Vorlesungsskript

Wahrscheinlichkeitstheorie für das Lehramt

Wintersemester 2021/2022

Martin Tautenhahn

Version vom 11. Oktober 2024

Universität Leipzig

Inhaltsverzeichnis

Abbildungsverzeichnis

Theoremverzeichnis

1 Diskrete Wahrscheinlichkeitsräume

In diesem Kapitel behandeln wir die Wahrscheinlichkeitstheorie auf höchstens abzählbar unendlichen Mengen. Dies hat den Vorteil, dass weniger theoretische Vorarbeit benötigt wird. Insbesondere benötigen wir keine Maßtheorie, da die kanonische σ-Algebra die Potenzmenge ist.

1.1 Grundbegriffe

Wir beginnen diesen Abschnitt mit einem einführenden Beispiel.

Beispiel 1.1 (Würfeln mit zwei fairen Würfeln)**.** Wir würfeln mit zwei fairen Würfeln und stellen uns folgende Frage:

> Wie groß ist die Wahrscheinlichkeit, dass die Augensumme größer oder gleich 10 ist?

Um diese Frage zu beantworten gibt es mindestens zwei Lösungswege:

Weg 1: Wir nehmen an, dass die Würfel unterscheidbar sind (z.B. gedanklich einfärben). Der Ereignisraum/Ergebnisraum, also die Menge aller möglichen Ergebnisse unseres Experimentes ergibt sich zu

$$\begin{aligned}\Omega &= \{\text{Mange aller geordneten Paare mit Einträgen von 1 bis 6}\}\\ &= \{(1,1),(1,2),\ldots,(6,5),(6,6)\}\\ &= \{(i,j) : i,j \in \{1,\ldots 6\}\}\,.\end{aligned}$$

Die Anzahl der Elemente von Ω ist $|\Omega| = 36$. Da es sich um zwei faire Würfel handelt besitzt jedes Paar dieselbe Wahrscheinlichkeit, und zwar $1/36$. Die günstigen Elemente aus Ω, nämlich diejenigen mit Augensumme größer oder gleich 10, sind

$$A = \{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\}.$$

Die Menge A besitzt $|A| = 6$ Elemente. Die Gesuchte Wahrscheinlichkeit ergibt sich somit zu

$$\mathbb{P}(A) = \frac{\text{Anzahl günstiger Fälle}}{\text{Anzahl möglicher Fälle}} = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}.$$

Weg 2: Wir nehmen an, dass die Würfel unterscheidbar sind. Der Ereignisraum bzw. Ergebnisraum ergibt sich in diesem Fall zu

$$\begin{aligned}\Omega &= \{\text{Menge der möglichen Augensummen}\}\\ &= \{2, \dots, 12\}.\end{aligned}$$

Achtung, die Wahrscheinlichkeitsfunktion $p : \Omega \to [0,1]$ ist in diesem Fall nicht die Gleichverteilung, vergleiche Tabelle 1.1.

ω	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(\{\omega\})$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Tabelle 1.1: Augensummen und deren Wahrscheinlichkeiten

Das Ereignis A ist in diesem Fall $A = \{10, 11, 12\}$ und die gesuchte Wahrscheinlichkeit ergibt sich zu

$$\mathbb{P}(A) = \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{6}{36} = \frac{1}{6}.$$

Wir fassen unsere Beobachtungen zusammen:

(a) Man definiert eine Menge Ω von Elementarereignissen.

(b) Dafür gibt es im Allgemeinen mehrere Möglichkeiten.

(c) Oft ist es günstig, wenn alle Elementarereignisse die selbe Wahrscheinlichkeit besitzen.

(d) Das Ereignis nach dem gefragt ist, identifiziert man mit einer Teilmenge $A \subset \Omega$.

(e) Die Wahrscheinlichkeit von A ist dann die Summe der Einzelwahrscheinlichkeiten der Elemente von A.

Dieses Konzept wollen wir jetzt mathematisch sauber definieren.

Definition 1.2 (Diskreter Wahrscheinlichkeitsraum)**.** Ein diskreter Wahrscheinlichkeitsraum (kurz W-Raum) ist ein Paar (Ω, p), bestehend aus einer endlichen oder höchstens abzählbaren Menge Ω und einer Abbildung $p : \Omega \to [0,1]$ mit der Eigenschaft

$$\sum_{\omega\in\Omega} p(\omega) = 1.$$

Für die in obiger Definition eingeführten Symbole führen wir folgende Sprechweisen ein:

- $\Omega \mathrel{\widehat{=}}$ Ereignisraum
- $\omega \in \Omega \mathrel{\widehat{=}}$ Elementarereignis (Atome)
- Teilmenge $A \subseteq \Omega \mathrel{\widehat{=}}$ Ereignis
- $p(\omega) \mathrel{\widehat{=}}$ Einzelwahrscheinlichkeit

Definition 1.3 (Induziertes Wahrscheinlichkeitsmaß)**.** Gegeben sei ein diskreter Wahrscheinlichkeitsraum (Ω, p). Die Abbildung $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ definiert durch

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$$

heißt das von p induzierte Wahrscheinlichkeitsmaß. Man nennt $\mathbb{P}$ auch Verteilung auf Ω.

Da $p(\omega) \geq 0$ für alle $\omega \in \Omega$ spielt die Reihenfolge der Summanden keine Rolle. Falls A (abzählbar) unendlich viele Elemente besitzt ist die obige Summe streng genommen zu interpretieren als

$$\sum_{\omega \in A} p(\omega) = \lim_{n \to \infty} \sum_{i=1}^{n} p(\omega_i),$$

wobei $\omega_1, \omega_2, \ldots$ eine Abzählung der Elemente von A ist. Definitionen 1.2 und 1.3 sind Spezielfälle eines allgemeinen und fundamentalen Konzeptes, welches wir erst später behandeln werden.

Lemma 1.4 (Kolmogorovsche Axiome)**.** *Sei (Ω, p) ein diskreter Wahrscheinlichkeitsraum. Dann hat das (induzierte) Wahrscheinlichkeitsmaß $\mathbb{P}$ beiden Eigenschaften*

(i) $\mathbb{P}(\Omega) = 1$. *(Normierung)*

(ii) Für (abzählbar viele) paarweise disjunkte Ereignisse A_i, $i \in \mathbb{N}$, gilt

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \mathbb{P}(A_i). \qquad (\sigma\text{-Additivität})$$

Ist umgekehrt Ω eine endliche oder abzählbare Menge, $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ eine Abbildung mit den Eigenschaften (i) und (ii), und $p : \Omega \to [0,1]$ definiert über

$$p(\omega) = \mathbb{P}(\{\omega\}),$$

so ist (Ω, p) ein diskreter Wahrscheinlichkeitsraum.

Wir führen folgende Sprechweisen für Ereignisse ein:

- $A, B, C \subseteq \Omega$: A, B, C sind Ereignisse,
- A: A tritt ein,
- $A^{\mathrm{c}} = \Omega \setminus A$: A tritt nicht ein,
- $A \cap B$: A und B treten ein,
- $A \cup B$: A oder B tritt ein (mindestens eins),
- $A \cap B = \emptyset$: A und B schließen einander aus,
- $A \subseteq B$: A impliziert B (A zieht B nach sich),
- $\emptyset$: unmögliches Ereignis,
- Ω: sicheres Ereignis.

