Wahrscheinlichkeitstheorie I

Prof. Dr. Wilhelm Stannat

Der vorliegende Text ist eine Zusammenfassung des ersten Kapitels der Vorlesung **Wahrscheinlichkeitstheorie I** im Sommersemester 2019 an der TU Berlin.

Korrekturen und Kommentare bitte per Email an `stannat@math.tu-berlin.de`

Inhaltsverzeichnis

1 Mathematische Modellierung von Zufallsexperimenten

1.1 Axiomatischer Aufbau der Wahrscheinlichkeitstheorie

Ziel dieses einleitenden Kapitels ist die Bereitstellung des grundlegenden Begriffsapparats zur mathematischen Beschreibung von Zufallsexperimenten. Unter einem **Zufallsexperiment** versteht man dabei zunächst einmal einen zeitlich wie örtlich fest umrissenen Vorgang mit unbestimmtem Ausgang. Beispiele hierfür sind

- Werfen eines Würfels oder Werfen einer Münze
- Zufälliges Ziehen von Kugeln aus einer Urne
- Kartenspiele

Daneben aber eben auch:

- Wahlergebnis der nächsten Europawahl
- Temperatur auf dem Alexanderplatz am 11. April 2019, 12:00
- Lebensdauern

Die Gesamtheit aller möglichen Ausgänge eines Zufallsexperiments heißt **Ergebnisraum** oder **Stichprobenraum** und wird mit Ω bezeichnet. Ein Element $\omega \in \Omega$ heißt **Ergebnis** oder **Stichprobe**. Es stellt einen möglichen Ausgang des zugrundeliegenden Zufallsexperiments dar.

Beispiele

(i) einmaliges Würfeln: $\Omega = \{1, 2, \ldots, 6\}$, $|\Omega| = 6$

(Hierbei bezeichnet $|\Omega|$ die **Mächtigkeit der Menge** Ω, also die Anzahl der Elemente in Ω.)

(ii) zweimaliges Würfeln:

$$\Omega = \{(i,j) : i, j \in \{1, \ldots, 6\}\} = \{1, 2, \ldots, 6\} \times \{1, 2, \ldots, 6\} = \{1, 2, \ldots, 6\}^2$$

also $|\Omega| = 36$.

(iii) Münzwurf: $\Omega = \{$ Kopf, Zahl $\}$.

(iv) Autos am Autobahndreieck Funkturm am 11. April 2019: $\Omega = \{0, 1, 2, 3, \ldots\} = \mathbb{N} \cup \{0\}$

(v) Temperatur in Grad Kelvin auf dem Alexanderplatz am 1. Juli 2016, 12 Uhr Mittags: $\Omega = [0, \infty[$ oder realistischer $[280, 310]$ $(0°C = 273.15°K)$

In den ersten vier Fällen sind die Ergebnisräume **endlich** oder **abzählbar unendlich**. Solche Ergebnisräume nennt man auch **diskret**. Im fünften Fall ist der Ergebnisraum nicht mehr abzählbar, sondern eine **kontinuierliche** Menge.

Ereignisse

Teilmengen $A \subset \Omega$ von Ω heißen **Ereignisse**. Die Gesamtheit aller Ereignisse ist also nichts anderes als $\mathcal{P}(\Omega)$, also die **Potenzmenge** von Ω (d.h. die Gesamtheit aller Teilmengen von Ω einschließlich der leeren Menge $\varnothing$ und der Menge Ω selber). Für einen endlichen Ergebnisraum gilt bekanntlich $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$. Unter den Ereignissen sind besonders hervorzuheben:

- das **sichere Ereignis** Ω,
- das **unmögliche Ereignis** $\varnothing$,
- die **Elementarereignisse** $\{\omega\}$ für $\omega \in \Omega$.

Beachte: Ereignisse sind Elemente der Potenzmenge $\mathcal{P}(\Omega)$ von Ω, also Teilmengen von Ω, während Ergebnisse Elemente von Ω sind. Insbesondere sind also das Ergebnis ω und das Elementarereignis $\{\omega\}$ voneinander zu unterscheiden.

Beispiele

(i) $A = \{1, 3, 5\} = \texttt{Augenzahl ungerade}$

(ii) $A = \{(5,6),(6,5),(6,6)\} = \texttt{Augensumme > 10}$

(iv) $A = \{40.000, 40.001, \ldots\} = \{n : n \geq 40.000\} = \texttt{ungewöhnlich hohes Verkehrsaufkommen}$

Die bekannten Mengenoperationen lassen sich als **Operationen auf Ereignissen** interpretieren:

$$A \cup B = A \texttt{ oder } B \texttt{ tritt ein}$$

$$A_1 \cup A_2 \cup \ldots \cup A_n =: \bigcup_{k=1}^n A_k = \texttt{mind. eines der } A_k \texttt{ tritt ein}$$

$$A \cap B = A \texttt{ und } B \texttt{ treten ein}$$

$$A_1 \cap A_2 \cap \ldots \cap A_n =: \bigcap_{k=1}^n A_k = \texttt{alle } A_k \texttt{ treten ein}$$

$$A^c := \Omega \setminus A := \{\omega \in \Omega : \omega \notin A\} = A \texttt{ tritt nicht ein}$$

A^c bezeichnet also das **Komplement** der Menge A (in Ω). Es gilt

$$\Omega^c = \varnothing \text{ und } \varnothing^c = \Omega\,.$$

Insbesondere im Falle kontinuierlicher Mengen ist es im allgemeinen unmöglich, jedem Ereignis in konsistenter Weise eine Wahrscheinlichkeit zuzuordnen (siehe Satz 1.15 unten) und man schränkt sich auf kleinere Mengensysteme ein, die unter den oben aufgezählten Mengenoperationen abgeschlossen sein sollen. Dies führt auf den Begriff der σ-**Algebra**.

