Kapitel 1

Mengensysteme, Wahrscheinlichkeitsmaße

Der Großteil der folgenden fundamentalen Begriffe sind schon aus den Vorlesungen 'Stochastische Modellbildung' und Analysis 3 bekannt:

Definition 1.1 *Eine Familie $\mathcal{A}$ von Teilmengen einer Menge Ω ('Grundgesamtheit') heisst* σ-Algebra, *falls die folgenden drei Bedingungen erfüllt sind:*

1. $\Omega \in \mathcal{A}$.
2. *Für jedes* $A \in \mathcal{A}$ *ist auch* $A^c \in \mathcal{A}$.
3. *Für* $A_1, A_2, A_3, \ldots \in \mathcal{A}$ *gilt auch* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Jedes $A \in \mathcal{A}$ *heisst* messbar *bzw.* Ereignis. $(\Omega, \mathcal{A})$ *heisst* Messraum.

Lemma 1.2 *Sei* $(\Omega, \mathcal{A})$ *ein Messraum und* $A, B, A_1, A_2, A_3, \ldots \in \mathcal{A}$. *Dann sind auch die folgenden Mengen messbar:*

1. $A \setminus B$.
2. $A \Delta B := (A \setminus B) \cup (B \setminus A)$ *(symmetrische Differenz).*
3. $\bigcap_{n=1}^{\infty} A_n$.
4. $\limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ *(limes superior der Folge* $(A_n)_{n\in\mathbb{N}}$*)*
5. $\liminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$ *(limes inferior der Folge* $(A_n)_{n\in\mathbb{N}}$*)*

Beweis: Übungsaufgabe

Bemerkung 1.3 Angelehnt an den reellen Fall heisst eine Folge $(A_n)_{n\in\mathbb{N}}$ messbarer Mengen *konvergent* genau dann, wenn $\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n$.

Satz 1.4 *Für jedes System* $\mathcal{E}$ *von Teilmengen von* Ω *existiert eine kleinste* σ*-Algebra* $\mathcal{A}_\sigma(\mathcal{E})$ *mit* $\mathcal{E} \subseteq \mathcal{A}_\sigma(\mathcal{E})$. *Wir nennen* $\mathcal{A}_\sigma(\mathcal{E})$ *die* von $\mathcal{E}$ erzeugte σ-Algebra.

Beweis: Übungsaufgabe. Es reicht zu zeigen, dass der beliebige Durchschnitt von σ-Algebren wieder eine σ-Algebra ist.

Beispiel 1.5 Die Vereinigung von σ-Algebren muss keine σ-Algebra sein: Betrachten wir $\Omega = [0,1]$ sowie $\mathcal{A}_1 = \{\emptyset, [0,1], [0,3/4], (3/4,1]\}$ und $\mathcal{A}_2 = \{\emptyset, [0,1], [0,1/4], (1/4,1]\}$ dann ist $\mathcal{A}_1 \cup \mathcal{A}_2$ offensichtlich keine σ-Algebra.

Definition 1.6 *Sei $(\Omega, \mathcal{A})$ ein Messraum. $P : \mathcal{A} \to [0,1]$ heisst Wahrscheinlichkeitsmaß genau dann wenn $P(\Omega) = 1$ gilt und P* σ-additiv *ist, i.e. für jede Folge $A_1, A_2, \ldots$ von messbaren, paarweise disjunkten Mengen gilt*

$$P\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} P(A_n). \tag{1.1}$$

Das Tripel $(\Omega, \mathcal{A}, P)$ heisst Wahrscheinlichkeitsraum.

Lemma 1.7 *Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, dann gilt:*

(a) $P(\emptyset) = 0$.

(b) Für $A, B \in \mathcal{A}$ mit $A \subseteq B$ gilt $P(A) \leq P(B)$.

(c) Für jede wachsende Folge $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ messbarer Mengen gilt $\lim_{n\to\infty} P(A_n) = P(\bigcup_{n=1}^{\infty} A_n)$ (Stetigkeit von unten).

(d) Für jede fallende Folge $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$ messbarer Mengen gilt $\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n)$ (Stetigkeit von oben).

Die Eigenschaften (b)-(d) besagen, dass jedes Wahrscheinlichkeitsmaß monoton und stetig von unten und oben ist.

