Vorwort (zur zweiten Auflage)

Für die zweite Auflage der Induktiven Statistik wurde der gesamte Text durchgesehen und alle Lösungen überprüft. Dies führte zu Modifikationen und Ergänzungen des Stoffes sowie der Übungsaufgaben, um die Transparenz der Darstellung zu steigern und weitere Anwendungsfelder der induktiven Statistik aufzuzeigen.

Bei den Übungsaufgaben wurden entdeckte Fehler beseitigt. Zur besseren Kontrolle des Selbststudiums sind jetzt die Lösungen aller Aufgaben, mit Ausnahme der Beweise, in einem eigenen Abschnitt angegeben.

Für die Überarbeitung erhielt ich von meinen Mitarbeitern, Frau Diplom Kauffrau Eva Plinta und Herrn Dr. Stephan Popp wertvolle Anregungen. Herr cand. M.A. Christian Haggert führte jede Textänderung mit großer Sorgfalt und Geduld in LaTeX durch. Frau Sabine Pings koordinierte die Arbeitsabläufe für die textliche Umsetzung äußerst effizient. Ihnen allen gilt mein besonderer Dank.

Nicht zuletzt danke ich all jenen Studierenden, die während ihres Studiums mit dem Buch „gearbeitet“ und viele nützliche Hinweise gegeben haben.

Essen, im Frühjahr 2009 — Walter Assenmacher

Vorwort (zur ersten Auflage)

Die universelle Verwendbarkeit statistischer Methoden bei der quantitativen Analyse eröffnet ihnen ein breites Anwendungsspektrum, das jedoch wegen des teilweise recht großen Rechenaufwands und der mathematischen Komplexität noch immer nicht voll ausgeschöpft ist. Mit der Verfügbarkeit statistischer Programmpakete gewinnen die Methoden vor allem der Induktiven Statistik sowohl in der Praxis als auch in der empirischen Forschung immer größere Bedeutung. Dieser Herausforderung muss die akademische Ausbildung Rechnung tragen. Da es unmöglich ist, die statistischen Methoden in ihrer ganzen Breite zu behandeln, sollte das Ausbildungsziel in der Bereitstellung eines soliden methodischen Grundwissens liegen, das zu einer sachadäquaten Methodenauswahl und kompetenten Interpretation quantitativer Ergebnisse befähigt.

Um dieses Ziel zu erreichen, versucht das vorliegende Lehrbuch zunächst Verständnis für Struktur und Logik induktiver statistischer Methoden zu vermitteln. Grundlegend für die eingeschlagene Konzeption ist daher nicht die Präsentation einer Vielfalt rezeptartig bereitgestellter Formeln, sondern die schrittweise Entwicklung derjenigen Methoden, die heute zur universitären Ausbildung in Induktiver Statistik zählen. Wie auch im vorangegangenen Band „Deskriptive Statistik" wird der Darstellung der statistischen Methoden viel Raum gewidmet, um so ihre inhaltliche und formale Struktur transparent zu machen. Hierfür ist es von Vorteil, wenn die hochschulübliche mathematische Propädeutik beherrscht wird. Die Ausführungen sind jedoch so gestaltet, dass der mehr anwendungsorientierte Leser die formalen Nachweise bestimmter Eigenschaften übergehen kann, ohne dadurch den inhaltlichen Zusammenhang zu verlieren. Jedoch erleichtern gerade Kenntnisse der formalen Struktur den Methodenzugang und die Adaption neuer Verfahren.

In jedem Kapitel werden grundlegende Begriffe in Definitionen und hergeleitete Ergebnisse in Sätzen oder nummerierten Gleichungen festgehalten. Zur schnelleren Orientierung – besonders beim Nachschlagen – sind zentrale Begriffe da, wo sie erläutert werden, durch Fettdruck hervorgehoben. Eine große Zahl nummerierter Beispiele im Text dient der Verdeutlichung der Ausführungen und Berechnungen. Zur Selbstkontrolle des Wissenstandes enden die meisten Abschnitte mit Übungsaufgaben, deren Lösungen man bis auf wenige Ausnahmen in einem Kapitel am Ende des Buches findet. Die Übungsaufgaben sind so nummeriert, dass die erste Ziffer das Kapitel, die zweite den Abschnitt und die dritte Ziffer die laufende Nummer der Gleichung angibt.