Definition 1.1. Es sei $\Omega \neq \varnothing$. Ein Mengensystem $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt σ-**Algebra** falls gilt:

(a) $\Omega \in \mathcal{A}$

(b) $A \in \mathcal{A} \quad \Longrightarrow \quad A^c \in \mathcal{A}$

(c) $(A_n)_{n\in\mathbb{N}} \in \mathcal{A} \quad \Longrightarrow \quad \bigcup_{n\geq 1} A_n \in \mathcal{A}$

Das Paar $(\Omega, \mathcal{A})$ heißt **messbarer Raum** oder **Raum der messbaren Ereignisse**. Eine Teilmenge $A \subseteq \Omega$ mit $A \in \mathcal{A}$ heißt $\mathcal{A}$-**messbar** (oder auch einfach nur messbar, wenn der Bezug zur zugrundeliegenden σ-Algebra $\mathcal{A}$ klar ist).

Lemma 1.2. Es sei $\mathcal{A}$ eine σ-Algebra in Ω. Dann gilt:

(i) $\varnothing \in \mathcal{A}$

(ii) $A, B \in \mathcal{A} \quad \Longrightarrow \quad A \cup B,\ A \cap B,\ A \backslash B \in \mathcal{A}$,

(iii) $(A_n)_{n\in\mathbb{N}} \in \mathcal{A} \quad \Longrightarrow \quad \bigcap_{n\geq 1} A_n \in \mathcal{A}$.

Beweis. (i) ist offensichtlich, da $\Omega \in \mathcal{A}$ und daher aufgrund von (b) $\varnothing = \Omega^c \in \mathcal{A}$.
(ii) Für die Folge $A_1 := A$, $A_2 := B$, $A_n := \varnothing$, $n \geq 3$, messbarer Ereignisse ($\in \mathcal{A}$) folgt aufgrund von (c), dass $A \cup B = \bigcup_{n\geq 1} A_n \in \mathcal{A}$.
Aus A, $B \in \mathcal{A}$ folgt hieraus weiterhin, wegen A^c, $B^c \in \mathcal{A}$, auch $A^c \cup B^c \in \mathcal{A}$ und daher

$$A \cap B = (A^c \cup B^c)^c \in \mathcal{A}\,.$$

Schließlich ergibt sich hieraus auch $A \backslash B = A \cap B^c \in \mathcal{A}$.
(iii) Übungsaufgabe. $\square$

Beispiel

(i) $\mathcal{P}(\Omega)$ ist eine σ-Algebra, und zwar die größtmögliche.

(ii) $\{\Omega, \varnothing\}$ ist eine σ-Algebra, und zwar die kleinstmögliche, **triviale σ-Algebra**.

(iii) Ist $A \subseteq \Omega$ beliebiges Ereignis, so ist $\{A, A^c, \Omega, \varnothing\}$ eine σ-Algebra, und zwar die kleinste σ-Algebra, die A enthält.

(iv) Ist $\mathcal{C} \subseteq \mathcal{P}(\Omega)$ ein beliebiges Mengensystem, so ist

$$\sigma(\mathcal{C}) := \bigcap_{\substack{\mathcal{C}\subseteq\mathcal{A} \\ \mathcal{A}\ \sigma\text{–Algebra auf }\Omega}} \mathcal{A}$$

eine σ-Algebra (Beweis!) und zwar die kleinste σ-Algebra, die $\mathcal{C}$ enthält. Sie heißt **die von $\mathcal{C}$ erzeugte σ-Algebra**. $\sigma(\mathcal{C})$ ist wohldefiniert, da der Schnitt nicht leer ist, da die Potenzmenge $\mathcal{P}(\Omega)$ eine σ-Algebra ist, die $\mathcal{C}$ enthält.
Weiterhin gelten die folgenden drei Aussagen:

(1) $\mathcal{C} \subseteq \sigma(\mathcal{C})$

(2) $\mathcal{C}_1 \subseteq \mathcal{C}_2 \Rightarrow \sigma(\mathcal{C}_1) \subseteq \sigma(\mathcal{C}_2)$

(3) $\mathcal{C}$ ist eine σ-Algebra genau dann wenn $\sigma(\mathcal{C}) = \mathcal{C}$.

Wahrscheinlichkeiten

Im nächsten Schritt wollen wir für jedes messbare Ereignis A eine Wahrscheinlichkeit $P(A)$ zwischen 0 und 1 festlegen. $P(A)$ soll ein Maß dafür sein, dass das Ereignis A eintritt:

- tritt A niemals ein, so setzt man $P(A) = 0$. Insbesondere $P(\varnothing) = 0$.
- tritt A sicher ein, so setzt man $P(A) = 1$. Insbesondere $P(\Omega) = 1$.

Zusätzlich sollte gelten: Sind A und B disjunkte Ereignisse, d.h. A und B enthalten keine gemeinsamen Ergebnisse, also $A \cap B = \varnothing$, so ist

$$P(A \cup B) = P(A) + P(B)\,. \tag{1.1}$$

Diese Eigenschaft von P bezeichnet man als **Additivität**.
Aus (1.1) folgt unmittelbar: sind $A_1, \ldots, A_n$ paarweise disjunkte Ereignisse, d.h. $A_k \cap A_l = \varnothing$ für $k \neq l$, so folgt:

$$P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n)\,. \tag{1.2}$$

Gilt schließlich auch für jede **unendliche** Folge $(A_n)_{n\in\mathbb{N}}$ paarweiser disjunkter Ereignisse

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \tag{1.3}$$

so spricht man von σ-**Additivität**.

Definition 1.3. Es sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Eine Abbildung

$$P : \mathcal{A} \to [0,1]$$

heißt **Wahrscheinlichkeitsmaß** auf $\mathcal{A}$, falls gilt:

(a) (Normiertheit) $P(\Omega) = 1$

(b) (σ-Additivität) Für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweiser disjunkter Ereignisse in $\mathcal{A}$ gilt

$$P\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} P(A_n)\,.$$

Lemma 1.4. (**Rechenregeln für** P)
Es seien P ein Wahrscheinlichkeitsmaß auf dem messbaren Raum $(\Omega, \mathcal{A})$ und $A, B, A_1, A_2, \ldots \in \mathcal{A}$. Dann gilt:

(i) $P(\varnothing) = 0$.