Beweis: Aussage (a) folgt direkt aus Gleichung (1.1).
(b): Wegen $B = A \cup (B \setminus A)$ und $A \cap (B \setminus A) = \emptyset$ folgt aus (1.1) unmittelbar (einfach $A_n := \emptyset$ für alle $n \geq 3$ setzen) $P(B) = P(A) + P(B \setminus A)$, also $P(B) - P(A) = P(B \setminus A) \geq 0$.
Zum Beweis von (c) kann wie folgt vorgegangen werden: Wir setzen $B_1 := A_1$ und, für jedes $n \geq 2$, $B_n := A_n \cap A_{n-1}^c$. Dann ist $B_n \in \mathcal{A}$ für jedes $n \in \mathbb{N}$, $B_i \cap B_j = \emptyset$ für $i \neq j$, und es gilt $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$. Anwendung der σ-Additivität liefert daher

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(B_n) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} P(A_n).$$

Beweis von (d): Übungsaufgabe ■

Bemerkung 1.8 A-priori ist unklar, warum man anstatt mit Wahrscheinlichkeitsmaßen auf σ-Algebren (oder schwächeren Strukturen) nicht gleich mit Wahrscheinlichkeitsmaßen auf Potenzmengen $\mathfrak{p}(\Omega)$ arbeitet. Die zwei Hauptgründe dafür sind die Folgenden:
Grund 1: Es ist unmöglich, (Wahrscheinlichkeits-) Maße mit natürlichen Eigenschaften auf ganz $\mathfrak{p}(\Omega)$ zu definieren. Angenommen, wir suchen eine Funktion $\mathcal{V} : \mathfrak{p}(\mathbb{R}^3) \to [0,\infty]$, die das 'Volumen' von Mengen beschreibt und daher mindestens folgende Bedingungen erfüllen sollte:

(V1) $\mathcal{V}(A \cup B) = \mathcal{V}(A) + \mathcal{V}(B)$ für alle $A, B \subseteq \mathbb{R}^3$ mit $A \cap B = \emptyset$

(V2) $\mathcal{V}([0,1]^3) = 1$

(V3) Für jede Bewegung $f : \mathbb{R}^3 \to \mathbb{R}^3$ und jedes $A \subseteq \mathbb{R}^3$ gilt $\mathcal{V}(f(A)) = \mathcal{V}(A)$.

Hausdorff zeigte 1905, dass es keine solche Funktion $\mathcal{V}$ gibt. Banach & Tarski bewiesen 1924 ein noch wesentlich stärkeres Resultat, das sog. 'Banach-Tarksi-Paradoxon'† (interessantes Thema für eine Bachelor Arbeit):

Satz 1.9 (Banach-Tarski Paradoxon, AC) *Angenommen $K, G \subseteq \mathbb{R}^3$ sind beschränkt und haben nichtleeres Inneres. Dann existieren paarweise disjunkte Mengen $K_1, \ldots, K_n$ mit $K = \bigcup_{i=1}^n K_i$ und Bewegungen $f_1, \ldots, f_n$ sodass $f_1(K_1), \ldots, f_n(K_n)$ paarweise disjunkt sind und $G = \bigcup_{i=1}^n f_i(K_i)$ gilt.*
Mit anderen Worten: Aus einem Tennisball K kann durch geschickte Zerlegung und Bewegung ein Fussball G gebastelt werden.

Grund 2: Wahrscheinlichkeitsmaße können eindeutig von 'guten' kleinen Klassen (Halbringe etc.) auf die erzeugten σ-Algebren fortgesetzt werden (vergleiche Satz 1.11 sowie [2, 4, 5]). Als Beispiel betrachten wir $\Omega = [0,1]$ und die Klasse $\mathfrak{h}$, definiert durch

$$\mathfrak{h} = \big\{(a,b] \subseteq [0,1] : a \leq b\big\} \tag{1.2}$$

Offensichtlich ist $\mathfrak{h}$ keine σ-Algebra, zum Beispiel ist weder $(0,1/4] \cup (1/2,1]$ noch $[0,1]$ ein Element von $\mathfrak{h}$. $\mathfrak{h}$ hat aber die folgenden Eigenschaften: (i) $\emptyset \in \mathfrak{h}$, (ii) $A, B \in \mathfrak{h}$ impliziert $A \cap B \in \mathfrak{h}$, (iii) Gelte $A, B \in \mathfrak{h}$ und $A \subseteq B$, dann existieren disjunkte Mengen $E_1, E_2 \in \mathfrak{h}$ mit $B \setminus A = E_1 \cup E_2$. $\mathfrak{h}$ ist ein sogenannter Halbring:

Definition 1.10 *Eine Familie $\mathfrak{h}$ von Teilmengen einer Menge Ω heisst* Halbring*, falls die folgenden drei Bedingungen erfüllt sind:*

1. *$\emptyset \in \mathfrak{h}$.*
2. *Für $A, B \in \mathfrak{h}$ gilt auch $A \cap B \in \mathfrak{h}$.*
3. *Für $A, B \in \mathfrak{h}$ mit $A \subseteq B$ gilt: Es existieren paarweise disjunkte Mengen $E_1, \ldots, E_n \in \mathfrak{h}$ sodass $B \setminus A = \bigcup_{i=1}^n E_i$.*

Es ist klar, wie Elementen von $\mathfrak{h}$ gemäß (1.2) eine Länge zugewiesen werden kann: Wir definieren einfach

$$\lambda((a,b]) := b - a \tag{1.3}$$

Wie kann dieses λ auf kompliziertere Mengen $A \in \mathcal{A}_\sigma(\mathfrak{h})$ erweitert werden? Naheliegende Idee: Überdecke A möglichst gut durch halboffene Intervalle, i.e. für $A \subseteq [0,1]$ definiere

$$\lambda(A) := \inf\left\{\sum_{i=1}^\infty \lambda(I_i) : I_1, I_2, \ldots \in \mathfrak{h},\, A \subseteq \bigcup_{i=1}^\infty I_i\right\} \tag{1.4}$$

Dann lässt sich zeigen (siehe Vorlesung Analysis 3/Maßtheorie), dass damit ein Wahrscheinlichkeitsmaß λ auf $\mathcal{A}_\sigma(\mathfrak{h})$ definiert ist. Im Folgenden werden wir λ als *Lebesgue'sches Maß* bezeichnen. Allgemeiner gilt folgender Satz:

†für mehr Hintergrundwissen sei auf die Vorlesung Analysis 3/Maßtheorie verwiesen

Satz 1.11 *Sei $(\Omega, \mathcal{A})$ ein Messraum und $\mathfrak{h}$ ein Halbring mit der Eigenschaft $\mathcal{A}_\sigma(\mathfrak{h}) = \mathcal{A}$. Weiters sei $\mu : \mathfrak{h} \to [0,1]$ σ-additiv[†]. Dann existiert ein eindeutiges Wahrscheinlichkeitsmaß P auf $(\Omega, \mathcal{A})$, für das $P(E) = \mu(E)$ für alle $E \in \mathfrak{h}$ gilt.*

Beweis: Übersteigt leider den Rahmen dieser Vorlesung, siehe Vorlesung Analysis 3/Maßtheorie sowie [2, 4, 5, 7]. Für Interessierte sei hier nur angemerkt, dass der Beweis der Fortsetzbarkeit auf der ganz elementaren, zuvor schon erwähnten Idee beruht: Approximiere (überdecke) die Menge A, von der die Fläche (das Volumen) berechnet werden soll, durch einfache Objekte (Rechtecke), deren Fläche (Volumen) bekannt ist.

Beispiel 1.12 Die folgenden Eigenschaften des Lebesgue-Maßes sind leicht zu zeigen (Übungsaufgabe) - $\mathfrak{h}$ bezeichnet dabei wieder den Halbring der halboffenen Intervalle gemäß (1.2):

(S1) Für jedes $[a,b] \subseteq (0,1]$ gilt $[a,b] \in \mathcal{A}_\sigma(\mathfrak{h})$

(S2) Für $[a,b] \subseteq (0,1]$ gilt $\lambda([a,b]) = b - a$

(S3) $\lambda(\mathbb{Q} \cap (0,1]) = 0$

Die im Folgenden für uns wichtigste σ-Algebra ist die sogenannte Borel'sche σ-Algebra:

Definition 1.13 *Gegeben sei ein metrischer Raum (Ω, d). Dann heißt die von der Familie $\mathcal{O}$ aller offenen Mengen erzeugte σ-Algebra die* σ-Algebra der Borelmengen bzw. Borel'sche σ-Algebra. *Wir bezeichnen sie im Folgenden mit $\mathcal{B}(\Omega)$.*

