Einführung in die Stochastik

Klaus Ritter

Darmstadt, SS 2008

Literatur

Insbesondere

H.-O. Georgii, *Stochastik*, 3. Auflage, de Gruyter, Berlin, 2007.

A. Irle, *Wahrscheinlichkeitstheorie und Statistik*, 2. Auflage, Teubner, Stuttgart, 2005.

U. Krengel, *Einführung in die Wahrscheinlichkeitstheorie und Statistik*, 8. Auflage, Vieweg, Braunschweig, 2005.

Inhaltsverzeichnis

Kapitel I

Grundbegriffe

Ein mathematisches Modell für ein zufälliges Phänomen besteht aus einem Wahrscheinlichkeitsraum und darauf definierten Zufallsvariablen.
Themen dieser einführenden Vorlesung sind

- allgemeine Begriffsbildung und Konstruktionen,
- Analyse und Simulation stochastischer Modelle,
- Analyse empirischer Daten.

1 Wahrscheinlichkeitsräume

Die möglichen *Ergebnisse* ω eines Zufallsexperimentes bilden den *Ergebnisraum* Ω. Im allgemeinen ist Ω eine nicht-leere Menge ohne weitere Struktur.

Beispiel 1.

(i) Würfeln, $\Omega := \{1, \ldots, 6\}$.

(ii) Anzahl Anrufe in Callcenter, $\Omega := \mathbb{N}_0$.

(iii) Wartezeit bei Anruf, $\Omega := \mathbb{R}_+ := [0, \infty[$.

(iv) Verlauf eines Aktienkurses über die Zeit von 0 bis T, $\Omega := C_+([0, T])$ Menge der nicht-negativen stetigen Funktionen auf $[0, T]$.

Für gewisse Teilmengen $A \subseteq \Omega$, genannt *Ereignisse*, definiert man die Wahrscheinlichkeit ihres Eintretens ($\omega \in A$). Die Menge $\mathfrak{A}$ aller Ereignisse in Ω bilden den *Ereignisraum*.

Beispiel 2.

(i) Würfeln, Ergebnis ist gerade Zahl, $A := \{2, 4, 6\}$.

(ii) Anzahl Anrufe in Callcenter, Kapazitätsgrenze K wird nicht überschritten, $A := \{0, \ldots, K\}$.

(iii) Wartezeit bei Anruf, Wartezeit liegt zwischen 1 und 2 (Minuten), $A := [1, 2]$.

(iv) Verlauf eines Aktienkurses, Kurs weicht von Anfangswert um nicht mehr als 10 (Euro) ab,
$$A := \{\omega \in C_+([0, T]) : \sup_{0 \le t \le T} |\omega(0) - \omega(t)| \le 10\}.$$

Nun: mengentheoretische Operationen mit Ereignissen.

Beispiel 3.

(i) Ereignis A oder Ereignis B tritt ein, $A \cup B$.

(ii) Ereignis A und Ereignis B treten ein, $A \cap B$.

(iii) Ereignis A tritt nicht ein, $A^{\mathrm{c}} := \Omega \setminus A$.

(iv) (mindestens) eines der Ereignisse $A_1, A_2, \ldots$ tritt ein, $\bigcup_{i=1}^{\infty} A_i$.

(v) alle Ereignisse $A_1, A_2, \ldots$ treten ein, $\bigcap_{i=1}^{\infty} A_i$.

Forderung: obige Operationen liefern wieder Ereignisse (Abschlußeigenschaft). Dazu führt man den Begriff der σ-Algebra ein.

Bezeichnung. Die *Potenzmenge* von Ω (Menge aller Teilmengen von Ω) wird mit $\mathfrak{P}(\Omega)$ und die *Mächtigkeit* (Anzahl der Elemente) einer endlichen Menge U wird mit $|U|$ bezeichnet.

Beispiel 4. Münzwurf, $\Omega := \{\mathrm{Z}, \mathrm{K}\}$,
$$\mathfrak{P}(\Omega) = \{\emptyset, \{\mathrm{Z}\}, \{\mathrm{K}\}, \{\mathrm{Z}, \mathrm{K}\}\}.$$

Beachte $\mathrm{Z} \notin \mathfrak{P}(\Omega)$, aber $\{\mathrm{Z}\} \in \mathfrak{P}(\Omega)$.

Satz 5. Für endliche Mengen Ω gilt
$$|\mathfrak{P}(\Omega)| = 2^{|\Omega|}.$$

Beweis. Durch Induktion über $n := |\Omega|$.
Verankerung: Gelte $|\Omega| = 0$. Dann $\Omega = \emptyset$ und $|\mathfrak{P}(\Omega)| = |\{\emptyset\}| = 1$.
Induktionsschritt: Gelte $|\Omega| = n + 1 \geq 1$ und sei die Behauptung für Mengen der Mächtigkeit n bereits bewiesen. Fixiere $\omega^* \in \Omega$. Dann

$$\begin{aligned} |\mathfrak{P}(\Omega)| &= |\{A \subseteq \Omega : \omega^* \in A\}| + |\{A \subseteq \Omega : \omega^* \notin A\}| \\ &= 2^n + 2^n = 2^{n+1}. \end{aligned}$$

□

Definition 6. $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ heißt σ-*Algebra* (in Ω), falls

(i) $\Omega \in \mathfrak{A}$,

(ii) $A \in \mathfrak{A} \Rightarrow A^{\mathrm{c}} \in \mathfrak{A}$,

(iii) $A_1, A_2, \ldots \in \mathfrak{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathfrak{A}$.

