Kapitel 1

Grundlagen aus der Wahrscheinlichkeitstheorie

1.1 Ereignisse, Zufallsvariablen und Wahrscheinlichkeiten

Worum geht es?

In vielen Situationen tritt „Zufall“ oder „Ungewissheit“ auf, in diesem Kapitel geht es uns darum, den üblichen mathematischen Rahmen kennen zu lernen, um solche Phänomene zu beschreiben.

Beispiele.

- Im Auftakt-Beispiel in Kapitel 0 haben wir ein zufälliges Pixel Z aus (einer diskretisierten Version von) $[0,1]^2$ gewählt.
- „Glücksspiel-Situationen“:

  Wir werfen eine Münze oder einen Würfel, spielen Lotto, schauen ein gut gemischtes Pokerblatt an, ...
- Wir befragen einen zufällig ausgewählten Teilnehmer einer Videokonferenz nach seinem Alter und danach, wieviel er letzte Woche für Online-Bestellungen ausgegeben hat.
- Wir übergeben n Zahlen in rein zufälliger Reihenfolge an einen Sortieralgorithmus;

  wir überprüfen zwei Polynome $f(x)$ und $g(x)$ vom Grad $d \gg 1$ auf Gleichheit, indem wir an k zufällig ausgewählten Stellen auf Gleichheit testen;

  wir betrachten einen Router, der Datenpakete an einen zufällig ausgewählten Nachbarknoten im Netzwerk schickt; ...

Mathematische Modellierung des Zufalls („Zutaten“)**:**

Ω (nicht-leere) Menge („Ergebnisraum“ oder „Stichprobenraum“ genannt), und $\mathcal{F} \subset 2^\Omega$ $(:= \{B : B \subset \Omega\})$.

$\omega \in \Omega$ heißt ein „Elementarereignis“,

$A \in \mathcal{F}$ nennen wir ein „Ereignis“,

insbesondere $\Omega \ldots$ „sicheres Ereignis“, $\varnothing \ldots$ „unmögliches Ereignis“

Vorstellung: der Zufall „wählt“ ein $\omega \in \Omega$ aus,
das Ereignis A „tritt ein“, wenn $\omega \in A$
(sonst „tritt es nicht ein“)

Operationen

(Mengenoperationen und ihre Interpretation für Ereignisse)

Für $A, B \in \mathcal{F}$

$A^c := \Omega \setminus A$ ... „A tritt nicht ein“
(${A^c}$ heißt Gegen- oder Komplementärereignis von A)

$A \cap B$ ... „A und B treten ein“

$A \cup B$ ... „A oder B treten ein“

$A \subset B$ ... „A impliziert B“ (falls A eintritt, so auch B)

A und B heißen *disjunkt*, wenn $A \cap B = \emptyset$.

Wir notieren auch $A \setminus B := A \cap B^c = \{\omega : \omega \in A, \omega \notin B\}$ („A tritt ein, aber B nicht“).

Forderungen für Ereignisse

Wir fordern, dass $\mathcal{F}$ erfüllt

$$\begin{aligned} &\textit{i)} \quad \emptyset \in \mathcal{F}, \\ &\textit{ii)} \quad A \in \mathcal{F} \;\Rightarrow\; A^c \in \mathcal{F}, \\ &\textit{iii)} \quad A_1, A_2, \dots \in \mathcal{F} \;\Rightarrow\; \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}. \end{aligned} \tag{1.1}$$

$\mathcal{F} \subset 2^{\Omega}$, das (1.1), *i)–iii)* genügt, heißt eine σ-*Algebra* (über Ω).

Insbesondere gilt dann auch

$\Omega = (\emptyset)^c \in \mathcal{F}$,

für $A_1, A_2, \dots \in \mathcal{F}$ ist $\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \in \mathcal{F}$

$A_1 \cup A_2 = A_1 \cup A_2 \cup \emptyset \cup \emptyset \cup \dots \in \mathcal{F}$ (de Morgan-Regeln)

Forderungen für Wahrscheinlichkeiten

Eine Abbildung $P : \mathcal{F} \to [0,1]$ mit

(N) $P(\Omega) = 1$ („Normierung") und

(A) $A_1, A_2, \dots \in \mathcal{F}$ paarweise disjunkt (1.2)

$$\Rightarrow P\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} P(A_n) \qquad \text{(„}\sigma\text{-Additivität")}$$

heißt ein *Wahrscheinlichkeitsmaß* (auf $(\Omega, \mathcal{F})$).