(ii) P ist (insbesondere) **endlich additiv**, d.h. für $A_1, \ldots, A_n$ paarweise disjunkt, ist

$$P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n) = \sum_{k=1}^{n} P(A_k)$$

(iii) (Komplementärwahrscheinlichkeit) $P(A^c) = 1 - P(A)$

(iv) (Monotonie) $A \subseteq B \implies P(A) \le P(B)$

(v) (Subadditivität) $A \subseteq \bigcup_n A_n \implies P(A) \le \sum_n P(A_n)$

Beweis. (i) Für die Folge $A_n = \emptyset$, $n = 1, 2, \ldots$, trivialerweise paarweise disjunkter Ereignisse gilt aufgrund der σ-Additivität von P wegen $\bigcup_{n\ge1} A_n = \emptyset$

$$P(\emptyset) = P\left(\bigcup_{n\ge1} A_n\right) = \sum_{n\ge1} P(\emptyset).$$

Wäre $P(\emptyset) > 0$, so wäre die rechte Seite unendlich, woraus sich ein Widerspruch ergibt. Es muss also $P(\emptyset) = 0$ gelten.
(ii) Mit $A_k := \emptyset$ für $k \ge n+1$ erhalten wir eine (unendliche) Folge paarweise disjunkter messbarer Ereignisse mit

$$\bigcup_{k\ge1} A_k = A_1 \cup \ldots \cup A_n$$

so dass aus der σ-Additivität von P folgt

$$\begin{aligned} P(A_1 \cup \ldots \cup A_n) &= P\left(\bigcup_{k\ge1} A_k\right) = \sum_{k\ge1} P(A_k) \\ &= \sum_{k=1}^{n} P(A_k). \end{aligned}$$

(iii) Da A und A^c disjunkt sind und $A \cup A^c = \Omega$, folgt aus der endlichen Additivität

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c).$$

(iv) Da $B = A \cup (B \cap A^c)$, sowie A und $B \cap A^c$ disjunkt, ergibt sich wieder aus der endlichen Additivität

$$P(B) = P(A) + P(B \cap A^c) \ge P(A).$$

(v) Definiere induktiv die Ereignisse $B_1 := A_1$ und $B_n := A_n \setminus (A_1 \cup \ldots \cup A_{n-1})$, $n = 2, 3, \ldots$. Dann gilt für alle n $B_n \subseteq A_n$ und $B_1 \cup \ldots \cup B_n = A_1 \cup \ldots \cup A_n$. Insbesondere sind die Ereignisse $(B_n)_{n\ge1}$ paarweise disjunkt. Aus der σ-Additivität von P folgt nun

$$P\left(\bigcup_n A_n\right) = P\left(\bigcup_n B_n\right) = \sum_n P(B_n) \le \sum_n P(A_n).$$

Ist allgemein $A \subseteq \bigcup_n A_n$, so folgt aus obiger Rechnung auch

$$P(A) \le P\left(\bigcup_n A_n\right) \le \sum_n P(A_n).$$

□

Beispiel 1.5. (i) Beim Würfeln mit einem fairen Würfel ist jede der sechs möglichen Augenzahlen gleichwahrscheinlich. Man setzt daher

$$P(\{\omega\}) = \frac{1}{6} \text{ für } \omega \in \Omega = \{1, 2, 3, 4, 5, 6\}.$$

Es folgt z.B.

$$P(\texttt{Augenzahl ungerade}) = P(\{1, 3, 5\}) = \frac{3}{6} = \frac{1}{2}.$$

(ii) Beim zweimaligen Würfeln mit einem fairen Würfel ist wiederum jedes der 36 Ergebnisse aus $\Omega = \{1,2,3,4,5,6\}^2$ gleichwahrscheinlich, also $P(\{\omega\}) = \frac{1}{36}\ \forall \omega \in \Omega$. Es folgt z.B.

$$P(\texttt{Augensumme} > 10) = P(\{(5,6),(6,5),(6,6)\}) = \frac{3}{36} = \frac{1}{12}\,.$$

Die drei Komponenten Ω, $\mathcal{A}$ und P werden nun im dem für die Wahrscheinlichkeitstheorie zentralen Begriff des Wahrscheinlichkeitsraumes zusammengefasst.

Definition 1.6. Es sei $\Omega \neq \varnothing$, $\mathcal{A}$ eine σ-Algebra in Ω und P ein Wahrscheinlichkeitsmaß auf $\mathcal{A}$. Dann heißt das Tripel $(\Omega, \mathcal{A}, P)$ **Wahrscheinlichkeitsraum**.

Zum Abschluss dieses Abschnittes leiten wir noch einige weitere Eigenschaften von Wahrscheinlichkeitsmaßen ab.

Satz 1.7. (Einschluss-Ausschluss-Prinzip, Formel von Sylvester)
Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Dann gilt für beliebige $A_1, \ldots, A_n \in \mathcal{A}$:

$$\begin{aligned} P(A_1 \cup \ldots \cup A_n) &= \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} P\left(\bigcap_{i \in I} A_i\right) \\ &= \sum_{j=1}^{n} (-1)^{j-1} \sum_{1 \leq i_1 < \ldots < i_j \leq n} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_j}) \end{aligned}$$

Insbesondere gelten für $A, B, C \in \mathcal{A}$ die Spezialfälle

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad + P(A \cap B \cap C)\,. \end{aligned}$$

Beweis. Durch Induktion nach n. Für $n = 2$ ist $A_1 \cup A_2 = A_1 \cup (A_2 \backslash A_1)$ eine Zerlegung in disjunkte Ereignisse, also

$$P(A_1 \cup A_2) = P(A_1) + P(A_2 \backslash A_1)\,. \tag{1.4}$$

Da außerdem $A_2 = (A_2 \backslash A_1) \cup (A_2 \cap A_1)$ eine Zerlegung von A_2 in disjunkte Ereignisse ist, gilt $P(A_2) = P(A_2 \backslash A_1) + P(A_2 \cap A_1)$ und damit

$$P(A_2 \backslash A_1) = P(A_2) - P(A_1 \cap A_2)\,. \tag{1.5}$$

Einsetzen von (1.5) in (1.4) liefert schließlich

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)\,. \tag{1.6}$$

Im Induktionsschluss von n auf $n+1$ wenden wir (1.6) an auf die beiden Mengen

$$A_1 \cup \ldots \cup A_n \quad \text{und} \quad A_{n+1}$$

und erhalten

$$\begin{aligned} P(A_1 \cup \ldots \cup A_{n+1}) &= P(A_1 \cup \ldots \cup A_n) + P(A_{n+1}) - P((A_1 \cup \ldots \cup A_n) \cap A_{n+1}) \\ &= P(A_1 \cup \ldots \cup A_n) + P(A_{n+1}) - P((A_1 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1})) \end{aligned} \tag{1.7}$$

