Wahrscheinlichkeitstheorie für Informatikstudien (Kapitel #1)

Siegfried Hörmann

TU Graz

WS 2020/21

Kursaufbau

1. Wahrscheinlichkeitsräume.
2. Laplace-Experimente.
3. Wichtige diskrete Verteilungen.
4. Zufallsvariablen.
5. Zufallsvektoren.
6. Momente.
7. Bedingte Wahrscheinlichkeiten.
8. Unabhängigkeit.
9. Konvergenz von Zufallsvariablen.
10. Der zentrale Grenzwertsatz.

Wahrscheinlichkeitsräume

1.1. Stichprobenraum.

1.2. Ereignisse.

1.3. Wahrscheinlichkeiten.

Stichprobenraum

In diesem Kapitel ist es unser Ziel, das Konzept eines Zufallsexperiments auf mathematische Art und Weise zu formulieren.

Im Großen und Ganzen kann jeder Prozess mit unsicherem Ausgang als ein Zufallsexperiment gesehen werden.

- Würfeln.
- Münzwurf.
- Tageshöchstwert in Graz in 2 Wochen.
- Der nächste Weltrekord auf der 100 Meter Sprintdistanz.
- ...

Der erste Schritt zu einem mathematisches Modell ist die Menge der möglichen Ergebnisse zu bestimmen.

Stichprobenraum

Der Stichprobenraum (Ereignisraum, Ergebnisraum...) ist die Menge aller möglichen Ergebnisse eines Zufallsexperiments. Wir nennen jene Menge Ω.

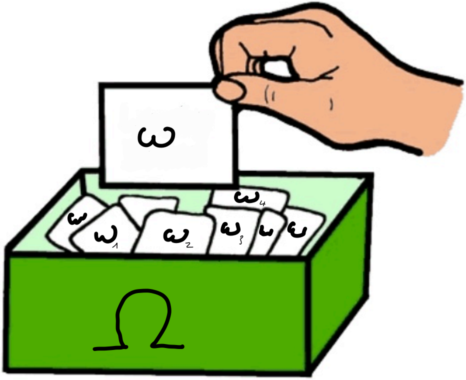
Ein Element $\omega \in \Omega$ nennen wir Elementarereignis.

Beispiel

- Einmal würfeln: $\Omega = \{1, \ldots, 6\}$.
- n-mal würfeln: $\Omega = \{1, \ldots, 6\}^n$.
- Eine Münze ∞ oft werfen: $\Omega = \{0, 1\}^\infty$.
- Temperatur in Graz in 2 Wochen: $\Omega = [-10, 30]$
  (oder etwa $\Omega = \mathbb{R}$?)

Stichprobenraum

Wir können uns also vorstellen, dass ein Zufallsexperiment dem *zufälligen Ziehen eines* $\omega \in \Omega$ entspricht. Natürlich muss das formalisiert werden!



Ereignisse

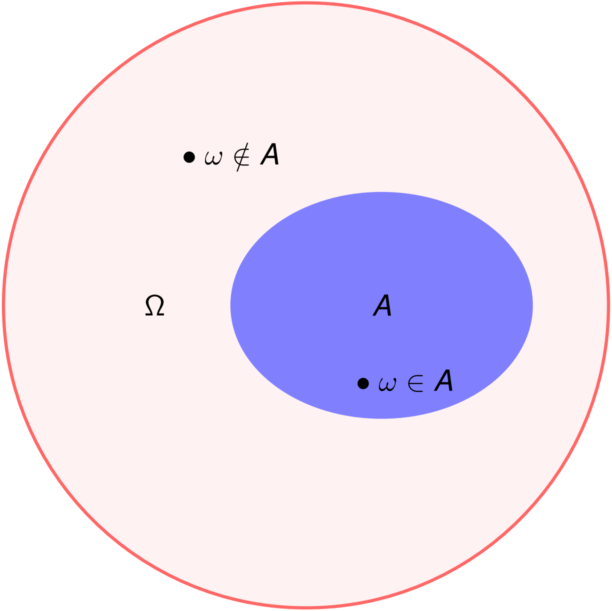
Oft sind wir nicht speziell daran interessiert, welches Elementarereignis das Resultat des Experimentes darstellt.

