Bausteine zur Vorlesung von Prof. Dr. Bernd Hofmann

Mathematik IV (Stochastik) für Informatiker

Fakultät für Mathematik der Technischen Universität Chemnitz

Sommersemester 2018

Dieser Text soll die Nacharbeit der Vorlesung erleichern und an Definitionen, Sätze, Zusammenhänge und Beispiele erinnern. Hinweise zu Tippfehlern und Unstimmigkeiten werden gern entgegengenommen.

Textstand: 03.04.2018.

Inhaltsverzeichnis

1 Einführung in die Wahrscheinlichkeitsrechnung 5

1.1 Wahrscheinlichkeitsräume 5

1.1.1 Rechnen mit zufälligen Ereignissen 6

1.1.2 Rechnen mit Wahrscheinlichkeiten 8

1.1.3 Grundformeln der Kombinatorik 10

1.2 Bedingte Wahrscheinlichkeiten 11

1.2.1 Multiplikationsregel, totale Wahrscheinlichkeit, Satz von BAYES 12

1.2.2 Stochastische Unabhängigkeit zufälliger Ereignisse 14

1.2.3 Methode der geometrischen Wahrscheinlichkeit 16

1.2.4 Ergänzende Beispiele zur Einführung von Wahrscheinlichkeiten 16

1.3 Zufallsgrößen und Verteilungsfunktionen 19

1.3.1 Einführung 19

1.3.2 Diskrete Zufallsgrößen 19

1.3.3 Stetige Zufallsgrößen 31

1.4 Das Gesetz der großen Zahlen und Grenzverteilungssätze 43

1.4.1 Das Gesetz der großen Zahlen 43

1.4.2 Grenzverteilungssätze 44

1.5 Mehrdimensionale Verteilungen 48

1.5.1 Einführung 48

1.5.2 Bedingte Verteilungen 53

1.5.3 Erwartungswertevektor, Kovarianzmatrix, Normalverteilung 54

2 Einführung in die mathematische Statistik 56

2.1 Grundbegriffe 56

2.2 Punktschätzung 57

2.3 Verteilungen wichtiger Stichprobenfunktionen 59

2.3.1 Quantile 59

2.3.2 Weitere stetige Verteilungen 60

2.3.3 Stichprobenfunktionen bei binomialverteilter Grundgesamtheit 62

2.3.4 Stichprobenfunktionen bei normalverteilter Grundgesamtheit 62
2.4 Bereichsschätzung 63
2.4.1 Konfidenzintervalle bei binomialverteilter Grundgesamtheit 63
2.4.2 Konfidenzintervalle bei normalverteilter Grundgesamtheit 64
2.4.3 Einseitige Konfidenzintervalle 66
2.5 Tests 66
2.5.1 Allgemeines Schema für Parametertests 66
2.5.2 Parametertests bei binomialverteilter Grundgesamtheit 67
2.5.3 Parametertests bei normalverteilter Grundgesamtheit 68
2.5.4 Vergleich zweier normalverteilter Grundgesamtheiten 71
2.5.5 χ^2-Test 72
2.6 Spezielle Schätzverfahren 73
2.6.1 Maximum-Likelihood-Methode 73
2.6.2 Momentenmethode 75
2.6.3 Methode der kleinsten Quadrate 76

1 Einführung in die Wahrscheinlichkeitsrechnung

Ziel dieses Kapitels ist die Einführung in mathematische Modelle zur Behandlung von zufallsbeeinflussten Vorgängen (Zufallssituationen). Dazu betrachten wir zunächst zwei einfache Beispiele.

Beispiel (Geburtstagsaufgabe)**.** In einem Raum befinden sich n Personen. Wie groß ist die Wahrscheinlichkeit, dass mindestens zwei der n Personen am selben Tag Geburtstag haben? Wie muss n gewählt werden, damit die Wahrscheinlichkeit dafür größer als $\frac{1}{2}$ ist? Wir werden diese Aufgabe in Abschnitt 1.1.3 lösen.

Beispiel. Zwei Personen A und B spielen mehrere Runden eines Spiels mit Geldeinsatz, bei dem jeder Spieler gleiche Gewinnchancen hat. Gesamtsieger ist, wer zuerst 4 Siege erreicht. Bei einem Stand von 3 : 1 für A wird das Spiel abgebrochen und die beiden Spieler teilen das Geld anhand der Wahrscheinlichkeit eines Gesamtsieges unter sich auf. Wie viel bekommt jeder? Lösung: Um Gesamtsieger zu werden, müsste B die nächsten 3 Spiele gewinnen: $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$. In allen anderen Fällen gewinnt A als erster 4 Spiele. Somit erhält Spieler A $\frac{7}{8}$ und Spieler B $\frac{1}{8}$ des Einsatzes.

Anwendung der Wahrscheinlichkeitsrechnung

- Statistische Qualitätskontrolle,
- Fehlerrechnung,
- Versicherungswesen,
- stochastische Finanzmathematik.

1.1 Wahrscheinlichkeitsräume

Zufallssituation: Eine Zufallssituation ist dadurch gekennzeichnet, dass sie (zumindest gedanklich) *beliebig oft wiederholbar* ist und das *Ergebnis absolut nicht vorhersagbar* ist.

Wahrscheinlichkeitsraum: Ein Wahrscheinlichkeitsraum ist die Zusammenfassung aller Teile eines mathematischen Modells zur Beschreibung einer Zufallssituation. Verschiedene Zufallssituationen führen im Allgemeinen auch auf verschiedene Wahrscheinlichkeitsräume.

Versuch: Ein Versuch ist die Realisierung einer Zufallssituation.