Wie jedes Buch hat auch dieses von der Hilfe anderer profitiert. Die Zusammenstellung der Übungsaufgaben und ihre Lösungen wurden zum größten Teil von meinen Mitarbeitern, Herrn Diplom Volkswirt Andreas Kunert und Herrn Diplom Volkswirt Stephan Popp, betreut. Beide lasen auch das Manuskript und gaben Hinweise, die der Lesbarkeit zugute kommen. Herr cand. rer. pol. Oliver Murschall war für die Grafiken und die LaTeX–Umsetzung

der Formeln zuständig. Er löste diese Aufgabe mit außergewöhnlichem Engagement und großer Geduld bei Änderungswünschen. Frau Ursula Schapals fertigte den Text in LaTeX wieder in gewohnt sorgfältiger und zuverlässiger Weise an. Ihnen allen gilt mein besonderer Dank.

Fast jede Produktion verursacht auch negative externe Effekte. Meiner Frau danke ich dafür, dass sie diese auch jetzt wieder mit großer Geduld ertragen hat.

Schließlich danke ich Herrn Dr. Werner Müller vom Springer–Verlag für die erneut angenehme und verständnisvolle Zusammenarbeit.

Essen, im März 2000 Walter Assenmacher

2 Grundlegende Begriffe und Sätze der Wahrscheinlichkeitsrechnung

2.1 Zufallsexperiment, Stichprobenraum und Ereignisse

Viele Vorgänge der uns umgebenden Realität sind durch zwei Eigenschaften gekennzeichnet:

(a) Sie besitzen verschiedene, sich gegenseitig ausschließende Ausgänge,[1] die alle bereits vor Beginn des Vorgangs bekannt sind;

(b) es ist nicht vorhersehbar, welcher Ausgang tatsächlich eintreten wird.

Durch zusätzliche Information über den Ablauf solcher Vorgänge lässt sich möglicherweise der Grad der Ungewissheit reduzieren, aber nicht beseitigen. Ein so charakteristisches Geschehen, für das man die möglichen Ausgänge kennt ohne zu wissen, welcher davon eintreten wird, heißt **Zufallsvorgang** bzw. **stochastischer Vorgang**. Beispiele für Zufallsvorgänge lassen sich leicht finden. Der Ausgang eines Fußballspiels ist ebenso ein Zufallsvorgang wie der Kurs einer Aktie am nächsten Tag oder die realisierte Augenzahl eines Würfelwurfes.

Ist ein Zufallsvorgang unverändert beliebig oft wiederholbar, liegt ein **Zufallsexperiment** vor. Die unveränderte Wiederholbarkeit beschreibt man auch als „unter gleichen Bedingungen“ bzw. „unter gleichen Randbedingungen“ wiederholbar. Dies impliziert, dass die Randbedingungen wie bei naturwissenschaftlichen Experimenten kontrolliert werden können. Durch ihre Kontrollierbarkeit ist sichergestellt, dass die Bedingungen, unter denen ein Zufallsexperiment stattgefunden hat, auch bei weiteren Durchführungen hätten eingehalten werden können. Damit gehören alle Zufallsvorgänge, die in einem fiktiven Sinne unter gleichen Bedingungen wiederholbar sind, zu den Zufallsexperimenten. Dies erlaubt es, auch solche Zufallsvorgänge als Zufallsexperimente aufzufassen, deren praktische Wiederholung „unter gleichen Bedingungen“ recht schwierig wäre.

Von den drei oben gegebenen Beispielen ist der Zufallsvorgang „Fußballspiel“ nicht unter gleichen Bedingungen wiederholbar; er stellt daher kein

[1] Diese Ausgänge heißen oft auch Ergebnisse.

Zufallsexperiment dar. Anders verhält es sich bei den beiden anderen Zufallsvorgängen. Während beim Würfelwurf eine Wiederholung unter gleichen Bedingungen intuitiv klar ist, wird dies beim Vorgang „Tageskurs einer Aktie“ unter Bezug auf die relevante ökonomische Theorie „im Prinzip“ möglich. Beide Vorgänge gehören deshalb zu den Zufallsexperimenten.