Nach Induktionsannahme gilt

$$P(A_1 \cup \ldots \cup A_n) = \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} P\left(\bigcap_{i\in I} A_i\right)$$

und

$$\begin{aligned} P\left((A_1 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1})\right) &= \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} P\left(\bigcap_{i\in I} (A_i \cap A_{n+1})\right) \\ &= \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} P\left(\bigcap_{i\in I\cup\{n+1\}} A_i\right). \end{aligned}$$

Einsetzen in (1.7) ergibt

$$\begin{aligned} P(A_1 \cup \ldots \cup A_{n+1}) &= P(A_1 \cup \ldots \cup A_n) + P(A_{n+1}) - P\left((A_1 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1})\right) \\ &= \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} P\left(\bigcap_{i\in I} A_i\right) + P(A_{n+1}) - \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} P\left(\bigcap_{i\in I\cup\{n+1\}} A_i\right) \\ &= \sum_{\varnothing \neq I \subseteq \{1,\ldots,n+1\}} (-1)^{|I|-1} P\left(\bigcap_{i\in I} A_i\right). \end{aligned}$$

□

Schließlich benötigen wir noch folgendes Ergebnis zu Stetigkeitseigenschaften von Wahrscheinlichkeitsmaßen aus der Maßtheorie, das wir ohne Beweis hier angeben.

Satz 1.8. Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $(A_n)_{n\in\mathbb{N}}$ eine Folge von Ereignissen aus $\mathcal{A}$. Dann gilt:

(i) Falls $A_n \uparrow A$, d.h. $A_1 \subseteq A_2 \subseteq A_2 \subseteq \ldots$ und $\bigcup_{n\geq 1} A_n = A$, so folgt

$$P(A_n) \uparrow P(A) \qquad \text{(Stetigkeit von unten)}.$$

(ii) Falls $A_n \downarrow A$, d.h. $A_1 \supseteq A_2 \supseteq A_2 \supseteq \ldots$ und $\bigcap_{n\geq 1} A_n = A$, so folgt

$$P(A_n) \downarrow P(A) \qquad \text{(Stetigkeit von oben)}.$$

1.2 Diskrete Wahrscheinlichkeitsräume

Ist der Ergebnisraum Ω **diskret**, d.h. die Menge Ω endlich oder höchstens abzählbar unendlich, so ist jede Teilmenge $A \subseteq \Omega$ abzählbar, etwa $A = \{\omega_1, \omega_2, \ldots\}$ und damit

$$A = \bigcup_n \{\omega_n\}$$

abzählbare Vereinigung gewisser Elementarereignisse. Ist also $\mathcal{A}$ eine σ-Algebra in Ω, die alle Elementarereignisse enthält, so enthält sie bereits alle Teilmengen von Ω und stimmt daher mit der Potenzmenge $\mathcal{P}(\Omega)$ überein. In diesem Sinne ist für diskrete Ergebnisräume Ω die Potenzmenge $\mathcal{P}(\Omega)$ die natürliche σ-Algebra und sie wird häufig nicht explizit mitangegeben. Ist weiterhin P ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$, so können wir für jede Teilmenge

$A = \{\omega_1, \omega_2, \ldots\} \subseteq \Omega$ die Wahrscheinlichkeit $P(A)$ aufgrund der σ-Additivität ausrechnen als den Wert der (möglicherweise unendlichen) Reihe

$$P(A) = \sum_n P(\{\omega_n\}) \; . \tag{1.8}$$

Hierbei kommt es auf die Reihenfolge der Aufzählung der Elemente von A nicht an, denn die Reihe (1.8) ist absolut konvergent. Daher schreiben wir im folgenden auch einfach

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) \; . \tag{1.9}$$

Insbesondere ist also im diskreten Fall jedes Wahrscheinlichkeitsmaß P auf $\mathcal{P}(\Omega)$ eindeutig durch die Wahrscheinlichkeiten $P(\{\omega\})$ der Elementarereignisse bestimmt. Durch $p(\omega) := P(\{\omega\})$ erhalten wir aus dem Wahrscheinlichkeitsmaß P eine Funktion $p : \Omega \to [0,1]$ mit der Eigenschaft $\sum_{\omega \in \Omega} p(\omega) = 1$. Umgekehrt liefert jede solche Funktion p durch (1.9) ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$ (siehe Satz 1.10) und damit eine sehr einfache Methode zur Konstruktion von Wahrscheinlichkeitsmaßen auf diskreten Ergebnisräumen.

Definition 1.9. Es sei $\Omega \neq \varnothing$ höchstens abzählbar. Eine Funktion $p : \Omega \to [0,1]$ mit $\sum_{\omega \in \Omega} p(\omega) = 1$ heißt **Zähldichte** oder auch **Wahrscheinlichkeitsfunktion**.

Satz 1.10. Es sei $\Omega \neq \varnothing$ höchstens abzählbar und p eine Zähldichte. Dann wird durch

$$P(A) := \sum_{\omega \in A} p(\omega) \, , A \subseteq \Omega \tag{1.10}$$

ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$ definiert.

Beweis. Nichtnegativität $P(A) \geq 0$ und Normiertheit $P(\Omega) = 1$ folgen direkt aus den Eigenschaften einer Zähldichte. Es bleibt also lediglich die σ-Additivität nachzuweisen. Dazu sei $(A_n)_{n \in \mathbb{N}}$ eine Folge paarweiser disjunkter Ereignisse und wir nehmen an, dass für alle $n \geq 1$ durch

$$A_n = \{\omega_{n1}, \omega_{n2}, \ldots\} = \{\omega_{nk}\}_{1 \leq k < a_n} \text{ mit } a_n = |A_n| + 1$$

eine Aufzählung aller Elemente von A_n gegeben ist. (Beachte, dass a_n unendlich sein könnte.) Dann ist $\{\omega_{nk} : n \geq 1, 1 \leq k < a_n\}$ eine Aufzählung aller Elemente von $A := \bigcup_{n \geq 1} A_n$, wobei jedes Element aus A genau einmal aufgezählt wird, da die Ereignisse $(A_n)_{n \in \mathbb{N}}$ paarweise disjunkt sind. Eine entsprechende Anordnung der Reihenglieder $p(\omega_{nk})$ ergibt

$$P(A) = \sum_{\omega \in A} p(\omega) = \sum_{n \geq 1} \sum_{1 \leq k < a_n} p(\omega_{nk}) = \sum_{n \geq 1} \sum_{\omega \in A_n} p(\omega) = \sum_{n \geq 1} P(A_n) \, .$$