Forderung: der Ereignisraum ist eine σ-Algebra.

Bemerkung 7. In der Regel betrachtet man $\mathfrak{A} := \mathfrak{P}(\Omega)$, falls Ω abzählbar ist. Dies ist im allgemeinen nicht mehr möglich, falls Ω überabzählbar ist, siehe Kapitel IV.

Lemma 8. Für jede σ-Algebra $\mathfrak{A}$ gilt

(i) $\emptyset \in \mathfrak{A}$,

(ii) $A, B \in \mathfrak{A} \Rightarrow A \cup B,\ A \cap B, A \setminus B \in \mathfrak{A}$,

(iii) $A_1, A_2, \ldots \in \mathfrak{A} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathfrak{A}$.

Beweis. Ad (i): Es gilt $\emptyset = \Omega^{\mathrm{c}} \in \mathfrak{A}$ nach Definition 6.(i), 6.(ii).
Ad (ii): Seien $A, B \in \mathfrak{A}$. Es gilt

$$A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \ldots \in \mathfrak{A}$$

nach (i) und Definition 6.(iii). Es gilt $A^{\mathrm{c}} \cup B^{\mathrm{c}} \in \mathfrak{A}$ nach (ii) und Definition 6.(ii) und somit auch

$$A \cap B = (A^{\mathrm{c}} \cup B^{\mathrm{c}})^{\mathrm{c}} \in \mathfrak{A}.$$

Man verwende $A \setminus B = A \cap B^{\mathrm{c}}$ um $A \setminus B \in \mathfrak{A}$ zu zeigen.
Ad (iii): Für $A_1, A_2, \ldots \in \mathfrak{A}$ gilt

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^{\mathrm{c}}\right)^{\mathrm{c}} \in \mathfrak{A}$$

nach Definition 6.(ii), 6.(iii). □

Bezeichnung. Mengen $A_1, A_2, \dots$ heißen *paarweise disjunkt* (p.d.), falls $A_i \cap A_j = \emptyset$ für $i \neq j$.

Im folgenden sei $\mathfrak{A}$ eine σ-Algebra in einer nicht-leeren Menge Ω.
Nun: Zuordung von Wahrscheinlichkeiten $P(A)$ zu den einzelnen Ereignissen $A \in \mathfrak{A}$. Dabei leitet uns folgende Vorstellung: Bei einer „großen" Anzahl von „unabhängigen" Wiederholungen des Zufallsexperimentes liegt die relative Häufigkeit des Eintretens von Ereignis A „nahe" bei $P(A)$.

Definition 9. $P : \mathfrak{A} \to [0, 1]$ heißt *Wahrscheinlichkeitsmaß* oder *Wahrscheinlichkeitsverteilung* (auf $\mathfrak{A}$), falls

(i) $P(\Omega) = 1$,

(ii) $A_1, A_2, \dots \in \mathfrak{A}$ p.d. $\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ (σ-*Additivität*).
Genauer: $\dots \Rightarrow \sum_{i=1}^{\infty} P(A_i)$ (absolut) konvergent und $\dots$

Beispiel 10. Die *Gleichverteilung* auf einer endlichen Menge Ω wird definiert durch (*Laplace-Annahme*)

$$P(A) := |A|/|\Omega|, \qquad A \subseteq \Omega.$$

Speziell gilt für alle $\omega \in \Omega$

$$P(\{\omega\}) = 1/|\Omega|.$$

Beh.: P ist Wahrscheinlichkeitsmaß auf $\mathfrak{A} := \mathfrak{P}(\Omega)$.

Beweis. Offensichtlich gilt $0 \leq P(A) \leq 1$ und $P(\Omega) = 1$.
Für $A_1, A_2, \dots \subseteq \Omega$ p.d. (notwendig: $A_i = \emptyset$ bis auf endlich viele i) gilt

$$\left|\bigcup_{i=1}^{\infty} A_i\right| = \sum_{i=1}^{\infty} |A_i|\,.$$

Dies zeigt die σ-Additivität. □

Definition 11. $(\Omega, \mathfrak{A}, P)$ heißt *Wahrscheinlichkeitsraum*, falls Ω eine nicht-leere Menge, $\mathfrak{A}$ eine σ-Algebra in Ω und P Wahrscheinlichkeitsmaß auf $\mathfrak{A}$ ist.

Beispiel 12. Ein stochastisches Modell für einmaliges Würfeln wird definiert durch

(i) $\Omega := \{1, \dots, 6\}$,

(ii) $\mathfrak{A} := \mathfrak{P}(\Omega)$,

(iii) P Gleichverteilung auf Ω.

Beispiel 13. Ein stochastisches Modell für Geschlecht eines Neugeborenen wird definiert durch

(i) $\Omega := \{\mathrm{W}, \mathrm{M}\}$,

(ii) $\mathfrak{A} := \mathfrak{P}(\Omega)$,

(iii) P definiert durch $P(\{\mathrm{W}\}) := 0.4863$.