Lemma 1.14 Wir betrachten die folgenden Mengensysteme in $\mathbb{R}$:

$$\begin{aligned}
\mathcal{E}_1 &= \big\{(a,b) \subseteq \mathbb{R} : a \leq b\big\}, & \mathcal{E}_1' &= \big\{(a,b) \subseteq \mathbb{R} : a \leq b \text{ und } a,b \in \mathbb{Q}\big\} \\
\mathcal{E}_2 &= \big\{(a,b] \subseteq \mathbb{R} : a \leq b\big\}, & \mathcal{E}_2' &= \big\{(a,b] \subseteq \mathbb{R} : a \leq b \text{ und } a,b \in \mathbb{Q}\big\} \\
\mathcal{E}_3 &= \big\{[a,b] \subseteq \mathbb{R} : a \leq b\big\}, & \mathcal{E}_3' &= \big\{[a,b] \subseteq \mathbb{R} : a \leq b \text{ und } a,b \in \mathbb{Q}\big\} \\
\mathcal{E}_4 &= \big\{(-\infty,b] \subseteq \mathbb{R} : b \in \mathbb{R}\big\}, & \mathcal{E}_4' &= \big\{(-\infty,b] \subseteq \mathbb{R} : b \in \mathbb{Q}\big\} \\
\mathcal{E}_5 &= \big\{(b,\infty) \subseteq \mathbb{R} : b \in \mathbb{R}\big\}, & \mathcal{E}_5' &= \big\{(b,\infty) \subseteq \mathbb{R} : b \in \mathbb{Q}\big\} \\
\mathcal{E}_6 &= \big\{K \subseteq \mathbb{R} : K \text{ kompakt}\big\} \\
\mathcal{E}_7 &= \big\{A \subseteq \mathbb{R} : A \text{ abgeschlossen}\big\}
\end{aligned} \tag{1.5}$$

Dann sind alle 12 Familien Erzeuger der Borel'sche σ-Algebra $\mathcal{B}(\mathbb{R})$ in $\mathbb{R}$.

Beweis: Wir zeigen nur, dass $\mathcal{E}_1$ und $\mathcal{E}_2$ die Borel'sche σ-Algebra $\mathcal{B}(\mathbb{R})$ erzeugen.
(i) Da jedes Element von $\mathcal{E}_1$ offen ist gilt $\mathcal{E}_1 \subseteq \mathcal{O}$ und damit $\mathcal{A}_\sigma(\mathcal{E}_1) \subseteq \mathcal{A}_\sigma(\mathcal{O}) = \mathcal{B}(\mathbb{R})$. Umgekehrt kann jede offene Menge $U \subseteq \mathbb{R}$ als höchstens abzählbare Vereinigung $\bigcup_{i=1}^\infty I_i$ von offenen Intervallen dargestellt werden. Wegen $\bigcup_{i=1}^\infty I_i \in \mathcal{A}_\sigma(\mathcal{E}_1)$ folgt damit $U \in \mathcal{A}_\sigma(\mathcal{E}_1)$, also $\mathcal{A}_\sigma(\mathcal{O}) \subseteq \mathcal{A}_\sigma(\mathcal{E}_1)$, und damit insgesamt $\mathcal{A}_\sigma(\mathcal{E}_1) = \mathcal{B}(\mathbb{R})$.
(ii) Wegen $(a,b] = \bigcap_{n=1}^\infty (a, b + 1/n)$ gilt $\mathcal{E}_2 \subseteq \mathcal{B}(\mathbb{R})$ und damit $\mathcal{A}_\sigma(\mathcal{E}_2) \subseteq \mathcal{B}(\mathbb{R})$. Da weiters $(a,b) = \bigcup_{n=1}^\infty (a, b - 1/n]$ gilt, folgt $\mathcal{E}_1 \subseteq \mathcal{A}_\sigma(\mathcal{E}_2)$, also auch $\mathcal{B}(\mathbb{R}) = \mathcal{A}_\sigma(\mathcal{E}_1) \subseteq \mathcal{A}_\sigma(\mathcal{E}_2)$. ■

