrm{d}y) = \int_E \mathbb{1}_B(X(y))\mu(\mathrm{d}y) = \int_E h(X(y))\mu(\mathrm{d}y).\end{aligned}$$

Schritt 2: Aufgrund der Linearität des Integrals gilt (5.2) für alle einfachen Funktionen.

Schritt 3: Ist $h \geqslant 0$ und h_n eine wachsende Folge einfacher Funktionen mit $h(x) = \lim_{n\to\infty} h_n(x), x \in F$, so ist auch $h_n \circ X$ eine wachsende Folge einfacher Funktionen und es gilt nach Definition des Integrals

$$\int_F h\mu^X(\mathrm{d}x) = \lim_{n\to\infty} \int_F h_n\mu^X(\mathrm{d}x) = \lim_{n\to\infty} \int_E (h_n \circ X)\mathrm{d}\mu = \int_E (h \circ X)\mathrm{d}\mu.$$

Schritt 4: (i) und (ii) für den allgemeinen Fall folgt aus der Zerlegung $h = h_+ - h_-$. □

Die hier verwendete Beweistechnik kommt häufig zum Einsatz und wird mitunter "Lifting-Methode" oder "maßtheoretische Induktion" genannt.

5.3 Zwei Beispiele

5.3.1 Integrale bezüglich diskreter Maße

Es sei nun $(E, \mathcal{A}, \mu)$ ein diskreter endlicher Maßraum, d.h. es gibt Folgen $(y_n : n \in M) \subseteq E, (q_n : n \in M) \subseteq (0, \infty)$ mit $M \subseteq \mathbb{N}$ und $\sum_{n\in M} q_n < \infty$, so dass

$$\mu(B) = \sum_{n\in M: y_n \in B} q_n \quad \text{für alle } B \subseteq \mathcal{A} := \mathscr{P}(E)$$

gilt. Offenbar ist jede Funktion $X\colon E \to \mathbb{R}$ $(\mathscr{P}(E), \mathscr{B}_{\mathbb{R}})$-messbar.

Beispiel 5.11 (Diracmaß)**.** Das Diracmaß (in 0) auf $(\mathbb{R}, \mathscr{P}(\mathbb{R}))$ ist gegeben durch

$$\delta_0(B) = \begin{cases} 1, & 0 \in B, \\ 0, & \text{sonst,} \end{cases} \quad B \in \mathscr{P}(\mathbb{R}).$$

Es hat Träger $Y = \{0\}$ und somit ist $\mathbb{R} \setminus \{0\}$ ein δ_0-Nullmenge.

Mit $Y := \{y_n : n \in M\}$ gilt wegen $\mu(Y^c) = 0$

$$X = \mathbb{1}_Y \cdot X \quad \mu\text{-fast überall.}$$

Satz 5.8 liefert

$$\int_E X\mathrm{d}\mu = \int_E \mathbb{1}_Y X\mathrm{d}\mu.$$

Zur Konstruktion einer Folge einfacher approximierender Funktionen, setzen wir $Y_m := \{y_n : n \in M, n \leqslant m\}$ und betrachten

$$y \mapsto \mathbb{1}_{Y_m}(y)X(y) = \sum_{n\in M: n\leqslant m} \mathbb{1}_{\{y_n\}}(y)X(y_n).$$

Dann gilt

$$\int_E \mathbb{1}_{Y_m} X \mathrm{d}\mu = \sum_{n\in M: n\leqslant m} X(y_n)q_n.$$

Ist $X \geqslant 0$, so folgt für $m \to \infty$

$$\int_E X \mathrm{d}\mu = \sum_{n\in M} X(y_n)q_n. \tag{5.3}$$

Für beliebige X benutzt man wieder die Zerlegung $X = X_+ - X_-$. Wir haben damit folgenden Satz bewiesen:

Satz 5.12. *Es sei $(E, \mathscr{P}(E), \mu)$ ein diskreter endlicher Maßraum mit $\mu(B) = \sum_{n\in M: y_n\in B} q_n$, $B \subseteq E$. Dann ist $X\colon E \to \mathbb{R}$ genau dann μ-integrierbar mit endlichem Integral $\int_E X \mathrm{d}\mu$ gegeben durch (5.3), wenn $\sum_{n\in M} |X(y_n)|q_n < \infty$.*

Aufgabe 5.13. Bestimmen Sie die erwartete Augensumme eines zweifachen Würfelwurfes.