Alle Ausgänge ω_i eines Zufallsvorganges bzw. -experimentes fasst man zu einem **Stichprobenraum** Ω zusammen.[2] Der Stichprobenraum ist formal eine Menge, deren Elemente die Ausgänge sind:

$$\Omega = \{\omega_1, \ldots \omega_m, \ldots\}\,. \tag{2.1}$$

Der Stichprobenraum Ω kann endlich oder unendlich viele Ausgänge enthalten. Lassen sich die unendlich vielen Ausgänge mit den natürlichen Zahlen $\mathbb{N}$ „abzählen“, bezeichnet man Ω als **abzählbar unendlich**;[3] gelingt dies nicht, heißt Ω **überabzählbar unendlich**. Der Zufallsvorgang „Werfen eines Würfels“ hat sechs mögliche Ausgänge; Ω ist endlich und lässt sich schreiben als:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_6\} = \{\omega_i, i = 1, \ldots, 6\} = \{1, 2, 3, 4, 5, 6\}.$$

Wirft man eine Münze, deren Seiten mit Zahl (Z) und Kopf (K) geprägt sind, so oft, bis zum ersten Mal der Ausgang „Zahl “ erscheint, lauten die möglichen Ausgänge:

$$\begin{aligned}
\omega_1 &= Z && \text{(zum ersten Mal „Zahl“ im ersten Wurf)},\\
\omega_2 &= KZ && \text{(zum ersten Mal „Zahl“ im zweiten Wurf)},\\
&\ \vdots\\
\omega_m &= \underbrace{K \ldots K}_{m-1}Z && \text{(zum ersten Mal „Zahl“ im } m\text{–ten Wurf)},\\
&\ \vdots
\end{aligned}$$

Der Stichprobenraum ist hier abzählbar unendlich:

$$\Omega = \{\omega_1, \ldots \omega_m, \ldots\}\,.$$

Die in Minuten gemessene Verspätung eines Zuges sei ein Zufallsvorgang mit Ausgängen des geschlossenen Intervalls [0 min, 10 min]. Bei unendlicher Messgenauigkeit sind überabzählbar viele Verspätungen möglich, da das Intervall

[2] Der Stichprobenraum wird auch Ergebnis- bzw. Ereignisraum genannt.
[3] Jeder endliche Stichprobenraum ist natürlich auch abzählbar.

[0;10] Teilmenge der reellen Zahlen $\mathbb{R}$ ist. Die reellen Zahlen sind, wie auch jede Teilmenge von ihnen, mächtiger als die natürlichen Zahlen und daher überabzählbar unendlich.[4]

Jede Teilmenge von Ω heißt **(Zufalls–)Ereignis**. Da eine Menge auch Teilmenge von sich selbst und die leere Menge $\emptyset$ Teilmenge jeder Menge ist, stellen die Mengen Ω und $\emptyset$ selbst Ereignisse von Ω dar. Ein Ereignis $A \subset \Omega$ tritt ein, wenn der Ausgang ω_i des Zufallsvorgangs Element von A ist: $\omega_i \in A$. Da ein Zufallsvorgang immer in einem Ausgang $\omega_i \in \Omega$ mündet, stellt Ω als Teilmenge interpretiert das sichere Ereignis dar. Analog hierzu heißt die leere Menge $\emptyset$ das unmögliche Ereignis, weil kein $\omega_i \in \Omega$ existiert, das Element der leeren Menge $\emptyset$ ist.

Teilmengen $\{\omega_i\}$, deren einziges Element ein Ausgang $\omega_i \in \Omega$ ist, heißen **Elementarereignisse**. Umfassen Teilmengen mehrere Ausgänge, nennt man sie **zusammengesetzte Ereignisse**. Z. B. ist beim „Wurf eines Würfels" der Ausgang: „Augenzahl 3 liegt oben" ein Elementarereignis und wird geschrieben als $\{3\}$; das Ereignis A: „gerade Augenzahl liegt oben" ist ein zusammengesetztes Ereignis, das als Menge in aufzählender Charakterisierung geschrieben lautet: $\{2, 4, 6\}$. Das Ereignis A ist eingetreten, wenn der Würfelwurf zu einer Augenzahl 2, 4 oder 6 führt.