$P(A)$ heißt / nennen wir die *Wahrscheinlichkeit des Ereignisses* A.

Definition 1.1. Ein Tripel $(\Omega, \mathcal{F}, P)$, wobei $\mathcal{F} \subset 2^{\Omega}$ die Forderungen (1.1), *i)–iii)* und P die Forderungen (1.2), (N) und (A) erfüllt, heißt ein *Wahrscheinlichkeitsraum*.

Beispiel und Definition 1.2 (Diskreter Wahrscheinlichkeitsraum).

Ω endlich oder abzählbar, $p : \Omega \to [0,1]$ Abbildung mit $\sum_{\omega\in\Omega} \underbrace{p(\omega)}_{\geq 0} = 1$.

$$P(A) := \sum_{\omega\in A} p(\omega), \quad A \subset \Omega$$

definiert ein Wahrscheinlichkeitsmaß auf $(\Omega, 2^\Omega)$.

p heißt die (Wahrscheinlichkeits-)*Gewichtsfunktion* von P,

$p(\omega)$ heißt das (Wahrscheinlichkeits-)*Gewicht* von ω.

(denn: (N): $P(\emptyset) = \sum_{\omega\in\emptyset} p(\omega) = 0$ (n. Def.)

(A): $A_1, A_2, \ldots$ paarw. disjunkt

$$P\Big(\dot{\bigcup_{n=1}^{\infty}} A_n\Big) = \sum_{\omega \in \dot{\bigcup}_{n=1}^{\infty} A_n} p(\omega) = \sum_{n=1}^{\infty} \overbrace{\sum_{\omega\in A_n} p(\omega)}^{= P(A_n)}$$

)

Beispiele 1.3. 1. (uniforme Wahl aus endlicher Menge, „Laplace-Experimente")

$\#\Omega < \infty$ und $p(\omega) = \frac{1}{\#\Omega}$ in Bsp. 1.2, z.B.

(a) (Wurf eines fairen 6er-Würfels)

$\Omega = \{1,2,\dots,6\}, \quad p(\omega) = \frac{1}{6}, \quad \omega \in \Omega \quad (\mathcal{F} = 2^\Omega)$

(b) (dreimaliger Wurf eines fairen 6er-Würfels)

$\{1,2,\dots,6\}^3$

$\Omega = \{(\omega_1,\omega_2,\omega_3) : \omega_i \in \{1,\dots,6\}, i=1,2,3\}, \quad p(\omega) = \frac{1}{6^3} = \frac{1}{216}$

$(\mathcal{F} = 2^\Omega)$

(c) (Auftakt-Beispiel aus Kapitel 0 mit Diskretisierung auf 32 Bit Genauigkeit)

$\Omega = \left\{ \frac{k}{2^{32}} : k = 0,1,\dots,2^{32}-1 \right\}^2, \quad p(\omega) = \frac{1}{(2^{32})^2} = \frac{1}{2^{64}}$

(d) (n verschiedene Eingaben in zufälliger Reihenfolge)

$\Omega = \{(x_1,x_2,\dots,x_n) : x_1,\dots,x_n \in \{1,2,\dots,n\}$, alle paarw. versch.$\}$

$(x_i \neq x_j$ für $i \neq j)$

$\#\Omega = n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1, \quad p(\omega) = \frac{1}{n!}$

Beispiele 1.3 (Fortsetzung).

2. (ein verfälschter Münzwurf) $\Omega = \{\mathrm{Kopf}, \mathrm{Zahl}\}$,

   $p(\mathrm{Kopf}) = 0.6 = 1 - p(\mathrm{Zahl})$

   (gew.: $\mathcal{F} = 2^{\Omega}$ mit dem zugeh. Ω)

3. (eine winziges Modell für Spam, das Sprache und Status einer Email betrachtet)

   $$\Omega = \{\texttt{Deutsch}, \texttt{Englisch}\} \times \{\texttt{Spam}, \texttt{keinSpam}\},$$

   $p\big((\texttt{Deutsch}, \texttt{Spam})\big) = 0.2$, $p\big((\texttt{Deutsch}, \texttt{keinSpam})\big) = 0.1$,

