3 Grundlagen der Wahrscheinlichkeitstheorie

3.1 Desiderata

Wir haben gesehen dass wir für die Wahrscheinlichkeitstheorie 2 große Desiderata haben:

1. Wir wollen (aussagen)logische Operationen für Ereignisse; und
2. wir möchten die logischen Operationen *numerisch interpretieren*, d.h. in numerische Funktionen verwandeln.

3.2 Boolesche Algebren

Logische Operationen können wir in Booleschen Algebren interpretieren:

Definition 1 *Sei M eine Menge. Ein Mengensystem $\mathcal{M} \subseteq \wp(M)$ ist eine Boolesche Algebra über M, falls*

1. *$M \in \mathcal{M}, \emptyset \in \mathcal{M}$;*
2. *falls $N \in \mathcal{M}$, dann ist auch $\overline{N} := M - N \in \mathcal{M}$;*
3. *falls $N_1, N_2 \in \mathcal{M}$, dann sind auch $N_1 \cup N_2 \in \mathcal{M}$.*

NB: die Definition impliziert dass falls $N_1, N_2 \in \mathcal{M}$, dann ist auch $N_1 \cap N_2 \in \mathcal{M}$, da $N_1 \cap N_2 = \overline{\overline{N_1} \cup \overline{N_2}}$. Unsere Definition betrifft eigentlich nur einen Spezialfall von Booleschen Algebren, nämlich solchen über Mengensystemen. Allerdings kann jede endliche Boolesche Algebra auf diesen Spezialfall reduziert werden.

Übung: Mengenlehre und Partitionen

- $M \cap N = \overline{\overline{M} \cup \overline{N}}$ (Interdefinierbarkeit 1)
- $M \cup N = \overline{\overline{M} \cap \overline{N}}$ (Interdefinierbarkeit 2)
- $\overline{\overline{M}} = M$ (doppeltes Komplement)
- $(M \cup N) \cap O = (M \cap O) \cup (N \cap O)$ (de Morgan)

Eine Partition einer Menge M ist eine Menge $X \subseteq \wp(M)$, d.h. $X = \{N_1, ..., N_i\}$ (im endlichen Fall), und es gilt:

1. $N_1 \cup ... \cup N_i = M$
2. für alle $N_i, N_j \in X$, entweder $N_i = N_j$ oder $N_i \cap N_j = \emptyset$.

3.3 Einige Beobachtungen

Wir haben bereits gesagt, dass Wahrscheinlichkeiten Zahlen in $[0, 1]$ sind, wobei wir die Korrespondenz haben

$0 \cong$ Unmöglichkeit
$1 \cong$ Sicherheit

Nun haben wir, aus logischen Gründen folgendes:

(1) $P(A) \leq P(A \text{ oder } B)$ und $P(B) \leq P(A \text{ oder } B)$

(In Zukunft schreiben wir: $P(A \cup B)$). Das ist klar: wann immer A eintritt, tritt auch A oder B ein, also ist das Ereignis wahrscheinlicher etc. Ebenso klar ist:

(2) $P(A \text{ und } B) \leq P(A)$ und $P(A \text{ und } B) \leq P(B)$

(In Zukunft schreiben wir: $P(A \cap B)$ oder einfach $P(AB)$). Das ist klar: die Wahrscheinlichkeit, dass sie bei Ihrer nächsten Radfahrt angefahren werden ist größer als die, dass sie angefahren werden und im Zuge dessen 50euro finden.

Gleichzeitig haben wir folgendes: sei $\perp$ ein Ereignis, das vollkommen unmöglich ist, z.B. Sie würfeln (mit einem handelsüblichen Würfel) eine 7. Dann haben wir natürlich:

(3) $P(A \cap \perp) = 0$; $P(A \cup \perp) = P(A)$

Also, in Worten: $\perp$ ist *absorbierend für Konjunktion* und *neutral für Disjunktion.*

Umgekehrt, sei $\top$ ein Ereignis, dessen Eintritt sicher ist, z.B. dass Sie eine Zahl zwischen 1 und 6 würfeln. Dann haben wir

