1 Mathematische Modellierung von Zufallsexperimenten

1.1 Axiomatischer Aufbau der Wahrscheinlichkeitstheorie

Ausgangspunkt der Wahrscheinlichkeitstheorie ist die mathematische Beschreibung von ***Zufallsexperimenten***, das heißt zeitlich wie örtliche fest umrissene Vorgänge mit unbestimmten Ausgang.

Beispiel 1.1

- Werfen eines Würfels oder einer Münze
- Zufälliges Ziehen einer Kugel aus einer Urne
- Kartenspiele
- Wahlergebnis der nächsten Europawahl
- Temperatur am Alexanderplatz am 11. April 2019 um 12:00 Uhr
- Lebensdauer von technischen Geräten

Die Gesamtheit aller möglichen Ausgänge/ Ergebnisse eines Zufallsexperimentes heißt ***Ergebnisraum*** oder ***Stichprobenraum*** (letzteres wird eher in der Statistik verwendet). Dies wird mit Ω bezeichnet. Ein Element $\omega \in \Omega$ heißt ***Ergebnis*** oder ***Stichprobe***.

Beispiel 1.2

- einmaligen Würfeln:
$$\Omega = \{1, ..., 6\}, \quad |\Omega| = 6$$
- zweimaligen Würfeln:
$$\Omega = \{(i,j) : i,j \in \{1,...,6\}\} = \{1,...,6\}\times\{1,...,6\} = \{1,...,6\}^2, \quad |\Omega| = 36$$
- Münzwurf: $\Omega = \{\text{Kopf, Zahl}\}$ oder $\Omega = \{0,1\}$.
- Anzahl der Autos am Funkturm am 11. April 2019: $\Omega = \mathbb{N} \cup \{0\}$
- Temperatur in Grad Kelvin am Alexanderplatz, d.h. $\Omega = [0, \infty)$ oder realistischer $\Omega = [270, 310]$

In den ersten drei Beispielen ist der Ergebnisraum *endlich*, im vorletzten Beispiel *abzählbar unendlich* und im letzten Beispiel *kontinuierlich* (d.h. überabzählbar unendlich). (i)-(iv) nennt man **diskrete** Ergebnisräume.

Ereignisse sind Teilmengen $A \subset \Omega$. Die Gesamtheit aller Ereignisse wird mit $\mathcal{P}(\Omega)$ bezeichnet (Potenzmenge). Besonders hervorzuheben sind

- das *sichere Ereignis*: Ω
- das *unmögliche Ereignis*: $\emptyset$
- die *Elementarereignisse*: $\{\omega\}$ für $\omega \in \Omega$

Beispiel 1.3

1. $A = \{1,3,5\} =$ "Augenzahl ungerade"
2. $A = \{(5,6),(6,5),(6,6)\} =$ "Augensumme > 10"
3. $A = \{n : n \geq 40000\} =$ "ungewöhnlich hohes Verkehrsaufkommen"

Operationen auf Ereignissen

1. $A \cup B = A$ oder B tritt ein
2. $A \cup ... \cup A_n = \bigcup_{k=1}^{n} A_k =$ mindestens eins der A_k tritt ein
3. $A \cap B$ = A und B treten ein
4. $A \cap ... \cap A_n = \bigcap_{k=1}^{n} A_k =$ alle A_k treten ein
5. $A^c := \Omega \setminus A = \{\omega : \omega \notin A\} = A$ tritt *nicht* ein ($\emptyset^c = \Omega, \Omega^c = \emptyset$)

Wichtig 1

Insbesondere im Falle kontinuierlicher Mengen ist es im allgemeinen unmöglich *jedem* Ereignis in konsistenter Weise Wahrscheinlichkeiten zuzuordnen (siehe Satz 1.15). Daher schränkt man sich auf *kleinere* Mengensysteme ein.

Dies führt auf den Begriff der ***σ-Algebra***.

Definition 1.1

Es sei $\Omega \neq \emptyset$. Ein Mengensystem $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt σ-Algebra, falls gilt

- $\Omega \in \mathcal{A}$
- $A \in \mathcal{A} \implies A^C = \Omega \setminus A \in \mathcal{A}$
- $(A_n)_{n\in\mathbb{N}} \in \mathcal{A} \implies \bigcup_{n\geq 1} A_n \in \mathcal{A}$

Das Paar $(\Omega, \mathcal{A})$ heißt ***messbarer Raum***. Eine Teilmenge $A \in \Omega$ heißt ***$\mathcal{A}$-messbar***, falls $A \in \mathcal{A}$