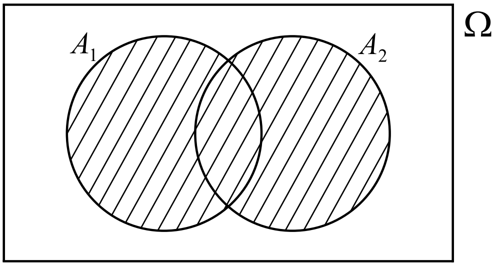
Die insgesamt möglichen Ereignisse eines Zufallsvorgangs findet man, indem alle Teilmengen für Ω gebildet werden. Die Zusammenfassung dieser Teilmengen führt bei endlichem oder abzählbar unendlichem Stichprobenraum Ω zur **Potenzmenge**, für die geschrieben wird: $PM(\Omega)$. Die Anzahl ihrer Elemente (=Teilmengen von Ω) beträgt 2^m, wobei m die Anzahl der Elemente des Stichprobenraumes Ω angibt. Damit liegt eine einfache Formel zur Berechnung der Anzahl möglicher Ereignisse für einen Zufallsvorgang mit endlichem Stichprobenraum Ω vor.

Beispiel 2.1 *Ein Zufallsvorgang hat den Stichprobenraum* $\Omega = \{1, 2, 3\}$*; wegen* $m = 3$ *beträgt die Anzahl der möglichen Ereignisse:* $2^3 = 8$*. Diese Ereignisse lauten:* $A_1 = \emptyset$*,* $A_2 = \{1\}$*,* $A_3 = \{2\}$*,* $A_4 = \{3\}$*,* $A_5 = \{1, 2\}$*,* $A_6 = \{1, 3\}$*,* $A_7 = \{2, 3\}$*,* $A_8 = \{1, 2, 3\} = \Omega$*. Die Potenzmenge wird gegeben durch:* $PM(\Omega) = \{A_1, \ldots, A_8\}$*.*

[4]Man sagt hierfür auch, die reellen Zahlen bilden ein Kontinuum.

Zwischen den möglichen Ereignissen können bestimmte Beziehungen vorliegen, die sinnvolle Unterscheidungen erlauben. Diese sollen am Beispiel 2.1 verdeutlicht werden. Das Ereignis $A_5 = \{1, 2\}$ tritt dann ein, wenn entweder das Ereignis $A_2 = \{1\}$ oder das Ereignis $A_3 = \{2\}$ eintritt. A_5 heißt daher Vereinigungsereignis und lässt sich schreiben als $A_5 = A_2 \cup A_3$. Verallgemeinert erhält man **Vereinigungsereignisse** V als: $V = \bigcup\limits_{j=1}^{n} A_j$. Für $n = 2$ ist V als schraffierte Fläche im Venn–Diagramm wiedergegeben, wobei das Rechteck den Stichprobenraum Ω festlegt.

Abb. 2.1: Vereinigungsereignis

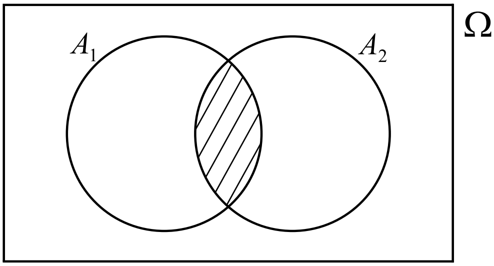


schraffierte Fläche: V.

Ist der Schnitt zweier Ereignisse A_1 und A_2 nicht leer, gilt also: $A_1 \cap A_2 = D \neq \emptyset$, so treten mit D auch die Ereignisse A_1 und A_2 ein. D heißt daher **Durchschnittsereignis**, das allgemein definiert ist als: $D = \bigcap\limits_{j=1}^{n} A_j$. D tritt ein, wenn alle A_j eintreten. Für $n = 2$ gibt Abbildung 2.2 das Durchschnittsereignis wieder.