5.3.2 Integrale bezüglich absolutstetiger Maße auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$

Ist $(E, \mathcal{A}) = (\mathbb{R}, \mathscr{B}_\mathbb{R})$ und λ das Lebesguemaß auf $\mathscr{B}_\mathbb{R}$, so schreiben wir auch

$$\int_\mathbb{R} X(y)\mathrm{d}y := \int_\mathbb{R} X(y)\lambda(\mathrm{d}y).$$

Aufgabe 5.14. Überprüfen Sie, ob die Funktionen f und g gegeben durch

$$\begin{aligned} f(y) &= 2\mathbb{1}_{[0,\frac{3}{4}]}(y) + 4\mathbb{1}_{[\frac{1}{4},1]}(y), \quad y \in [0,1], \\ g(y) &= 2\mathbb{1}_{[0,\frac{1}{4}]}(y) + 6\mathbb{1}_{[\frac{1}{4},\frac{3}{4}]}(y) + 4\mathbb{1}_{(\frac{3}{4},1]}(y), \quad y \in [0,1], \end{aligned}$$

auf dem Intervall $[0,1]$ gleich sind. Berechnen Sie die Integrale $\int_0^1 f(y)\mathrm{d}y$ und $\int_0^1 g(y)\mathrm{d}y$.

Definition 5.15. Es seien μ ein σ-endliches Maß auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ und $f\colon \mathbb{R} \to \mathbb{R}_+$ eine nichtnegative Funktion mit

$$\mu((a,b]) = \int_{(a,b]} f(x)\lambda(\mathrm{d}x) = \int_{(a,b]} f(x)\mathrm{d}x =: \int_a^b f(x)\mathrm{d}x, \quad -\infty \leqslant a < b \leqslant \infty.$$

Dann heißt das Maß μ *absolutstetig (bezüglich des Lebesguemaßes)* und f heißt *(Lebesgue-)Dichte* von μ.

Beispiel 5.16.

(i) Das Lebesguemaß λ auf einem Intervall $[a,b]$ hat die Dichte $f(x) = \mathbb{1}_{[a,b]}(x), x \in \mathbb{R}$.

(ii) Für festes $m \in \mathbb{R}$ und $\sigma^2 > 0$ ist die Dichte des so genannten Gaußschen Maßes μ (Normalverteilung $\mathcal{N}(m,\sigma^2)$) gegeben durch

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{(x-m)^2}{2\sigma^2}\Big), \quad x \in \mathbb{R}.$$

Für die Verteilungsfunktion $F\colon \mathbb{R} \to \mathbb{R}_+$ jedes endlichen absolutstetigen Maßes μ auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ gilt dann

$$F(x) = \mu((-\infty, x]) = \int_{-\infty}^{x} f(y)\mathrm{d}y, \quad x \in \mathbb{R}.$$

Ist f eine (stückweise) stetige Funktion, gilt nach dem *Hauptsatz der Differenzial- und Integralrechnung* $F'(x) = f(x)$ für λ-fast alle $x \in \mathbb{R}$.

Satz 5.17. *Es sei μ ein σ-endliches Maß auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ mit Dichte $f\colon \mathbb{R} \to \mathbb{R}_+$. Falls $h\colon \mathbb{R} \to \mathbb{R}$ eine Borel-messbare Funktion ist mit $h \geqslant 0$ oder $\int |h(y)|\mu(\mathrm{d}y) < \infty$, so gilt*

$$\int_\mathbb{R} h(y)\mu(\mathrm{d}y) = \int_\mathbb{R} hf\mathrm{d}\lambda = \int_\mathbb{R} h(y)f(y)\mathrm{d}y.$$

Insbesondere gilt $\mu(B) = \int_B f(y)\mathrm{d}y$ für alle $B \in \mathscr{B}_\mathbb{R}$.

Beweis. Man verwendet die maßtheoretische Induktion, wobei die Aussage für $h = \mathbb{1}_{(a,b]}$ nach Definition einer Dichte gilt. □

5.4 L^p-Räume*

Im Folgenden sei $(E, \mathcal{A}, \mu)$ stets ein σ-endlicher Maßraum.

Definition 5.18. Für $p \geqslant 1$ bezeichnen wir mit $\mathcal{L}^p = \mathcal{L}^p(E, \mathcal{A}, \mu)$ die Menge aller reellwertigen Borel-messbaren Funktionen $X\colon E \to \mathbb{R}$ mit

$$\int_E |X(y)|^p \mu(\mathrm{d}y) < \infty.$$

Für jedes $X \in \mathcal{L}^p$ definieren wir

$$\|X\|_p := \Big(\int_E |X|^p \mathrm{d}\mu\Big)^{1/p}.$$

Ferner definieren wir $\mathcal{L}^\infty := \mathcal{L}^\infty(E, \mathcal{A}, \mu)$ als Menge aller Borel-messbaren Funktionen $X\colon E \to \mathbb{R}$ mit

$$\|X\|_\infty := \sup_{y \in E} |X(y)| < \infty.$$

Die $\mathcal{L}^p$-Räume erfüllen zwei wesentliche Ungleichungen.