Mit Ω bezeichnen wir die Ergebnismenge eines Versuchs, d.h. die Menge aller möglichen Ergebnisse, und mit ω ein konkretes Ergebnis, also $\omega \in \Omega$. Dabei gehen wir davon aus, dass jedem Ergebnis eines Versuchs eindeutig ein Element ω der Ergebnismenge Ω zugeordnet ist.

Beispiel. Beim Werfen eines idealen Würfels ist $\Omega = \{1, 2, 3, 4, 5, 6\}$ eine endliche Ergebnismenge mit 6 möglichen verschiedenen Ergebnissen.

Beispiel. Ein Fahrzeug kann mit einem begrenzten Treibstoffvorrat nur eine bestimmte Strecke zurücklegen. Somit ist $\Omega = \{\omega \in \mathbb{R} : \omega \geq 0\}$ eine überabzählbar unendliche Ergebnismenge.

Beispiel. Es soll der Zustand von n elektrischen Geräten überprüft werden. Wir bezeichnen den Zustand des i-ten Gerätes mit

$$\omega_i = \begin{cases} 1, & \text{Gerät in Ordnung} \\ 0, & \text{Gerät defekt} \end{cases}.$$

Somit ist $\Omega = \big\{(\omega_1, \ldots, \omega_n) \in \mathbb{R}^n : \omega_i \in \{0, 1\}\big\}$ eine endliche Ergebnismenge mit n Elementen.

Beispiel. Bei der Bestimmung der Lebensdauer von n Lampen, wobei ω_i die Lebensdauer der i-ten Lampe bezeichnet, ist die Ergebnismenge $\Omega = \{(\omega_1, \ldots, \omega_n) \in \mathbb{R}^n : \omega_i \geq 0\}$ überabzählbar unendlich.

Definition 1.1.1. Ein *zufälliges Ereignis* ist eine Teilmenge $A \subset \Omega$ der Ergebnismenge. Man sagt, dass das Ereignis A eingetreten ist, wenn das Versuchsergebnis ω in A liegt, d.h. wenn $\omega \in A$ gilt.

Nicht jede Teilmenge $A \subset \Omega$ muss sich als zufälliges Ereignis betrachten lassen, aber alle zufälligen Ereignisse sind Teilmengen von Ω.

Beispiel. Wir betrachten beim Würfeln mit einem idealen Würfel das Ereignis „gerade Zahl gewürfelt“. Haben also $A = \{2, 4, 6\}$.

Beispiel. Für ein Fahrzeug mit begrenztem Treibstoffvorrat interessiert das Ereignis „fährt mindestens 150 km“. Haben dann $A = \{\omega \in \mathbb{R} : \omega \geq 150\}$.

Beispiel. Bei der Überprüfung von n Geräten sollen „mindestens 2 in Ordnung“ sein, d.h. $A = \{(\omega_1, \ldots, \omega_n) \in \mathbb{R}^n : \omega_i \in \{0, 1\}$ und $\omega_1 + \ldots + \omega_n \geq 2\}$.

Beispiel. Die mittlere Brenndauer von n Lampen soll „zwischen 500 und 5000 Stunden“ betragen, d.h. $A = \{(\omega_1, \ldots, \omega_n) \in \mathbb{R}^n : 500 \leq \frac{\omega_1 + \ldots + \omega_n}{n} \leq 5000\}$.

1.1.1 Rechnen mit zufälligen Ereignissen

Oder-Ereignis: Das Ereignis $C =$ „A oder B“ tritt ein, wenn entweder A oder B oder beide eintreten, d.h. $C = A \cup B$.

Und-Ereignis: Das Ereignis $C =$ „A und B“ tritt ein, wenn sowohl A als auch B eintritt, d.h. $C = A \cap B$.

Komplementärereignis (Gegenereignis): Das Ereignis $C = \overline{A} =$ „nicht A“ tritt ein, wenn A nicht eintritt, d.h. $C = \Omega \setminus A$.

Sicheres Ereignis: $A = \Omega$.

Unmögliches Ereignis: $A = \overline{\Omega} = \varnothing$.

Elementarereignis: $A = \{\omega\}$, d.h. A enthält genau ein Element der Ergebnismenge Ω.

Unvereinbare Ereignisse: A und B heißen unvereinbar, wenn $A \cap B = \varnothing$ gilt.

Man kann „oder“ und „und“ auch auf endlich bzw. abzählbar unendlich viele Ereignisse anwenden:

$$\bigcup_{i=1}^{n} A_i \quad \text{und} \quad \bigcap_{i=1}^{n} A_i \qquad \text{bzw.} \qquad \bigcup_{i=1}^{\infty} A_i \quad \text{und} \quad \bigcap_{i=1}^{\infty} A_i.$$

Definition 1.1.2. Eine Menge $\mathfrak{A}$ von Ereignissen, d.h. von Teilmengen der Ergebnismenge Ω, heißt bezogen auf eine feste Zufallssituation *Ereignisfeld* (auch *Ereignisalgebra* oder *σ-Algebra*), wenn die folgenden drei Bedingungen erfüllt sind:

- $\Omega \in \mathfrak{A}$ (sicheres Ereignis gehört dazu),
- $A \in \mathfrak{A} \Rightarrow \overline{A} \in \mathfrak{A} \quad \forall A \in \mathfrak{A}$ (mit Ereignis gehört auch Komplementärereignis dazu),
- $A_1, A_2, \ldots \in \mathfrak{A} \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathfrak{A} \quad \forall A_1, A_2, \ldots \in \mathfrak{A}$ (abzählbare Vereinigungen ebenfalls).