Tritt das Ereignis A_1 genau dann ein, wenn A_2 nicht eintritt, so sind die beiden Ereignisse zueinander komplementär. A_1 heißt **Komplementärereignis** oder kurz Komplement und lässt sich schreiben als: $A_1 = \bar{A}_2$. Natürlich ist auch A_2 Komplementärereignis zu A_1: $A_2 = \bar{A}_1$. Im Beispiel 2.1 ist das Ereignis $A_2 = \{1\}$ das Komplement zu dem Ereignis $A_7 = \{2, 3\}$: $A_2 = \bar{A}_7$. Umgekehrt gilt auch: $A_7 = \bar{A}_2$. Wegen der in der Definition enthaltenen Bedingung „genau dann…“ muss für ein Ereignis A und für sein Komplement

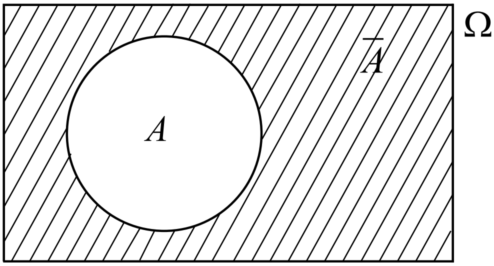
Abb. 2.2: Durchschnittsereignis



schraffierte Fläche: D.

$\bar{A}$ gelten: $A \cup \bar{A} = \Omega$ und $A \cap \bar{A} = \emptyset$. Dies kommt auch in Abbildung 2.3 zum Ausdruck.

Abb. 2.3: Komplementärereignis



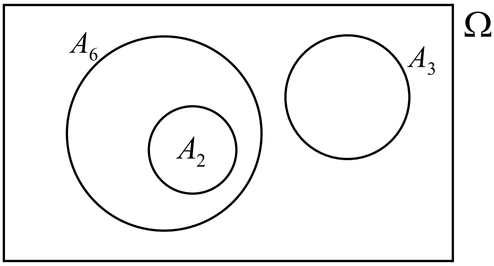
schraffierte Fläche: $\bar{A}$

Zwei Ereignisse A_1 und A_2 heißen **disjunkt**, wenn ihr Schnitt leer ist[5]: $A_1 \cap A_2 = \emptyset$. Komplementäre Ereignisse sind daher immer auch disjunkt, die Umkehrung gilt aber nicht. So sind im Beispiel 2.1 die Ereignisse $A_2 = \{1\}$ und $A_3 = \{2\}$ zwar disjunkt, aber nicht komplementär. Denn wenn A_3 nicht eintritt, folgt nicht notwendigerweise das Eintreten von A_2, sondern es könnte auch A_4=$\{3\}$ eintreten.

[5]In der Terminologie von Ereignissen lautet diese Definition: ... wenn ihr Durchschnittsereignis D das unmögliche Ereignis $\emptyset$ ist.

Ein Ereignis A_1 stellt ein **Teilereignis** von A_2 dar, wenn jeder Ausgang eines Zufallsvorgangs, der zu A_1 gehört, auch in A_2 liegt, A_2 aber mindestens einen Ausgang ω_i enthält, der nicht auch in A_1 enthalten ist. A_1 ist eine echte Teilmenge von A_2 : $A_1 \subset A_2$. In Abbildung 2.4 repräsentieren die Kreise die Ereignisse A_2, A_3 und A_6 des Beispiels 2.1.

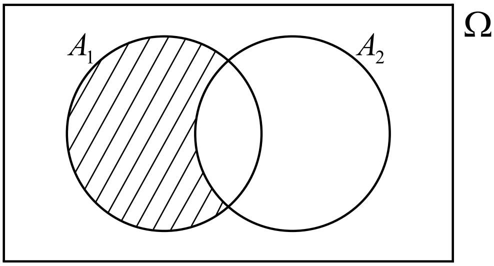
Abb. 2.4: Teilereignisse und disjunkte Ereignisse



A_2 : Teilereignis von A_6,
(A_2, A_3) und (A_3, A_6) : disjunkte Ereignisse.

Schließlich kann noch ein **Differenzereignis** $A_1 \setminus A_2$ definiert werden. Dieses Ereignis tritt dann ein, wenn der Ausgang des Zufallsvorgangs in A_1, aber nicht in A_2 liegt (vgl. Abbildung 2.5).

Abb. 2.5: Differenzereignis (relatives Komplement)



schraffierte Fläche: Differenzereignis.

