7.6 Grundlagen der Wahrscheinlichkeitstheorie

Die Stochastik ist paradigmatisch für den *Modellbildungskreislauf*. Hier sollen vom Zufall abhängige Vorgänge (wie z.B. Glücksspiele) modelliert werden. Der Zufall soll mit dem Begriff der Wahrscheinlichkeit erfasst werden. Als erste Näherung kann man den *frequentistischen Wahrscheinlichkeitsbegriff* heranziehen, der auf Richard von Mises (1883-1953) zurückgeht. Ihm liegt die Erfahrungstatsache zugrunde, dass sich die relative Häufigkeit eines Ergebnisses in einem vom Zufall abhängigen Experiment mit wachsender Anzahl an Wiederholungen des Experimentes stabilisiert (*Empirisches Gesetz großer Zahlen*). Den Grenzwert, gegen den die relative Häufigkeit mit wachsender Wiederholungszahl strebt, interpretiert man dann als *Wahrscheinlichkeit* des Ergebnisses. Dieser Erfahrungstatsache folgend kann man relative Häufigkeiten als Schätzwerte für Wahrscheinlichkeiten nehmen. Eine rigorose Definition des Wahrscheinlichkeitsbegriffs gelang von Mises nicht, da eine mathematische Präzisierung des Stabilisierens der relativen Häufigkeiten fehlschlug. Aus diesem Grund wählt man heute den auf Andrei Nikolajewitsch Kolmogorow (1903-1983) zurückgehenden *axiomatischen* Zugang zur Wahrscheinlichkeitstheorie.

Thema der Stochastik sind Ereignisse (und die Wahrscheinlichkeit ihres Eintretens. Unsere Ausgangsfragen sind:

(i) Was kann passieren?

(ii) Mit welcher Wahrscheinlichkeit treten diese oder jene Ereignisse ein?

Unabhängig von der konkreten betrachteten Situation wird die *Menge aller möglichen Ereignisse* (also die Antwort auf die erste Frage) in der Wahrscheinlichkeitstheorie mit Ω bezeichnet.

Beispiele 7.6.1. Wir betrachten einige Situationen und modellieren das zugehörige Ω:

(i) *Ein Münzwurf.* Eine Münze wird geworfen, und als Ereignis tritt entweder "Kopf" oder "Zahl" ein. Wenn wir zur Vereinfachung das eine mit "0" notieren und das andere mit "1", ist die Menge aller möglichen Ereignisse in diesem Fall $\Omega = \{0, 1\}$.

(ii) *Würfelwurf.* $\Omega = \{1, 2, 3, 4, 5, 6\}$.

(iii) *Mehrere Münzwürfe.* In dieser Situation fassen wir jede n-elementige Folge von Nullen und Einsen (alias Köpfen und Zahlen) als ein Ereignis auf; es ist also

$$\Omega = \big\{(x_1, x_2, \dots, x_n) \;\big|\; x_i \in \{0, 1\}\big\} = \{0, 1\}^n$$

die Menge aller möglichen Ereignisse. Beim zweifachen Münzwurf also

$$\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$$

D.h. das Element $(2, 3) \in \Omega$ bedeutet: erst wurde eine 2 und dann eine 3 gewürfelt.

(iv) *Unendlich viele Münzwürfe.* Hier gilt

$$\Omega = \{(x_i)_{i\in\mathbb{N}} \,|x_i \in \{0, 1\}\} \,.$$

Im Gegensatz zum einfachen oder n-fachen Münzwurf ist Ω hier nicht endlich.

(v) *Zufallszahl zwischen* 0 *und* 1. $\Omega = [0, 1]$.

Die Fokussierung auf spezielle Fragestellungen erfolgt durch die Betrachtung sogenannter *Ereignisse.* Ein Ereignis A ist eine Teilmenge des Grundraums Ω, also $A \subset \Omega$.

Ein *Elementarereignis* ist ein Element $\omega \in \Omega$.

Man sagt „das Ereignis A tritt ein", wenn ein ω mit $\omega \in A$ beobachtet wird.

Da im Allgemeinen nicht jede Teilmenge von Ω als Ereignis sinnvoll sein muss (z.B. im Modell für den unendlichen Münzwurf), werden *gewisse* Teilmengen als Ereignisse ausgezeichnet, indem sie in einem *Ereignis-System*

$$\mathcal{A} := \{A \subseteq \Omega \,|\, A \text{ ist ein Ereignis } \} \subseteq \mathcal{P}(\Omega).$$

aufgelistet werden. Hier ist $\mathcal{P}(\Omega)$ die Potenzmenge von Ω. Wir verlangen, dass ein Ereignissystem $\mathcal{A}$ *abgeschlossen* gegenüber allen abzählbaren Mengenoperationen (etwa Durchschnitt, Vereinigung, Komplement, Mengedifferenzen,...) ist, d.h. führt man diese Operationen mit Ereignissen aus $\mathcal{A}$ durch, so erhält man wieder Ereignisse aus $\mathcal{A}$.

In den meisten Fällen wählen wir $\mathcal{A} = \mathcal{P}(\Omega)$.