Lemma 1.5 (Eigenschaften von $\mathbb{P}$)**.** *Sei (Ω, p) ein diskreter Wahrscheinlichkeitsraum. Dann hat das zugehörige Wahrscheinlichkeitsmaß $\mathbb{P}$ folgende Eigenschaften:*

(a) Für alle $A \subset \Omega$ gilt $\mathbb{P}(A^{\mathrm{c}}) = 1 - \mathbb{P}(A)$.

(b) $\mathbb{P}(\varnothing) = 0$.

(c) Für alle $A, B \subset \Omega$ mit $A \subset B$ gilt $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(d) Für alle $A, B \subset \Omega$ gilt $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.

(e) Für alle $A, B \subset \Omega$ gilt $\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$

(f) Für $A_i \subset \Omega$, $i \in \mathbb{N}$, gilt

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \mathbb{P}(A_i).$$

(g) Falls $A_i \subseteq \Omega$, $i \in \mathbb{N}$, und $A \subset \Omega$ mit $A_i \searrow A$ oder $A_i \nearrow A$, so gilt

$$\lim_{i \to \infty} \mathbb{P}(A_i) = \mathbb{P}(A).$$

Wir schreiben $A_i \nearrow A$, falls „$(A_i)_i$ fallend gegen A konvergiert“, das heißt $A_i \subset A_{i+1}$ für $i \in \mathbb{N}$ und $\cup_{i \in \mathbb{N}} A_i = A = \sup_{i \in \mathbb{N}} A_i$. Analog bedeutet $A_i \searrow A$, dass $A_i \supset A_{i+1}$ für $i \in \mathbb{N}$ und $\cap_{i \in \mathbb{N}} A_i = A = \inf_{i \in \mathbb{N}} A_i$.

Beweis von Lemma 1.5. Aussagen (a)–(f) folgen sofort aus den Eigenschaften von Lemma 1.4 und sind Übungsaufgabe.

Für den Beweis von (g) beginnen wir mit dem Fall $A_i \nearrow A$. Wir zerlegen A mittels

$$A = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \dots$$

und erhalten

$$A = \dot\bigcup_{i=1}^{\infty}(A_i \setminus A_{i-1}) \quad \text{mit} \quad A_0 = \emptyset.$$

Folglich gilt nach Lemma 1.4

$$\mathbb{P}(A) = \mathbb{P}\left(\dot\bigcup_{i=1}^{\infty}(A_i \setminus A_{i-1})\right) = \sum_{i=1}^{\infty}\mathbb{P}(A_i \setminus A_{i-1}) = \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{P}(A_i \setminus A_{i-1})$$

Da $A_{i-1} \subset A_i$ folgt $A_i \cap A_{i-1} = A_{i-1}$. Aus Eigenschaft (d) folgt also $\mathbb{P}(A_i \setminus A_{i-1}) = \mathbb{P}(A_i) - \mathbb{P}(A_{i-1})$. Folglich gilt (Teleskopsumme)

$$\mathbb{P}(A) = \lim_{n\to\infty}\sum_{i=1}^{n}[\mathbb{P}(A_i) - \mathbb{P}(A_{i-1})] = \lim_{n\to\infty}[\mathbb{P}(A_n) - \mathbb{P}(A_0)].$$

Da $\mathbb{P}(A_0) = \mathbb{P}(\emptyset) = 0$ folgt die Behauptung.

Wir betrachten nun den Fall $A_i \searrow A$. In diesem folgt unmittelbar $A_i^{\mathrm{c}} \subset A_{i+1}^{\mathrm{c}}$ sowie $\cap_{i=1}^{\infty} A_i^{\mathrm{c}} = Ac$. Also gilt $A_i^{\mathrm{c}} \nearrow A^{\mathrm{c}}$. Aus dem ersten Teil des Beweises folgt also

$$\mathbb{P}(A^{\mathrm{c}}) = \lim_{n\to\infty}\mathbb{P}(A_n).$$

Unter Benutzung dieser Identität erhalten wir

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^{\mathrm{c}}) = 1 - \lim_{n\to\infty}\mathbb{P}(A_n^{\mathrm{c}}) = \lim_{n\to\infty}(1 - \mathbb{P}(A_n^{\mathrm{c}})) = \lim_{n\to\infty}\mathbb{P}(A_n). \qquad \square$$

1.2 Laplacesche Wahrscheinlichkeitsräume

Bei einem Laplaceschen Wahrscheinlichkeitsraum fordern wir, dass die Menge der Elementarereignisse Ω eine endliche Menge ist und jedem Elementarereignis dieselbe Wahrscheinlichkeit zugeordnet wird. Die Funktion $p : \Omega \to [0,1]$ ist also gegeben durch

$$p(\omega) = \frac{1}{|\Omega|}$$

und erfüllt offensichtlich die (für einen diskreten Wahrscheinlichkeitsraum geforderte) Bedingung $\sum_{\omega\in\Omega} p(\omega) = 1$. Sei nun $A \subset \Omega$ ein beliebiges Ereignis. Dann folgt nach Definition 1.3

$$\mathbb{P}(A) = \sum_{\omega\in A} p(\omega) = \sum_{\omega\in A}\frac{1}{|\Omega|} = \frac{|A|}{|\Omega|} = \frac{\text{„Günstige“}}{\text{„Mögliche“}}.$$

Beispiel 1.6 (Würfeln mit einem Würfel)**.** Wir wählen $\Omega = \{1, 2, \ldots, 6\}$ und $p(\omega) = 1/|\Omega|$ für alle $\omega \in \Omega$. Dann ist (Ω, p) ein Laplacescher Wahrscheinlichkeitsraum mit $|\Omega| = 6$. Wir betrachten die Ereignisse $A = \{\text{Das Ergebnis ist eine } 6\}$ und $B = \{\text{Das Ergebnis ist gerade}\}$. Dann gilt

- $A = \{6\}$, $\mathbb{P}(A) = \frac{|A|}{|\Omega|} = 1/6$,
- $B = \{2, 4, 6\}$, $\mathbb{P}(B) = 3/6 = 1/2$.

Beispiel 1.7 (Würfeln mit zwei Würfeln)**.** Wir wählen

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

und $p(\omega) = 1/|\Omega|$ für alle $\omega \in \Omega$. Dann ist (Ω, p) ein Laplacescher Wahrscheinlichkeitsraum mit $|\Omega| = 36$. Wir betrachten die Ereignisse $A = \{\text{Augensumme } 10\}$, $B = \{\text{Das ergebnis ist ein Pasch}\}$ und $C = \{\text{Beide Zahlen sind gerade}\}$. Dann gilt

- $A = \{(4, 6), (6, 4), (5, 5)\}$, $\mathbb{P}(A) = 3/36 = 1/12$,
- $B = \{(1, 1), \ldots, (6, 6)\}$, $\mathbb{P}(B) = 6/36 = 1/6$,
- $C = \{(2, 2), (2, 4), (2, 6), (4, 4), (4, 2), (4, 6), (6, 2), (6, 4), (6, 6)\}$, $\mathbb{P}(C) = 1/4$.