Man nennt $A_1 \setminus A_2$ auch das **relative Komplement** zu A_2 bezüglich A_1. Das Differenzereignis lässt sich mengentheoretisch angeben als: $A_1 \setminus A_2 = A_1 \cap \bar{A}_2$. Im Beispiel 2.1 folgt für $A_6 = \{1,3\}$ und $A_5 = \{1,2\}$ das Differenzereignis $A_6 \setminus A_5$ als:

$$A_6 \setminus A_5 = A_6 \cap \bar{A}_5 = \{1,3\} \cap \{3\} = \{3\}.$$

Nur der Ausgang $\omega_3 = 3$ führt dazu, dass $A_6 \setminus A_5$ eintritt. Man sieht an diesem Beispiel leicht, dass im allgemeinen $A_1 \setminus A_2$ und $A_2 \setminus A_1$ verschiedene Ereignisse sind.

Jedes zusammengesetzte Ereignis A kann in disjunkte Teilereignisse $A_j \neq \emptyset$ so zerlegt werden, dass gilt: $A = \bigcup_{j=1}^{n} A_j$. Das Ereignis A ergibt sich jetzt als disjunkte Vereinigung. In Beispiel 2.1 lässt sich das Ereignis A_8 in $A_2 = \{1\}$ und $A_7 = \{2,3\}$, aber auch in $A_2 = \{1\}$, $A_3 = \{2\}$ und $A_4 = \{3\}$ zerlegen. Beim zweiten Fall wurde A_8 in Elementarereignisse zerlegt. Die Zerlegung eines Ereignisses A in Elementarereignisse heißt **kanonische Darstellung**: Jedes Ereignis ergibt sich eindeutig als Vereinigungsereignis von Elementarereignissen:[6] $A = \bigcup_{j=1}^{n} \{\omega_j\}$.

Da auch der Stichprobenraum Ω zu den zusammengesetzten Ereignissen gehört, kann auch er in (Teil–) Ereignisse A_j zerlegt werden. Gilt für diese Zerlegung:

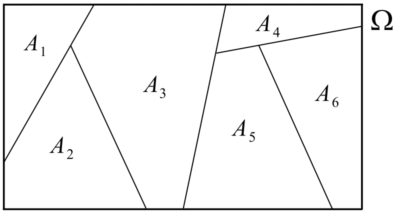
$$\begin{array}{ll} (1) & \Omega = A_1 \cup \ldots \cup A_n = \bigcup_{j=1}^{n} A_j, \\ (2) & A_i \cap A_j = \emptyset \quad \text{für } i \neq j, \\ (3) & A_j \neq \emptyset \quad \text{für } j = 1, \ldots, n, \end{array}$$

bildet die Menge $\{A_1, \ldots, A_n\}$ ein **vollständiges System von Ereignissen**, auch **vollständiges Ereignissystem** bzw. **vollständige Zerlegung** genannt. Abbildung 2.6 stellt ein vollständiges Ereignissystem des Stichprobenraumes Ω dar.

Für jeden Stichprobenraum existieren meistens mehrere vollständige Ereignissysteme. So lassen sich für den Stichprobenraum $\Omega = \{1,2,3\}$ des Beispiels

[6] Elementarereignisse selbst liegen bereits in kanonischer Darstellung vor.

Abb. 2.6: Vollständiges Ereignissystem



2.1 die vollständigen Ereignissysteme $\{A_1, A_2, A_3\}$, $\{A_1, A_5\}$, $\{A_2, A_6\}$ oder $\{A_3, A_4\}$ angeben.

Übungsaufgaben zu 2.1

2.1.1 Geben Sie den Stichprobenraum Ω für folgende Zufallsvorgänge an!

a) Dreimaliges Werfen einer Münze, wobei die Reihenfolge der Ausgänge beachtet wird.

b) Wurf dreier Münzen, wenn nur interessiert, wie oft „Kopf“ oben liegt.

c) Ziehen von vier Spielkarten aus einem Satz von 32 Karten, wenn die Anzahl an gezogenen „Damen“ gezählt wird.

2.1.2 Geben Sie für den Zufallsvorgang (b) der Aufgabe 2.1.1 die Potenzmenge $PM(\Omega)$ an! Wie nennt man in der Statistik die Elemente der Potenzmenge? Welche Ausgänge des Stichprobenraumes unter 2.1.1 (a) korrespondieren mit dem Element $\{0, 2\} \in PM(\Omega)$?