Letzteres wurde empirisch ermittelt als relative Häufigkeit unter den 25 171 123 Lebendgeburten in D in den Jahren 1970–1999. Siehe Hesse (2003, p. 23).

Beispiel 14. Hard core model der Physik, Gleichverteilung auf einer sehr großen Menge „unbekannter“ Mächtigkeit.
Betrachte ein Gitter

$$G = \{1, \ldots, m\}^d$$

mit $d = 3$. Mit Ω_0 bezeichnen wir die Menge aller Abbildungen (Konfigurationen)

$$\varphi : G \to \{0, 1\}.$$

Interpretation: $\varphi(x) = 1$ gdw. der Gitterpunkt $x \in G$ besetzt ist. Es gilt $|\Omega_0| = 2^{(m^d)}$.
Zwei Gitterpunkte $x, y \in G$ heißen benachbart, falls sie sich in genau einer Koordinate unterscheiden, d.h., falls $\sum_{i=1}^{d} |x_i - y_i| = 1$. Eine Konfiguration $\varphi \in \Omega_0$ heißt zulässig, falls benachbarte Gitterpunkte $x, y \in G$ nicht zugleich besetzt sind, also $\varphi(x) \cdot \varphi(y) = 0$ erfüllen.
Man studiert die Gleichverteilung auf der Menge Ω der zulässigen Konfigurationen. Siehe Häggström (2002).

Beispiel 15. Ein (fragwürdiges) stochastisches Modell für Pfeiltreffer auf Dartscheibe mit Radius $r > 0$ wird definiert durch

(i) $\Omega := \{\underbrace{(x, y)}_{=:\omega} \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}$,

(ii) eine „σ-Algebra $\mathfrak{A}$ in Ω, deren Elementen ein ‘Flächeninhalt’ $\lambda(A)$ zugeordnet werden kann“, siehe Kapitel IV.1 und Kapitel IV.2,

(iii) $P(A) := \lambda(A)/(\pi r^2)$.

Beachte: Es gilt $P(\{\omega\}) = 0$ für alle $\omega \in \Omega$.

Hilfreiche Intuition zu Wahrscheinlichkeiten: relative Anzahlen, siehe Beispiel 10, oder normierte Flächeninhalte, siehe Beispiel 15.
Im folgenden bezeichne $(\Omega, \mathfrak{A}, P)$ stets einen Wahrscheinlichkeitsraum.
Nun: Eigenschaften von Wahrscheinlichkeitsmaßen.

Satz 16. Für $A, B \in \mathfrak{A}$ gilt

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$ (*Additivität*),

(ii) $A \subseteq B \Rightarrow P(B) = P(A) + P(B \setminus A)$,

(iii) $A \subseteq B \Rightarrow P(A) \leq P(B)$ (*Monotonie*),

(iv) $P(A^{\mathrm{c}}) = 1 - P(A)$,

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

Beweis. Seien $A, B \in \mathfrak{A}$.
Ad(i): Im Fall $A \cap B = \emptyset$ gilt

$$A \cup B = A \cup B \cup \emptyset \cup \emptyset \ldots$$

mit p.d. Mengen und somit

$$P(A \cup B) = P(A) + P(B) + \sum_{i=1}^{\infty} P(\emptyset).$$

Da $P(A \cup B) < \infty$, folgt $P(\emptyset) = 0$ und weiter die Behauptung.
Ad (ii): Im Fall $A \subseteq B$ gilt

$$B = A \cup (B \setminus A).$$

mit disjunkten Mengen A und $B \setminus A$. Verwende (i).
Ad (iii): Verwende (ii) und $P(B \setminus A) \geq 0$.
Ad (iv): Verwende (ii) mit $B = \Omega$, also $P(B) = 1$.
Ad (v): Verwende

$$A \cup B = A \cup (B \setminus A) = A \cup \big(B \setminus (A \cap B)\big)$$

zusammen mit (i) und (ii). □

Satz 17. Seien $A_1, A_2, \ldots \in \mathfrak{A}$. Dann gilt

(i) $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i) \in [0, \infty]$ (*σ-Subadditivität*),

(ii) aus $A_1 \subseteq A_2 \subseteq \ldots$ folgt $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n)$ (*σ-Stetigkeit von unten*),

(iii) aus $A_1 \supseteq A_2 \supseteq \ldots$ folgt $P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n)$ (*σ-Stetigkeit von oben*).

Beweis. Seien $A_1, A_2, \ldots \in \mathfrak{A}$. Setze $A_0 := \emptyset$ und $B_i = A_i \setminus \left(\bigcup_{j=0}^{i-1} A_j\right)$ für $i \geq 1$. Lemma 8 sichert $B_i \in \mathfrak{A}$. Beachte, daß $\bigcup_{i=1}^\infty A_i = \bigcup_{i=1}^\infty B_i$ mit p.d. Mengen B_i.
Ad (i): Mit der σ-Additivität und der Monotonie von P folgt

$$P\left(\bigcup_{i=1}^\infty A_i\right) = P\left(\bigcup_{i=1}^\infty B_i\right) = \sum_{i=1}^\infty P(B_i) \leq \sum_{i=1}^\infty P(A_i). \tag{1}$$