Satz 1.1.3 (Rechenregeln). *Für zufällige Ereignisse A und B bzw. $A_1, A_2, \ldots$ gilt:*

- $\begin{array}{l} A \cup B = B \cup A \\ A \cap B = B \cap A \end{array}$ *(Kommutativität),*
- $\begin{array}{l} (A \cup B) \cup C = A \cup (B \cup C) \\ (A \cap B) \cap C = A \cap (B \cap C) \end{array}$ *(Assoziativität),*
- $\begin{array}{l} (A \cup B) \cap C = (A \cap C) \cup (B \cap C) \\ (A \cap B) \cup C = (A \cup C) \cap (B \cup C) \end{array}$ *(Distributivität),*
- $\begin{array}{l} \overline{A \cup B} = \overline{A} \cap \overline{B} \\ \overline{A \cap B} = \overline{A} \cup \overline{B} \end{array}$ *(*De Morgan*'sche Regeln),*
- $\begin{array}{ll} \overline{\bigcup\limits_{i=1}^{n} A_i} = \bigcap\limits_{i=1}^{n} \overline{A_i}, & \overline{\bigcup\limits_{i=1}^{\infty} A_i} = \bigcap\limits_{i=1}^{\infty} \overline{A_i} \\ \overline{\bigcap\limits_{i=1}^{n} A_i} = \bigcup\limits_{i=1}^{n} \overline{A_i}, & \overline{\bigcap\limits_{i=1}^{\infty} A_i} = \bigcup\limits_{i=1}^{\infty} \overline{A_i} \end{array}$ *(verallgemeinerte* De Morgan*'sche Regeln),*
- $A \cup \varnothing = A, \quad A \cap \varnothing = \varnothing,$
- $A \cup \Omega = \Omega, \quad A \cap \Omega = A.$

Falls $\omega \in A \Rightarrow \omega \in B$ für alle $\omega \in A$ gilt, so sagen wir „A zieht B nach sich“ und schreiben $A \subset B$. Dazu äquivalente mathematische Beschreibungen sind

$$A \subset B \quad \Leftrightarrow \quad A \cap B = A \quad \Leftrightarrow \quad A \cup B = B \quad \Leftrightarrow \quad A \cap \overline{B} = \varnothing \quad \Leftrightarrow \quad \overline{B} \subset \overline{A}.$$

Venn-Diagramme sind hilfreich bei der Illustration des Ereignis-Kalküls!

1.1.2 Rechnen mit Wahrscheinlichkeiten

Definition 1.1.4. Sei $A \in \mathfrak{A}$ ein festes Ereignis innerhalb einer Zufallssituation. Wir Bezeichnen dann mit n die Anzahl der ausgeführten Versuche, mit n_A die Anzahl der Versuche, bei denen A eingetreten ist, und mit

$$H_n = H_n(A) = \frac{n_A}{n}$$

die *relative Häufigkeit* für das Eintreten von A bei n Versuchen. Erfahrungsgemäß strebt $H_n(A)$ unter Verwendung eines speziell zugeschnittenen Grenzwertbegriffs gegen eine feste Zahl, die Wahrscheinlichkeit $P(A)$ für das Eintreten von A:

$$P(A) = \lim_{n \to \infty} H_n(A).$$

Beispiel. Beim Werfen einer Münze wird das Ereignis $A =$ „Kopf liegt oben“ betrachtet. Schon in den vergangenen Jahrhunderten galten Münzexperimente als interessant. Sie wurden z.B. von Comte de Buffon (1707-1788) und K. Pearson (1857-1936) durchgeführt.

	n	n_A	$H_n(A)$	$P(A)$
Buffon	4040	2048	0,5069	0,5
Pearson	24000	12012	0,5005	0,5

Eigenschaften der relativen Häufigkeit

- Offenbar ist $0 \leq H_n(A) \leq 1$.
- Ebenfalls offenbar ist $H_n(\Omega) = 1$.
- Für unvereinbare Ereignisse A und B, d.h. $A \cap B = \varnothing$, addieren sich offenbar die relativen Häufigkeiten:

$$H_n(A \cup B) = \frac{n_{A \cup B}}{n} = \frac{n_A + n_B}{n} = \frac{n_A}{n} + \frac{n_B}{n} = H_n(A) + H_n(B).$$

Diese Eigenschaften der relativen Häufigkeit bilden die Grundlage des Axiomensystems zum Rechnen mit Wahrscheinlichkeiten nach Kolmogorov (veröffentlicht 1933 im Springer-Verlag in seinem in deutscher Sprache verfassten Buch *Grundbegriffe der Wahrscheinlichkeitsrechnung*).

Definition 1.1.5 (Kolmogorov'sches Axiomensystem)**.** Gegeben sei eine Zufallssituation, die durch eine Ergebnismenge Ω und ein Ereignisfeld $\mathfrak{A}$ beschrieben wird. Jedem $A \in \mathfrak{A}$ ist dann eindeutig eine reelle Zahl $P(A)$, die Wahrscheinlichkeit für das Eintreten von A, zugeordnet. Dabei gelten die folgenden Axiome.

A1: Es gelte $0 \leq P(A) \leq 1$.

A2: Es gelte $P(\Omega) = 1$.

A3: Für paarweise unvereinbare Ereignisse $A_i \in \mathfrak{A}$, d.h. $A_i \cap A_j = \varnothing$ mit $i \neq j$ gelte stets

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

Folgerungen aus den KOLMOGOROV'schen Axiomen Für beliebige Ereignisse $A, B \in \mathfrak{A}$ gilt:

- $P(\varnothing) = 0$.