Für die Ereignisse B und C gibt es elegante Lösungen mittels bedingter Wahrscheinlichkeit oder Unabhängigkeit. Dies werden wir später kennenlernen.

Beispiel 1.8 (n-maliger Münzwurf)**.** Wir wählen

$$\Omega = \{(a_1, \ldots, a_n)\colon a_i \in \{0, 1\}\} = \{0, 1\}^n$$

und $p(\omega) = 1/|\Omega|$ für alle $\omega \in \Omega$. Dann ist (Ω, p) ein Laplacescher Wahrscheinlichkeitsraum mit $|\Omega| = 2^n$. Wir betrachten die Ereignisse $A = \{\text{Alle Ergebnisse gleich}\}$, $B = \{\text{Kopf und Zahl abwechselnd}\}$ und $C = \{\text{Kopf und Zahl gleich oft}\}$. Dann gilt:

- $A = \{(0, 0, \ldots, 0), (1, 1, \ldots, 1)\}$, $\mathbb{P}(A) = 2/2^n = 1/2^{n-1}$. F ür $n = 1$ folgt $\mathbb{P}(A) = 1$. Im Grenzwert $n \to \infty$ erhalten wir $\lim_{n\to\infty} \mathbb{P}(A) = 0$.
- $B = \{(1, 0, 1, \ldots), (0, 1, 0, \ldots)\}$, $\mathbb{P}(A) = 2/2^n = 1/2^{n-1}$.
- Falls n ungerade ist folgt $C = \emptyset$ und somit $\mathbb{P}(C) = 0$. Wir nehmen nun an, dass n gerade ist und setzen $k = n/2$. Dann ist

$$\begin{aligned} C &= \{\text{alle Vektoren der Länge } 2k\text{, die je } k\text{-mal den Eintrag } 1 \\ &\qquad \text{und } k\text{-mal den Eintrag } 0 \text{ besitzen}\} \\ &= \Big\{a \in \Omega\colon \sum_{i=1}^{n} a_i = k = n/2\Big\}. \end{aligned}$$

Die anzahl der Elemente von C entspricht also die Anzahl der Möglichkeiten k Einsen auf $2k$ Plätze zu verteilen und ergibt sich zu

$$|C| = \binom{2k}{k} = \frac{(2k)!}{k!k!},$$

vgl. Abschnitt 1.3. Die gesuchte Wahrscheinlichkeit ist also

$$\mathbb{P}(A) = \frac{\binom{2k}{k}}{2^{2k}}.$$

Für $k = 1$ (also $n = 2$) ist $\mathbb{P}(C) = 1/2$. Für $k \to \infty$ ist $\lim_{k\to\infty} \mathbb{P}(C) = 0$. Letzteres kann man mit Hilfe der Stirlingschen Formel zeigen.

Beispiel 1.9 (Zwei Kartenstapel)**.** Gegeben sind $2n$ Spielkarten, darunter sind genau 2 Joker ($n \geq 1$). Wir bilden 2 gleichgroße Stapel und fragen uns nach der Wahrscheinlichkeit vom Ereignis

$$A = \{\text{Beide Joker befinden sich im selben Stapel}\}.$$

Wir besprechen im Folgenden zwei Lösungen.

L1: Als Ergebnismenge wählen wir die Menge der möglichen Aufteilungen von $2n$ Karten auf 2 Stapel gleicher größe, also

$$\Omega = \{\{a_1, a_2, \dots, a_n\} \colon a_i \in \{1, 2, \dots, 2n\},\ a_i \neq a_j,\ \text{für } i \neq j\}.$$

Dies entspricht der Auswahl von n Karten (für den linken Stapel) aus $2n$ Karten. Daher ist

$$|\Omega| = \binom{2n}{n},$$

vgl. Abschnitt 1.3. Damit beide Joker im gleichen Stapel sind gibt es zwei Möglichkeiten: beide links oder beide rechts. Sind beide Joker im linken Stapel gesetzt, so gibt es

$$\binom{2n-2}{n-2}$$

Möglichkeiten den linken Stapel auf n Karten aufzufüllen. Sind beide Joker im rechten Stapel gesetzt, so gibt es

$$\binom{2n-2}{n}$$

Möglichkeiten den linken Stapel auf n Karten aufzufüllen. Also ist

$$|A| = \binom{2n-2}{n-2} + \binom{2n-2}{n} = 2\binom{2n-2}{n}, \quad \text{da} \quad \binom{n}{k} = \binom{n}{n-k}.$$

Für die gesuchte Wahrscheinlichkeit folgt somit

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{2(2n-2)!n!n!}{n!(n-2)!(2n)!} = \frac{2n(n-1)}{2n(2n-1)} = \frac{n-1}{2n-1}.$$

Wir interpretieren noch drei Spezialfälle:

$n = 1$: $\mathbb{P}(A) = 0$,

$n = 2$: $\mathbb{P}(A) = 1/3$ (Überlege selbst, es gibt genau 2 günstige aus 6 möglichen Elementarereignissen),

$n \to \infty$: $\mathbb{P}(A) \to 1/2$.

L2: Wir wählen für Ω die Menge der möglichen Plätze für die Joker, also

$$\Omega = \{(i,j)\colon i,j \in \{1,2,\ldots,2n\},\ i \neq j\}.$$

Dann ist $|\Omega| = 2n(2n-1)$. Das gesuchte Ereignis besitzt die Darstellung

$$A = \{(i,j \in \Omega\colon i,j \leq n)\} \cup \{(i,j \in \Omega\colon i,j \geq n+1)\}.$$

Da obige Vereinigung disjunkt ist folgt $|A| = n(n-1) + n(n-1) = 2n(n-1)$. Die gesuchte Wahrscheinlichkeit ist in Übereinstimmung zur ersten Lösung

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{2n(n-1)}{2n(2n-1)} = \frac{n-1}{2n-1}.$$

1.3 Urnenmodelle/Kombinatorik

Um die Anzahl der Elemente der Ereignisse in den vorausgegangenen Beispielen zu berechnen ist es oft hilfreich sich sogenannter Urnenmodelle zu bedienen. Vier Grundfragen der Kombinatorik werden meistens mit Urnenmodellen beschrieben. Eine Urne enthält n durchnummerierte Kugeln. Es werden nacheinander k Kugeln aus der Urne gezogen. Es ergeben sich 4 separate Fälle:

- Die gezogene Kugel wird jedesmal in die Urne zurückgelegt oder auch nicht.
- Die Reihenfolge/Anordnung spielt eine Rolle oder auch nicht.

In diesem Abschnitt diskutieren wir die Anzahl der verschiedenen Ziehungsergebnisse als Funktion von n und k. Übliche Namen der vier Grundaufgaben sind:

(I) Ziehen mit Zurücklegen und mit Anordnung.

(II) Ziehen ohne Zurücklegen und mit Anordnung.

(III) Ziehen ohne Zurücklegen und ohne Anordnung.