Ad (ii): Aus $A_1 \subseteq A_2 \subseteq \ldots$ folgt $P(B_i) = P(A_i) - P(A_{i-1})$, siehe Satz 16.(ii), und weiter

$$\sum_{i=1}^n P(B_i) = P(A_n).$$

Verwende (1) um

$$P\left(\bigcup_{i=1}^\infty A_i\right) = \lim_{n\to\infty}\left(\sum_{i=1}^n P(B_i)\right) = \lim_{n\to\infty} P(A_n)$$

zu erhalten.
Ad (iii): Verwende (ii) und Komplementbildung. □

2 Bedingte Wahrscheinlichkeiten und Unabhängigkeit

Gegeben sei ein Wahrscheinlichkeitsmaß P auf einer σ-Algebra $\mathfrak{A}$. Falls bekannt ist, daß ein Ereignis $B \in \mathfrak{A}$ eingetreten ist, kann man zu einem neuen Wahrscheinlichkeitsmaß auf $\mathfrak{A}$ übergehen. Idee: „Restriktion" auf B und Normierung.

Definition 1. Für $A, B \in \mathfrak{A}$ mit $P(B) > 0$ heißt

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

bedingte Wahrscheinlichkeit von A gegeben B.

Bemerkung 2. $P(\cdot \mid B)$ ist ein Wahrscheinlichkeitsmaß auf $\mathfrak{A}$ mit $P(B \mid B) = 1$.

Beispiel 3. Für die Gleichverteilung P auf einer endlichen Menge Ω und $\emptyset \neq B \subseteq \Omega$ sowie $A \subseteq \Omega$ gilt

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

d.h. $P(A \mid B)$ ist der relative Anteil von Elementen aus A in B. Somit ist $P(\cdot \mid B)$ die „Gleichverteilung" auf B. Analog in Beispiel 1.15.

Speziell: einmaliges Würfeln (wie üblich modelliert) und $B := \{1, 5, 6\}$. Dann

$$P(\{\omega\} \mid B) = \begin{cases} 1/3, & \text{falls } \omega \in B \\ 0, & \text{falls } \omega \notin B. \end{cases}$$

Beispiel 4. Betrachte zwei weiße (1, 2) und drei schwarze (3, 4, 5) Kugeln, ziehe zwei Kugeln ohne Zurücklegen. Gesucht ist die Wahrscheinlichkeit dafür, daß die 2. Kugel schwarz ist, falls die 1. Kugel weiß war.
Modell: Gleichverteilung auf

$$\Omega := \{(\omega_1, \omega_2) \in \{1, \dots, 5\}^2 : \omega_1 \neq \omega_2\}.$$

Für

$$A := \{(\omega_1, \omega_2) \in \Omega : \omega_2 \geq 3\},$$
$$B := \{(\omega_1, \omega_2) \in \Omega : \omega_1 \leq 2\}$$

gilt (wie erwartet)

$$P(A \mid B) = \frac{|A \cap B|}{|B|} = \frac{6}{8} = \frac{3}{4}.$$

Satz 5. Für p.d. Mengen $B_1, \dots, B_n \in \mathfrak{A}$ gelte

$$\forall i: \ P(B_i) > 0$$

und

$$\bigcup_{i=1}^{n} B_i = \Omega.$$

Dann folgt für jedes $A \in \mathfrak{A}$ (Formel von der totalen Wahrscheinlichkeit)

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i).$$

Falls zusätzlich $P(A) > 0$ gilt, so folgt (Formel von Bayes)

$$P(B_i \mid A) = \frac{P(A \mid B_i) \cdot P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j) \cdot P(B_j)}.$$

Analoge Aussagen gelten für abzählbar viele Mengen B_i.

Beweis. Zur totalen Wahrscheinlichkeit: Es gilt

$$A = \bigcup_{i=1}^{n} (A \cap B_i)$$

mit p.d. Mengen $A \cap B_i$. Somit folgt

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i).$$

Zur Formel von Bayes: Es gilt für jedes $i \in \{1, \dots, n\}$

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{P(A)} \cdot \frac{P(B_i)}{P(B_i)} = \frac{P(A \mid B_i) \cdot P(B_i)}{P(A)}.$$

Die Behauptung ergibt sich nun mit Hilfe der Formel von der totalen Wahrscheinlichkeit. □

Beispiel 6. Situation

- 3 Maschinen, $i = 1, 2, 3$,
- Anteil an Tagesproduktion, $r_i = 60\%, 30\%, 10\%$,
- Anteil defekter Produkte pro Maschine, $d_i = 1\%, 2\%, 3\%$.

Fragen

- Mit welcher Wahrscheinlichkeit ist ein zufällig gewähltes Produkt defekt?
- Mit welcher Wahrscheinlichkeit stammt ein defektes Produkt von Maschine 1?

Modell

(i) $\Omega := \{(1,+), (1,-), (2,+), (2,-), (3,+), (3,-)\}$,

(ii) $\mathfrak{A} := \mathfrak{P}(\Omega)$,

(iii) P definiert durch

$$P(\{i,-\}) := r_i \cdot d_i, \qquad P(\{i,+\}) := r_i \cdot (1 - d_i).$$

Für $D := \{(1,-), (2,-), (3,-)\}$ und $M_i := \{(i,+), (i,-)\}$ folgt

$$P(M_i) = r_i, \qquad P(D \mid M_i) = \frac{r_i \cdot d_i}{r_i} = d_i.$$

Man erhält

$$P(D) = d_1 r_1 + d_2 r_2 + d_3 r_3 = \frac{3}{200}$$

und

$$P(M_1 \mid D) = \frac{P(D \mid M_1) \cdot P(M_1)}{P(D)} = \frac{200}{3} \cdot d_1 r_1 = \frac{2}{5}.$$

Häufig wie in diesem Beispiel: Modellierung durch Vorgabe bedingter Wahrscheinlichkeiten, etwa bei Markov-Ketten.