2.1.3 Für das Zufallsexperiment „Ziehen einer Spielkarte“ aus einem Satz mit 32 Karten sind folgende Ereignisse definiert:

$$A_1 = \text{gezogene Karte ist rot, } A_2 = \text{gezogene Karte ist ein König,}$$
$$A_3 = \text{gezogene Karte ist ein Ass.}$$

Zeigen Sie die Beziehungen zwischen den Ereignissen durch Aufzählen der Elemente und anhand eines Venn-Diagramms auf!

a) Geben Sie das Komplementärereignis zu A_1 an!

b) Bestimmen Sie folgende Durchschnittsereignisse: $A_1 \cap A_2$, $A_2 \cap A_3$ und $A_1 \cap A_2 \cap A_3$!

c) Wie sieht das Vereinigungsereignis von A_2 und A_3 aus?

d) Wie lautet das Differenzereignis von A_1 und $(A_2 \cup A_3)$?

2.1.4 Ein Stichprobenraum Ω ist definiert als die Menge der natürlichen Zahlen von 1 bis 10. Drei seiner Ereignisse lauten:

$$A_1 = \{1{,}3{,}7{,}8{,}10\},\ A_2 = \{5{,}6{,}9\},\ A_3 = \{2{,}4{,}5{,}6{,}9{,}10\}.$$

Kennzeichnen Sie die folgenden Ereignisse durch Aufzählen ihrer Elemente sowie mit Venn–Diagrammen!

a) $(A_1 \cup A_2) \cap A_3$,	*b)* $(A_1 \setminus A_2) \cap A_3$,	*c)* $A_1 \cup \bar{A}_1$,
d) $A_1 \cap A_2 \cap A_3$,	*e)* $(A_1 \cap A_3) \cup (A_2 \cap A_3)$,	*f)* $A_1 \cap \bar{A}_1$,
g) $(\bar{A}_1 \cap A_2) \cap A_3$,	*h)* $\overline{A_2 \cap A_3}$,	*i)* $A_2 \cap A_3$,
j) $\bar{A}_1 \cup \bar{A}_3$,	*k)* $\bar{A}_1 \cap \bar{A}_2$,	*l)* $(A_3 \setminus A_2) \cup A_1$!

Stellt die Menge $\{A_1, A_2, A_3\}$ ein vollständiges Ereignissystem dar?

2.1.5 Geben Sie die folgenden Ereignisse in Mengenschreibweise an!

a) „Ereignisse A und B treten ein, aber Ereignis C tritt nicht ein“,

b) „Ereignisse A, B und C treten ein“,

c) „Ereignisse A, B oder C treten ein“,

d) „Ereignisse B und A treten nicht ein, aber Ereignis C tritt ein“.

2.1.6 Die Ereignisse A und B sind nicht disjunkt. Zeigen Sie, dass $A \setminus B \cup (A \cap B)$ eine disjunkte Vereinigung für A und $B \setminus A \cup (A \cap B)$ eine disjunkte Vereinigung für B ist.

2.2 Ereignis– und σ–Algebra

Bei überabzählbar unendlichen Stichprobenräumen ist es notwendig, sich auf einen Teil der möglichen Ereignisse eines Zufallsvorgangs zu beschränken. Aber auch bei endlichen bzw. abzählbar unendlichen Stichprobenräumen ist eine solche Einschränkung vorteilhaft und bei vielen praktischen Fällen auch ausreichend, da selbst bei endlichem Stichprobenraum die Potenzmenge $PM(\Omega)$ sehr umfangreich sein kann.

Die Auswahl der Ereignisse zur Bildung eines Teilmengensystems muss so erfolgen, dass die mit den ausgewählten Ereignissen durchgeführten Mengenoperationen „Durchschnitt“ und „Vereinigung“ wiederum Ereignisse liefern, die ebenso wie die jeweiligen Komplementärereignisse zu dem Teilmengensystem gehören. Ein so gebildetes Teilmengensystem heißt **Boole'sche (Mengen-)Algebra** bzw. **Ereignisalgebra** $\mathcal{A}$ oder kurz **Algebra** und ist definiert als:

Definition 2.1 *Ein Teilmengensystem ist eine Ereignisalgebra $\mathcal{A}$, wenn folgende Bedingungen erfüllt sind:*

(1) Das sichere Ereignis Ω ist Element von $\mathcal{A}$: $\Omega \in \mathcal{A}$.