Satz 5.19 (Hölder-Ungleichung)**.** *Es seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ oder $p = 1, q = \infty$. Dann gilt für $X \in \mathcal{L}^p(E, \mathcal{A}, \mu)$ und $Y \in \mathcal{L}^q(E, \mathcal{A}, \mu)$, dass $X \cdot Y \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ und*

$$\int_E (X \cdot Y)\mathrm{d}\mu \leqslant \int_E |X \cdot Y|\mathrm{d}\mu \leqslant \|X\|_p \cdot \|Y\|_q.$$

Daraus ergeben sich zwei wichtige Spezialfälle:

Korollar 5.20.

*(i) Für $X, Y \in \mathcal{L}^2(E, \mathcal{A}, \mu)$ gilt $\|X \cdot Y\|_1 \leqslant \|X\|_2\|Y\|_2$ (*Cauchy-Schwarz-Ungleichung*).*

*(ii) Ist μ ein Wahrscheinlichkeitsmaß und $X \in \mathcal{L}^2(E, \mathcal{A}, \mu)$, so gilt $\|X\|_1 \leqslant \|X\|_2$ (*Einbettung von $\mathcal{L}^2$ in $\mathcal{L}^1$*).*

Die zweite Ungleichung ist folgende:

Satz 5.21 (Minkowski-Ungleichung)**.** *Falls $p \in [1, \infty]$ und $X, Y \in \mathcal{L}^p(E, \mathcal{A}, \mu)$, so gilt*

$$\|X + Y\|_p \leqslant \|X\|_p + \|Y\|_p.$$

Aus der Minkowski-Ungleichung ergibt sich, dass für jedes $p \in [1, \infty]$ die Menge $\mathcal{L}^p(E, \mathcal{A}, \mu)$ einen *linearen Raum* bildet, d.h.

$$X, Y \in \mathcal{L}^p, \alpha, \beta \in \mathbb{R} \quad \Rightarrow \quad \alpha X + \beta Y \in \mathcal{L}^p.$$

Die Operation $X \mapsto \|X\|_p$ erfüllt also die Eigenschaften:

(i) für $X \in \mathcal{L}^p$ und $\alpha \in \mathbb{R}$ gilt $\|\alpha X\|_p = |\alpha| \cdot \|X\|_p$ (*Skalierung*) und

(ii) für $X, Y \in \mathcal{L}^p$ gilt $\|X + Y\|_p \leqslant \|X\|_p + \|X\|_p$ (*Dreiecksungleichung*).

Damit $\|\cdot\|_p$ eine *Norm* bildet, benötigen wir noch die wünschenswerte Eigenschaft

(iii) $\|X\|_p = 0$ gilt genau dann, wenn $X = 0$ richtig ist.

Diese ist jedoch nicht erfüllt, da aus $\|X\|_p = 0$ nur $\mu(\{y \in E : X(y) \neq 0\}) = 0$ folgt. Um diese Eigenschaft dennoch zu erhalten, fasst man alle Borel-messbaren Funktionen auf E, die μ-fast überall gleich sind, zu einer *Äquivalenzklasse* zusammen:

$$[X] := \big\{Y : (E, \mathcal{A}) \to (\mathbb{R}, \mathscr{B}_{\mathbb{R}}) \text{ messbar } : \mu(\{X \neq Y\}) = 0\big\}.$$

Mit der Definition

$$\alpha[X] + b[Y] := [\alpha X + \beta Y]$$

wird die Menge aller Äquivalenzklassen ein linearer Raum mit Nullelement $[0] = \{X : \mu(\{X \neq 0\}) = 0\}$. Da für $p \in [1, \infty)$ das Integral $\int |X|^p \mathrm{d}\mu$ nicht von Änderungen von X auf einer μ-Nullmenge abhängt, ist das Integral $\int |Y|^p \mathrm{d}\mu$ für alle $Y \in [X]$ einer Äquivalenzklasse gleich und wir können definieren

$$\|[X]\|_p := \Big(\int_E |Y|^p \mathrm{d}\mu\Big)^{1/p} \quad \text{für einen beliebigen Repräsentanten } Y \in [X].$$

Der lineare Raum aller Äquivalenzklassen $[X]$ auf $(E, \mathcal{A}, \mu)$ mit $\|[X]\|_p < \infty$ heißt $L^p = L^p(E, \mathcal{A}, \mu)$. Auf L^p erfüllt nun $\|\cdot\|_p$ die drei Normeigenschaften (i) bis (iii). Aus Gründen der Vereinfachung und weil Verwechslungen kaum möglich sind, nennt man $L^p, p \in [1, \infty)$, den Raum der p-integrierbaren Funktionen, meint aber eigentlich den Raum der entsprechenden Äquivalenzklassen. Durch ein ähnliches Vorgehen kann man auch $\mathcal{L}^\infty$durch Betrachtung von Äquivalenzklassen zu einem normierten Raum L^∞ abändern.