Nun: ein zentraler Begriff der Wahrscheinlichkeitstheorie.

Definition 7. Zwei Ereignisse $A, B \in \mathfrak{A}$ heißen *unabhängig*, falls

$$P(A \cap B) = P(A) \cdot P(B).$$

Bemerkung 8. Im Falle $P(B) > 0$ gilt

$$A, B \text{ unabhängig} \Leftrightarrow P(A \mid B) = P(A).$$

Beispiel 9. Einmaliges Würfeln (wie üblich modelliert) und

$$B := \{1, 2, 3, 4\}, \qquad A_1 := \{2, 4, 6\}, \qquad A_2 := \{1\}.$$

Dann gilt

$$P(A_1 \mid B) = \frac{1}{2}, \qquad P(A_1) = \frac{1}{2},$$

d.h. A_1, B unabhängig. Ferner gilt

$$P(A_2 \mid B) = \frac{1}{4}, \qquad P(A_2) = \frac{1}{6},$$

d.h. A_2, B nicht unabhängig.

Beispiel 10. Zweimaliger Wurf einer fairen Münze,

$$\Omega := \{(\mathrm{Z}, \mathrm{Z}), (\mathrm{Z}, \mathrm{K}), (\mathrm{K}, \mathrm{Z}), (\mathrm{K}, \mathrm{K})\},$$

$\mathfrak{A} := \mathfrak{P}(\Omega)$ und P Gleichverteilung auf Ω. Betrachte

$$\begin{aligned}
A_1 &:= \{(\mathrm{Z}, \mathrm{Z}), (\mathrm{Z}, \mathrm{K})\} && \text{1. Wurf Z,}\\
A_2 &:= \{(\mathrm{Z}, \mathrm{K}), (\mathrm{K}, \mathrm{K})\} && \text{2. Wurf K,}\\
A_3 &:= \{(\mathrm{Z}, \mathrm{K}), (\mathrm{K}, \mathrm{Z})\} && \text{Würfe verschieden.}
\end{aligned}$$

Es gilt $|A_i| = 2$ und $|A_i \cap A_j| = 1$ für $i \neq j$. Also

$$A_1,\ A_2 \text{ unabh.}, \qquad A_1,\ A_3 \text{ unabh.}, \qquad A_2,\ A_3 \text{ unabh.}$$

Im folgenden sei $I = \{1, \ldots, n\}$ oder $I = \mathbb{N}$.

Definition 11. Eine Folge $(A_i)_{i \in I}$ von Ereignissen heißt *unabhängig*, falls für jede endliche Menge $\emptyset \neq J \subseteq I$ gilt

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j).$$

Vgl. Definition 7 im Spezialfall $|I| = |J| = 2$.

Bemerkung 12. Falls $(A_i)_{i\in I}$ unabhängig, so folgt die *paarweise Unabhängigkeit*

$$\forall j_1, j_2 \in I,\ j_1 \neq j_2 :\ A_{j_1}, A_{j_2} \text{ unabhängig.}$$

Die Umkehrung ist falsch, wie das folgende Beispiel zeigt.

Beispiel 13. Zweimaliger Wurf einer fairen Münze, siehe Beispiel 10. Die Folge (A_1, A_2, A_3) ist nicht unabhängig, da

$$|A_1 \cap A_2 \cap A_3| = 1.$$

Alternative Argumentation: es gilt$P(A_3|A_1 \cap A_2) = 1$ aber $P(A_3) = 1/2$, siehe Bemerkung 14.

Bemerkung 14. Gelte $P(A_i) > 0$ für alle $i \in I$. Dann ist $(A_i)_{i\in I}$ genau dann unabhängig, wenn für alle endlichen Mengen $\emptyset \neq J_1, J_2 \subseteq I$ mit $J_1 \cap J_2 = \emptyset$ gilt

$$P\Big(\bigcap_{j_1\in J_1} A_{j_1} | \bigcap_{j_2\in J_2} A_{j_2}\Big) = P\Big(\bigcap_{j_1\in J_1} A_{j_1}\Big).$$

Beweis. $\boxed{\text{ÜBUNG}}$ □

Gängige Sprechweise: $A_1, A_2, \ldots$ sind unabhängig statt $(A_1, A_2, \ldots)$ ist unabhängig.

3 Reellwertige Zufallsvariablen

Oft interessiert man sich (nur) für spezielle Aspekte eines Zufallsexperimentes. Dazu betrachtet man Abbildungen $\Omega \to \mathbb{R}$.