Lemma 1.1

Es sei $\mathcal{A}$ eine σ-Algebra. Dann gilt

1. $\Omega \notin \mathcal{A}$
2. $A, B \in \mathcal{A} \implies A \cup B, A \cap B, A \setminus B \in \mathcal{A}$.
3. $(A_n)_{n\in\mathbb{N}} \in \mathcal{A} \implies \bigcap_{n\geq 1} A_n \in \mathcal{A}$

Proof. Siehe Skript. □

Beispiel 1.4

1. $\mathcal{P}(\Omega)$ ist eine σ-Algebra und zwar die *größtmögliche*.
2. $\{\emptyset, \Omega\}$ ist σ-Algebra und zwar die kleinstmögliche, *triviale* σ-Algebra.
3. Ist $A \subset \Omega$ ein Ereignis, so ist $\{\emptyset, \Omega, A, A^c\}$ eine σ-Algebra und zwar die kleinste, die A enthält.
4. **Hüllenoperator:** Ist $\mathcal{C} \subset \mathcal{P}(\Omega)$ ein beliebiges Mengensystem, so ist
$$\sigma(\mathcal{C}) = \bigcap_{\substack{\mathcal{A}\ \sigma\text{-Algebra auf } \Omega \\ \text{mit } \mathcal{C} \subset \mathcal{A}}} \mathcal{A}$$
eine σ-Algebra (Beweis!). Sie ist die kleinste die $\mathcal{C}$ enthält. Sie heißt die ***von $\mathcal{C}$ erzeugte σ-Algebra***. $\sigma(\mathcal{C})$ ist wohldefiniert, da der Schnitt nicht leer ist (siehe $\mathcal{P}(\Omega)$).

Es gelten für den Hüllenoperator:

1. $\mathcal{C} \subset \sigma(\mathcal{C})$
2. $\mathcal{C}_1 \subset \mathcal{C}_2 \implies \sigma(\mathcal{C}_1) \subset \sigma(\mathcal{C}_2)$

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

Wahrscheinlichkeiten

Nicht gottgegeben!!! Im nächsten Schritt wollen wir für jedes messbare Ereignis A eine Wahrscheinlichkeit $P(A)$ angeben zwischen 0 und 1. $P(A)$ soll ein Maß dafür sein, dass A auftritt:

- tritt A niemals ein, so setzen wir $P(A) = 0$,
- tritt A sicher sein, so setzen wir $P(A) = 1$.

Zusätzlich sollte gelten: Sind A und B *disjunkte* Ereignisse, d.h. $A \cap B = \emptyset$, so ist

$$P(A \cup B) = P(A) + P(B) \quad . \qquad \text{(Additivität)}$$

Unmittelbare Folgerung $A_1, ..., A_n$ paarweise disjunkte Ereignisse, so gilt

$$P(A_1 \cup ... \cup A_n) = P(A_1) + ... + P_(A_n).$$

Definition 1.2: Wahrscheinlichkeitsmaß

Sei (Ω, A) ein messbarer Raum. EIne Abbildung $P : \mathcal{A} \to [0,1]$ heißt ***Wahrscheinlichkeitsmaß*** auf $\mathcal{A}$, falls gilt

1. **Normiertheit**: $P(\Omega) = 1$
2. **σ-Additivität**: Für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweiser disjunkter Ereignisse in $\mathcal{A}$ gilt
$$P(\bigcup_{n\geq 1} A_n) = \sum_{n\geq 1} P(A_n).$$

Die Begriffe Wahrscheinlichkeitsmaß und σ-Algebra sind absolut essentiell.

Lemma 1.2: Rechenregeln für P

Sei P ein Wahrscheinlichkeitsmaß auf $\Omega, \mathcal{A}$ und $A, B, A_1, A_2, ... \in \mathcal{A}$. Dann gilt

1. $P(\emptyset) = 0$
2. P ist insbesondere *endlich additiv*, d.h. für paarweise disjunkte $A_1, ..., A_n$ gilt
$$P(A_1 \cup ... \cup A_n) = P(A_1) + ... + P(A_n).$$
3. **Komplementärwahrscheinlichkeit** $P(A^c) = 1 - P(A)$
4. **Monotonie** (die wichtigste Eigenschaft): $A \subset B \implies P(A) \leq P(B)$
5. **Subadditivität**: $A \subset \bigcup_n A_n \implies P(A) \leq \sum_n P(A_n)$.

Proof. 1. Betrachte die triviale Folge $A_n := \emptyset$ für $n = 1, 2, ...$. Offensichtlich sind die $A_n, n \geq 1$ paarweise disjunkt. Wegen $\bigcup_n A_n = \bigcup \emptyset = \emptyset$ und aus der σ-Additivität folgt jetzt
$$P(\emptyset) = P(\bigcup_n A_n) = \sum_n P(\emptyset).$$
Wäre $P(\emptyset) > 0$, so würde $\sum P(\emptyset)$ divergieren. Also $P(\emptyset) \in [0,1] \implies P(\emptyset) = 0$.