□

Satz 1.10 besagt also insbesondere, dass für diskrete Ergebnisräume durch (1.10) eine $1-1$-Beziehung zwischen Zähldichten und Wahrscheinlichkeitsmaßen auf $\mathcal{P}(\Omega)$ gegeben ist. Für die vollständige mathematische Beschreibung des Wahrscheinlichkeitsraumes reicht also in diesem Spezialfall die Angabe des Paares (Ω, p) aus.

Beispiel 1.11. n-maliges Werfen einer fairen Münze.

$$\Omega = \{(i_1, i_2, \ldots, i_n) : i_j \in \{0,1\}, 1 \leq j \leq n\} = \{0,1\}^n \, .$$

Hierbei steht 0 für das Ereignis `Kopf` und 1 für das Ereignis `Zahl`. "fair" bedeutet, dass alle Ergebnisse gleichwahrscheinlich sind, also

$$p(\omega) = \frac{1}{|\Omega|} = \frac{1}{2^n}\,.$$

Für $k \in \{1, 2, \dots, n\}$ ist dann

$$\begin{aligned} &P(\texttt{im k-ten Münzwurf "Kopf" werfen}) \\ &\quad = |\{(i_1, \dots, i_{k-1}, 0, i_{k+1}, \dots, i_n) : i_j \in \{0,1\}, 1 \le j \le n\}| \cdot 2^{-n} \\ &\quad = 2^{n-1} \cdot 2^{-n} = \frac{1}{2}\,. \end{aligned}$$

Das obige Beispiel ist Spezialfall eines Laplaceschen Wahrscheinlichkeitsraumes.

Laplacescher Wahrscheinlichkeitsraum

Ist Ω eine endliche Menge, so definiert

$$p(\omega) := \frac{1}{|\Omega|}\,, \quad \omega \in \Omega$$

eine Zähldichte auf Ω. Für die Wahrscheinlichkeit $P(A)$ eines beliebigen Ereignisses folgt hieraus sofort

$$P(A) = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|}\,. \tag{1.11}$$

$P(A)$ heißt **Laplace-Wahrscheinlichkeit von** A. Da jedes Elementarereignis gleichwahrscheinlich ist, spricht man von P auch als der **Gleichverteilung auf** Ω**.**

Die Berechnung der Wahrscheinlichkeit $P(A)$ in (1.11) führt auf das Problem der **Abzählung der Elemente in** A, also auf ein **Abzählproblem**. Die wichtigsten Abzählprobleme sollen im folgenden anhand von einfachen **Urnenmodellen** illustriert werden:

Eine Urne enthalte n unterscheidbare Kugeln $1, 2, \dots, n$. Wir unterscheiden dann das k-malige Ziehen einer Kugel aus der Urne mit/ohne Zurücklegen, wobei es auf die Reihenfolge der gezogenen Kugeln ankommt/nicht ankommt:

(1) **in Reihenfolge mit Zurücklegen**

$$\Omega = \{\omega = (x_1, \dots, x_k) : x_i \in \{1, \dots, n\}\}\,, |\Omega| = n^k$$

d.h., ein Ergebnis $\omega = (x_1, \dots, x_k)$ ist ein **k-Tupel**, d.h. eine geordnete Menge der Länge k, wobei x_i für die Nummer der i-ten gezogenen Kugel steht.

(2) **in Reihenfolge ohne Zurücklegen**

$$\Omega = \{\omega = (x_1, \dots, x_k) : x_i \in \{1, \dots n\}, x_i \neq x_j \text{ für } i \neq j\}$$

$$|\Omega| = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1) = \frac{n!}{(n-k)!}\,.$$

Zur Erinnerung: Fakultätsfunktion

$$m! := m(m-1)\cdot(m-2)\cdot\ldots\cdot 2\cdot 1 = \Pi_{k=1}^{m} k\,,\ \text{und } 0! := 1\,.$$

Insbesondere

$$n! = n\cdot(n-1)! = n\cdot(n-1)\cdot(n-2)! = \ldots = n\cdot(n-1)\cdot\ldots\cdot(n-k+1)\cdot(n-k)!\,,$$

also

$$\frac{n!}{(n-k)!} = n\cdot(n-1)\cdot\ldots\cdot(n-k+1)\,.$$

Für $k =$ n erhält man als Spezialfall

$$|\Omega| = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$$

$n!$ ist also gleich der Anzahl aller möglichen Anordnungen (oder auch **Permutationen**) der n-elementigen Menge $\{1,\ldots,n\}$.

(3) **ohne Reihenfolge ohne Zurücklegen**

$$\Omega = \{\omega = \{x_1,\ldots,x_k\} : x_i \in \{1,2,\ldots,n\}, x_i \neq x_j \text{ für } i \neq j\}$$

Im Unterschied zum Ziehen in Reihenfolge werden nun alle k-Tupel $(x_1,\ldots,x_k)$, die zu derselben Menge der gezogenen Kugeln führen, zu einem Elementarereignis zusammengefasst. Insgesamt gibt es $k!$ solcher Tupel (das entspricht also gerade der Anzahl der Permutationen der Menge der k gezogenen Kugeln), also erhalten wir insgesamt

$$\frac{n!}{(n-k)!}\cdot\frac{1}{k!} = \binom{n}{k}$$

Ergebnisse. Es gilt also

$$|\Omega| = \binom{n}{k}.$$

Insbesondere: $\binom{n}{k}$ ist gleich der Anzahl aller k-elementigen Teilmengen aus einer n-elementigen Grundmenge.