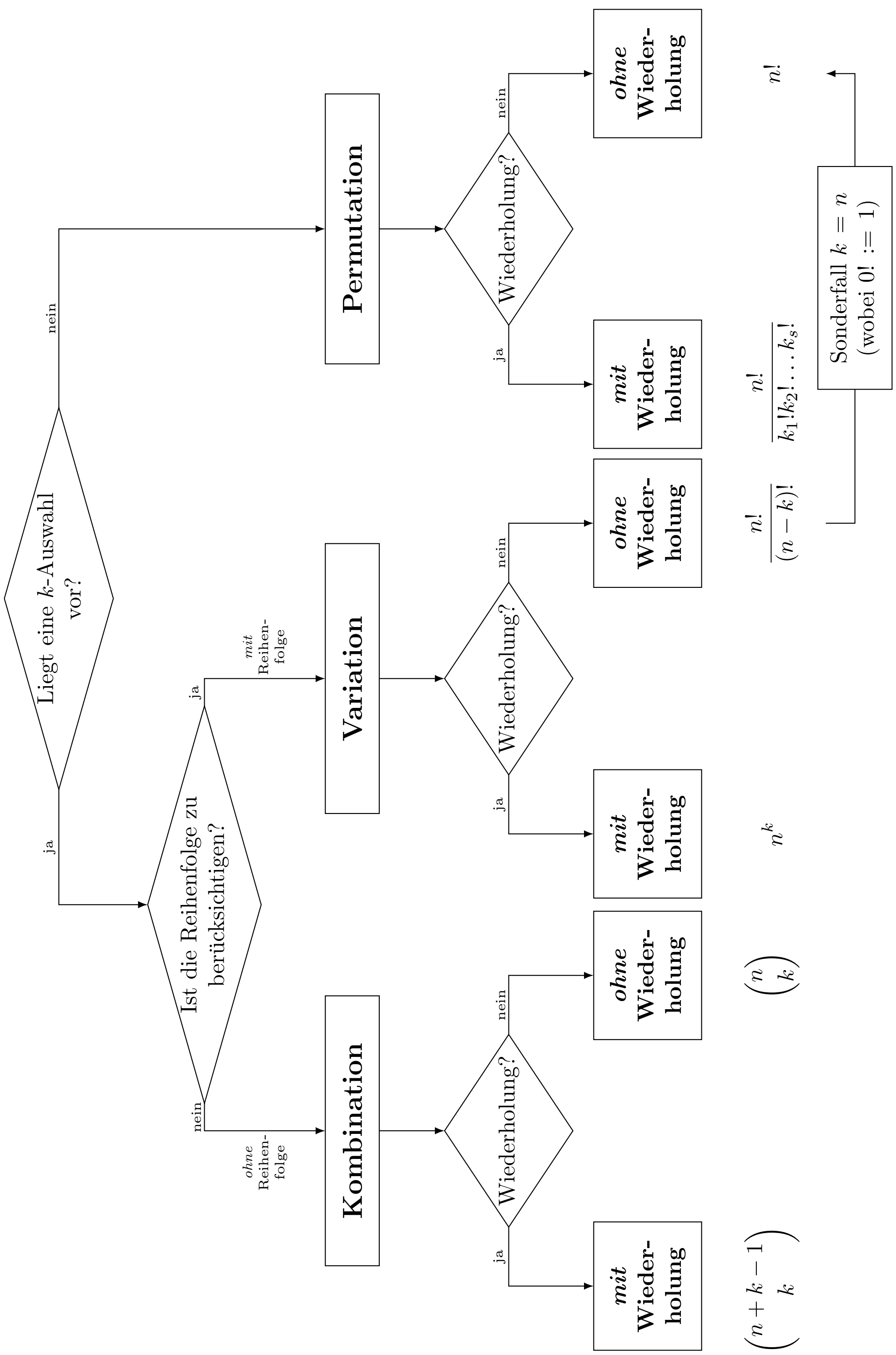

Abbildung 1.1: Flussdiagramm zur Kombinatorik

(IV) Ziehen mit Zurücklegen und ohne Anordnung.

Abbildung 1.1 gibt einen Überblick über diese vier Fälle.

Formalisierung: Die Urne idendifiziert man mit der Menge $A = \{a_1, a_2, \ldots, a_n\}$. Man kann sich den Index j als Zahl vorstellen, die eine Kugel beschriftet, dafür identifizieren wird manchmal A mit $\{1, 2, \ldots, n\}$. Dann entspricht jeder Urnenzug dem ziehen einer Zahl zwischen 1 und n.

Alternative Denkweise (Platz-Anordnungs-Modell): Die Menge A entspricht dabei n Plätzen (z.B Schälchen). In jedes kann ich ein oder mehrere Plastik-Chips legen. Durch jeden der Züge 1 bis k wird bestimmt, wohin der aktuelle Chip gelegt wird. Zum Beispiel

- mit Zurücklegen: Ein Platz kann mit mehreren Chips belegt werden.
- ohne Zurücklegen: höchstens ein Chip in einem Schälchen.
- mit Anordnung : die Chips sind nummeriert bzw. notiert man auf jedem Chip die Zahl, wann er gezogen wurde.

(I) Mit zurücklegen, mit Anordnung (Variation)

Seien $k, n \in \mathbb{N}$. Wie viele k-Tupel $(b_1, b_2, \ldots, b_k)$ mit Komponenten $b_j \in A = a_1, \ldots, a_n$ für jedes $j \in 1, \ldots, k$ gibt es? Die Antwort ist

$$n^k$$

Beweis. Es ist

$$\{(b_1, \ldots, b_k)\colon b_j \in A \text{ für } j = \{1, \ldots, k\}\} = A \times A \times \ldots \times A = A^k.$$

Da $|A^k| = |A|^k$ (Übungsaufgabe) folgt die Behauptung. □

Beispiel 1.10 (Variation mit Wiederholung)**.** (a) Wie viele verschiedene Informationen kann man mit k Bits darstellen?

(b) Wie viele Telefonnummern mit k Stellen gibt es? Jede Stelle der Telefonnummer zieht eine der Ziffern $0, \ldots, 9$.

(c) Ein Alphabet enthalte n Zeichen. Wie viele (nicht notwendigerweise sinnvoll aussprechbare) Wörter der Länge k gibt es?

(d) Sei $|A| = k$ und $|B| = n$. Wie viele Abbildungen $f : A \to B$ gibt es?

(e) Auf wieviele (auch sehr ungerechte) Weisen kann man k Geschenke auf n Personen verteilen? (Personen können mehrere Geschenke bekommen.)

(II) Ohne Zurücklegen, mit Anordnung (Variation)

Wie viele k-Tubel $(b_1, b_2, \dots, b_k)$ gibt es, deren Komponenten $b_j \in A = \{1, \dots, k)$ alle verschieden sind? Also wie viele Elemente besitzt

$$\{(b_1, b_2, \dots, b_k) \colon b_i \in A,\ b_i \neq b_j \text{ für } i \neq j\}?$$

Hier muss neben $n, k \in \mathbb{N}$ offensichtlich $k \leq n$ gelten. Die Antwort ist

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

Beweis. Beim 1. Zug habe ich n Kugeln zur Auswahl. Beim 2. Zug $n-1$ Kugeln. Und so weiter, bis ich beim k-ten Zug $n-(k-1)$ Kugeln zur Auswahl habe. □

Beispiel 1.11 (Variation ohne Wiederholung)**.** (a) Wie viele Worte der Länge k eines n-elementigen Alphabets kann man bilden, falls kein Buchstabe doppelt vorkommen soll?