Den Zusammenhang zwischen der Beschreibung von Ereignissen in der realen Welt und im Modell zeigt die folgende Tabelle:

Reale Welt	**Modell**
A und B treten (gleichzeitig) ein	$\omega \in A \cap B$
A oder B tritt ein (oder beide)	$\omega \in A \cup B$
Entweder A oder B tritt ein	$\omega \in (A \setminus B) \cup (B \setminus A)$
B tritt ein, A aber nicht	$\omega \in B \setminus A$
A tritt nicht ein	$\omega \in A^C$
A und B schließen sich aus	$A \cap B = \emptyset$ (A und B sind *disjunkt.*)
Mindestens ein A_i tritt ein	$\omega \in \bigcup_{i \geq 1} A_i$
Alle A_i treten ein	$\omega \in \bigcap_{i \geq 1} A_i$

Beispiele 7.6.2. Wir betrachten Situationen aus dem obigen Beispiel.

(i) *Ein Münzwurf.* "1 tritt ein": $A = \{1\}$.

(ii) *Würfelwurf.* $\Omega = \{1, 2, 3, 4, 5, 6\}$. Es wird eine gerade Zahl gewürfelt $A = \{2, 4, 6\}$.

(iii) *Mehrere Münzwürfe.* "Es treten genau k Einsen auf":

$$A = \Big\{(x_1, \dots, x_n) \in \{0,1\}^n \;\Big|\; \sum_{i=1}^{n} x_i = k\Big\}.$$

(iv) Unendlich viele Münzwürfe. "Die relative Häufigkeit der 1 ist p":

$$A = \{(x_i)_{i\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}} | \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = p\}.$$

(v) *Zufallszahl zwischen* 0 *und* 1. "Es tritt eine Zahl aus $[a, b]$ auf": $A = [a, b] \subseteq \Omega = [0, 1]$.

7.7 Zufallsexperimente und Wahrscheinlichkeitsräume

Definition 7.7.1. Ein Paar (Ω, p) heißt *Zufallsexperiment*, falls

(i) Ω eine *abzählbare* (d.h. endliche oder abzählbar unendliche), nicht-leere Menge (*Ergebnismenge, Grundraum, Stichprobenraum*) und

(ii) $p : \Omega \to [0, 1]$ eine Funktion mit

$$\sum_{\omega\in\Omega} p(\omega) = 1 \qquad \textit{(Wahrscheinlichkeitsfunktion, Wahrscheinlichkeitsdichte)}$$

ist.

Ein Element $\omega \in \Omega$ heißt *Ergebnis bzw.Elementarereignis* und $p(\omega)$ heißt *Wahrscheinlichkeit* des Ergebnisses bzw. desElementarereignisses ω.

Beispiele 7.7.2.

Definition 7.7.3. Ein Tripel $(\Omega, \mathcal{A}, P)$ heißt *Wahrscheinlichkeitsraum*, falls gilt:

(1) Ω ist eine nicht-leere Menge (*Ergebnismenge*)

(2) $\mathcal{A}$ ist ein Ereignis-System (bei uns in der Regel $\mathcal{A} = \mathcal{P}(\Omega)$)

(3) P ist ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$, d.h. eine Funktion $P : \mathcal{A} \to [0, 1]$ mit folgenden Eigenschaften:

(N) *Normiertheit*: $P(\Omega) = 1$

(A) *σ-Additivität*:
Für paarweise disjunkte Ereignisse $A_1, A_2, ... \in \mathcal{A}$ (d.h. $A_i \cap A_j = \emptyset$ für $i \neq j$) gilt:

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Man sagt, dass $P(A)$ die *Wahrscheinlichkeit* von A ist (bzw. die Wahrscheinlichkeit ist, dass A eintritt).

Satz 7.7.4. *Für ein Zufallsexperiment (Ω, p) wird durch $\mathcal{A} := \mathcal{P}(\Omega)$ und*

$$P(A) := \sum_{\omega\in A} p(\omega) \qquad \text{für} \quad A \in \mathcal{A}$$

ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ definiert.

Satz 7.7.5. *(Eigenschaften von Wahrscheinlichkeitsmaßen) Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Dann gelten:*

(1) Endliche Additivität*: Für paarweise disjunkte Ereignisse $A_1, A_2, ...A_n \in \mathcal{A}$ gilt:*

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$$

(2) $P(\emptyset) = 0$

(3) Für alle $A \in \mathcal{A}$ gilt: $P(A^C) = 1 - P(A)$

(4) Für alle $A, B \in \mathcal{A}$ gilt: $P(B \setminus A) = P(B) - P(A \cap B)$

(5) Für alle $A, B \in \mathcal{A}$ gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(6) Für alle $A, B \in \mathcal{A}$ gilt: $A \subset B \quad \Rightarrow \quad P(A) \leq P(B)$

(7) Endliche Subadditivität*: Für $A_1, ..., A_n \in \mathcal{A}$ gilt:*

$$P(A_1 \cup ... \cup A_n) \leq P(A_1) + ... + P(A_n)$$

(8) Subadditiviät*: Für $A_1, A_2, ... \in \mathcal{A}$ gilt: $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$*