6 Konvergenz messbarer Funktionen

Es seien $(E, \mathcal{A}, \mu)$ ein σ-endlicher Maßraum und $X_n \colon E \to \mathbb{R}, n \geqslant 1$, eine Folge reellwertiger Borel-messbarer Funktionen. In diesem Kapitel werden wir mehrere Möglichkeiten kennenlernen, eine Konvergenz der Folge $(X_n)_{n \geqslant 1}$ gegen eine Funktion $X \colon E \to \mathbb{R}$ zu definieren. Eine erste Art haben wir bereits kennengelernt:

Definition 6.1. Eine Folge von Funktionen $X_n \colon E \to \mathbb{R}$ *konvergiert punktweise* gegen eine Funktion $X \colon E \to \mathbb{R}$, falls

$$\lim_{n \to \infty} X_n(y) = X(y) \quad \text{für alle } y \in E.$$

Die *Grenzfunktion* X ist eindeutig bestimmt. Sie ist Borel-messbar, wenn alle X_n Borel-messbar sind, vgl. Korollar 4.5.

6.1 Konvergenz μ-fast-überall

Definition 6.2. Eine Folge $X_n \colon E \to \mathbb{R}, n \geqslant 1$, Borel-messbarer Funktionen heißt *μ-fast-überall konvergent* (kurz: μ-f.ü.) gegen eine Borel-messbare Funktion X auf $(E, \mathcal{A}, \mu)$, falls es eine μ-Nullmenge $N \in \mathcal{A}$ gibt, mit

$$\lim_{n\to\infty} X_n(y) = X(y) \quad \text{für alle } y \in N^c,$$

kurz: $X_n \to X$ μ-f.ü.

Satz 6.3. *Sei $X_n \colon E \to \mathbb{R}, n \geqslant 1$, eine Folge Borel-messbarer Funktionen.*

(i) Aus $X_n \to X$ μ-f.ü. und $X_n \to \widetilde{X}$ μ-f.ü. folgt $X = \widetilde{X}$ μ-f.ü.. Aus $X_n \to X$ μ-f.ü. und $X_n = \widetilde{X}_n$ μ-f.ü. für alle $n \geqslant 1$ folgt $\widetilde{X}_n \to X$ μ-f.ü..

(ii) Ist $\varphi \colon \mathbb{R}^2 \to \mathbb{R}$ stetig und gilt $X_n \to X$ μ-f.ü. sowie $\widetilde{X}_n \to \widetilde{X}$ μ-f.ü., so folgt $\varphi(X_n, \widetilde{X}_n) \to \varphi(X, \widetilde{X})$ μ-f.ü., insbesondere gilt

$$X_n \cdot \widetilde{X}_n \to X \cdot \widetilde{X} \quad \mu\text{-f.ü.}, \quad \frac{X_n}{\widetilde{X}_n} \to \frac{X}{\widetilde{X}} \quad \text{auf } \{\widetilde{X} \neq 0\} \ \mu\text{-f.ü.},$$
$$\alpha X_n + \beta \widetilde{X}_n \to \alpha X + \beta \widetilde{X} \quad \mu\text{-f.ü.}.$$

Aufgabe 6.4. Auf $([0,1], \mathscr{B}_{[0,1]}, \lambda_{[0,1]})$ mit Lebesguemaß $\lambda_{[0,1]}$ auf $[0,1]$ betrachte man die Folge $X_n(y) = y^n, y \in [0,1], n \geqslant 1$. Gegen welche Funktion $X \colon [0,1] \to \mathbb{R}$ konvergiert diese punktweise? Konvergiert die Folge $(X_n)_{n\geqslant 1}$ $\lambda_{[0,1]}$-f.ü. gegen $\widetilde{X}(y) = \mathbb{1}_{\{1\}}(y), y \in [0,1]$?

Es gilt folgende Charakterisierung:

Lemma 6.5. *Für eine Folge Borel-messbarer Funktionen $(X_n)_{n\geqslant 1}$ gilt $X_n \to X$ μ-f.ü. genau dann, wenn*

$$\lim_{n\to\infty} \mu\big(\big\{y \in E : \sup_{k\geqslant n} |X_k(y) - X(y)| \geqslant \varepsilon\big\}\big) = 0 \quad \text{für alle } \varepsilon > 0.$$

6.2 Konvergenzsätze

Ein großer Vorteil in der Arbeit mit dem oben eingeführten Integralbegriff ist die Möglichkeit, Grenzwerte von Funktionen im Sinne der μ-f.ü.-Konvergenz und Integrale über diese Funktionen bezüglich des Maßes μ unter relativ allgemeinen Bedingungen "vertauschen" zu können, wie folgende Sätze zeigen (die Beweise findet man z.B. in Küchler [3, Satz 6.6]).