[†] i.e. für paarweise disjunkte Mengen $A_1, A_2, \ldots$ in $\mathfrak{h}$ mit $\bigcup_{n=1}^\infty A_n \in \mathfrak{h}$ gilt Gleichung (1.1)

Bemerkung 1.15 Entsprechende Aussagen gelten auch im allgemeinen Fall von $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, siehe [2, 4, 5]. So ist zum Beispiel die Klasse $\mathcal{I}_d$ aller halboffenen d-dimensionalen Rechtecke

$$\mathcal{I}_d := \Big\{ \bigtimes_{i=1}^{d} (a_i, b_i] : \forall i \in \{1, \dots, d\} : a_i \leq b_i \Big\} \tag{1.6}$$

ein Halbring, der $\mathcal{B}(\mathbb{R}^d)$ erzeugt, i.e. es gilt $\mathcal{A}_\sigma(\mathcal{I}_d) = \mathcal{B}(\mathbb{R}^d)$.

Folgerung 1.16 *Es existiert genau ein Wahrscheinlichkeitsmaß λ_d auf $([0,1]^d, \mathcal{B}([0,1]^d))$, das*

$$\lambda_d\Big(\bigtimes_{i=1}^{d} (a_i, b_i] \Big) = \prod_{i=1}^{d} (b_i - a_i)$$

*für alle $a_i, b_i \in [0,1]$ und $a_i \leq b_i$ erfüllt. Dieses Maß geht auf Henri Lebesgue zurück und trägt den Namen d-*dimensionales *Lebesgue Maß.*

Beweis: Die Familie $\mathcal{I}_d$ aller d-dimensionalen Rechtecke der Form $\bigtimes_{i=1}^{d} (a_i, b_i]$ ist ein Halbring und λ_d ist σ-additiv auf $\mathcal{I}_d$ (beides selbst verifizieren !). Damit ist Satz 1.11 anwendbar. ■

Bemerkung 1.17 Die Definition von λ_d kann auch unschwer auf ganz $\mathcal{B}(\mathbb{R}^d)$ erweitert werden (siehe Analysis 3). Das entsprechenden Maß (für das $\lambda_d(\mathbb{R}^d) = \infty$ gilt) werden wir ebenfalls mit λ_d bezeichnen.

Kapitel 2

Zufallsvariable

In der Vorlesung 'Stochastische Modellbildung' haben Sie zu einem großen Teil diskrete Zufallsvariable X, i.e. Abbildungen $X : (\Omega, \mathcal{A}, P) \to \mathbb{R}$ mit abzählbarem Wertebereich $W_X = \{\alpha_1, \alpha_2, \ldots\}$, betrachtet, die Wahrscheinlichkeiten $P(X = \alpha_i)$ berechnet und nachfolgend verwendet. Dabei wurde immer implizit angenommen, dass $P(X = \alpha_i)$ wohldefiniert ist, i.e. dass $\{\omega \in \Omega : X(\omega) = \alpha_i\}$ ein Element der σ-Algebra ist, auf der P definiert ist. Mit anderen Worten: Grundannahme war

$$X^{-1}(\{\alpha_i\}) := \{\omega \in \Omega : X(\omega) = \alpha_i\} \in \mathcal{A} \quad \text{für jedes } \alpha_i \in W_X. \tag{2.1}$$

Falls Bedingung (2.1) erfüllt ist, gilt für jede Menge $E \subseteq W_X$ auch $X^{-1}(E) := \{\omega \in \Omega : X(\omega) \in E\} \in \mathcal{A}$ weil

$$\begin{aligned} X^{-1}(E) &= \{\omega \in \Omega : X(\omega) \in E\} = \bigcup_{e \in E} \{\omega \in \Omega : X(\omega) = e\} \\ &= \bigcup_{e \in E} \underbrace{X^{-1}(\{e\})}_{\in \mathcal{A}} \in \mathcal{A}. \end{aligned}$$

Sei $X : (\Omega, \mathcal{A}, P) \to \mathbb{R}$ nun eine beliebige Abbildung, die den Ausgang eines Zufallsexperiments beschreibt, beispielsweise die Lebensdauer eines elektronischen Bauteiles. Um die Güte des Bauteils bewerten zu können, interessiert man sich insbesondere für die Wahrscheinlichkeit dafür, dass der Bauteil höchstens bis zum Zeitpunkt T funktioniert, i.e. für

$$P(X^{-1}(-\infty, T])) = P(X \leq T).$$

Mengen der Form

$$X^{-1}((-\infty, T]) = \{\omega \in \Omega : X(\omega) \leq T\}$$

müssen also $X^{-1}((-\infty, T]) \in \mathcal{A}$ erfüllen.