  Beweis. Seien $A_1 = \varnothing$ und $A_2 = \varnothing$ unvereinbare Ereignisse ($A_1 \cap A_2 = \varnothing$). Haben dann unter Verwendung von $A_1 \cup A_2 = \varnothing$ und Axiom A3

  $$P(A_1 \cup A_2) = P(A_1) + P(A_2) \quad \Rightarrow \quad P(\varnothing) = 2P(\varnothing) = 0. \qquad \square$$

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

  Beweis. Es gilt $A \cup \overline{A} = \Omega$ und $A \cap \overline{A} = \varnothing$. Axiom A3 liefert also $P(\Omega) = P(A) + P(\overline{A})$ und aus Axiom A2 folgt somit $1 = P(A) + P(\overline{A})$. □

- $P(A \cap \overline{B}) = P(A) - P(A \cap B)$.

  Beweis. Es gilt $A = A \cap \Omega = A \cap (B \cup \overline{B}) = (A \cap B) \cup (A \cap \overline{B})$. Aus Axiom A3 folgt somit $P(A) = P(A \cap B) + P(A \cap \overline{B})$. □

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

  Beweis. Es gilt $A \cup B = (A \cap \overline{B}) \cup B = (A \cap \overline{A \cap B}) \cup B$ und somit ist nach dem vorhergehenden Punkt und Axiom A3

  $$P(A \cup B) = P(A \cap \overline{A \cap B}) + P(B) = P(A) - P(A \cap B) + P(B). \qquad \square$$

Methode der klassischen Wahrscheinlichkeit Für eine endliche Ergebnismenge Ω mit N Elementen $\omega_1, \ldots, \omega_N$ und

$$P(\{\omega_1\}) = \cdots = P(\{\omega_N\}) = \frac{1}{N} \qquad \text{(LAPLACE-Annahme)}$$

gilt mit $A = \{\omega_{i_1}, \ldots, \omega_{i_M}\}$

$$P(A) = \frac{M}{N} = \frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}}.$$

Beispiel. Beim Würfeln mit einem idealen Würfel wird das Ereignis $A =$ „Primzahl gewürfelt“ betrachtet. Haben dann $A = \{2, 3, 5\}$, $M = 3$, $N = 6$ und somit $P(A) = \frac{3}{6} = \frac{1}{2}$.

Definition 1.1.6. Durch die Ergebnismenge Ω, das Ereignisfeld $\mathfrak{A}$ und das Wahrscheinlichkeitsmaß P sei eine Zufallssituation gegeben. Das Tripel $(\Omega, \mathfrak{A}, P)$ heißt *Wahrscheinlichkeitsraum* dieser Zufallssituation.

1.1.3 Grundformeln der Kombinatorik

Aus einem Gefäß mit n Kugeln, die (z.B. durch Nummerierung) unterscheidbar sind, sollen m Kugeln entnommen werden. Uns interessiert die Anzahl der möglichen Ergebnisse. Dabei berücksichtigen wir, ob die Reihenfolge eine Rolle spielt und ob eine entnommene Kugel vor der Entnahme der nächsten wieder zurückgelegt wird.

- Variationen (Reihenfolge wichtig):

  Anzahl mit Zurücklegen[1] $= {}^wV_n^m = n^m$,

  Anzahl ohne Zurücklegen $= V_n^m = n(n-1)(n-2)\dots(n-m+1) = \frac{n!}{(n-m)!}$.

- Kombinationen (Reihenfolge unwichtig):

  Anzahl mit Zurücklegen $= {}^wC_n^m = \binom{n+m-1}{m}$,

  Anzahl ohne Zurücklegen $= C_n^m = \binom{n}{m} = \frac{n!}{m!(n-m)!}$.

Beispiel. In einem Raum befinden sich n Personen, von denen keine am 29. Februar Geburtstag hat. A_n sei das Ereignis, dass mindestens 2 der n Personen am gleichen Tag Geburtstag haben, wobei wir davon ausgehen, dass alle 365 möglichen Tage gleichwahrscheinlich sind. Wir betrachten $\overline{A_n}$ und berechnen dann $P(A_n)$ durch $P(A_n) = 1 - P(\overline{A_n})$. Wir ziehen unter Beachtung der Reihenfolge und mit Zurücklegen n Tage aus 365 und erhalten somit $N = 365^n$ mögliche Fälle. Ziehen von n Tagen aus 365 ohne Zurücklegen und unter Beachtung der Reihenfolge ergibt $M = 365 \cdot 364 \cdot \ldots \cdot (365 - n + 1)$ günstige Ergebnisse. Haben somit

$$P(\overline{A_n}) = \frac{M}{N} = \frac{365 \cdot 364 \cdot \ldots \cdot (366-n)}{365^n}.$$

Für konkrete n erhalten wir die folgenden Wahrscheinlichkeiten.

n	$P(A_n)$	n	$P(A_n)$	n	$P(A_n)$
1	0	15	0,253	50	0,970
2	0,003	20	0,411	60	0,9941
3	0,008	22	0,476	70	0,99916
4	0,016	23	0,507	80	0,999914
5	0,027	30	0,706	90	0,9999938
10	0,117	40	0,891	100	0,99999969

Ab $n = 23$ Personen im Raum ist die Wahrscheinlichkeit, dass zwei am gleichen Tag Geburtstag haben, also größer als 50 %.

Beispiel. Es werden 6 aus 49 nummerierten Kugeln ohne Zurücklegen und ohne Beachtung der Reihenfolge gezogen. Zuvor wird ein Tipp abgegeben, welche Kugeln dies sein werden. A_k mit $k = 0, \ldots, 6$ bezeichne das Ereignis, dass genau k Kugeln richtig getippt wurden. Wir erhalten

$$P(A_k) = \frac{\text{günstige Fälle}}{\text{mögliche Fälle}} = \frac{\binom{6}{k}\binom{43}{6-k}}{\binom{49}{6}} = \frac{\binom{6}{k}\binom{43}{6-k}}{13983816}.$$

Dies ergibt die folgenden Wahrscheinlichkeiten.