(4) $P(A \cap \top) = P(A)$; $P(A \cup \top) = 1$

Also gilt: $\top$ ist absorbierend für Disjunktion, und neutral für Konjunktion. Nun haben wir, nach Annahme:

(5) $P(\top) = 1;\ P(\bot) = 0$

Wir suchen also Operationen, für die 1, 0 jeweils neutral bzw. absorbierend sind. Das wird erfüllt von den Operationen + und $\cdot$:

(6) $n + 0 = n \qquad n \cdot 0 = 0$

Ebenso haben wir:

(7) $n \cdot m \leq n$ und $n \cdot m \leq m$, für $n, m \in [0, 1]$,

sowie:

(8) $n \cdot 1 = n \qquad n + 1 \geq 1$

sowie:

(9) $n \leq n + m$ und $m \leq n + m$, für $n, m \in [0, 1]$

Wir haben also folgende Korrespondenz:

Konjunktion $\cong \cdot$
Disjunktion $\cong +$

Das Problem ist, dass sich in dem einfachen Fall die Wahrscheinlichkeiten *nicht* auf 1 aufsummieren. Wir haben eine Korrespondenz, aber das ist noch zu einfach gedacht. Das sieht man auch an folgendem Beispiel:

(10) $P(A \cap A) = P(A) \neq P(A) \cdot P(A)$

sowie

(11) $P(A \cup A) = P(A) \neq P(A) + P(A)$

Konjunktion und Disjunktion sind also **idempotent**, im Gegensatz zur Addition und Multiplikation. Die Materie ist also durchaus komplex; es gibt allerdings eine wunderbar elegante Lösung, die uns mit allen nötigen Rechenregeln versorgt.

3.4 Definition von Wahrscheinlichkeitsräumen

Folgende Definition stammt von Kolmogorov, und ist das Ergebnis langer Überlegungen und Dispute.

Definition 2 *Ein Wahrscheinlichkeitsraum ist ein Tripel* $(\Omega, \mathfrak{A}, P)$*, wobei* $\mathfrak{A} \subseteq \wp(\Omega)$ *eine Boolesche Algebra ist, und* $P : \mathfrak{A} \to [0, 1]$ *eine Wahrscheinlichkeitsfunktion, so dass*

1. $P(\Omega) = 1$*;*

2. $P(\emptyset) = 0$*, und*

3. *falls* $A_1, A_2, ..., A_n$ *paarweise disjunkt sind, dann ist*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

Zur Erklärung: mit **paarweise disjunkt** meint man: für alle i, j so dass $1 \leq i, j \leq n$, falls $i \neq j$, dann ist $A_i \cap A_j = \emptyset$.

Die Bedingung der Booleschen Algebra ist wie folgt zu verstehen: falls $A, B \subseteq \Omega$ Ereignisse sind, die eine Wahrscheinlichkeit haben, dann haben auch die Ereignisse $A \cup B$ (d.h.: A oder B trifft ein), $A \cap B$ (d.h. beide A und B treffen ein) und $\overline{A}$ (d.h. A trifft nicht ein) eine Wahrscheinlichkeit.

3.5 Ereignisse und Ergebnisse

Wir nennen eine Menge $A \subseteq \Omega$ ein **Ereignis**; wir nennen $a \in \Omega$ ein **Ergebnis**. Meistens entspricht ein Ergebnis a einem Ereignis $\{a\}$. Aber nicht immer ist das intuitiv: nehmen wir an, wir würfeln mit zwei Würfeln, wobei unsere Ergebnisse die Form haben

$$\langle m, n\rangle$$

Nun ist "der erste Wurf ist eine 2" kein Ergebnis, sondern ein Ereignis, nämlich das Ereignis

$$\{\langle 2, 1\rangle, ..., \langle 2, 6\rangle\}$$

Daher weisen wir Wahrscheinlichkeiten normalerweise Ereignissen zu, nicht Ergebnissen.