Manchmal wollen wir lieber wissen, ob ein *zufällig gewähltes* ω spezielle Eigenschaften hat.

- Einmal würfeln: $\Omega = \{1, \ldots, 6\}$.
  Frage: Würfeln wir eine gerade Zahl?
  Formal: $\omega \in \{2, 4, 6\}$?
- Temperatur in 2 Wochen: $\Omega = \mathbb{R}$.
  Frage: Werden wir Temperaturen unter 0 haben?
  Formal: $\omega \in (-\infty, 0)$?

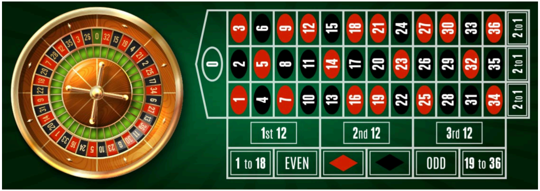
Wir nennen $A \subset \Omega$ ein *Ereignis*. Wir sagen, dass das Ereignis A eintritt, falls $\omega \in A$.

Ereignisse



Ereignisse

Beispiel (Roulette)



37 Felder:

- 18x Schwarz, 18x Rot, 1x Grün.

> Wir drehen 2 Mal. Schreibe folgende Dinge formal auf:
>
> - Ω.
> - A_1 = 2x Rot.
> - A_2 = zuerst 1–12 dann 13–24.
> - A_3 = genau einmal Schwarz.

Ereignisse

Beispiel (Bridge)

52 Karten:

- 4 Farben (Pik, Herz, Karo, Treff) zu je
- 13 Karten (Ass, König, Dame, Bube, 10, 9, 8, 7, 6, 5, 4, 3, 2)

Wir ziehen 2 Karten. Schreibe folgende Dinge formal auf:

- Ω.
- A_1 = 2x Herz.
- A_2 = genau ein Ass und genau einmal Herz.

Ereignisse

Beispiel (Würfeln)

Angenommen wir würfeln 10x. In diesem Falle wählen wir $\Omega = \{(x_1, \ldots, x_{10}) \in \{1, \ldots, 6\}^{10}\}$. Wir betrachten das Ereignis
A : Summe der Augenzahlen ist 40.

Für welche i ist $A \subset A_i$?

$$A_1 = \{(x_1, \ldots, x_{10}) \in \Omega \colon x_1 + \cdots + x_{10} \leq 30\}.$$

$$A_2 = \{(x_1, \ldots, x_{10}) \in \Omega \colon \max_{1 \leq i \leq 10} x_i \geq 4\}.$$

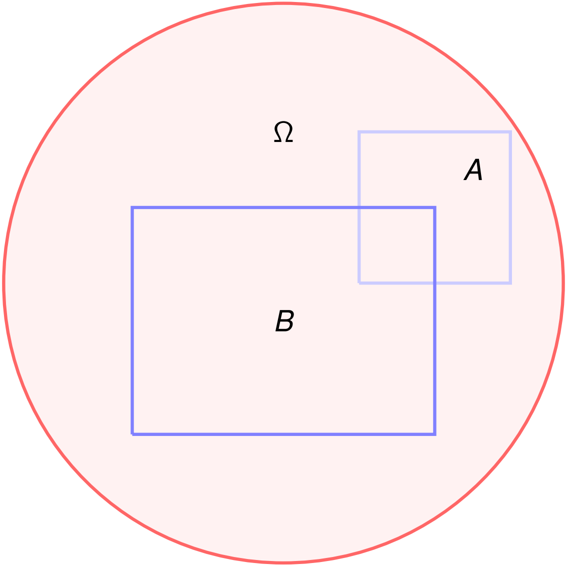
$$A_3 = \{(x_1, \ldots, x_{10}) \in \Omega \colon x_1 + \cdots + x_9 \in [34, 40]\}.$$