[1] w steht für „mit Wiederholung“

k	günstige Fälle	$P(A_k)$	k	günstige Fälle	$P(A_k)$
0	6096454	0,436	4	13545	0,0009686
1	5775588	0,413	5	258	0,00001845
2	1851150	0,132	6	1	0,000000072
3	246820	0,018			

1.2 Bedingte Wahrscheinlichkeiten

In bestimmten Zufallssituationen kann es vorkommen, dass sich die Wahrscheinlichkeit für das Eintreten eines Ereignisses ändert, wenn man beachtet, dass ein anderes Ereignis bereits eingetreten ist.

Als Motivation für die Einführung bedingter Wahrscheinlichkeiten betrachten wir die relative Häufigkeit: Es werden n Versuche durchgeführt. Die relative Häufigkeit für das Eintreten von Ereignis A als Folge des Eintretens von B berechnet sich durch

$$H_n(A|B) := \frac{n_{A\cap B}}{n_B} = \frac{\frac{n_{A\cap B}}{n}}{\frac{n_B}{n}} = \frac{H_n(A\cap B)}{H_n(B)}.$$

Definition 1.2.1. Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum und seien A und B zwei zufällige Ereignisse mit $P(B) > 0$. Dann heißt die Größe

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

bedingte Wahrscheinlichkeit für das Eintreten von Ereignis A unter der Bedingung, dass B bereits eingetreten ist.

Beispiel. Mit einem idealen Würfel werden zwei Würfe ausgeführt. Wir betrachten die beiden Ereignisse

$$A = \text{„erster Wurf ist eine 6“} = \{(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\},$$

$$B = \text{„Augensumme beider Würfe ist 8“} = \{(6,2),(5,3),(4,4),(3,5),(2,6)\}.$$

Offensichtlich ist $P(A) = \frac{6}{36} = \frac{1}{6}$, $P(B) = \frac{5}{36}$ und $P(A\cap B) = \frac{1}{36}$. Somit ist $P(A|B) = \frac{1}{5}$.

Rechenregeln für bedingte Wahrscheinlichkeiten

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

  Beweis.

$$\begin{aligned} P(A|C) + P(\overline{A}|C) &= \frac{P(A\cap C)}{P(C)} + \frac{P(\overline{A}\cap C)}{P(C)} \\ &= \frac{P((A\cap C)\cup(\overline{A}\cap C))}{P(C)} = \frac{P((A\cup\overline{A})\cap(A\cup C)\cap(C\cup\overline{A})\cap(C\cup C))}{P(C)} \\ &= \frac{P(\Omega\cap(A\cup C)\cap(C\cup\overline{A})\cap C)}{P(C)} = \frac{P(\Omega\cap C\cap C)}{P(C)} = \frac{P(C)}{P(C)} = 1. \end{aligned}$$

$\square$

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

  Beweis.
  $$\begin{aligned} P(A \cup B|C) &= \frac{P((A \cup B) \cap C)}{P(C)} = \frac{P((A \cap C) \cup (B \cap C))}{P(C)} \\ &= \frac{P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)}{P(C)} \\ &= \frac{P(A \cap C)}{P(C)} + \frac{P(B \cap C)}{P(C)} - \frac{P((A \cap B) \cap C)}{P(C)} \\ &= P(A|C) + P(B|C) - P(A \cap B|C). \end{aligned}$$ □

- $P(C|C) = 1$

  Beweis.
  $$P(C|C) = \frac{P(C \cap C)}{P(C)} = \frac{P(C)}{P(C)} = 1.$$ □

Beispiel. Torsten durchsucht 7 gleichgroße CD-Stapel nach einer ganz bestimmten CD. Die Wahrscheinlichkeit, dass die gesuchte CD überhaupt in einem der Stapel vorhanden ist, sei $\frac{4}{5}$. Er hat bereits 6 Stapel erfolglos durchsucht. Wie groß ist die Wahrscheinlichkeit, die CD im 7. Stapel zu finden? A_i mit $i = 1, \ldots, 7$ sei das Ereignis „CD im i-ten Stapel“, wobei $P(A_1) = \ldots = P(A_7)$. Haben dann

$$P(A_i) = \frac{1}{7}(P(A_1) + \ldots + P(A_7)) = \frac{1}{7}P(A_1 \cup \ldots \cup A_7) = \frac{1}{7} \cdot \frac{4}{5} = \frac{4}{35}$$

und somit

$$P(A_7|\overline{A_1} \cap \ldots \cap \overline{A_6}) = \frac{P(A_7 \cap \overline{A_1} \cap \ldots \cap \overline{A_6})}{P(\overline{A_1} \cap \ldots \cap \overline{A_6})} = \frac{P(A_7)}{1 - P(A_1 \cup \ldots \cup A_6)} = \frac{\frac{4}{35}}{1 - 6 \cdot \frac{4}{35}} = \frac{4}{11}.$$

1.2.1 Multiplikationsregel, totale Wahrscheinlichkeit, Satz von Bayes

Stellt man die Formel für die bedingte Wahrscheinlichkeit $P(A|B)$ (siehe Definition 1.2.1) nach $P(A \cap B)$ um, so erhält man die *einfache Multiplikationsregel*:

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

Beispiel. An einer Universität schließen 70 % eines Jahrgangs das Fach Mathematik wenigstens mit der Note 3 ab (Ereignis B). Unter diesen Studenten erreichen 25 % sogar eine der Noten 1 oder 2 (Ereignis A). Mit welcher Wahrscheinlichkeit schließt ein beliebig ausgewählter Student des Jahrgangs das Fach mit 1 oder 2 ab? Lösung: Mit $P(B) = \frac{7}{10}$ und $P(A|B) = \frac{1}{4}$ liefert die Multiplikationsregel

$$P(A) = P(A \cap B) = P(A|B)P(B) = \frac{1}{4} \cdot \frac{7}{10} = \frac{7}{40} = 0{,}175.$$