Frage 2.1 Für welche weiteren Menge $E \subseteq \mathbb{R}$ sollte $P(X^{-1}(E))$ für das Beispiel der Lebensdauer X bekannt sein? Wen interessiert $P(X^{-1}(E))$ für eine sehr komplizierte Menge E, zum Beispiel die Cantor Menge?

Das folgende Resultat zeigt, dass aus der Eigenschaft $X^{-1}((-\infty, T]) \in \mathcal{A}$ für jedes $T \in \mathbb{R}$ schon $X^{-1}(E) \in \mathcal{A}$ für jede Borel-Menge $E \in \mathcal{B}(\mathbb{R})$ folgt - allgemeiner gilt folgendes wichtiges Resultat:

Satz 2.2 *Gegeben sei eine Abbildung* $f : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ *sowie ein Erzeuger* $\mathcal{E}_2$ *von* $\mathcal{A}_2$. *Gilt* $f^{-1}(E_2) \in \mathcal{A}_1$ *für jedes* $E_2 \in \mathcal{E}_2$ *dann folgt* $f^{-1}(A_2) \in \mathcal{A}_1$ *für jedes* $A_2 \in \mathcal{A}_2$.

Beweis: Wir beweisen das Resultat in drei Schritten:
(i) Leicht nachzuprüfen (Übungsaufgabe): Für jedes beliebige $f : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ sind die Mengensysteme

$$f^{-1}(\mathcal{A}_2) := \big\{ f^{-1}(B) : B \in \mathcal{A}_2 \big\}, \quad f(\mathcal{A}_1) := \big\{ B \subseteq \Omega_2 : f^{-1}(B) \in \mathcal{A}_1 \big\} \tag{2.2}$$

wieder σ-Algebren.
(ii) Wir zeigen, dass für jedes Mengensystem $\mathcal{S}_2 \subseteq \mathfrak{p}(\Omega_2)$ und jedes $f : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ die folgende Gleichheit gilt:

$$\mathcal{A}_\sigma\big(f^{-1}(\mathcal{S}_2)\big) = f^{-1}\big(\mathcal{A}_\sigma(\mathcal{S}_2)\big) \tag{2.3}$$

Einerseits gilt offensichtlich $f^{-1}(\mathcal{S}_2) \subseteq f^{-1}(\mathcal{A}_\sigma(\mathcal{S}_2))$, also $\mathcal{A}_\sigma\big(f^{-1}(\mathcal{S}_2)\big) \subseteq \mathcal{A}_\sigma\big(f^{-1}(\mathcal{A}_\sigma(\mathcal{S}_2))\big) = f^{-1}(\mathcal{A}_\sigma(\mathcal{S}_2))$. Für den Beweis der anderen Inklusion betrachten wir das Mengensystem $\mathcal{C} := f\Big(\mathcal{A}_\sigma\big(f^{-1}(\mathcal{S}_2)\big)\Big)$. Nach (i) ist $\mathcal{C}$ eine σ-Algebra und für jedes $B \in \mathcal{S}_2$ gilt offensichtlich $f^{-1}(B) \in f^{-1}(\mathcal{S}_2) \subseteq \mathcal{A}_\sigma\big(f^{-1}(\mathcal{S}_2)\big)$, also $B \in \mathcal{C}$ und damit $\mathcal{A}_\sigma(\mathcal{S}_2) \subseteq \mathcal{C}$. Anwendung von f^{-1} liefert schliesslich $f^{-1}\big(\mathcal{A}_\sigma(\mathcal{S}_2)\big) \subseteq f^{-1}(\mathcal{C}) \subseteq \mathcal{A}_\sigma\big(f^{-1}(\mathcal{S}_2)\big)$.
(iii) Satz 2.2 kann nun wie folgt bewiesen werden: Aus $f^{-1}(E_2) \in \mathcal{A}_1$ für jedes $E_2 \in \mathcal{E}_2$ folgt unter Verwendung von (ii) $f^{-1}(\mathcal{A}_2) = f^{-1}(\mathcal{A}_\sigma(\mathcal{E}_2)) = \mathcal{A}_\sigma(f^{-1}(\mathcal{E}_2)) \subseteq \mathcal{A}_\sigma(\mathcal{A}_1) = \mathcal{A}_1$, also die gewünschte Eigenschaft. ■