Ereignisse

Indem wir elementare Mengenlehre anwenden, können wir Ereignisse kombinieren um neue zu erzeugen und Relationen ausdrücken:

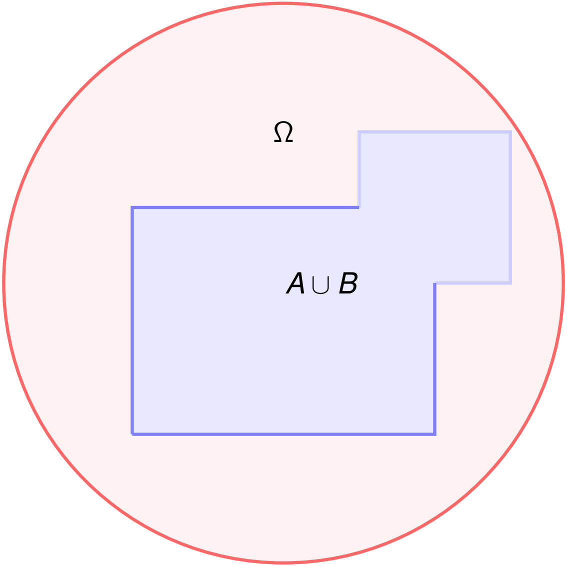
- A oder B tritt ein: $\omega \in A \cup B$.
- A und B treten ein: $\omega \in A \cap B$.
- A tritt ein, aber B tritt nicht ein: $\omega \in A \backslash B$.
- A tritt nicht ein: $\omega \in A^C$.
- Ereignis A impliziert Ereignis B: $A \subset B$.

Ereignisse



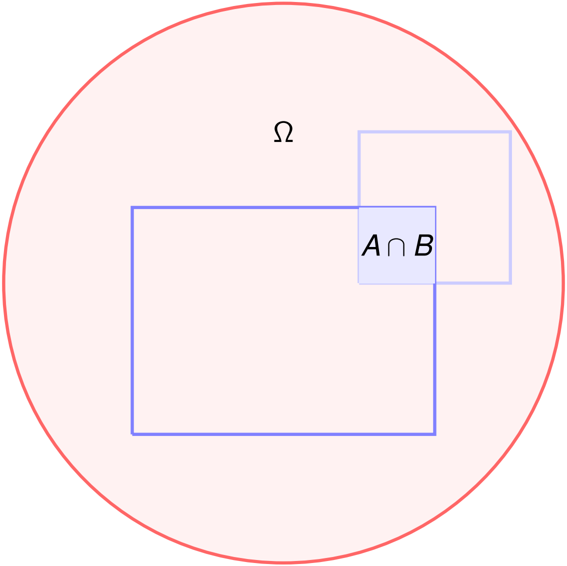
Ereignisse

A oder B tritt ein.



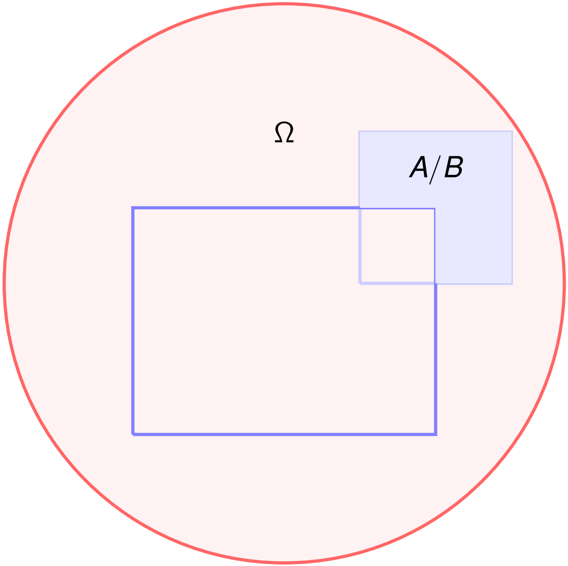
Ereignisse

A und B treten ein.