(b) Eine Mensaküche kann n Gerichte kochen (nur eins am Tag). Wie viele Wochenspeisepläne ohne Wiederholung gibt es?

(c) Sechs Marathonläufer kommen nacheinander ins Ziel. Wie viele mögliche Reihenfolgen gibt es?

(d) Beim Pferdetoto muss man die Reihenfolge der ersten 3 Pferde tippen. Wie viele Mo?glichkeiten gibt es bei 15 Pferden?

(e) Sei $|A| = k$, $|B| = n$ und $n \geq k$. Wie viele injektive Abbildungen $f : A \to B$ gibt es?

(III) Ohne Zurücklegen, ohne Anordnung (Kombination)

Wie viele k-Tubel $(b_1, b_2, \dots, b_k)$ gibt es, deren Komponenten $b_j \in A = \{1, \dots, k)$ alle verschieden und der größe nach geordnet sind? Also wie viele Elemente besitzt

$$\{(z_1, z_2, \dots, z_k) \colon 1 \leq z_1 < z_2 < \ldots < z_{k-1} < z_k \leq n\}$$

Satz 1.12 (Kombination ohne Wiederholung)**.** *Es gibt genau*

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

k-elementige Teilmenge von einer n-elementigen Menge.

Beweis. Im Fall einer Variation ohne Wiederholung (d.h. Komponenten b_j verschieden aber nicht der größe nach geordnet), gibt es nach (II) genau

$$\frac{n!}{(n-k)!}$$

Möglichkeiten. Nun aber können die k ausgewählten Elemente ihrerseits auf $k!$ verschiedene Weisen angeordnet werden. Wenn diese verschiedenen Anordnungen allesamt keine Rolle spielen, müssen wir das erhaltene Ergebnis noch einmal durch $k!$ teilen und erhalten damit nur noch

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

Möglichkeiten. □

Beispiel 1.13 (Kombination ohne Wiederholung)**.** (a) In der Bundesliga gebe es n Mannschaften. Wie viele Paarungen gibt es?

(b) Ein Verein mit n Mitgliedern wählt ein k-köpfigen Vorstand. Wie viele verschiedene Vorstände sind möglich?

(c) Beim Elfmeterschießen werden 5 von 11 Spieler vom Trainer ausgewählt. Wie viele Möglichkeiten hat der Trainer, falls es nicht auf die Reihenfolge ankommt?

(d) Wieviele Möglichkeiten gibt es, 6 Zahlen aus 49 auszuwa?hlen?

(e) Eine Bibliothek hat n Bücher. Sie dürfen auf einmal höchstens k Bücher ausleihen. Wie viele Möglichkeiten gibt es?

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{k} = \sum_{j=0}^{k} \binom{n}{j}$$

(f) Wie viele Möglichkeiten gibt es, 12 Plätze mit 4 blauen, 3 roten und 5 gru?nen Chips zu belegen?

$$\binom{12}{4} \cdot \binom{8}{3} \cdot \binom{5}{5}.$$

Allgemeine Frage zu (f). In einer Urne sind k_1 Kugeln vom Typ 1, k_2 Kugeln vom Typ 2, ..., k_m Kugeln vom Typ m.

Satz 1.14. *Es gibt*

$$\frac{n!}{k_1!k_2! \cdot \ldots \cdot k_m!}$$

Möglichkeiten, eine Menge mit $n = k_1 + k_2 + \ldots + k_m$ *Elementen in* m *Teilmengen mit* k_1 *bzw.* k_2 *bzw.* $\ldots$ *bzw.* k_m *Elementen zu unterteilen.*

Beispiel 1.15 (Permutation mit Wiederholung)**.** Eine Klasse mit 26 Schuülern will einen 4-stimmigen Kanon singen. Die Gruppen sollen moöglichst gleich stark sein, als 7, 7, 6, 6. Wie viele Möglichkeiten gibt es, Gruppen zu bilden?

(IV) Mit Zurücklegen, ohne Anordnung (Kombination)

Wie viele k-Tubel $(b_1, b_2, \dots, b_k)$ gibt es, deren Komponenten $b_j \in A = \{1, \dots, k\}$ größe nach geordnet sind, benachbarte Elemente aber gleich sein dürfen? Also wie viele Elemente besitzt

$$\{(z_1, z_2, \dots, z_k) \colon 1 \leq z_1 \leq z_2 \leq \dots \leq z_{k-1} \leq z_k \leq n\}?$$

Die Antwort ist

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

Auf den etwas länglichen Beweis verzichten wir an dieser Stelle. Stattdessen verweisen wir auf [Geo09, Abschnitt 2.2].

Beispiel 1.16 (Kombination mit Wiederholung)**.** (a) Wie viele Kombinationen gibt es $k = 5$ Gummibärchen aus einer Tüte Gummibärchen mit $n = 5$ verschiedenen Farben auszuwählen? Antwort ist

$$\binom{5+5-1}{5} = \binom{9}{5} = 126.$$

(b) Wie viele verschiedene Würfe sind mit drei Würfeln möglich? Die Antwort ist

$$\binom{6+3-1}{3} = \binom{8}{3} = 56.$$

2 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Oft sind bei der Frage nach der Wahrscheinlichkeit von Ereignissen schon Vorinformationen vorhanden. So zum Beispiel bei Lebensversicherungen (Alter, Vorerkrankungen, ...) oder beim Skat (eigene Karten). In diesem Kapitel fragen wir uns nach der Wahrscheinlichkeit eines Ereignisses A unter der Annahme, dass ein Ereignis B bereits eingetreten ist. Eine solche Wahrscheinlichkeit heißt *bedingte Wahrscheinlichkeit*. Eng verwoben mit der bedingten Wahrscheinlichkeit ist das Konzept der *Unabhängigkeit*, welches wir im zweiten Teil dieses Kapitels kennenlernen.

Während des gesamten Kapitels sei (Ω, p) ein diskreter Wahrscheinlichkeitsraum.

2.1 Bedingte Wahrscheinlichkeiten

Wir beginnen mit einem motivierenden Beispiel.

Beispiel 2.1 (Anzahl der Raucher in der Gesellschaft)**.** Bei einer Befragung einer Grundgesamtheit Ω (z.B. alle Einwohner Berlins) betrachten wir folgende Ereignisse:

$$A = \{\text{zufällig ausgewählte Person raucht}\}$$
$$B = \{\text{ausgewählte Person ist zwischen 21-30 Jahre alt}\}$$

Gesucht ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Person raucht, unter der Annahme, dass diese Person zwischen 21 und 30 Jahre alt ist. Für diese Wahrscheinlichkeit werden wir $\mathbb{P}(A \mid B)$ schreiben.

Die gesuchte Wahrscheinlichkeit enspricht dem Anteil der Raucher under den 21 bis 30-jährigen. Damit ist (unter Annahme der Gleichverteilung)

$$\mathbb{P}(A|B) = \frac{|\{\text{Raucherin ; weiblich ; 21-30 Jahre}\}|}{|\{\text{weiblich ; 21-30 Jahre}\}} = \frac{\frac{|(A\cap B)|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

In der Tat wird die bedingte Wahrscheinlichkeit genau so definiert.