Definition 2.3 *Seien* $(\Omega_1, \mathcal{A}_1)$ *und* $(\Omega_2, \mathcal{A}_2)$ *Messräume. Eine Abbildung* $f : \Omega_1 \to \Omega_2$ *heisst* $\mathcal{A}_1 - \mathcal{A}_2$ messbar *genau dann wenn für jedes* $A_2 \in \mathcal{A}_2$ *das vollständige Urbild*

$$f^{-1}(A_2) := \{\omega \in \Omega_1 : f(\omega) \in A_2\} \in \mathcal{A}_1 \tag{2.4}$$

erfüllt. Eine Abbildung $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *heisst* d-dimensionaler Zufallsvektor *falls sie* $\mathcal{A} - \mathcal{B}(\mathbb{R}^d)$ *messbar ist; im Fall* $d = 1$ *nennen wir ein solches* X *Zufallsvariable. Eine Zufallsvariable heisst* diskret *wenn sie höchstens abzählbar viele Werte annimmt, sie heisst* Indikatorvariable *wenn es eine Menge* $A \in \mathcal{A}$ *gibt, sodass*

$$X(\omega) = \mathbf{1}_A(\omega) := \begin{cases} 1 & \textit{wenn } \omega \in A \\ 0 & \textit{wenn } \omega \notin A \end{cases} \tag{2.5}$$

für alle $\omega \in \Omega$ *gilt.*

Beispiel 2.4 Wir betrachten $\Omega = \{1, 2, 3\}$ und $\mathcal{A} = \{\emptyset, \Omega, \{1, 2\}, \{3\}\}$, sowie die folgenden zwei Abbildungen: $X, Y : \Omega \to \mathbb{R}$, definiert durch $X(1) = X(2) = 0, X(3) = 1$ und $Y(1) = 1, Y(2) = 2, Y(3) = 3$. Dann ist X eine Zufallsvariable, Y jedoch nicht (warum ?).

Beispiel 2.5 Auch aus der Analysis bekannte 'komplizierte', nicht-Riemann-integrierbare Funktionen können Zufallsvariable sein. Betrachten wir zum Beispiel $\Omega = [0, 1]$, $\mathcal{A} = \mathcal{B}([0, 1])$, $P = \lambda$ und setzen $\mathbb{Q}_0 = \mathbb{Q} \cap [0, 1]$ sowie $X = \mathbf{1}_{\mathbb{Q}_0^c}$. Dann ist X sogar eine Indikatorvariable und wir erhalten für $\mathbb{E}(X)$

$$\mathbb{E}(X) = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = P(X = 1) = \lambda(\mathbb{Q}_0^c) = 1 - \lambda(\mathbb{Q}_0) = 1.$$

Wäre 1 nicht auch der intuitiv richtige Wert für $\int_{[0,1]} \mathbf{1}_{\mathbb{Q}_0^c}(t)\, dt$?

Nachdem es gemäß Satz 2.2 reicht, Erzeuger zu betrachten, erhalten wir sofort folgendes Resultat:

Folgerung 2.6 *Eine Abbildung $X : (\Omega, \mathcal{A}, P) \to \mathbb{R}$ ist eine Zufallsvariable genau dann wenn für eines der in Lemma 1.14 erwähnten Mengensysteme $\mathcal{E}$ gilt, dass $X^{-1}(E) \in \mathcal{A}$ für jedes $E \in \mathcal{E}$.*

Bemerkung 2.7 Beachten Sie die Ähnlichkeit zwischen dem Begriff der stetigen Abbildung (Urbild jeder offenen Menge ist offen) und dem Begriff der messbaren Abbildung (Urbild jeder messbaren Menge ist messbar).