Satz 1.2.2 (erweiterte Multiplikationsregel)**.** *Seien* $A_1, A_2, \ldots, A_n$ *Ereignisse aus dem Ereignisfeld* $\mathfrak{A}$ *einer festen Zufallssituation mit* $P(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$*. Dann gilt*

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

Beweis. Der Satz folgt durch vollständige Induktion aus der einfachen Multiplikationsregel:

$$\begin{aligned} P(A_1 \cap \ldots \cap A_n) &= P(A_n|A_1 \cap \ldots \cap A_{n-1})P(A_1 \cap \ldots \cap A_{n-1}) \\ P(A_1 \cap \ldots \cap A_{n-1}) &= P(A_{n-1}|A_1 \cap \ldots \cap A_{n-2})P(A_1 \cap \ldots \cap A_{n-2}) \\ &\vdots \\ P(A_1 \cap A_2 \cap A_3) &= P(A_3|A_1 \cap A_2)P(A_1 \cap A_2) \\ P(A_1 \cap A_2) &= P(A_2|A_1)P(A_1). \end{aligned}$$

□

Definition 1.2.3. Eine Menge von Ereignissen $B_1, B_2, \ldots, B_n$ heißt *vollständiges Ereignissystem*, wenn gilt:

- $B_1 \cup B_2 \cup \ldots \cup B_n = \bigcup_{i=1}^{n} B_i = \Omega$,
- die Ereignisse sind paarweise unvereinbar (disjunkt), d.h. $B_i \cap B_j = \varnothing$ für $i \neq j$.

Satz 1.2.4 (Satz von der totalen Wahrscheinlichkeit)**.** *Sei die Menge der Ereignisse* $B_1, \ldots, B_n$ *ein vollständiges Ereignissystem. Dann gilt für ein beliebiges (anderes) Ereignis* A *die Formel*

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

Beweis. Es gilt

$$A = A \cap \Omega = A \cap (B_1 \cup \ldots \cup B_n) = (A \cap B_1) \cup \ldots \cup (A \cap B_n)$$

und da die Ereignisse $A \cap B_i$ $(i = 1, \ldots, n)$ paarweise unvereinbar sind, können wir die Multiplikationsregel anwenden und erhalten

$$P(A) = P((A \cap B_1) \cup \ldots \cup (A \cap B_n)) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$

□

Satz 1.2.5 (Satz von Bayes)**.** *Sei die Menge der Ereignisse* $B_1, \ldots, B_n$ *ein vollständiges Ereignissystem. Dann gilt für ein beliebiges (anderes) Ereignis* A *mit* $P(A) > 0$ *und für* $j = 1, \ldots, n$ *die Formel*

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)} \qquad \text{(Bayes'sche Formel).}$$

Beweis. Der Satz ist eine direkte Folgerung aus der Multiplikationsregel und aus dem Satz von

der totalen Wahrscheinlichkeit:

$$P(A \cap B_j) = P(A|B_j)P(B_j) = P(B_j|A)P(A) \quad \Rightarrow \quad P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)}.$$

□

Beispiel. Aus der Jahresstatistik einer großen deutschen Pannenhilfsorganisation geht hervor, dass bei vorgefundenen Schäden im Bereich der Motorausfälle die folgende Schadenstypenverteilung zu verzeichnen war:

- 50 % Störungen der Zündanlage (davon 50 % vor Ort behoben),
- 30 % Störungen der Kraftstoffzufuhr (davon 30 % vor Ort behoben),
- 20 % sonstige Störungen (davon 5 % vor Ort behoben).

Uns interessiert nun, wie viel Prozent der Motorausfälle vor Ort behoben werden konnten und wie sich die vor Ort behobenen Motorausfälle auf die einzelnen Schadensarten verteilen.

Wir bezeichnen die Ereignisse einer Störung mit B_1 (Zündanlage), B_2 (Kraftstoffzufuhr) und B_3 (sonstige Störungen) und das Ereignis, dass ein Motorausfall vor Ort behoben werden konnte mit A. Die Ereignisse B_1, B_2, B_3 bilden ein vollständiges Ereignissystem. Aus dem Satz von der totalen Wahrscheinlichkeit erhalten wir

$$P(A) = \underbrace{P(A|B_1)}_{0{,}5}\underbrace{P(B_1)}_{0{,}5} + \underbrace{P(A|B_2)}_{0{,}3}\underbrace{P(B_2)}_{0{,}3} + \underbrace{P(A|B_3)}_{0{,}05}\underbrace{P(B_3)}_{0{,}2} = 0{,}35,$$

d.h. in 35 % aller Fälle konnte vor Ort geholfen werden. Die Aufteilung der vor Ort behobenen Motorausfälle auf die einzelnen Schadensarten erhalten wir mit dem Satz von Bayes:

$$\begin{aligned} P(B_1|A) &= \frac{P(A|B_1)P(B_1)}{P(A)} = 0{,}714, \\ P(B_2|A) &= \frac{P(A|B_2)P(B_2)}{P(A)} = 0{,}257, \\ P(B_3|A) &= \frac{P(A|B_3)P(B_3)}{P(A)} = 0{,}029. \end{aligned}$$

1.2.2 Stochastische Unabhängigkeit zufälliger Ereignisse

Definition 1.2.6. Zwei Ereignisse $A, B \in \mathfrak{A}$ heißen *stochastisch unabhängig*, wenn gilt

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

Bemerkung. „Stochastisch unabhängig“ bedeutet also $P(A|B) = P(A)$ und $P(B|A) = P(B)$, d.h. der Zufallscharakter der Ereignisse A und B beeinflusst sich nicht.