Definition 2.2 (Bedingte Wahrscheinlichkeit)**.** Seien $A, B \subset \Omega$ zwei Ereignisse mit $\mathbb{P}(B) > 0$. Dann heißt

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

bedingte Wahrscheinlichkeit von A unter B (oder von A gegeben B).

Satz 2.3 (Eigenschaften der bedingten Wahrscheinlichkeit)**.** *(a) Sei $B \subset \Omega$ ein Ereignis mit $\mathbb{P}(B) > 0$. Dann erfüllt $\mathbb{P}(\cdot \mid B) : \mathcal{P}(\Omega) \to [0,1]$ die Kolmogorovschen Axiome (vgl. Lemma 1.4), d.h.*

- *$\mathbb{P}(\Omega \mid B) = 1$ und*
- *für alle Folgen paarweise disjunkter Ereignisse $(A_i)_{i\in\mathbb{N}}$ gilt*

$$\mathbb{P}\left(\dot{\bigcup_{i=1}^{\infty}} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

(b) ***Satz der Totalen Wahrscheinlichkeit:*** *Sei $A \subset \Omega$ ein Ereignis und $(B_i)_{i\in\mathbb{N}}$ eine Folge paarweise disjunkter Ereignisse mit $\Omega = \cup_{i=1}^{\infty} B_i$ sowie $\mathbb{P}(B_i) > 0$ für alle $i \in \mathbb{N}$. Dann gilt*

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i).$$

(c) ***Satz von Bayes:*** *Seien $A, B \subset \Omega$ Ereignisse mit $\mathbb{P}(A) > 0$, und $(B_i)_{i\in\mathbb{N}}$ eine Folge paarweise disjunkter Ereignisse mit $\Omega = \cup_{i=1}^{\infty} B_i$ sowie $\mathbb{P}(B_i) > 0$ für alle $i \in \mathbb{N}$. Dann gilt*

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\sum_{i=1}^{\infty} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)}.$$

Beweis. Einfach □

Beispiel 2.4 (Test auf eine seltene Krankheit)**.** Fakten:

- Krankheit liegt bei 0,5% der Bevölkerung vor
- Test ist positiv bei 99% der Kranken
- Test ist positiv bei 2% der Gesunden

Fragestellung: Mit welcher WK ist eine positiv getestete Person tatsächlich krank?
Modellierung:

$$\Omega = B_1 \dot{\cup} B_2 \text{ mit } \begin{cases} B_1 \widehat{=} \{\text{Person ist krank}\} \\ B_2 \widehat{=} \{\text{Person ist gesund}\} \end{cases}$$

Weiterhin sei A das Ereignis, dass der Test positiv ist. Gesucht ist $\mathbb{P}(B_1 \mid A)$. Es gilt

$$\mathbb{P}(B_1) = 0.005, \quad \mathbb{P}(A|B_1) = 0.99, \quad \mathbb{P}(A|B_2) = 0.02.$$

Mit dem Satz von Bayes folgt

$$\begin{aligned}\mathbb{P}(B_1|A) &= \frac{\mathbb{P}(B_1)\mathbb{P}(A|B_1)}{\mathbb{P}(B_1)\mathbb{P}(A|B_1) + \mathbb{P}(B_2)\mathbb{P}(A|B_2)} \\ &= \frac{0.005 \cdot 0.99}{0.005 \cdot 0.99 + 0.995 \cdot 0.02} \approx 19.9\%\end{aligned}$$

Die Wahrscheinlichkeit mit der die Krankheit mit dem Test nicht entdeckt wird ist

$$\mathbb{P}(A^c|B_1) = 1 - \mathbb{P}(A|B_1) = 0.01.$$

Meistens wird bei einem positiven medizinischen Testergebnis ein zweiter (spezifischerer) Test durchgeführt.

Lemma 2.5 (Multiplikationsformel)**.** *Sei $n \in \mathbb{N}$ und $A_1, \dots, A_n \subset \Omega$ Ereignisse mit $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$. Dann gilt*

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 \setminus A_1)\mathbb{P}(A_3 \setminus A_1 \cap A_2) \dots \mathbb{P}(A_n \setminus A_1 \cap \dots \cap A_{n-1}).$$

Beweis. Einfach □

Beispiel 2.6 (Skat: Jeder besitzt genau ein Ass)**.** Ein Skatblatt besteht aus 32 Karten, darunter viert Ässer. Jeder Spieler bekommt zehn Karten, zwei Karten werden beiseite gelegt und bilden den sogenannten Skat. Für $i \in \{1, 2, 3\}$ betrachten wir das Ereignis

$$A_i = \{i\text{-ter Spieler hat genau ein Ass}\}.$$

Dann ist

$$A = \{\text{jeder Spieler hat genau ein Ass}\} = A_1 \cap A_2 \cap A_3$$

und nach der Multiplikationsformel gilt

$$\begin{aligned}\mathbb{P}(A_1 \cap A_2 \cap A_3) &= \mathbb{P}(A_1)\mathbb{P}(A_2 \setminus A_1)\mathbb{P}(A_3 \setminus A_1 \cap A_2) \\ &= \frac{\binom{4}{1}\binom{28}{9}}{\binom{32}{10}} \cdot \frac{\binom{3}{1}\binom{19}{9}}{\binom{22}{10}} \cdot \frac{\binom{2}{1}\binom{10}{9}}{\binom{12}{10}} \approx 5.56\%.\end{aligned}$$

2.2 Unabhängigkeit von Ereignissen

Wie kann man mathematisch sauber definieren, dass zwei Ereignisse A und B unabhängig sind? Unabhängigkeit bedeutet anschaulich, dass das Eintreten von B keinerlei Einfluss auf das Eintreten von A hat. Ergo wäre $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ eine guter Ansatz. Dieser Ansatz hat allerdings folgende Nachteile:

- man braucht $\mathbb{P}(B) > 0$,
- nicht symmetrisch in A und B,
- kann nicht ohne Weiteres auf mehr als zwei Ereignisse verallgemeinert werden.