Satz 6.6 (Monotone Konvergenz, Satz von Beppo Levi)**.** *Es seien $(X_n)_{n\geqslant 1}$ und X nichtnegative Borel-messbare Funktionen auf einem σ-endlichem Maßraum $(E, \mathcal{A}, \mu)$. Falls $0 \leqslant X_n(y) \leqslant X_{n+1}(y) \uparrow X(y)$ für alle $y \in E$ bis auf eine μ-Nullmenge, so gilt*

$$\lim_{n\to\infty} \int_E X_n \mathrm{d}\mu = \int_E \lim_{n\to\infty} X_n \mathrm{d}\mu = \int_E X \mathrm{d}\mu.$$

Beweis. Aus der Monotoniebedingung an die Folge $(X_n)_{n\geqslant 1}$ folgt einerseits $0 \leqslant \int_E X_n \mathrm{d}\mu \leqslant \int_E X \mathrm{d}\mu$ und anderseits, dass $(\int_E X_n \mathrm{d}\mu)_{n\geqslant 1}$ eine nichtfallende Folge ist. Folglich besitzt die Folge der Integrale einen Grenzwert für den gilt

$$\lim_{n\to\infty} \int_E X_n \mathrm{d}\mu \leqslant \int_E X \mathrm{d}\mu.$$

Es bleibt also nur noch die umgekehrte Richtung $\int_E X \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_E X_n \mathrm{d}\mu$ zu zeigen.

Hierfür sei $(Y_m)_{m\geqslant 1}$ eine Folge einfacher Funktionen die X approximiert, also $Y_m(y) \uparrow X(y)$ für alle $y \in E$. Für jedes beliebige $c > 1$ und die Mengen

$$B_{n,m} := \big\{cX_n \geqslant Y_m\big\} = \big\{y \in E : cX_n(y) \geqslant Y_m(y)\big\}$$

gilt $B_{n,m} \in \mathcal{A}$, $B_{n,m} \subseteq B_{n+1,m}, n \geqslant 1$, sowie $\bigcup_{n\geqslant 1} B_{n,m} = E$ für alle $m \geqslant 1$. Daher gilt für die Folge einfacher Funktionen $(Y_m \cdot \mathbb{1}_{B_{n,m}})_{n\geqslant 1}$, dass

$$Y_m \cdot \mathbb{1}_{B_{n,m}}(y) \leqslant cX_n(y) \quad \text{und} \quad Y_m \cdot \mathbb{1}_{B_{n,m}}(y) \uparrow Y_m(y) \quad \text{für alle } y \in E.$$

Die Definition des Integrals ergibt

$$\int_E Y_m \mathrm{d}\mu = \lim_{n\to\infty} \int_E Y_m \mathbb{1}_{B_{n,m}} \mathrm{d}\mu \leqslant \lim_{n\to\infty} c \int_E X_n \mathrm{d}\mu \quad \text{für alle } m \geqslant 1.$$

Da $c > 1$ beliebig war, folgt $\int_E Y_m \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_E X_n \mathrm{d}\mu$ also

$$\int_E X \mathrm{d}\mu = \lim_{m\to\infty} \int_E Y_m \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_E X_n \mathrm{d}\mu. \qquad \square$$

Satz 6.7 (Dominierte Konvergenz, Satz von Lebesgue)**.** *Es sei $(E, \mathcal{A}, \mu)$ ein σ-endlicher Maßraum und $X_n \colon E \to \mathbb{R}, n \geqslant 1$, eine Folge Borel-messbarer Funktionen, die μ-fast überall gegen eine Borel-messbare Funktion $X \colon E \to \mathbb{R}$ konvergiert. Gibt es ein $Y \in \mathcal{L}^1(E, \mathcal{A}, \mu)$, so dass $|X_n| \leqslant Y$ μ-f.ü. für alle $n \geqslant 1$, dann ist $X_n \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ für alle $n \geqslant 1$ und es gilt*

$$\lim_{n\to\infty} \int_E X_n \mathrm{d}\mu = \int_E \lim_{n\to\infty} X_n \mathrm{d}\mu = \int_E X \mathrm{d}\mu.$$

Auf die Bedingung $|X_n| \leqslant Y$ μ-f.ü. kann im Allgemeinen nicht verzichtet werden, wie folgende Aufgabe illustriert:

Aufgabe 6.8. Auf dem Maßraum $([0,1], \mathscr{B}_{[0,1]}, \lambda_{[0,1]})$ mit Lebesguemaß $\lambda_{[0,1]}$, betrachten wir die Folge

$$X_n \colon \mathbb{R} \to \mathbb{R}_+, \quad X_n(y) = n\mathbb{1}_{[0,1/n]}(y), y \in \mathbb{R}.$$

(i) Gegen welche Funktion $X \colon \mathbb{R} \to \mathbb{R}$ konvergiert die Folge $(X_n)_{n\geqslant 1}$ $\lambda_{[0,1]}$-f.ü.?