Satz 1.2.7. *Wenn die Ereignisse A und B stochastisch unabhängig sind, so sind auch die Ereignisse $\overline{A}$ und $\overline{B}$ stochastisch unabhängig.*

Beweis. Es gilt

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

d.h. A und $\overline{B}$ sind stochastisch unabhängig, woraus analog die stochastische Unabhängigkeit von $\overline{A}$ und $\overline{B}$ folgt. $\square$

Beispiel. Ein elektrisches Gerät besteht aus zwei Bauteilen T_1 und T_2, bei denen unabhängig voneinander Defekte auftreten können. Wir betrachten die stochastisch unabhängigen Ereignisse A_1 = „Bauteil T_1 funktioniert" mit $P(A_1) = p_1$ und A_2 = „Bauteil T_2 funktioniert" mit $P(A_2) = p_2$.

Eine *Serienschaltung* der beiden Bauteile funktioniert, wenn sowohl T_1 als auch T_2 funktioniert:

$$P(A_1 \cap A_2) = P(A_1)P(A_2) = p_1 p_2.$$

Für das Zahlenbeispiel $p_1 = p_2 = 0{,}9$ ergibt dies eine Wahrscheinlichkeit für das Funktionieren des Geräts von 0,81.

Eine *Parallelschaltung* funktioniert, wenn T_1 oder T_2 oder beide funktionieren:

$$P(A_1 \cup A_2) = 1 - P(\overline{A_1 \cup A_2}) = 1 - P(\overline{A_1} \cap \overline{A_2}) = 1 - (1 - p_1)(1 - p_2).$$

Für das Zahlenbeispiel $p_1 = p_2 = 0{,}9$ ergibt dies eine Wahrscheinlichkeit für das Funktionieren des Geräts von 0,99.

Definition 1.2.8. Die Ereignisse $A_1, \ldots, A_n$ heißen *vollständig stochastisch unabhängig*, wenn für jede Auswahl von k Ereignissen aus den n gegebenen gilt:

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}).$$

Aus dieser Definition ergibt sich für vollständig stochastisch unabhängige Ereignisse der folgende wichtige Zusammenhang:

$$\begin{aligned} P(A_1 \cup \cdots \cup A_n) &= 1 - P(\overline{A_1 \cup \cdots \cup A_n}) = 1 - P(\overline{A_1} \cap \cdots \cap \overline{A_n}) \\ &= 1 - (1 - P(A_1)) \cdots (1 - P(A_n)). \end{aligned}$$

Beispiel. Paarweise stochastisch unabhängige Ereignisse müssen nicht gleichzeitig vollständig stochastisch unabhängig sein. Sei z.B. $\Omega = \{1, 2, 3, 4\}$ mit

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = \frac{1}{4} \quad \text{und} \quad A = \{1, 2\},\ B = \{1, 3\},\ C = \{2, 3\}.$$

Dann ist $P(A) = P(B) = P(C) = \frac{1}{2}$ und

$$\begin{aligned} P(A \cap B) &= P(\{1\}) = \frac{1}{4} = P(A)P(B), \\ P(A \cap C) &= P(\{2\}) = \frac{1}{4} = P(A)P(C), \\ P(B \cap C) &= P(\{3\}) = \frac{1}{4} = P(B)P(C). \end{aligned}$$

Also sind die Ereignisse A, B, C paarweise stochastisch unabhängig, aber

$$P(A \cap B \cap C) = P(\varnothing) = 0 \neq \frac{1}{8} = P(A)P(B)P(C),$$

d.h. die Ereignisse A, B, C sind nicht vollständig stochastisch unabhängig.

1.2.3 Methode der geometrischen Wahrscheinlichkeit

Die Methode der geometrischen Wahrscheinlichkeit ist ein Spezialfall der klassischen Wahrscheinlichkeit (siehe Abschnitt 1.1.2). Die Ergebnismenge Ω ist überabzählbar unendlich und verkörpert ein geometrisches Objekt, d.h. eine Menge von Punkten in der Ebene oder im Raum, wobei die folgenden zwei Bedingungen gelten:

- Ω lässt sich als geometrisches Objekt mit endlichem Inhalt darstellen,
- Teilmengen von Ω mit gleichem Inhalt sind gleiche Wahrscheinlichkeiten zugeordnet.

Exemplarisch kann das Schießen auf eine Dartscheibe Ω betrachtet werden. Das Ereignis A tritt ein, wenn ein bestimmtes Feld (z.B. innerer Ring) getroffen wird, d.h. A ist die Fläche dieses Feldes als Teilmenge von Ω. Dann gilt

$$P(A) = \frac{\text{günstige Fälle}}{\text{mögliche Fälle}} = \frac{\text{Inhalt von } A}{\text{Inhalt von } \Omega}.$$

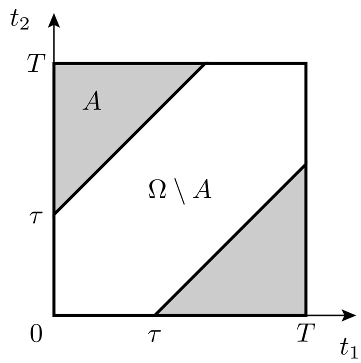
Dabei gibt es Ereignisse A mit Inhalt 0 (und damit $P(A) = 0$), die aber keine unmöglichen Ereignisse sind, z.B. Kurvenstücke in der Ebene oder Flächenstücke im Raum.