2. Setze $A_k := \emptyset$ für $k \geq n+1$ und wir erhalten somit eine Folge paarweiser disjunkter Ereignisse $(A_k)_{k\geq 1}$ mit $\bigcup_{k\geq 1} A_k = A_1 \cup ... \cup A_n \implies P(A_1 \cup ... \cup A_n) = P(\bigcup_{k\geq 1} A_k) = \sum_{k\geq 1} P(A_k) = P(A_1) + ... + P(A_n)$, denn $P(A_k) = P(\emptyset) = 0$ für $k \geq n+1$.

3. Sei $A \in \mathcal{A}$. Dann sind A, A^c paarweise disjunkt und $A \cup A^c = \Omega$. Damit $1 = P(\Omega) = P(A \dot\cup A^c) = P(A) + P(A^c) \implies P(A^c) = 1 - P(A)$. Der Vereinigungsoperator $\dot\cup$ bedeutet, dass man zwei disjunkte Mengen vereinigt.

4. $A \subset B$, d.h. $B = A \dot\cup B \setminus A \implies P(B) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0} \geq P(A)$.

5. Definiere induktiv $B_1 := A_1, B_2 = A_2 \setminus A_1 = A_2 \setminus B_1$. Also definiere einfach für alle $n \geq 2$: $B_n := A_n \setminus (A_1 \cup ... \cup A_{n-1})$. Dann sind die B_n paarweise disjunkte Ereignisse und $B_1 \cup ... \cup B_n = A_1 \cup ... \cup A_n$ für alle $n \in \mathbb{N}$. Damit
$$P(\bigcup_n A_n) = P(\bigcup_n B_n) = \sum_n P(B_n) \leq \sum_n P(A_n)$$
und ist $A \subset \bigcup_n A_n \implies P(A) \leq P(\bigcup_n A_n) \leq \sum_n P(A_n)$.

□

Definition 1.3: Wahrscheinlichkeitsraum

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

Theorem 1.1: Einschluss-Ausschluss-Prinzip, Formel von Sylvester

Es sei $(\Omega, \mathcal{A}, P$ ein Wahrscheinlichkeitsraum. Dann gilt für beliebige $A_1, ... A_n \in \mathcal{A}$:
$$\begin{aligned} P(A_1 \cup ... \cup A_n) &= \sum_{\emptyset \neq I \subset \{1,...,n\}} (-1)^{|I|-1} \cdot P(\bigcap_{i \in I} A_i) \\ &= \sum_{j=1}^{n} (-1)^{j-1} \sum_{1 \leq i_1 < i_2 < ... < i_j \leq n} P(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_j}), \end{aligned}$$

Spezialfälle

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Proof. Induktion nach n: Induktionsanfang $n = 1$. Für $n = 2$: Es gilt $A_1 \cup A_2 = A_1 \dot\cup (A_2 \setminus A_1) \implies P(A_1 \cup A_2) = P(A_1) + P(A_2 \setminus A_1)$ und $A_2 = (A_2 \setminus A_1) \dot\cup (A_2 \cap A_1) \implies P(A_2) = P(A_2 \setminus A_1) + P(A_2 \setminus A_1)$. Damit folgt
$$P(A_1 \cup A_2) = P(A_1) + P(A_2 \setminus A_1) = P(A_1) + P(A_2) - P(A_1 \cap A_2). \qquad (1)$$

Induktionsschluss: $n \rightsquigarrow n+1$ Wende (1) auf $A_1 \cup ... \cup A_n$ und A_{n+1} an. Wir erhalten
$$P(A_1 \cup ... \cup A_{n+1}) = P(A_1 \cup ... \cup A_n) + P(A_{n+1}) + P(A_{n+1}) + P((A_1 \cup ... \cup A_n) \cap A_{n+1}) \qquad (2)$$
$$= P(A_1 \cup ... \cup A_n) + P(A_{n+1}) + P(A_{n+1}) + P((A_1 \cap A_{n+1}) \cup ... \cup (A_n \cap A \qquad (3)$$

Nach Induktionsannahme gilt
$$P(A_1 \cup ... \cup A_n) = \sum_{\emptyset \neq I \subset \{1,...,n\}} (-1)^{|I|-1} P(\bigcup_{i \in I} A_i)$$
und $P((A_1 \cap A_{n+1}) \cup ... \cup (A_n \cap A_{n+1})) = \sum_{\emptyset \neq I \subset \{1,...,n\}} (-1)^{|I|-1} P(\bigcap_{i \in I} (A_i \cap A_{n+1})) = \sum_{\emptyset \neq I \subset \{1,...,n\}} (-1)^{|I|-1} P(\bigcap_{i \in I \cup \{n+1\}} A_1)$.
Einsetzen in (**??**) ergibt:
$$P(A_1 \cup ... \cup A_{n+1}) = \sum_{\emptyset \neq I \subset \{1,...,n\}} (-1)^{|I|-1} P(\bigcap_{i \in I} A_i) + P(A_{n+1}) - \sum_{\emptyset \neq I \subset \{1,...,n\}} (-1)^{|I|-1} P(\bigcap_{i \in I \cup \{n+1}$$
und wir sind fertig.