Definition 2.7 (Bedingte Wahrscheinlichkeit)**.** (a) Zwei Ereignisse $A \subset \Omega$ und $B \subset \Omega$ heißen unabhängig, falls

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

(b) Sei I eine beliebige Indexmenge (mit $|I| \geq 2$). Eine Familie $(A_i)_{i\in I}$ von Ereignissen heißt unabhängig, falls für jede endliche Teilmenge $J \subset I$ die folgende Formel gilt:

$$\mathbb{P}\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} \mathbb{P}(A_j).$$

Beispiel 2.8 (n-facher Würfelwurf)**.** Sei $n \in \mathbb{N}$, $\Omega = \{1, 2, \ldots, 6\}^n$ und $p\colon \Omega \to [0, 1]$ die Gleichverteilung, d.h. $p(\omega) = 1/|\Omega|$ für $\omega \in \Omega$. Für $i \in I = \{1, 2, \ldots, n\}$ betrachten wir das Ereignis

$$A_i = \{i\text{-ter Wurf zeigt } 6\} = \{(a_1, a_2, \ldots, a_n) \in \Omega\colon a_i = 6\}.$$

Dann ist die Familie $(A_i)_{i\in I}$ unabhängig. In der Tat, sei $J = \{n_1, n_2, \ldots n_k\} \subset I$ eine beliebige k-elementige Teilmenge von I. Dann ist

$$A_{n_1} \cap \ldots \cap A_{n_k} = \{(a_1, a_2, \ldots, a_n) \in \Omega\colon a_{n_1} = \ldots = a_{n_k} = 6\}$$

und

$$|A_{n_1} \cap \ldots \cap A_{n_k}| = 6^{n-k}.$$

Folglich gilt

$$\mathbb{P}(A_{n_1} \cap \ldots \cap A_{n_k}) = \frac{|A_{n_1} \cap \ldots \cap A_{n_k}|}{|\Omega|} = \frac{6^{n-k}}{6^n} = 6^{-k}.$$

Da ferner $|A_i| = 6^{n-1}$ folgt

$$\mathbb{P}(A_i) = \frac{6^{n-1}}{6^n} = 6^{-1}$$

für jedes $i \in I$. Da $|J| = k$ folgt schließlich

$$\mathbb{P}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} \mathbb{P}(A_i).$$

Beispiel 2.9 (n-facher Münzwurf)**.** Sei $n \in \mathbb{N}$, $\Omega = \{0,1\}^n$ und $p\colon \Omega \to [0,1]$ die Gleichverteilung, d.h. $p(\omega) = 1/|\Omega|$ für $\omega \in \Omega$. Für $i \in I = \{1,2,\dots,n\}$ betrachten wir das Ereignis

$$A_i = \{i\text{-ter Wurf zeigt 1 (=Zahl)}\} = \{(a_1,\dots,a_n) \in \Omega\colon a_i = 1\}.$$

Dann ist die Familie $(A_i)_{i\in I}$ unabhängig. Der Beweis dieser Behauptung erfolgt analog zu Beispiel 2.8.

Beispiel 2.10 (Unabhängigkeit bei der Urne)**.** Wir betrachten eine Urne mit $s \geq 1$ schwarzen und $w \geq 1$ weißen Kugeln- Wir ziehen zwei Kugeln nacheinander aus der Urne und betrachten die Ereignisse

$$A_1 = \{\text{erste Kugel ist weiß}\} \quad \text{und} \quad A_2 = \{\text{zweite Kugel ist weiß}\}.$$

Beim ziehen mit Zurücklegen sind A_1 und A_2 unabhängig. Beim ziehen ohne Zurücklegen sind A_1 und A_2 nicht unabhängig. Einen Beweis dieser Aussagen überlassen wir dem Leser als Übungsaufgabe.

Bemerkung 2.11 (Bemerkungen zur Unabhängigkeit)*.* (i) Jede Teilfamilie einer Familie unabhängiger Ereignisse ist unabhängig.

(ii) Es gibt den schwächeren Begriff der *paarweisen Unabhängigkeit*. Eine Familie $(A_i)_{i\in I}$ von ereignissen heißt paarweise unabhängig, falls für alle $i, j \in I$ mit $i \neq j$ gilt

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j),$$

d.h. falls A_i und A_j unabhängig sind.

Offensichtlich gilt

$$(A_i)_{i\in I} \text{ unabhängig} \quad \Rightarrow \quad (A_i)_{i\in I} \text{ paarweise unabhängig.}$$

Die Umkehrung diese Aussage gilt nicht. Dazu betrachten wir folgendes Beispiel. Sei $\Omega = \{1,2,3,4\}$ und $p(\omega) = 1/4$ für $\omega \in \Omega$. Wir betrachten die Ereignisse

$$A = \{1,2\}, \quad B = \{2,3\}, \quad \text{und} \quad C = \{1,3\}.$$

Dann sind diese drei Ereignisse paarweise unabhängig, denn

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(\{2\}) = \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2} = \mathbb{P}(A)\mathbb{P}(B),\\
\mathbb{P}(A \cap C) &= \mathbb{P}(\{1\}) = \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2} = \mathbb{P}(A)\mathbb{P}(C),\\
\mathbb{P}(B \cap C) &= \mathbb{P}(\{3\}) = \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2} = \mathbb{P}(B)\mathbb{P}(C).
\end{aligned}$$

Allerdings sind diese drei Ereignisse nicht unabhängig, denn

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\emptyset) = 0 \neq \frac{1}{8} = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

(iii) A unabhängig von sich selbst genau dann wenn $\mathbb{P}(A) \in \{0, 1\}$, denn

$$\mathbb{P}(A \cap A) = \mathbb{P}(A)\mathbb{P}(A) \quad \Leftrightarrow \quad \mathbb{P}(A) = (\mathbb{P}(A))^2 .$$

(iv) Unabhängigkeit von Ereignissen kann sogar trotz Kausalität vorliegen. Dazu betrachten wir folgendes Beispiel. Sei

$$\Omega = \{1, 2, \ldots, 6\} \times \{1, 2, \ldots, 6\} \quad \text{und} \quad p(\omega) = \frac{1}{36} \text{ für alle } \omega \in \Omega.$$

Wir beschreiben also das Würfeln mit zwei Würfeln. Wir betrachten die Ereignisse

$$A = \{\text{Augensumme ist } 7\} = \{(i, j) \in \Omega \colon i + j = 7\}.$$

und

$$B = \{\text{Der erste Würfel zeigt } 6\} = \{(i, j) \in \Omega \colon i = 6\}.$$

Dann gilt

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{1}{6}, \quad \mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{1}{6} \quad \text{und} \quad \mathbb{P}(A \cap B) = \frac{|A \cap B|}{|\Omega|} = \frac{1}{6}.$$

Demnach sind die Ereignisse A und B unabhängig. Aber A ist kausal abhängig von B, da der Wurf des ersten Würfels die Summe der Augenzahlen mitbestimmt.

(v) Man verwechsle nicht

- Disjunktheit ($A \cap B = \emptyset$) mit
- Unabhängigkeit ($\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$).

Sind A und B disjunkt, so schließen sie einander aus. Das ist in der Regel eine recht starke Art der Abhängigkeit. Für disjunkte und unanbhängige Ereignisse A und B muss nämlich gelten:

$$\mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0 \quad \Rightarrow \quad \mathbb{P}(A) = 0 \text{ oder } \mathbb{P}(B) = 0.$$

Das heißt, mindestens eines der Ereignisse ist unmöglich.