Beispiel. Eine Funkstation sendet zu zwei zufälligen Zeiten t_1 und t_2 im Zeitintervall $[0, T]$ je ein punktförmiges Signal aus. Ein Empfänger kann diese beiden Signale getrennt empfangen, wenn für ihre Zeitdifferenz $|t_1 - t_2| \geq \tau > 0$ gilt. Wie groß ist die Wahrscheinlichkeit, dass die beiden Signale getrennt empfangen werden können? Wir haben

$$\Omega = \{(t_1, t_2) \in [0, T] \times [0, T]\}$$

und betrachten das Ereignis

$$A = \{(t_1, t_2) \in \Omega : |t_1 - t_2| \geq \tau\}.$$



Dann gilt

$$P(A) = \frac{\text{Fläche von } A}{\text{Fläche von } \Omega} = \frac{(T-\tau)^2}{T^2} = \left(\frac{T-\tau}{T}\right)^2 = \left(1 - \frac{\tau}{T}\right)^2.$$

1.2.4 Ergänzende Beispiele zur Einführung von Wahrscheinlichkeiten

Beispiel. In einem Rechnerpool befinden sich 75 Computer an festen Standorten. Genau einmal im Jahr wird der Pool modernisiert. Jeder Computer wird regulär nach zwei Jahren durch

ein moderneres Modell ersetzt. Falls jedoch ein Computer im ersten Jahr durch mindestens fünf erforderliche Reparaturen auffällt, so wird er bereits nach einem Jahr ausgetauscht. Das Ereignis A = „Computer fällt im ersten Jahr auf" habe die Wahrscheinlichkeit p und A_k = „Computer wird im k-ten Jahr ausgetauscht" habe die Wahrscheinlichkeit p_k. Wie groß ist die Wahrscheinlichkeit p_k, dass an einem fixierten Standort im k-ten Jahr ein Computeraustausch stattfindet?

Es gilt $A_k = \overline{A_{k-1}} \cup (A_{k-1} \cap A)$ und somit ist

$$\begin{aligned} p_k = P(A_k) &= P(\overline{A_{k-1}}) + P(A_{k-1} \cap A) = 1 - P(A_{k-1}) + \underbrace{P(A|A_{k-1})}_{p} P(A_{k-1}) \\ &= 1 - p_{k-1} + pp_{k-1} = 1 - (1-p)p_{k-1} = 1 + (p-1)p_{k-1}. \end{aligned}$$

Aus dieser Rekursionsvorschrift ergibt sich

$$\begin{aligned} p_1 &= p = 1 + (p-1) \qquad \text{(ohne Rekursionsformel)} \\ p_2 &= 1 + (p-1) + (p-1)^2 \\ p_3 &= 1 + (p-1) + (p-1)^2 + (p-1)^3 \\ &\vdots \\ p_k &= \sum_{i=0}^{k} (p-1)^i \qquad \text{(endliche Summe einer geometrischen Reihe)} \\ &= \frac{1-(p-1)^{k+1}}{1-(p-1)} = \frac{1-(p-1)^{k+1}}{2-p}. \end{aligned}$$

Für $p = 0{,}1$ ergeben sich die folgenden Wahrscheinlichkeiten.

k	1	2	3	4	5	10	20	21	∞
p_k	0,1	0,91	0,181	0,837	0,247	0,691	0,584	0,474	0,526

Beispiel. Wir betrachten die Geschlechterverteilung bei der Geburt von Zwillingen. Dabei sind die Ereignisse $K_1 \cap K_2$ (zwei Knaben), $K_1 \cap M_2$ (erst Knabe, dann Mädchen), $M_1 \cap M_2$ (zwei Mädchen) und $M_1 \cap K_2$ (erst Mädchen, dann Knabe) möglich. Es sind die folgenden statistischen Informationen bekannt:

- Die Wahrscheinlichkeit, dass bei einer Geburt ein Knabe zur Welt kommt, beträgt 51 %, d.h. $P(K_1) = P(K_2) = 0{,}51$.
- Bei Zwillingsgeburten ist die Wahrscheinlichkeit gleichgeschlechtlicher Zwillinge 64 %, d.h. $P((K_1 \cap K_2) \cup (M_1 \cap M_2)) = 0{,}64$.
- Bei einer Zwillingsgeburt sind $K_1 \cap M_2$ und $M_1 \cap K_2$ gleichwahrscheinliche Ereignisse, d.h. $P(K_1 \cap M_2) = P(M_1 \cap K_2)$.

Wie groß ist die Wahrscheinlichkeit, dass der zweite geborene Zwilling ein Knabe ist, wenn der zuerst geborene Zwilling auch ein Knabe war?

Wir haben $\Omega = (K_1 \cap K_2) \cup (M_1 \cap M_2) \cup (K_1 \cap M_2) \cup (M_1 \cap K_2)$. Zusammen mit der zweiten und dritten statistischen Information folgt daraus

$$P(K_1 \cap M_2) = P(M_1 \cap K_2) = \frac{1-0{,}64}{2} = 0{,}18$$

und unter Verwendung von $P(K_1) = P((K_1 \cap M_2) \cup (K_1 \cap K_2))$ und der ersten statistischen Information ergibt sich

$$P(K_1 \cap K_2) = P(K_1) - P(K_1 \cap M_2) = 0{,}51 - 0{,}18 = 0{,}33.$$

Somit ist unsere gesuchte Wahrscheinlichkeit

$$P(K_2|K_1) = \frac{P(K_1 \cap K_2)}{P(K_1)} = \frac{0{,}33}{0{,}51} = 0{,}647.$$

Weiterhin gilt

$$P(M_2|M_1) = 0{,}633, \qquad P(M_2|K_1) = 0{,}353, \qquad P(K_2|M_1) = 0{,}367.$$