Bezeichnung. Die *Indikatorfunktion* $1_U : V \to \mathbb{R}$ einer Teilmenge $U \subseteq V$ ist definiert durch

$$1_U(x) := \begin{cases} 1, & \text{falls } x \in U \\ 0, & \text{falls } x \notin U. \end{cases}$$

Beispiel 1. Anzahl Anrufe in Callcenter an Tagen $1, \ldots, n$. Dazu sei

$$\Omega := \mathbb{N}_0^n := \{\underbrace{(\omega_1, \ldots, \omega_n)}_{=:\omega} : \omega_i \in \mathbb{N}_0 \text{ für } i = 1, \ldots, n\}$$

und $\mathfrak{A} := \mathfrak{P}(\Omega)$. Die Anzahl Anrufe an Tag i ist gegeben durch

$$X_i(\omega) := \omega_i$$

und die Gesamtanzahl der Anrufe ist gegeben durch

$$X(\omega) := \sum_{i=1}^{n} \omega_i$$

Mittels

$$Y_i(\omega) := 1_{\{K+1,\dots\}}(\omega_i)$$

wird beschrieben, ob an Tag i die Kapazitätsgrenze K überschritten wurde. Die Anzahl der Tage an, an denen die Kapazitätsgrenze K überschritten wurde, ist gegeben durch

$$Y(\omega) := \sum_{i=1}^{n} 1_{\{K+1,\dots\}}(\omega_i).$$

Definition 2. Eine Abbildung $X : \Omega \to \mathbb{R}$ heißt (reellwertige) *Zufallsvariable* (auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$), falls

$$\forall x \in \mathbb{R} : \ \{\omega \in \Omega : X(\omega) \leq x\} \in \mathfrak{A}.$$

Manchmal zugelassen: Funktionswerte $\pm\infty$. Siehe auch Lemma 8 und Lemma IV.3.1.

Bemerkung 3.

(i) Zufallsvariablen sind Abbildungen!

(ii) Für Zufallsvariablen sind die Wahrscheinlichkeiten $P(\{\omega \in \Omega : X(\omega) \leq x\})$ wohldefiniert.

(iii) Im Falle $\mathfrak{A} = \mathfrak{P}(\Omega)$ ist jede Abbildung $\Omega \to \mathbb{R}$ eine Zufallsvariable.

Bezeichnung. Suggestive Kurzschreibweise

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}$$

für $X : \Omega \to \mathbb{R}$ und $B \subseteq \mathbb{R}$ sowie

$$\{X \leq x\} := \{\omega \in \Omega : X(\omega) \leq x\} = \{X \in \,]{-\infty}, x]\}$$

für $x \in \mathbb{R}$. Analog mit „=“ usw. Also: Urbilder von Mengen, $\{X \in B\} = X^{-1}(B)$.

Beispiel 4. Callcenter

$$\Omega := \mathbb{N}_0^n, \qquad \mathfrak{A} := \mathfrak{P}(\Omega), \qquad X_i(\omega) := \omega_i.$$

Es gilt (Verknüpfung von Zufallsvariablen)

$$X = \sum_{i=1}^{n} X_i, \qquad Y_i = 1_{\{K+1,\dots\}} \circ X_i, \qquad Y = \sum_{i=1}^{n} Y_i.$$

Gängige Schreibweise $1_A(X_i)$ für $1_A \circ X_i$.
Spezielle Ereignisse: Gesamtanzahl der Anrufe liegt zwischen 1000 und 2000,

$$\{1000 \leq X \leq 2000\},$$

Kapazitätsgrenze K wurde nie überschritten,

$$\bigcap_{i=1}^{n} \{X_i \leq K\} = \bigcap_{i=1}^{n} \{Y_i = 0\} = \{Y = 0\}.$$

Definition 5. Die *Verteilungsfunktion* $F_X : \mathbb{R} \to [0,1]$ einer Zufallsvariable X auf $(\Omega, \mathfrak{A}, P)$ ist definiert durch

$$F_X(x) := P(\{X \leq x\}).$$

Wir werden sehen, daß die wahrscheinlichkeitstheoretische Struktur einer Zufallsvariable vollständig durch ihre Verteilungsfunktion beschrieben wird.

Gleichheit von Zufallsvariablen in folgendem Sinn (und damit schwächer als die Gleichheit von Abbildungen).

Definition 6. Zwei Zufallsvariablen X auf $(\Omega, \mathfrak{A}, P)$ und X' auf $(\Omega', \mathfrak{A}', P')$ heißen *identisch verteilt*, falls $F_X = F_{X'}$.

Beispiel 7. Betrachte die Gleichverteilung P auf $\Omega := \{1, \ldots, n\}$ und setze

$$X(\omega) := \omega.$$

Dann gilt

$$P(\{X = x\}) = \begin{cases} 1/n, & \text{falls } x \in \{1, \ldots, n\} \\ 0, & \text{sonst} \end{cases}$$

Betrachte den Wahrscheinlichkeitsraum $(\Omega', \mathfrak{A}', P')$ zur Modellierung des Pfeiltreffers auf einer Dartscheibe, siehe Beispiel 1.15. Definiere $A'_0 = \{(0,0)\}$ und für $x = 1, \ldots, n$ die Sektoren

$$A'_x := \{\rho(\cos\alpha, \sin\alpha) : \rho \in \,]0, r]\,,\ \alpha \in \,](x-1)/n \cdot 2\pi, x/n \cdot 2\pi]\}.$$

Sei $X'(\omega')$ der getroffene Sektor, d.h.

$$X'(\omega') := x,$$

falls $\omega' \in A'_x$. Dann gilt für $x = 1, \ldots, n$

$$P'(\{X' = x\}) = P'(A'_x) = \lambda(A'_x)/\lambda(\Omega') = 1/n,$$

sowie $P'(\{X' = x\}) = 0$ für $x \in \mathbb{R} \setminus \{1, \ldots, n\}$. Also

$$\forall x \in \mathbb{R} :\ P(\{X = x\}) = P'(\{X' = x\}). \tag{1}$$

Beh.: X und X' sind identisch verteilt.