Alternative Darstellung von Ω: Unter allen k-Tupeln, die zur selben Menge $\{x_1,\ldots,x_k\}$ führen, gibt es genau ein Tupel $(x_{(1)},\ldots,x_{(k)})$, in dem die Elemente ihrer Größe nach angeordnet sind:

$$x_{(1)} < x_{(2)} < \ldots < x_{(k)}\,.$$

Wir können daher auch schreiben

$$\Omega = \{(x_1,\ldots,x_k) \,:\, x_i \in \{1,\ldots,n\}\,, x_1 < x_2 < \ldots < x_k\}\,.$$

(4) **ohne Reihenfolge mit Zurücklegen**

Analog zu (3) ordnen wir wieder die Nummern der gezogenen Kugeln der Größe nach an:

$$x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(k)} \tag{1.12}$$

wobei wegen des Zurücklegens Kugeln mehrfach gezogen werden können.

Durch Übergang von $x_{(i)}$ zu $x_{(i)} + i - 1$ erhält man aus (1.12) eine streng monoton aufsteigende Folge

$$x_{(1)} < x_{(2)} + 1 < x_{(3)} + 2 < \ldots < x_{(k)} + k - 1 .$$

Wir erhalten als Stichprobenraum in diesem Falle also

$$\Omega = \{(x_1, \ldots, x_k) : x_i \in \{1, \ldots, n, n+1, \ldots, n+k-1\}, x_1 < x_2 < \ldots < x_k\} .$$

Für die Mächtigkeit $|\Omega|$ von Ω ergibt sich nach (3)

$$|\Omega| = \binom{n+k-1}{k} .$$

Die Anzahl der Möglichkeiten in den vier verschiedenen Fällen nochmals im Überblick

	mit Zurücklegen	ohne Zurücklegen
in Reihenfolge	n^k	$\frac{n!}{(n-k)!}$
ohne Reihenfolge	$\binom{n+k-1}{k}$	$\binom{n}{k}$

Einige einfache Anwendungen

Satz 1.12. (Binomischer Lehrsatz)

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} .$$

Beweis. Zunächst schreibe

$$(x+y)^n = (x_1 + y_1) \cdot (x_2 + y_2) \cdot \ldots \cdot (x_n + y_n)$$

mit $x_i = x$ und $y_i = y$. Beim Ausmultiplizieren tritt der Term $x^k y^{n-k}$ immer dann auf, wenn in k Klammern der Faktor x_i und in $n-k$ Klammern der Faktor y_i gewählt wird, also in genau $\binom{n}{k}$ Fällen (vgl. Abzählmethode (3)). □

Korollar 1.13. (i) $\sum_{k=0}^{n} \binom{n}{k} = 2^n$

(ii) $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$

(iii) $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$.

Beweis. (i) Setze in 1.12 $x = y = 1$.
(ii) Setze in 1.12 $x = -1$, $y = 1$.
(iii) Aus

$$n(x+y)^{n-1} = \frac{d}{dx}(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}\frac{d}{dx}\left(x^k y^{n-k}\right) = \sum_{k=0}^{n} k\binom{n}{k}x^{k-1}y^{n-k}$$

folgt die Behauptung, indem man $x = y = 1$ setzt. □

1.3 Stetige Wahrscheinlichkeitsräume

Für viele Zufallsexperimente kann der Ergebnisraum nicht diskret gewählt werden (z.B. bei Wartezeiten, Lebensdauern, Lufttemperatur,...), sondern er muss vielmehr als kontinuierliche Menge, z.B. als Intervall $[a,b]$ oder gleich ganz $\mathbb{R}$ oder $\mathbb{R}^d$ gewählt werden. Kontinuierliche Ergebnisräume ergeben sich natürlicherweise auch bei unendlichen Wiederholungen von Zufallsexperimenten.

Beispiel 1.14. ∞-viele faire Münzwürfe
Der zugehörige Ergebnisraum ist der Raum der unendlichen $0-1$-Folgen

$$\Omega = \{(\omega_n)_{n\in\mathbb{N}} \mid \omega_n \in \{0,1\}\} = \{0,1\}^{\mathbb{N}}$$

und damit eine überabzählbare Menge. Wir können die ersten n Münzwürfe $\{\omega_1,\dots,\omega_n\}$ als Teilmenge von Ω auffassen, indem wir die zugehörige Zylindermenge

$$\{\omega_1,\dots,\omega_n\} \times \{0,1\} \times \{0,1\} \times \dots \subseteq \Omega$$

betrachten.

Gibt es nun ein Wahrscheinlichkeitsmaß P auf $\mathcal{P}(\Omega)$, so dass für jede solche Zylindermenge

$$P\left(\{\omega_1,\dots,\omega_n\} \times \{0,1\} \times \{0,1\} \times \dots\right) = 2^{-n} \tag{1.13}$$

gilt, d.h. also dass die Wahrscheinlichkeit eines Ereignisses, welches nur von den ersten n Münzwürfen abhängt, gerade der Wahrscheinlichkeit dieses Ereignisses bezüglich des in Beispiel 1.11 betrachteten Wahrscheinlichkeitsmaßes für n faire Münzwürfe entspricht? Die überraschende Antwort ist **nein**, wie Vitali im Jahre 1905 zeigen konnte. Dies ist der Kern einiger Paradoxa der Wahrscheinlichkeitsrechnung, wie z.B. das Bertrandsche Paradoxon (siehe Beispiel 1.19).

Der später gefundene Ausweg besteht in der Verkleinerung der betrachteten Ereignissysteme und in der Tat gibt es ein Wahrscheinlichkeitsmaß P mit der Eigenschaft (1.13) auf einer echt kleineren σ-Algebra $\mathcal{A}$ als der Potenzmenge $\mathcal{P}(\Omega)$, die andererseits aber immer noch alle, von endlich vielen Münzwürfen erzeugten, Zylindermengen enthält.

Zurück zum bereits erwähnten Satz von Vitali (siehe [Ge02]).

Satz 1.15. Es sei $\Omega = \{0,1\}^{\mathbb{N}}$. Dann gibt es **kein** Wahrscheinlichkeitsmaß P auf $\mathcal{P}(\Omega)$ mit

$$P(T_n(A)) = P(A) \text{ für alle } A \subseteq \Omega \text{ und } n \geq 1\,. \tag{1.14}$$

Hierbei ist

$$T_n(\omega_1,\omega_2,\dots) = (\omega_1,\dots,\omega_{n-1},1-\omega_n,\omega_{n+1},\dots)$$

und $T_n(A) = \{T_n((\omega_n)_{n\in\mathbb{N}}) \mid (\omega_n)_{n\in\mathbb{N}} \in A\}$ das Bild von A unter der Abbildung T_n.