(ii) Berechnen Sie die Integrale $\int_\mathbb{R} X_n \mathrm{d}\lambda_{[0,1]}, n \geqslant 1$, und $\int_\mathbb{R} X \mathrm{d}\lambda_{[0,1]}$.

Der Vollständigkeit halber zitieren wir noch

Lemma 6.9 (Lemma von Fatou*)**.** *Ist $(X_n)_{n\geqslant 1}$ eine Folge nichtnegativer Borel-messbarer Funktionen auf einem σ-endlichem Maßraum $(E, \mathcal{A}, \mu)$, dann gilt*

$$\int_E \liminf_{n\to\infty} X_n \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int_E X_n \mathrm{d}\mu.$$

6.3 Weitere Konvergenzarten*

Eine schwächere Form der Konvergenz als die μ-f.ü.-Konvergenz einer Folge $X_n \colon E \to \mathbb{R}, n \geqslant 1$, Borel-messbarer Funktionen auf einem σ-endlichen Maßraum $(E, \mathcal{A}, \mu)$ ist folgende

Definition 6.10. Die Folge $(X_n)_{n\geqslant 1}$ *konvergiert nach Maß* μ gegen eine Borel-messbare Funktion $X \colon E \to \mathbb{R}$, falls gilt

$$\lim_{n\to\infty} \mu\big(\{y \in E : |X_n(y) - X(y)| > \varepsilon\}\big) = 0 \quad \text{für jedes } \varepsilon > 0.$$

Wir schreiben kurz $X_n \xrightarrow{\mu} X$. Ist μ ein Wahrscheinlichkeitsmaß, dann wird diese Konvergenz auch *stochastische Konvergenz* genannt.

Um Konvergenz nach Maß mit Konvergenz fast überall in Beziehung zu setzen erwähnen wir folgendes Resultat:

Satz 6.11. *Auf einem σ-endlichen Maßraum $(E, \mathcal{A}, \mu)$ sei $(X_n)_{n\geqslant 1}$ eine Folge reellwertiger Borel-messbarer Funktionen.*

(i) *Aus* $\lim_{n\to\infty} X_n = X$ *μ-f.ü. folgt* $\lim_{n\to\infty} X_n = X$ *nach Maß* μ.

(ii) *Wenn* $\lim_{n\to\infty} X_n = X$ *nach Maß* μ*, dann gibt es eine Teilfolge* $(n_k)_{k\geqslant 1}$ *natürlicher Zahlen mit* $\lim_{k\to\infty} X_{n_k} = X$ *μ-f.ü..*

Vor allem durch den zentralen Grenzwertsatz ist eine weitere Konvergenzart wichtig in der Wahrscheinlichkeitstheorie:

Definition 6.12. Es seien $\mu_n, n \geqslant 1$, und μ endliche Maße auf $(\mathbb{R}, \mathscr{B}_\mathbb{R})$. Wir sagen, dass $(\mu_n)_{n\geqslant 1}$ *schwach* oder *in Verteilung* gegen μ *konvergiert,* geschrieben $\mu_n \overset{w}{\to} \mu$ oder $\mu_n \overset{d}{\to} \mu$, falls für jede stetige beschränkte Funktion $f\colon \mathbb{R} \to \mathbb{R}$ gilt

$$\lim_{n\to\infty} \int_\mathbb{R} f \mathrm{d}\mu_n = \int f \mathrm{d}\mu.$$

Man kann zeigen, dass schwache Konvergenz auch durch punktweise Konvergenz der entsprechenden Verteilungsfunktionen charakterisiert werden kann:

$$\mu_n \overset{d}{\to} \mu \quad \Longleftrightarrow \quad \lim_{n\to\infty} F_{\mu_n}(x) = F_\mu(x) \quad \text{für alle } x \in \mathbb{R} \text{ in denen } F \text{ stetig ist.}$$

Außerdem gilt

Lemma 6.13. *Es sei* $(X_n)_{n\geqslant 1}$ *eine Folge reellwertiger Borel-messbarer Funktionen auf einem endlichen Maßraum* $(E, \mathcal{A}, \mu)$ *mit Konvergenz nach Maß* $X_n \overset{\mu}{\to} X$ *für eine Borel-messbare Funktion* $X\colon E \to \mathbb{R}$*. Dann gilt die schwache Konvergenz* $\mu^{X_n} \overset{d}{\to} \mu^X$ *der Bildmaße.*

Schließlich erwähnen wir noch die Konvergenz im L^p-Sinn.