Beweis. Für $x \in \mathbb{R}$, $M := \,]-\infty, x]$ und $D := \{1, \ldots, n\}$ gilt

$$\begin{aligned} P(\{X \leq x\}) &= P(\{X \in M\} \cap \{X \in D\}) \\ &= \sum_{y \in M \cap D} \underbrace{P(\{X = y\})}_{=P'(\{X'=y\})} = P'(\{X' \leq x\}). \end{aligned}$$

□

Siehe auch Lemma III.4.4.

Warnung: aus (1) folgt im allgemeinen nicht, daß X und X' identisch verteilt sind.

Bezeichnung. $\mathfrak{M} := \{M \subseteq \mathbb{R} : M \text{ oder } M^c \text{ Intervall}\}$.

Lemma 8. Für jede Zufallsvariable X auf $(\Omega, \mathfrak{A}, P)$ gilt

$$\forall M \in \mathfrak{M} : \ \{X \in M\} \in \mathfrak{A}.$$

Beweis. Im folgenden seien $a, b \in \mathbb{R}$ mit $a \leq b$. Wir beweisen die Aussage in Teilschritten.

0. Für $M = \mathbb{R}$ gilt $\{X \in M\} = \Omega \in \mathfrak{A}$.

1. Für $M =]-\infty, b]$ gilt $\{X \in M\}$ nach Definition einer Zufallsvariable.

2. Für $M =]a, \infty[$ gilt $M = \mathbb{R} \setminus]-\infty, a]$, also

$$\{X \in M\} = \{X \in \mathbb{R}\} \setminus \{X \in]-\infty, a]\} \in \mathfrak{A}.$$

3. Für $M =]-\infty, b[$ gilt $M = \bigcup_{n=1}^{\infty}]-\infty, b - 1/n]$, also

$$\{X \in M\} = \bigcup_{i=1}^{\infty} \{X \in]-\infty, b - 1/n]\} \in \mathfrak{A}.$$

4. Für $M = [a, \infty[$ gilt $M = \mathbb{R} \setminus]-\infty, a[$, also

$$\{X \in M\} = \{X \in \mathbb{R}\} \setminus \{X \in]-\infty, a[\} \in \mathfrak{A}.$$

5. Beschränkte Intervalle sind Durchschnitte der unter 1.–4. betrachteten Intervalle, und es gilt

$$\{X \in I_1 \cap I_2\} = \{X \in I_1\} \cap \{X \in I_2\}.$$

Demnach gilt die Aussage für alle Intervalle, unter Benutzung der Definition der σ-Algebra auch für deren Komplemente. □

Satz 9. Zwei Zufallsvariablen X auf $(\Omega, \mathfrak{A}, P)$ und X' auf $(\Omega', \mathfrak{A}', P')$ sind genau dann identisch verteilt, wenn

$$\forall M \in \mathfrak{M} : \ P(\{X \in M\}) = P'(\{X' \in M\}).$$

Beweis (und Rechentechnik). Zu zeigen ist nur „⇒“. Wir gehen wieder in obigen Teilschritten vor.

1. Für $M :=]a, b]$ gilt

$$\begin{aligned} P(\{X \in M\}) &= P(\{X \leq b\} \setminus \{X \leq a\}) \\ &= P(\{X \leq b\}) - P(\{X \leq a\}) \\ &= P'(\{X' \leq b\}) - P'(\{X' \leq a\}) = P'(\{X' \in M\}). \end{aligned}$$

2. Für $M :=]a, b[$ gilt $M = \bigcup_{n=1}^{\infty}]a, b - 1/n]$. Also nach Satz 1.17.(ii) und 1.

$$\begin{aligned} P(\{X \in M\}) &= P\left(\bigcup_{n=1}^{\infty} \{X \in]a, b - 1/n]\}\right) \\ &= \lim_{n\to\infty} P\left(\{X \in]a, b - 1/n]\}\right) \\ &= \lim_{n\to\infty} P'\left(\{X' \in]a, b - 1/n]\}\right) = P'(\{X' \in M\}). \end{aligned}$$

3. Für $M := \{a\}$ gilt

$$M = \bigcap_{n=1}^{\infty}]a - 1/n, a + 1/n[\,.$$

Also nach Satz 1.17.(iii) und 2.

$$\begin{aligned} P(\{X \in M\}) &= \lim_{n\to\infty} P\left(\{X \in]a - 1/n, a + 1/n[\}\right) \\ &= \lim_{n\to\infty} P'\left(\{X' \in]a - 1/n, a + 1/n[\}\right) = P'(\{X' \in M\}). \end{aligned}$$

Für alle weiteren Typen von Mengen $M \in \mathfrak{M}$ nutze man die σ-Additivität, Additivität und die Rechenregel für Komplemente. □

Bemerkung 10. Gemäß Satz 9 bestimmt die Verteilungsfunktion F_X die Wahrscheinlichkeiten $P(\{X \in M\})$ für $M \in \mathfrak{M}$ eindeutig. Siehe auch Satz IV.3.6.