(vi) Die Gültigkeit der Produktformel für gewisse $J \subset I$ impliziert nicht die Gültigkeit der Produktformel für kleinere Teilmengen J. Dazu betrachten wir ein Beispiel, nämlich einen dreimaligen Münzwurf. Sei

$$\Omega = \{0, 1\}^3, \quad \text{und} \quad p(\omega) = \frac{1}{8} \text{ für alle } \omega \in \Omega.$$

Wir betrachten die Ereignisse

$$\begin{aligned} A &= \{\text{mindestens zwei mal Kopf}\}, \\ B &= \{\text{beim ersten Wurf Kopf}\} \quad \text{und} \\ C &= \{\text{zweiter und dritter Wurf gleich}\}. \end{aligned}$$

Dann folgt $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2$ und $\mathbb{P}(A \cap B \cap C) = 1/8 = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$. Jedoch ist

$$\mathbb{P}(A \cap B) = \frac{3}{8} \neq \mathbb{P}(A)\mathbb{P}(B).$$

Satz 2.12. *Es sei I eine beliebige Indexmenge.*

(a) Sei $(A_i)_{i \in I}$ eine Familie von unabhängigen Ereignissen, $k \notin I$ und $A_k \subset \Omega$ mit $\mathbb{P}(A_k) \in \{0, 1\}$. Dann ist auch die Familie $(A_i)_{i \in I \cup \{k\}}$ unabhängig.

(b) Sei $(A_i)_{i \in I}$ eine Familie von unabhängigen Ereignissen, für alle $i \in I$ sei $B_i \in \{A_i, A_i^{\mathrm{c}}, \emptyset, \Omega\}$. Dann ist auch $(B_i)_{i \in I}$ unabhängig.

Sei nun $n \in \mathbb{N}$, $I_n = \{1, 2, \ldots, n\}$ und $(A_i)_{i \in I_n}$ eine Familie von Ereignissen.

(c) Die Ereignisse $(A_i)_{i \in I_n}$ sind genau dann unabhängig, wenn für alle Wahlen $B_i \in \{A_i, A_i^{\mathrm{c}}\}$, $i \in I$, gilt

$$\mathbb{P}\left(\bigcap_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} \mathbb{P}(B_i).$$

(d) Insbesondere sind die Ereignisse $A_1, \ldots, A_n \subset \Omega$ genau dann unabhängig, wenn $A_1^{\mathrm{c}}, \ldots, A_n^{\mathrm{c}}$ unabhängig sind.

Beweis. Muss noch getippt werden. □

2.3 Produkträume

Der Begriff der Unabhängigkeit ist eng verknüpft mit Produktäumen von mehreren Wahrscheinlichkeitsräumen. Man denke an nacheinander und unabhängig von einander ausgeführten Zufallsexperimenten, die jeweils durch einen Wahrscheinlichkeitsraum beschrieben werden. Die gesamte Versuchsreihe dieser n Experimente wird in natürlicher Weise durch den Produktraum beschrieben.

Definition 2.13 (Produktwahrscheinlichkeitsraum)**.** Es sei $n \in$ und (Ω_1, p_1), (Ω_1, p_2), …, (Ω_n, p_n) diskrete Wahrscheinlichkeitsräume. Auf der Produktmenge

$$\Omega = \Omega_1 \times \ldots \times \Omega_n = \{(\omega_1, \ldots, \omega_n) \colon \omega_1 \in \Omega_1, \ \ldots, \omega_n \in \Omega_n\}$$

definieren wir $p : \Omega \to [0,1]$ durch

$$p(\omega) = \prod_{i=1}^{n} p_i(\omega_i).$$

Dann ist (Ω, p) ein diskreter Wahrscheinlichkeitsraum, der Produktwahrscheinlichkeitsraum der Räume $(\Omega_1, p_1), \ldots, (\Omega_n, p_n)$ genannt wird. Er wird auch mit $\otimes_{i=1}^{n}(\Omega_i, p_i)$ bezeichnet. Falls die Räume $(\Omega_1, p_1), \ldots, (\Omega_n, p_n)$ identisch sind, so schreiben wir auch $(\Omega, p) = (\Omega_1, p_1)^{\otimes n}$.

Als Übungsaufgabe zeige man, dass (Ω, p) tatsächlich ein diskreter Wahrscheinlichkeitsraum ist, das heißt, für alle $\omega \in \Omega$ gilt $p(\omega) \geq 0$ sowie

$$\sum_{\omega \in \Omega} p(\omega) = 1.$$

Satz 2.14 (Satz zur Unabhängigkeit bei Produkträumen)**.** *Sei $n \in \mathbb{N}$, (Ω_1, p_1), … , (Ω_n, p_n) diskrete Wahrscheinlichkeitsräume, und $A_1 \subset \Omega_1, \ldots, A_n \subset \Omega_n$ Ereignisse. Für $i \in \{1, 2, \ldots, n\}$ sei weiterhin*

$$\hat{A}_i = \{(\omega_1, \ldots, \omega_n) \in \Omega \colon \omega_i \in A_i\}.$$

Dann sind die Ereignisse $\hat{A}_1, \ldots, \hat{A}_n \subset \Omega$ unabhängig (im Produktraum (Ω, p)).

Die Ereignisse $\hat{A}_i$ sind nichts weiter als eine Art Einbettung der Ereignisse A_i in den Produktraum.

Beweis von Satz 2.14. Sei $\mathbb{P}$ das von p induzierte Wahrscheinlichkeitsmaß und $\mathbb{P}_i$, $i \in \{1, \ldots, n\}$, die von p_i induzierten Wahrscheinlichkeitsmaße. Sei $J \subset \{1, 2, \ldots, n\}$ beliebig. Dann gilt

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{j \in J} \hat{A}_j\right) &= \mathbb{P}\left(\omega \in \Omega \colon \omega_j \in A_j \text{ für alle } j \in J\right) \\
&= \sum_{\substack{\omega \in \Omega: \\ \omega_j \in A_j, j \in J}} p(\omega) \\
&= \sum_{\substack{\omega_1, \ldots, \omega_n: \\ \omega_j \in A_j, j \in J}} p_1(\omega_1) \cdot \cdots \cdot p_n(\omega_n).
\end{aligned}$$

Den letzten Ausdruck können wir mit der Notation $B_j = A_j$ falls $j \in J$ und $B_j = \Omega_j$ falls $j \notin J$ zusammenfassen als

$$\sum_{\omega_1 \in B_1} p_1(\omega_1) \cdots \sum_{\omega_n \in B_n} p_n(\omega_n) = \prod_{i=1}^{n} \sum_{\omega_i \in B_i} p_i(\omega_i) = \prod_{i=1}^{n} \mathbb{P}_i(B_i) = \prod_{j \in J} \mathbb{P}_j(A_j),$$

da ja $\mathbb{P}_i(B_i) = \mathbb{P}_i(\Omega_i) = 1$ für $j \notin J$. Man rechnet leicht nach, dass

$$\hat{A}_i = \Omega_1 \times \ldots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \ldots \times \Omega_n, \text{ und somit} \mathbb{P}(\hat{A}_i) = \mathbb{P}_i(A_i)$$

gilt. Damit ist die Produktformel

$$\mathbb{P}\left(\bigcap_{j \in J} \hat{A}_j\right) = \prod_{j \in J} \mathbb{P}(\hat{A}_j)$$

gezeigt und der Beweis beendet. $\square$

Kommentare für Vorlesenden

- In Abschnitt 4 koennte man Theorie und Beispiele zu bedingten Erwartungswerten aufnehmen.
- Ende Abschnitt 4.5. Lemma zur besten linearen Vorhersage passend.

Literaturverzeichnis

[Geo09] H.-O. Georgii. *Stochastik*. Walter de Gruyter, Berlin, 2009.