Definition 6.14. Für einen endlichen Maßraum $(E, \mathcal{A}, \mu)$ und $p \geqslant 1$ seien $(X_n)_{n\geqslant 1} \subseteq L^p(E, \mathcal{A}, \mu)$ eine Folge von integrierbaren Funktionen (Äquivalenzklassen) und $X \in L^p(E, \mathcal{A}, \mu)$. Dann *konvergiert* X_n gegen X im *L^p-Sinn,* falls $\lim_{n\to\infty} \|X_n - X\|_p = 0$.

L^p-Konvergenz impliziert Konvergenz nach Maß. Aus dem Satz über dominierte Konvergenz folgt außerdem, dass μ-f.ü.-Konvergenz die L^p-Konvergenz impliziert, falls es eine Funktion $Y \in L^p(E, \mathcal{A}, \mu)$ gibt mit $|X_n| \leqslant Y$ für alle $n \geqslant 1$ (allgemein folgt allerdings aus f.ü.-Konvergenz keine L^p-Konvergenz).

7 Produktmaße

In der Wahrscheinlichkeitstheorie spielt der Begriff der Unabhängigkeit für Ereignisse bzw. Zufallsvariablen eine zentrale Rolle. Die gemeinsame Verteilung voneinander unabhängiger Zufallsgrößen ist von spezialer Gestalt, es ist eine sogenannte Produktverteilung oder, in der Sprache der Maßtheorie, ein Produktmaß. Letzteres wird in diesem Abschnitt eingeführt. Ziel ist der Satz von Fubini mit dem Integrale bezüglich Produktmaßen auf einfachere Integrale zurückgeführt werden können.

Es seien $(E, \mathcal{A})$ und $(F, \mathcal{B})$ zwei messbare Räume. Wir haben bereits den Produktraum

$$E \times F = \{(x, y) : x \in E, y \in F\}$$

und die Produkt-σ-Algebra

$$\mathcal{A} \otimes \mathcal{B} = \sigma(\mathcal{A} \times \mathcal{B}) = \sigma\big(\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}\big)$$

kennen gelernt.

Beispiel 7.1.

(i) Ist $(E, \mathcal{A}) = (\mathbb{R}^n, \mathscr{B}_{\mathbb{R}^n})$ und $(F, \mathcal{B}) = (\mathbb{R}^m, \mathscr{B}_{\mathbb{R}^m})$, so ist $E \times F = \mathbb{R}^{n+m}$ und $\mathscr{B}_{\mathbb{R}^n} \otimes \mathscr{B}_{\mathbb{R}^m} = \mathscr{B}_{\mathbb{R}^{n+m}}$.

(ii) Sind E und F endliche Räume, dann gilt $\mathscr{P}(E) \otimes \mathscr{P}(F) = \mathscr{P}(E \times F)$.

Definition 7.2. Ist $X \colon E \times F \to \mathbb{R}$ eine Funktion, so heißt für jedes feste $x \in E$ die Funktion

$$F \ni y \mapsto X(x, y) \in \mathbb{R}$$

der *Schnitt von X am Punkt $x \in E$*. Analog ist der Schnitt von X am Punkt $y \in F$ definiert als $x \mapsto X(x, y)$.

Lemma 7.3. *Es sei $X \colon E \times F \to \mathbb{R}$ eine $(\mathcal{A} \otimes \mathcal{B}, \mathscr{B}_{\mathbb{R}})$-messbare Abbildung. Dann ist für jedes $x \in E$ (bzw. $y \in F$) der Schnitt $F \ni y \mapsto X(x, y) \in \mathbb{R}$ (bzw. $E \ni x \mapsto X(x, y) \in \mathbb{R}$) eine $(\mathcal{B}, \mathscr{B}_{\mathbb{R}})$-messbare (bzw. $(\mathcal{A}, \mathscr{B}_{\mathbb{R}})$-messbare) Abbildung.*

Der Beweis erfolgt wieder über maßtheoretische Induktion, vgl. Küchler [3, Aussage 7.3].

Sind nun $(E, \mathcal{A}, \mu)$ und $(F, \mathcal{B}, \nu)$ Maßräume mit σ-endlichen Maßen μ und ν, definiert

$$\kappa(A \times B) := \mu(A) \times \nu(B), \quad A \in \mathcal{A}, B \in \mathcal{B}$$

eine Mengenfunktion κ auf $\mathcal{A} \times \mathcal{B}$.