Im folgenden sei $I := \{1, \ldots, n\}$ oder $I := \mathbb{N}$. Wir betrachten eine Folge $(X_i)_{i\in I}$ von Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$.

Definition 11. $X_1, X_2 \ldots$ heißen *identisch verteilt*, falls für alle $i, j \in I$ die Zufallsvariablen X_i und X_j identisch verteilt sind.

Definition 12. $(X_i)_{i\in I}$ heißt *unabhängig*, falls für jede Folge $(x_i)_{i\in I}$ in $\mathbb{R}$ die Folge $(\{X_i \le x_i\})_{i\in I}$ von Ereignissen unabhängig ist.

Gängige Sprechweise: $X_1, X_2, \ldots$ sind unabhängig statt $(X_1, X_2, \ldots)$ ist unabhängig.

Bemerkung 13. $(X_i)_{i\in I}$ ist genau dann unabhängig, wenn für alle endlichen Mengen $J \subseteq I$ mit $|J| \ge 2$ und Folgen $(x_j)_{j\in J}$ in $\mathbb{R}$ gilt

$$P\left(\bigcap_{j\in J} \{X_j \le x_j\}\right) = \prod_{j\in J} P(\{X_j \le x_j\}).$$

Beispiel 14. Zweimaliger Münzwurf, siehe Beispiel 2.10. Betrachte für $i = 1, 2$ die Zufallsvariablen $X_i(\omega) := 1_{\{Z\}}(\omega_i)$.

Beh.: X_1 und X_2 sind unabhängig und identisch verteilt.

Beweis. Es gilt

$$\{X_i \leq x\} = \begin{cases} \emptyset, & \text{für } x < 0 \\ \{\omega \in \Omega : \omega_i = K\}, & \text{für } 0 \leq x < 1 \\ \Omega, & \text{für } x \geq 1. \end{cases}$$

Also

$$P(\{X_i \leq x\}) = \begin{cases} 0 & , \text{ falls } x < 0 \\ \frac{1}{2} & , \text{ falls } 0 \leq x < 1 \\ 1 & , \text{ falls } x \geq 1. \end{cases}$$

Insbesondere sind X_1 und X_2 identisch verteilt.
Für die Unabhängigkeit ist noch zu zeigen, daß

$$P(\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}) = P(\{X_1 \leq x_1\}) \cdot P(\{X_2 \leq x_2\}). \tag{2}$$

Klar: (2) gilt, falls $x_1 < 0, x_2 < 0, x_1 \geq 1$ oder $x_2 \geq 1$. Für $x_i \in [0, 1[$ gilt

$$\begin{aligned} P(\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}) &= P(\{(K, K)\}) = \frac{1}{4} \\ &= P(\{X_1 \leq x_1\}) \cdot P(\{X_2 \leq x_2\}). \end{aligned}$$

□

Satz 15. $(X_i)_{i \in \mathbb{N}}$ ist genau dann unabhängig, wenn

$$\forall n \in \mathbb{N}\ \forall M_1, \dots, M_n \in \mathfrak{M} : P\left(\bigcap_{j=1}^{n} \{X_j \in M_j\}\right) = \prod_{j=1}^{n} P(\{X_j \in M_j\}).$$

Analog für $I := \{1, \dots, n\}$. Siehe auch Satz IV.3.10.

Beweis. $\boxed{\text{Übung}}$ □

Bemerkung 16. Stochastische Modelle beruhen sehr häufig auf einer unabhängigen Folge von identisch verteilten Zufallsvariablen. Abkürzung: *iid* für independent and identically distributed.

Literatur

H.-O. Georgii, *Stochastik*, 3. Auflage, de Gruyter, Berlin, 2007.

Ch. M. Grinstead, J. L. Snell, *Introduction to Probability*,
http://www.dartmouth.edu/~chance/teaching_aids/
books_articles/probability_book/book.html

O. Häggström, *Finite Markov Chains and Algorithmic Applications*, Cambridge Univ. Press, Cambridge, 2002.

Ch. Hesse, *Angewandte Wahrscheinlichkeitstheorie*, Vieweg, Braunschweig, 2003.

A. Irle, *Wahrscheinlichkeitstheorie und Statistik*, 2. Auflage, Teubner, Stuttgart, 2005.

U. Krengel, *Einführung in die Wahrscheinlichkeitstheorie und Statistik*, 8. Auflage, Vieweg, Braunschweig, 2005.

J. Lehn, H. Wegmann, *Einführung in die Statistik*, 5. Auflage, Teubner, Stuttgart, 2004.

D. Meintrup, St. Schäffler, *Stochastik*, Springer, Berlin, 2005.

T. Müller-Gronbach, E. Novak, K. Ritter, *Monte-Carlo-Algorithmen*, erscheint bei Springer, Berlin, 2008.

A. N. Širjaev, *Wahrscheinlichkeit*, Deutscher Verlag der Wissenschaften, Berlin, 1988.
A. N. Shiryayev, *Probability*, Springer, New York, 1984.

Definitionen und Bezeichnungen