3.6 Die Komplement-Regel

Wir kommen nun zu den Rechenregeln. Die Regel für die Berechnung des Komplementes $P(\overline{A})$ aus $P(A)$ lautet wie folgt:

(1) $P(\overline{A}) = 1 - P(A)$

Das lässt sich sehr einfach ableiten: wir haben

1. $P(A \cup \overline{A}) = P(\Omega) = 1$ und
2. $A \cap \overline{A} = \emptyset$;

also:

$$\begin{aligned} 1 &= P(A \cup \overline{A}) \\ &= P(A) + P(\overline{A}) \\ \Leftrightarrow 1 - P(A) &= P(\overline{A}) \end{aligned}$$

3.7 Die Summenregel

Die Summenregel erlaubt es uns, die logische Disjunktion rechnerisch aufzulösen. Die Summenregel lautet:

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

Intuitiv bedeutet das: um die Wahrscheinlichkeit einer Disjunktion zu berechnen, reicht es die Wahrscheinlichkeiten zu addieren, wenn man nur die Wahrscheinlichkeitsmasse abzieht, die auf die Konjunktion beider Ereignisse entfällt (illustrierbar mittels Venn-Diagramm). Das lässt sich wie folgt ableiten aus den Axiomen:

$$\begin{aligned} P(A \cup B) &= P(A \cup (B \cap \overline{A})) && \text{(Mengenlehre)} \\ &= P(A) + P(B \cap \overline{A}) && \text{(Disjunkte Mengen)} \\ &= P(A) + P(B \cap (\overline{A} \cup \overline{B})) && \text{(Mengenlehre)} \\ &= P(A) + P(B \cap (\overline{A \cap B})) && \text{(Mengenlehre)} \\ &= P(A) + P(\overline{\overline{B \cap (\overline{A \cap B})}}) && \text{(Mengenlehre)} \\ &= P(A) + P(\overline{\overline{B} \cup (A \cap B)}) && \text{(Mengenlehre)} \\ &= P(A) + (1 - P(\overline{B} \cup (A \cap B))) && \text{(Subtraktionsregel)} \end{aligned}$$

$$
\begin{aligned}
&= P(A) + (1 - P(\overline{B}) + P(A \cap B)) && \text{(Disjunkte Mengen)}\\
&= P(A) + (1 - (1 - P(B)) + P(A \cap B)) && \text{(Disjunkte Mengen)}\\
&= P(A) + (1 - (1 - P(B))) - P(A \cap B)) && \text{(Arithmetik)}\\
&= P(A) + (1 - 1) + P(B) - P(A \cap B)) && \text{(Arithmetik)}\\
&= P(A) + P(B) - P(A \cap B)) && \text{(Arithmetik)}
\end{aligned}
$$

3.8 Die Produktregel

Um die Konjunktion sinnvoll zu interpretieren, brauchen wir die Definition der *bedingten Wahrscheinlichkeit*. Wir definieren

(3) $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Nehmen Sie nun an wir suchen die Wahrscheinlichkeit $P(A \cap B)$; wir bekommen sie durch eine ganz simple Termumformung:

(4) $$P(A \cap B) = P(A|B)P(B)$$

Da $P(A \cap B) = P(B \cap A)$ ($\cap$ ist kommutativ), bekommen wir also die Produktregel:

(5) $$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Intuitiv ist das wie folgt zu verstehen: wenn A und B eintreffen, dann bedeutet dass das 1. A eintrifft und 2. unter dieser Voraussetzung B eintrifft (oder umgekehrt). Wichtig ist: $P(A|B)$ sagt nichts über zeitliche Reihenfolge! So ist die Formel intuitiv richtig. Wir werden später noch mehr zum Konzept der bedingten Wahrscheinlichkeit erfahren.

Diese Umformung mag einem auf Anhieb nicht sehr hilfreich erscheinen, allerdings ist sie eines der zentralen Gesetze. Denn in der Praxis kennen wir oft bedingte Wahrscheinlichkeiten besser als unbedingte, so dass wir uns das Gesetz leicht zunutze machen können. Es gibt übrigens noch eine allgemeinere Form der Produktregel:

(6) $$P(A \cap B|X) = P(A|BX)P(B|X) = P(B|AX)P(A|X)$$

Das generalisiert die letzte Formel, da X beliebig (auch leer) sein kann.