Definition 7.7.6. Ein Zufallsexperiment (Ω, p) heißt *Laplace-Versuch* (oder *Laplace-Experiment*), wenn Ω endlich und $p : \Omega \to [0,1]$ die sogenannte *Gleichverteilung* (oder *Laplace'sche Wahrscheinlichkeit(sfunktion)*) ist, die jedem Ergebnis diesselbe Wahrscheinlichkeit, nämlich

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

zuordnet, wobei $|\Omega|$ gleich der Anzahl der Elemente in Ω bezeichnet. Der zugehörige Wahrscheinlichkeitsraum $(\Omega, \mathcal{P}(\Omega), P)$ heißt *Laplace'scher Wahrscheinlichkeitsraum* (oder *Laplace-Modell*). Im Laplace-Modell gilt für ein Ereignis $A \subset \Omega$:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der für } A \text{ günstigen Ergebnisse}}{\text{Anzahl aller möglichen Ergebnisse}}$$

In Laplace-Modellen ermittelt man Wahrscheinlichkeiten häufig durch geschicktes Abzählen.

Einschub 7.7.7. ...

7.8 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Offenbar ist es möglich, dass durch das Eintreten eines Ereignisses B die Wahrscheinlichkeit eines Ereignisses A verändert werden kann. Gesucht ist in diesem Zusammenhang die Wahrscheinlichkeit von A unter der Bedingung, dass B eingetreten ist. Um dies zu modellieren führen wir bedingte Wahrscheinlichkeiten ein. Die bedingte Wahrscheinlichkeit kann als Neueinschätzung der Wahrscheinlichkeit von A interpretiert werden, wenn die Information vorliegt, dass das Ereignis B bereits eingetreten ist.

Definition 7.8.1. Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $B \in \mathcal{A}$ mit $P(B) > 0$. Dann heißt

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

die *bedingte Wahrscheinlichkeit von A gegeben B.*

Beispiele 7.8.2. Fairer Würfelwurf.

Bedingte Wahrscheinlichkeiten kann man wie folgt mithilfe von *Wahrscheinlichkeitsbäumen* visualisieren:

Einschub 7.8.3. ...

Für das Rechnen mit bedingten Wahrscheinlichkeiten gelten die folgenden Sätze:

Satz 7.8.4. *(Multiplikations- oder Pfadregel)*
Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A_1, ..., A_n \in \mathcal{A}$ mit $P(A_1 \cap ... \cap A_{n-1}) > 0$. Dann gilt:

$$P(A_1 \cap ... \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot ... \cdot P(A_n|A_1 \cap ... \cap A_{n-1}).$$

Einschub 7.8.5. ...

Satz 7.8.6. *(von der totalen Wahrscheinlichkeit)*
Sei $(\Omega, \mathcal{A}, P)$ ein W.raum, $A \in \mathcal{A}$ und $B_i \in \mathcal{A}$, $i \geq 1$, eine Partition *von Ω, d.h. es gelten:*

(i) $B_i \cap B_j = \emptyset$ *für alle* $i \neq j$,

(ii) $\bigcup_{i \geq 1} B_i = \Omega$.

Ist weiterhin $P(B_i) > 0$ für alle $i \geq 1$, so gilt:

$$P(A) = \sum_{i \geq 1} P(A|B_i) \cdot P(B_i).$$

Einschub 7.8.7. ...

Beispiele 7.8.8. Satz von der totalen W.keit und Pfadregel im Spezialfall $\Omega = B \cup B^C$.

Beispiele 7.8.9. In einer Urne befinden sich zwei blaue und zwei rote Kugeln. Es wird zweimal aus gezogen. Wird im ersten Zug eine rote Kugel gezogen, so wird sie wieder zurückgelegt. Dann wird ein zweites Mal gezogen. Wird im ersten Zug eine blaue Kugel gezogen, so wird sie nicht zurückgelegt und dann ein zweites Mal gezogen.

Satz 7.8.10. *(von Bayes) Sei* $(\Omega, \mathcal{A}, P)$ *ein W.raum und* $B_i \in \mathcal{A}$, $i \geq 1$ *eine Partition von* Ω *mit* $P(B_i) > 0$ *für alle* $i \geq 1$. *Dann gilt für alle* $j = 1, 2, ...$ *und alle* $A \in \mathcal{A}$*:*

$$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i \geq 1} P(A|B_i) \cdot P(B_i)}.$$

Bemerkung 7.8.11. Speziell für die Partition (B, B^C) bedeutet der Satz von Bayes:

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A|B) \cdot P(B) + P(A|B^C) \cdot P(B^C)}$$

Beispiele 7.8.12. In Deutschland sei eine Person von 1000 an einer bestimmten Krankheit erkrankt. Zur Diagnose der Krankheit wird ein Test verwendet, der in 99% der Fälle, in denen die Krankheit vorliegt, ein positives Testergebnis liefert. Bei einer nicht-erkrankten Person gibt der Test allerdings in 2% der Fälle fälschlicherweise ein positives Testergebnis aus.
Mit welcher Wahrscheinlichkeit ist eine Person, die ein positives Testergebnis erhalten hat, tatsächlich krank?