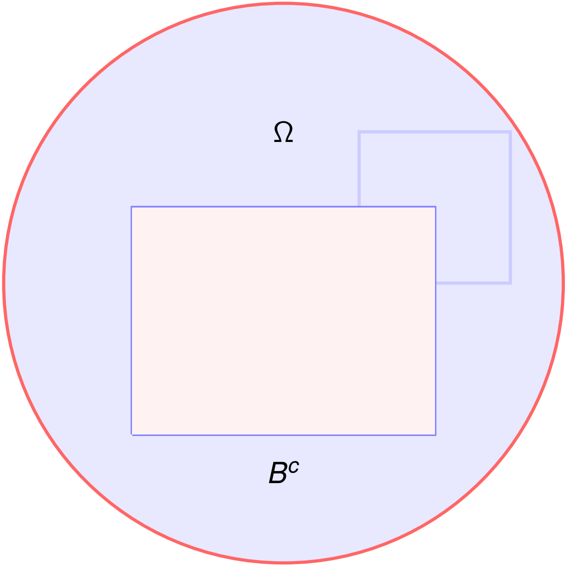
Ereignisse

A tritt ein, aber B nicht.



Ereignisse

B tritt nicht ein.



Ereignisse

Beispiel

Wir haben 52 Bridge Karten. Wir ziehen zufällig 3 Karten mit Zurücklegen.

Sei A_i das Ereignis in welchem die i-te gezogene Karte ein Ass ist.

- Wie können wir Ω definieren?
- Was ist A_i?
- Drücke die folgendenen Ereignisse als Kombination der A_i aus:
  1. Wir ziehen mindestens ein Ass.
  2. Wir ziehen kein Ass.
  3. Wir ziehen 3 Asse.
  4. Wir ziehen genau ein Ass.

Beispiel

Wir haben erneut 52 Bridge Karten und ziehen dieses Mal zufällig 5 davon ohne Zurücklegen.

Sei A_i das Ereignis, in welchem i Asse in der Stichprobe sind, $i = 1, 2, 3, 4$.

- Wie können wir Ω definieren?
- Was ist A_i?
- Drücke die folgendenen Ereignisse als Kombination der A_i aus:
  1. Wir ziehen mindestens ein Ass.
  2. Wir ziehen kein Ass.
  3. Wir ziehen genau 3 Asse.

Beispiel

Angenommen wir werfen eine Münze ∞ oft, also $\Omega = \{0,1\}^\infty$.

Was beschreibt das Ereignis

$$A_i = \{\omega = (\omega_1, \omega_2, \ldots) \colon \sum_{k=1}^{i} \omega_k \leq i - \sum_{k=1}^{i} \omega_k\}.$$

Was beschreibt das Ereignis

$$B_n = \bigcap_{i=1}^{n-1} A_i \cap A_n^c \,?$$

Ereignisse

Was beschreibt das Ereignis

$$A = \bigcup_{n \geq 1} \left(\bigcap_{i=1}^{n-1} A_i \cap A_n^c \right) ?$$

Wahrscheinlichkeiten

Das Hauptziel der Wahrscheinlichkeitstheorie ist es, Ereignissen eine "Wahrscheinlichkeit” zuzuordnen.

Das bedeutet, dass wir eine Funktion P suchen mit

$$A \mapsto P(A) \in [0, 1].$$

Ist zum Beispiel

$$P(A) = 0.3$$

dann sagen wir entweder, ”A hat Wahrscheinlichkeit 0.3” oder ”A hat 30% Wahrscheinlichkeit”.

Wahrscheinlichkeiten

Natürlich macht nicht jede Funktion Sinn in diesem Kontext.