Lemma 2.8 *Jede stetige Abbildung $f : \mathbb{R}^d \to \mathbb{R}^\rho$ ist Borel-messbar.*

Beweis: Stetigkeit von f impliziert, dass das Urbild jeder offenen Menge offen ist, es gilt also $f^{-1}(V) \in \mathcal{B}(\mathbb{R}^d)$ für jedes offene $V \subseteq \mathbb{R}^\rho$. Anwendung von Satz 2.2 liefert sofort das gewünschte Resultat. ■

Das folgende einfache Resultat werden wir im Folgenden häufig verwenden:

Satz 2.9 *Seien $X_1, X_2, \ldots X_d$ Zufallsvariable auf $(\Omega, \mathcal{A}, P)$ und $f : \mathbb{R}^d \to \mathbb{R}^\rho$ $\mathcal{B}(\mathbb{R}^d) - \mathcal{B}(\mathbb{R}^\rho)$-messbar (kurz: Borel-messbar). Dann ist sowohl $X = (X_1, X_2, \ldots, X_d)$ als auch $f \circ X$ ein Zufallsvektor.*

Beweis: Gemäß Bemerkung 1.15 ist $\mathcal{I}_d$ ein Erzeuger von $\mathcal{B}(\mathbb{R}^d)$. Für $\times_{i=1}^d (a_i, b_i]$ gilt aber offensichtlich

$$X^{-1}\Big(\bigtimes_{i=1}^{d} (a_i, b_i]\Big) = \bigcap_{i=1}^{d} X_i^{-1}((a_i, b_i]) \in \mathcal{A}.$$

Anwendung von Satz 2.2 liefert also sofort die Messbarkeit von X. Um die zweite Aussage zu beweisen, reicht es offensichtlich, zu zeigen, dass die Verkettung zweier messbarer Funktionen $g : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ und $f : (\Omega_2, \mathcal{A}_2) \to (\Omega_3, \mathcal{A}_3)$ wieder messbar ist, was leicht zu zeigen ist: Für $C \in \mathcal{A}_3$ gilt

$$(f \circ g)^{-1}(C) = \big\{\omega \in \Omega_1 : f(g(\omega)) \in C\big\} = \big\{\omega \in \Omega_1 : g(\omega) \in f^{-1}(C)\big\} = g^{-1}\big(\underbrace{f^{-1}(C)}_{\in \mathcal{A}_2}\big),$$

also $(f \circ g)^{-1}(C) \in g^{-1}(\mathcal{A}_2) \subseteq \mathcal{A}_1$. ■

Nachdem, wie gesagt, jede stetige Abbildung $f : \mathbb{R}^d \to \mathbb{R}$ Borel-messbar ist erhalten wir insbesondere:

Beispiel 2.10 *Seien $X_1, X_2, \ldots X_d$ Zufallsvariable auf $(\Omega, \mathcal{A}, P)$ und $a_1, a_2, \ldots, a_d \in \mathbb{R}$. Dann sind die folgenden Funktionen ebenfalls Zufallsvariable auf $(\Omega, \mathcal{A}, P)$:*

- $\sum_{i=1}^d a_i X_i$
- $\min(X_1, \ldots, X_d)$, $\max(X_1, \ldots, X_d)$
- $\prod_{i=1}^d X_i$

- $\sum_{i=1}^{d}(X_i - a_i)^2$

Frage 2.11 Gibt es überhaupt Abbildungen $X : (\Omega, \mathcal{A}, P) \to \mathbb{R}$, die keine Zufallsvariable sind?

Satz 2.12 *Jeder d-dimensionale Zufallsvektor $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ induziert durch*

$$P^X(B) := P(X^{-1}(B)) = P\big(\{\omega \in \Omega : X(\omega) \in B\}\big) \quad \text{für } B \in \mathcal{B}(\mathbb{R}^d) \tag{2.6}$$

ein Wahrscheinlichkeitsmaß P^X auf $\mathcal{B}(\mathbb{R}^d)$). Wir nennen P^X im Folgenden die Verteilung *von X und schreiben im Falle von $P^X = \mu$ auch $X \sim \mu$.*

Beweis: Übungsaufgabe

Definition 2.13 *Für jeden d-dimensionale Zufallsvektor $X = (X_1, \ldots, X_d) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ heisst die Funktion $F_X : \mathbb{R}^d \to [0,1]$, definiert durch*

$$F_X(x_1, \ldots, x_d) := P\Big(\bigcap_{i=1}^{d} X_i^{-1}((-\infty, x_i]\Big) \tag{2.7}$$

Verteilungsfunktion *von X.*