Die Eigenschaft (1.14) des Wahrscheinlichkeitsmaßes P drückt die Fairness und die Unabhängigkeit der einzelnen Münzwürfe aus und ist für Zylindermengen gleichwertig zu (1.13).

Beweis. Betrachte auf Ω die folgende Äquivalenzrelation: $\omega \sim \omega'$ genau dann wenn $\omega_n = \omega'_n$ für alle bis auf endlich viele n. Nach dem Auswahlaxiom existiert eine Menge $A \subset \Omega$, die aus jeder Äquivalenzklasse genau ein Element enthält. Weiterhin sei $\mathcal{I} = \{I \subseteq \mathbb{N} \mid |I| < \infty\}$ das System aller endlichen Teilmengen von $\mathbb{N}$. Beachte, dass $\mathcal{I}$ ein abzählbares Mengensystem ist. Für eine endliche Indexmenge $I = \{n_1, \ldots n_k\}$ sei $T_I := T_{n_1} \circ \ldots \circ T_{n_k}$ der Flip der Münzwürfe $n_1, \ldots, n_k$. Dann gilt:

(i) $\Omega = \bigcup_{I \in \mathcal{I}} T_I(A)$, denn zu jedem $\omega \in \Omega$ gibt es ein $\omega' \in A$ mit $\omega \sim \omega'$ und somit eine endliche Indexmenge $I \subseteq \mathbb{N}$ mit $T_I(\omega') = \omega$.

(ii) Die Mengen $(T_I(A))_{I \in \mathcal{I}}$ sind paarweise disjunkt, denn aus $T_I(A) \cap T_{I'}(A) \neq \varnothing$ für zwei Indexmengen I, I' ergibt sich die Existenz von ω, $\omega' \in A$ mit $T_I(\omega) = T_{I'}(\omega')$, woraus dann aber
$$\omega \sim T_I(\omega) = T_{I'}(\omega') \sim \omega' \text{ also } \omega \sim \omega'$$
folgt. Es ergibt sich $\omega = \omega'$, da A aus jeder Äquivalenzklasse nur ein Element enthält. Mithin folgt $T_I(\omega) = T_{I'}(\omega)$ und damit $I = I'$.

Aus diesen beiden Eigenschaften folgt dann schließlich
$$1 = P(\Omega) = \sum_{I \in \mathcal{I}} P(T_I(A)) = \sum_{I \in \mathcal{I}} P(A)$$
was einen Widerspruch ergibt, da die rechte Seite nur 0 oder ∞ sein kann. □

Der Satz von Vitali verdeutlicht also die Notwendigkeit, im kontinuierlichen Fall im Vergleich zur Potenzmenge $\mathcal{P}(\Omega)$ echt kleinere Ereignissysteme zu betrachten. Wir wollen uns im folgenden auf $\Omega = \mathbb{R}^d$ als Ergebnisraum beschränken und hierauf das Mengensystem
$$\mathcal{C} := \{]a, b] =]a_1, b_1] \times \ldots \times]a_d, b_d] \mid a = (a_1, \ldots, a_d), b = (b_1, \ldots, b_d) \in \mathbb{R}^d, a \leq b\}$$
aller nach links offenen Intervalle betrachten. Hierbei schreiben wir $a \leq b$ genau dann wenn $a_i \leq b_i$ für $1 \leq i \leq d$. Anschließend betrachten wir nun die von $\mathcal{C}$ erzeugte σ-Algebra und bezeichnen diese mit $\mathcal{B}(\mathbb{R}^d)$. $\mathcal{B}(\mathbb{R}^d)$ heißt **Borelsche σ-Algebra** und in der Maßtheorie zeigt man, dass diese echt kleiner ist als $\mathcal{P}(\mathbb{R}^d)$ und dass sie alle offenen und alle abgeschlossenen Teilmengen $A \subseteq \mathbb{R}^d$ enthält.

An die Stelle der Zähldichte tritt im stetigen Fall eine Dichtefunktion $f : \mathbb{R}^d \to [0, \infty)$ im Sinne der folgenden Definition:

Definition 1.16. Eine Riemann-integrierbare (allgemeiner: Lebesgue-integrierbare) Funktion $f : \mathbb{R}^d \to [0, \infty)$ mit $\int_{R^d} f(x)\,dx = 1$ heißt **Dichte**.

Analog zu Satz 1.10 gilt dann

Satz 1.17. Es sei f eine Dichte auf $\mathbb{R}^d$. Dann wird durch
$$P(A) := \int_A f(x)\,dx \quad \text{für } A \subseteq \mathcal{B}(\mathbb{R}^d)\,. \tag{1.15}$$
ein Wahrscheinlichkeitsmaß P auf $\mathcal{B}(\mathbb{R}^d)$ definiert.

Analog zum Beweis von Satz 1.10 im diskreten Fall sind Nichtnegativität $P(A) \geq 0$ und Normiertheit $P(\Omega) = 1$ offensichtlich, und es bleibt nur, die σ-Additivität zu zeigen. Der vollständige Beweis hierzu findet sich in Büchern zur Maß- und Integrationstheorie (siehe [Ba91], [Kl06]), wird aber für das weitere Verständnis der Vorlesung Wahrscheinlichkeitstheorie I vorerst nicht benötigt.