1. Eine passende Funktion sollte einige logische und technische Anforderungen erfüllen, wie z.B.

$$A \subset B \Longrightarrow P(A) \leq P(B).$$

2. In dieser Klasse der Funktionen, welche die technischen Anforderungen (gleich dazu mehr) erfüllen, müssen wir also ein Maß finden, welches für das Problem am besten geeignet ist: Modellierung.

Wir beschäftigen uns zuerst mit Punkt 1.

Zunächst beobachten wir, dass die Klasse der Ereignisse (Teilmengen von Ω) welche wir betrachten wollen, nicht zu klein sein sollte. Zum Beispiel, sind A und B zulässige Ereignisse, so wollen wir, dass auch $A \cup B$, $A \cap B$, etc. zulässige Ereignisse sind.

> Wenn eine Klasse von Ereignissen $\mathcal{A}$ abgeschlossen ist gegenüber jeder belieben Abfolge an Mengenoperationen, dann nennt man $\mathcal{A}$ eine σ-Algebren.

Wir gehen darauf nicht näher ein, und verwenden vorerst $\mathcal{A} = \mathcal{P}(\Omega)$ die Potenzmenge (Menge aller Teilmengen von Ω).

Wahrscheinlichkeiten

Wir sind nun am Punkt angelangt, dass wir jedem Ereignis $A \in \mathcal{A}$ eine Wahrscheinlichkeit $P(A) \in [0, 1]$ zuweisen wollen.

Wir setzen für P folgendes voraus:

(N) $P(\Omega) = 1$.

(A) Sind $(A_i \colon i \geq 1)$ paarweise disjunkt, so gilt

$$P\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} P(A_i).$$

Hier steht (N) für Normierung und (A) für σ-Additivität.

Wahrscheinlichkeiten

Eine Abbildung $P : \mathcal{A} \to [0, 1]$ die (N) und (A) erfüllt, heißt Wahrscheinlichkeitsmaß. Das Tripel $(\Omega, \mathcal{A}, P)$ wird Wahrscheinlichkeitsraum genannt.

Frequentistische Interpretation: Angenommen wir können das Experiment unter denselben Gegebenheiten immer und immer wieder wiederholen. Dann gilt

$$\frac{\text{Häufigkeit mit der } A \text{ eintritt}}{\text{Anzahl der Wiederholungen}} \to P(A).$$

Dies wird in einem späteren Kapitel mithilfe des Gesetzes der Großen Zahlen nachgewiesen.

Wahrscheinlichkeiten

Satz

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und seien $A, B, A_1, A_2, \ldots$ Ereignisse in $\mathcal{A}$. Dann gelten unter anderem:

1. $P(\emptyset) = 0$;
2. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$;
3. $A \subset B \Longrightarrow P(A) \leq P(B)$;
4. $P(\bigcup_{i \geq 1} A_i) \leq \sum_{i \geq 1} P(A_i)$;
5. $A_n \nearrow A \Longrightarrow P(A_n) \to P(A)$;
6. $A_n \searrow A \Longrightarrow P(A_n) \to P(A)$.

- 3. ist die Monotonie;
- 4. ist die σ-Subadditivität;
- 5. ist die Stetigkeit von unten;
- 6. ist die Stetigkeit von oben.

Beweis.

1. Wegen Voraussetzung (A) für Wahrscheinlichkeitsmasse:

$$P(\emptyset) = P(\emptyset \cup \emptyset \cup \cdots) = \sum_{i \geq 1} P(\emptyset) \Longrightarrow P(\emptyset) = 0.$$

2. Ähnlich wie in 1. lässt sich das leicht für disjunkte Ereignisse sehen. Im Allgemeinen

$$\begin{aligned} P(A \cup B) &= P(A \cup (B \cap A^c) \cup \emptyset \cup \emptyset \cup \cdots) \\ &= P(A) + P(B \cap A^c) + 0 + 0 + \cdots \\ &= P(A) + \underbrace{P(B \cap A^c) + P(B \cap A)}_{P(B)} - P(B \cap A). \end{aligned}$$

3.–6. Übung.