   $p\big((\texttt{Englisch}, \texttt{Spam})\big) = 0.6$, $p\big((\texttt{Englisch}, \texttt{keinSpam})\big) = 0.1$

4. (Anzahl Würfe, bevor beim wiederholten fairen Münzwurf zum ersten Mal Kopf kommt)

   $\Omega = \mathbb{N}_0$, $p(n) = (1/2)^n \cdot (1/2) = 2^{-n-1}$ für $n \in \Omega$

   $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$

   $$\sum_{\omega \in \Omega} p(\omega) = \sum_{n=0}^{\infty} 2^{-n-1} = \frac{1}{2} \underbrace{\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n}_{= \frac{1}{1-\frac{1}{2}} = 2} = 1$$

   ($n \mathrel{\hat{=}}$ Anz. geworfener „Zahl“, bevor zu ersten Mal „Kopf“ kommt“)

Bemerkung 1.4. Wahrscheinlichkeitsräume $(\Omega, \mathcal{F}, P)$ gemäß Def. 1.1 werden heute, nicht zuletzt wegen des einflussreichen Werks von A.N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, 1933, oft als „Basisobjekt“ der mathematischen Modellierung von Zufallsvorgängen verwendet. Insoweit ist die explizite Wahl eines geeigneten Wahrscheinlichkeitsraums (oft) ein Teil der Arbeit, wenn ein umgangssprachlich formuliertes Problem in Mathematik übersetzt werden soll.

Man sollte sich bewusst sein, dass es dabei i.A. viele mögliche Wahlen gibt und dass die Formulierung mit Zufallsvariablen oftmals einen für die Intuition und für das Argumentieren sehr angenehmen Zugang ergibt (s.u.).

Siehe auch Abschnitt 1.2 („Kleingedrucktes“) der Vorlesungsnotizen für Hintergrund und weitere Diskussion.

Lemma 1.5. Für $A, B, A_1, A_2, \dots \in \mathcal{F}$ gilt

$$P(\varnothing) = 0 \tag{1.3}$$

$$P(A \cup B) + P(A \cap B) = P(A) + P(B) \quad \text{(endliche Additivität)} \tag{1.4}$$

insbesondere $P(A) + P(A^c) = 1$

$$A \subset B \Rightarrow P(A) \le P(B) \qquad \text{(Monotonie)} \tag{1.5}$$

zudem gilt auch

$$P\Big(\bigcup_{n=1}^{\infty} A_n\Big) \le \sum_{n=1}^{\infty} P(A_n) \qquad (\sigma\text{-Subadditivität}) \tag{1.6}$$

falls $A_n \nearrow_{n\to\infty} A$ (d.h. $A_1 \subset A_2 \subset \dots$ mit $A = \bigcup_{n=1}^{\infty} A_n$)
oder $A_n \searrow_{n\to\infty} A$ (d.h. $A_1 \supset A_2 \supset \dots$ mit $A = \bigcap_{n=1}^{\infty} A_n$),
so gilt $P(A) = \lim_{n\to\infty} P(A_n)$ $\qquad$ (σ-Stetigkeit) $\qquad$ (1.7)

Zu (1.3): $P(\emptyset) = P(\emptyset \cup \emptyset \cup \emptyset \cup \dots) = \sum_{n=1}^{\infty} P(\emptyset) \Rightarrow P(\emptyset) = 0$

Zu (1.4): Falls $A \cap B = \emptyset$, so gilt $P(A \cup B) = P(A \cup B \cup \emptyset \cup \emptyset \dots)$
$= P(A) + P(B) + 0 + \dots$,
d.h. (1.4) gilt.

Allg.: $A \cup B = (A \setminus B) \,\dot\cup\, (B \setminus A) \,\dot\cup\, (A \cap B)$

Lemma 1.5. Für $A, B, A_1, A_2, \dots \in \mathcal{F}$ gilt

$$P(\varnothing) = 0 \tag{1.3}$$

$$P(A \cup B) + P(A \cap B) = P(A) + P(B) \quad \text{(endliche Additivität)} \tag{1.4}$$

insbesondere $P(A) + P(A^c) = 1$

$$A \subset B \Rightarrow P(A) \le P(B) \checkmark \quad \text{(Monotonie)} \tag{1.5}$$

zu (1.5): $B = A \,\dot\cup\, (B \setminus A)$

$\Rightarrow P(B) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0}$

zudem gilt auch

$$P\Big(\bigcup_{n=1}^{\infty} A_n\Big) \le \sum_{n=1}^{\infty} P(A_n) \quad (\sigma\text{-Subadditivität}) \tag{1.6}$$

zu (1.6): $A_n' := A_n \setminus (A_1 \cup A_2 \cup \dots \cup A_{n-1})$

falls $A_n \nearrow_{n\to\infty} A$ (d.h. $A_1 \subset A_2 \subset \dots$ mit $A = \bigcup_{n=1}^{\infty} A_n$)