Beispiel 1.18. Gleichverteilung auf $\Omega = [a, b]$
Das d-dimensionale Volumen des Intervalls $[a, b]$ ist bekanntlich gleich $\prod_{i=1}^{d}(b_i - a_i)$. Daher ist die Funktion

$$f(x) = \begin{cases} \left(\prod_{i=1}^{d}(b_i - a_i)\right)^{-1} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

eine Dichte. Das zugehörige Wahrscheinlichkeitsmaß P auf $\mathcal{B}(\mathbb{R}^d)$ heißt **Gleichverteilung** auf $[a, b]$. Bezeichnen wir für eine Borelmenge $A \subseteq [a, b]$ ihr d-dimensionales Volumen mit $|A|$, so ergibt sich analog zum diskreten Fall (1.11) die Formel

$$P(A) = \frac{|A|}{|[a, b]|}. \tag{1.16}$$

Allgemeiner kann man zu jeder Borelmenge $\Omega \subseteq \mathbb{R}^d$ mit endlichem d-dimensionalem Volumen $|\Omega|$ über die Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{|\Omega|} & \text{für } x \in \Omega \\ 0 & \text{sonst} \end{cases}$$

die Gleichverteilung auf Ω definieren. Analog zu (1.16) ergibt sich auch in diesem Fall

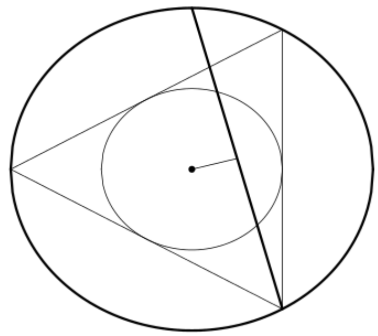
$$P(A) = \frac{|A|}{|\Omega|} \quad \text{für jede Borelmenge } A \subseteq \Omega. \tag{1.17}$$

Beispiel 1.19. Bertrandsches Paradox (siehe [Ge02])
In einem Kreis mit Radius $r > 0$ werde "zufällig" eine Sehne gezogen. Wie groß ist die Wahrscheinlichkeit, dass sie länger ist als die Seite des einbeschriebenen gleichseitigen Dreiecks?

Die Antwort auf diese Frage hängt davon ab, was unter "zufällig" verstanden wird.

Variante 1: Die Sehne ist durch ihren Mittelpunkt eindeutig bestimmt, solange dieser nicht gerade der Kreismittelpunkt ist, was man vernachlässigen kann. Als zugrundeliegenden Ergebnisraum wählen wir daher $\Omega_1 = \{x \in \mathbb{R}^d \mid |x| \leq r\}$ und hierauf die Gleichverteilung P_1. Das Ereignis, dass die Sehne länger ist als die Seite des einbeschriebenen gleichseitigen Dreiecks, tritt genau dann ein, wenn der Abstand des Mittelpunktes der Sehne zum Kreismittelpunkt kleiner als der Durchmesser des Inkreises des einbeschriebenen gleichseitigen Dreiecks ist, also $A_1 = \{x \in \Omega_1 \mid |x| \leq \frac{r}{2}\}$ (siehe Graphik). Da $|\Omega_1| = \pi r^2$ und $|A_1| = \pi \left(\frac{r}{2}\right)^2$ ergibt sich in diesem Fall als gesuchte Wahrscheinlichkeit

$$P(A_1) = \frac{|A_1|}{|\Omega_1|} = \frac{\pi \left(\frac{r}{2}\right)^2}{\pi r^2} = \frac{1}{4}.$$



Bertrandsches Paradoxon: Variante 1

Variante 2: Die Sehne wird auch durch ihre beiden Endpunkte eindeutig festgelegt. Aufgrund der Drehsymmetrie können wir einen Endpunkt fest wählen, und damit den zweiten Endpunkt durch seinen Winkel $\in \Omega_2 :=]0, 2\pi[$ mit dem fest gewählten Endpunkt eindeutig festlegen und betrachten hierauf die Gleichverteilung P_2. Das in Frage stehende Ereignis entspricht in dieser Variante dem Ereignis $A_2 =]\frac{2\pi}{3}, \frac{4\pi}{3}[$ und als gesuchte Wahrscheinlichkeit ergibt sich

$$P(A_2) = \frac{|A_2|}{|\Omega_2|} = \frac{\frac{2\pi}{3}}{2\pi} = \frac{1}{3}\,.$$

Variante 3: Die Sehne wird auch durch ihren Abstand zum Mittelpunkt und ihrer Richtung festgelegt. Aufgrund der Drehsymmetrie können wir die Richtung vernachlässigen und somit $\Omega_3 := [0, r[$ als Ergebnisraum wählen. Bei dieser Wahl ist $A_3 = [0, \frac{r}{2}[$ das betrachtete Ereignis und als gesuchte Wahrscheinlichkeit ergibt sich

$$P(A_3) = \frac{|A_3|}{|\Omega_3|} = \frac{\frac{r}{2}}{r} = \frac{1}{2}.$$

Literaturverzeichnis

[Ba91] H. Bauer, Maß- und Integrationstheorie, 2. Auflage, Walter de Gruyter, Berlin, 1991.

[Ba02] H. Bauer, Wahrscheinlichkeitstheorie, 5. Auflage, Walter de Gruyter, Berlin, 2002.

[Be00] E. Behrends, Introduction to Markov Chains - with Special Emphasis on Rapid Mixing, Vieweg, Braunschweig, 2000.

[Bi86] P. Billingsley, Probability and Measure, Wiley, New York, 1986.

[Bu88] P. Bundschuh, Einführung in die Zahlentheorie, Springer, Berlin, 1988.

[Ch78] K. L. Chung, Elementare Wahrscheinlichkeitstheorie und stochastische Prozesse, Springer, Berlin, 1978.

[Fe68] W. Feller, An Introduction to Probability Theory and its Applications, Vol. I, 3rd ed., Wiley, New York, 1968.

[Ge02] H.-O. Georgii, Stochastik. Einführung in die Wahrscheinlichkeitstheorie und Statistik, Walter de Gruyter, Berlin, 2002.

[H02] O. Häggström, Finite Markov Chains and Algorithmic Applications, Cambridge University Press, 2002.

[Ha64] H. Hasse, Vorlesungen über Zahlentheorie, Springer, Berlin, 1964.

[Kl06] A. Klenke, Wahrscheinlichkeitstheorie, Springer, Berlin, 2006.

[Kr02] U. Krengel, Einführung in die Wahrscheinlichkeitstheorie und Statistik, 6. Auflage, Vieweg, Braunschweig, 2002.

[No97] J. Norris, Markov Chains, Cambridge University Press, 1997.

[Pr00] H. Pruscha, Vorlesungen über Mathematische Statistik, Vieweg, Braunschweig, 2000.

[Se81] E. Seneta, Non-Negative Matrices and Markov Chains, Springer, New York, 1981.