Satz 7.4. *κ ist eine nichtnegative, σ-endliche und σ-additive Mengenfunktion auf $\mathcal{A} \times \mathcal{B}$ und kann auf eindeutige Weise zu einem σ-endlichen Maß auf der Produkt-σ-Algebra fortgesetzt werden. Dieses wird* Produktmaß *aus μ und ν genannt und mit $\mu \otimes \nu$ bezeichnet.*

Beweisskizze. Die kleinste Algebra die $\mathcal{A} \times \mathcal{B}$ enthält ist gegeben durch

$$\mathcal{C}^0 := \Big\{ \bigcup_{k=1}^{n} C_k : C_k = (A_k \times B_k) \in \mathcal{A} \times \mathcal{B}, k = 1, \dots, n, \text{ paarweise disjunkt}, n \in \mathbb{N} \Big\}.$$

Für jedes Element $\bigcup_{k=1}^{n} (A_k \times B_k) \in \mathcal{C}^0$ können wir κ (eindeutig) fortsetzen mittels

$$\overline{\kappa}\Big(\bigcup_{k=1}^{n} (A_k \times B_k) \Big) = \sum_{k=1}^{n} \kappa(A_k \times B_k) = \sum_{k=1}^{n} \mu(A_k)\nu(B_k).$$

Nach dem Satz von Caratheodory, können wir $\overline{\kappa}$ eindeutig von $\mathcal{C}^0$ auf $\sigma(\mathcal{C}^0) = \sigma(\mathcal{A} \times \mathcal{B}) = \mathcal{A} \otimes \mathcal{B}$ fortsetzen, sofern κ (und damit auch $\overline{\kappa}$) σ-additiv und σ-endlich ist. Letzteres muss gezeigt werden, vgl. Küchler [3, Aussage 7.4]. □

Die Maße μ und ν nennt man die *Komponenten* des Produktmaßes $\mu \otimes \nu$. Der folgende Satz findet häufig Anwendung. Er besagt, dass man die Reihenfolge der Integration für Produktmaße vertauschen kann, sofern alle Integrale wohldefiniert sind.

Satz 7.5 (Satz von Fubini). *Für jede $(\mathcal{A} \otimes \mathcal{B}, \mathscr{B}_{\mathbb{R}})$-messbare Funktion $X \colon E \times F \to \mathbb{R}$, die entweder nichtnegativ ist oder ein endliches Integral bzgl. $\mu \otimes \nu$ besitzt, sind die Funktionen*

$$E \ni x \mapsto \int_F X(x, y)\nu(\mathrm{d}y) \in \mathbb{R} \quad \textit{bzw.} \quad F \ni y \mapsto \int_E X(x, y)\mu(\mathrm{d}x) \in \mathbb{R}$$

μ-fast überall (bzw. ν-fast überall) endlich, $(\mathcal{A}, \mathscr{B}_{\mathbb{R}})$-messbar (bzw. $(\mathcal{B}, \mathscr{B}_{\mathbb{R}})$-messbar) und es gilt

$$\begin{aligned} \int_{E \times F} X(x, y)(\mu \otimes \nu)(\mathrm{d}x, \mathrm{d}y) &= \int_E \Big(\int_F X(x, y)\nu(\mathrm{d}y) \Big) \mu(\mathrm{d}x) \\ &= \int_F \Big(\int_E X(x, y)\mu(\mathrm{d}x) \Big) \nu(\mathrm{d}y), \end{aligned} \tag{7.1}$$

wobei die Seiten dieser Gleichung entweder alle endlich oder alle unendlich sind.

Bemerkung 7.6.

(i) Aus der Existenz der iterierten Integrale auf der rechten Seite von (7.1) folgt weder die Existenz des Integrals $\int_{E\times F} X \mathrm{d}(\mu\otimes\nu)$ noch ihre Gleichheit.

(ii) Ist dagegen $X \geqslant 0$ und ist $E \ni x \mapsto \int_F X(x,y)\nu(\mathrm{d}y)$ eine μ-fast überall endliche Funktion, so existiert das Integral $\int_{E\times F} X \mathrm{d}(\mu\otimes\nu)$ und es gilt (7.1) (Satz von Tonelli).

(iii) Wenn $X(x,y) = G(x)\cdot H(y), x \in E, y \in F$ für zwei Funktionen $G \in \mathcal{L}^1(E,\mathcal{A},\mu)$ und $H \in \mathcal{L}^1(F,\mathcal{B},\nu)$, dann gilt

$$\int_{E\times F} X \mathrm{d}(\mu\otimes\nu) = \int_E G \mathrm{d}\mu \cdot \int_F H \mathrm{d}\nu.$$

Letztere Formel ist in der Stochastik wichtig. In der dortigen Sprache besagt sie: Sind G und H zwei stochastisch unabhängige Zufallsgrößen mit endlichem Erwartungswert, so gilt $E[G\cdot H] = E[G]\cdot E[H]$.

Literatur

[1] Elstrodt, J. (2007). *Maß- und Integrationstheorie.* Springer, Berlin Heidelberg.

[2] Georgii, H.-O. (2007). *Stochastik. Einführung in die Wahrscheinlichkeitstheorie und Statistik.* de Gruyter, Berlin.

[3] Küchler, U. (2013). *Maßtheorie für Statistiker.* Vorlesungsskript, Berlin.