$\Rightarrow A_n' \subset A_n$ und $\bigcup_{n=1}^{\infty} A_n = \dot{\bigcup_{n=1}^{\infty}} A_n'$

oder $A_n \searrow_{n\to\infty} A$ (d.h. $A_1 \supset A_2 \supset \dots$ mit $A = \bigcap_{n=1}^{\infty} A_n$),

so gilt $P(A) = \lim_{n\to\infty} P(A_n)$ (σ-Stetigkeit) (1.7)

$A \cup B = (A \setminus B) \,\dot\cup\, (B \setminus A) \,\dot\cup\, (A \cap B)$

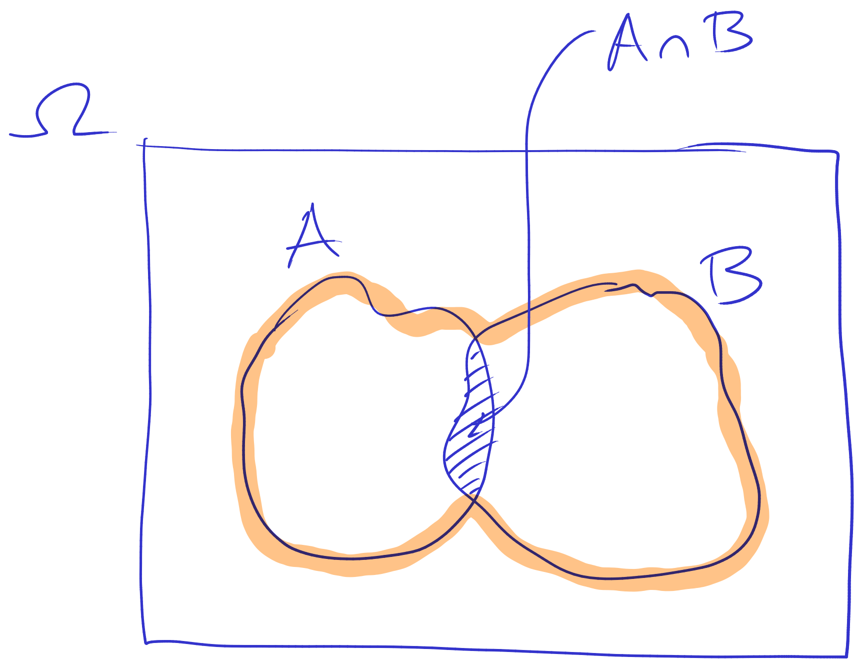
$\Rightarrow P(A \cup B) + P(A \cap B) = P(A \setminus B) + P(B \setminus A) + 2P(A \cap B)$

$= (P(A \setminus B) + P(A \cap B)) + (P(B \setminus A) + P(A \cap B))$

$= P(A) + P(B)$

Bemerkung 1.6 (Einschluss-Ausschluss-Formel)**.** Formel (1.4) aus Lemma 1.5 kann man auch folgendermaßen schreiben

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



Bemerkung 1.6 (Fortsetzung: allgemeine Form der Einschluss-Ausschluss-Formel, „Siebformel“)**.**

Allgemein gilt für $A_1, A_2, \dots, A_n \in \mathcal{F}$

$$
\begin{aligned}
P\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i,j=1\\ i\neq j}}^{n} P(A_i \cap A_j) \\
+ \sum_{\substack{i,j,k=1\\ i\neq j, i\neq k, j\neq k}}^{n} P(A_i \cap A_j \cap A_k) \\
\pm \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)
\end{aligned}
$$

d.h. knapp ausgedrückt

$$
P\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{\emptyset \neq J \subset \{1,\dots,n\}} (-1)^{(\#J)-1} P\Big(\bigcap_{i\in J} A_i\Big) \qquad (1.8)
$$

`[z.B. induktiv, siehe auch Notizen, Abschnitt 1.2.3]`