□

Beispiel 1.5

Wie groß ist die Wahrscheinlichkeit bei 10-maligen Würfeln nicht alle 6 Augenzahlen zu würfeln?

$E_i := i$ kommt in den 10 Würfen nicht vor, $i = 1, ..., 6$
Gesuchte Wahrscheinlichkeit: $P(E_1 \cup ... \cup E_6)$
Es gilt: $P(E_1 \cap ... \cap E_i) = \frac{(6-i)^{10}}{6^{10}}$. Allgemeiner $P(\bigcap_{i \in I} E_i) = \frac{(6-|I|)^{10}}{6^{10}}$.

Das heißt

$$P(E_1 \cup ... \cup E_6) = \sum_{j=1}^{6} (-1)^{j-1} \sum_{1 \le i_1 < .. < i_j \le 6} \frac{(6-j)^{10}}{6^{10}} = \sum_{j=1}^{6} (-1)^{j-1} \binom{6}{j} \frac{(6-j)^{10}}{6^{10}}$$

Theorem 1.2

Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $(A_n)_n$ eine Folge von Ereignissen in $\mathcal{A}$. Dann gilt

1. $A_n \uparrow A$, d.h. $A_1 \subset A_2 \subset ...$ und $\bigcup_n A_n = A$, so folgt
$$\lim_{n\to\infty} P(A_n) = P(A) \qquad \text{(Stetigkeit von unten)}$$
2. $A_n \downarrow A$, d.h. $A_1 \supset A_2 \supset ...$ und $\bigcap_n A_n = A$, so folgt
$$\lim_{n\to\infty} P(A_n) = P(A) \qquad \text{(Stetigkeit von oben)}$$

Proof. Ohne Beweis. □

1.2 Diskrete Wahrscheinlichkeitsräume

Ist Ω diskret, so ist Teilmenge $A \subset \Omega$ abzählbar. Also

$$A = \{w_1, w_2, ...\} \quad \text{und damit} \quad A = \bigcup_n \{w_n\}$$

Ist $\mathcal{A}$ eine σ-Algebra in Ω mit $\{w\} \in \mathcal{A}, \forall w \in \Omega$ und P ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$, so folgt

$$P(A) = \sum_n P(\{w_n\}) = \sum_{w\in A} P(\{w\})$$

d.h. P ist eindeutig bestimmt durch die Wahrscheinlichkeiten $p(w) := P(\{w\})$. p definiert eine Funktion $p : \Omega \to [0,1]$ heißt ***Zähldichte***, d.h. eine Funktion $p : \Omega \to [0,1]$ mit $\sum_{w\in\Omega} p(w) = 1$. Ist umgekehrt p eine Zähldichte, so definiert

$$P(A) := \sum_{w\in A} p(w), \quad A \in \mathcal{P}(\Omega) \tag{4}$$

ein Wahrscheinlichkeitsmaß auf *ganz* $P(\Omega)$.

Wichtig 2

Im diskreten Fall ist $\mathcal{P}(\Omega)$ die natürliche σ-Algebra und daher ersetzt man im diskreten Falle $(\Omega, \mathcal{P}(\Omega), P)$ einfach durch (Ω, p) (siehe Satz 1.10 im Skript).

Satz 1.10 besagt also, dass für diskrete Ergebnisräume durch (4) eine 1-1 Beziehung zwischen Zähldichten und Wahrscheinlichkeitsmaßen auf $P(\Omega)$ gegeben ist.

Beispiel 1.6: n-maliges Werfen einer fairen Münze

Sei $\Omega = \{(i_1, ..., i_n) : i_j \in \{0,1\}, 1 \le j \le n\}, |\Omega| = 2^n$. Damit $p(w) = \frac{1}{|\Omega|} = \frac{1}{2^n}$. Für $k \in \{1, ..., n\}$ gilt dann beispielsweise:

$$P(\text{im k-ten Münzwurf "Kopf" werfen}) = \frac{|\{(i_1, ..., i_{k-1}, 1, i_{k+1}, ..., i_n) : i_j = 0,1\}|}{|\Omega|} = \frac{2^{n-1}}{2^n} = \frac{1}{2}$$