(2) Ist das Ereignis A Element von $\mathcal{A}$, so gilt dies auch für das Komplementärereignis $\bar{A}$: $A \in \mathcal{A} \to \bar{A} \in \mathcal{A}$.

(3) Sind zwei Ereignisse A_1 und A_2 Elemente von $\mathcal{A}$, dann ist auch ihre Vereinigung $A_1 \cup A_2$ Element von $\mathcal{A}$:

$$A_1, A_2 \in \mathcal{A} \to \mathcal{A}_1 \cup \mathcal{A}_2 \in \mathcal{A}.$$

Eine Ereignisalgebra weist vier wichtige Eigenschaften auf, die als Satz 2.1 zusammengefasst sind:

Satz 2.1 *(a) Das unmögliche Ereignis ist Element von $\mathcal{A}$: $\emptyset \in \mathcal{A}$.*

(b) Aus $A_1, A_2 \in \mathcal{A}$ folgt: $A_1 \cap A_2 \in \mathcal{A}$.

(c) Aus $A_1, A_2 \in \mathcal{A}$ folgt: $A_1 \setminus A_2 \in \mathcal{A}$.

(d) Sind A_j, $j = 1, \ldots, n$ Elemente von $\mathcal{A}$, so gilt dies auch für das Vereinigungsereignis $\bigcup\limits_{j=1}^{n} A_j$ und das Durchschnittsereignis $\bigcap\limits_{j=1}^{n} A_j$.

Die Aussagen (a) bis (d) dieses Satzes lassen sich leicht beweisen. Wegen der Bedingung (1) und (2) der Definition 2.1 ist das zu Ω komplementäre Ereignis $\bar{\Omega}$ Element von $\mathcal{A}$. Da die Vereinigung eines beliebigen Ereignisses mit seinem

Komplementärereignis immer das sichere Ereignis Ω ergibt, gilt: $\Omega \cup \bar{\Omega} = \Omega$ und somit $\bar{\Omega} = \emptyset$. Damit ist Aussagen (a) bewiesen.

Der Beweis der Aussage (b) kann wie folgt erbracht werden:
Aus $A_1, A_2 \in \mathcal{A}$ folgt wegen Bedingung (2): $\bar{A}_1, \bar{A}_2 \in \mathcal{A}$.Aus $\bar{A}_1, \bar{A}_2 \in \mathcal{A}$ folgt wegen Bedingung (3): $\bar{A}_1 \cup \bar{A}_2 \in \mathcal{A}$. Wegen des **de Morgan'schen Gesetzes**[7] gilt: $\bar{A}_1 \cup \bar{A}_2 = \overline{A_1 \cap A_2}$; das Ereignis $\overline{A_1 \cap A_2}$ gehört somit auch zu der Ereignisalgebra $\mathcal{A}$. Dann ist aber wegen Bedingung (2) der Definition 2.1 auch das hierzu komplementäre Ereignis $A_1 \cap A_2$ Element von $\mathcal{A}$.

Zum Beweis der Aussage (c) schreibt man für das Differenzereignis $A_1 \backslash A_2$: $A_1 \setminus A_2 = A_1 \cap \bar{A}_2$. Wegen $A_1 \in \mathcal{A}$ und $\bar{A}_2 \in \mathcal{A}$ folgt aus der bewiesenen Aussage (b): $A_1 \cap \bar{A}_2 \in \mathcal{A}$.

Aussage (d) schließlich kann durch sukzessive Anwendung von Bedingung (3) der Definition 2.1 gezeigt werden. Aus $A_1, A_2 \in \mathcal{A}$ folgt: $A_1 \cup A_2 \in \mathcal{A}$. Aus $(A_1 \cup A_2)$, $A_3 \in \mathcal{A}$ folgt: $(A_1 \cup A_2) \cup A_3 = A_1 \cup A_2 \cup A_3 \in \mathcal{A}$. Fährt man so fort, erhält man $\bigcup\limits_{j=1}^{n} A_j \in \mathcal{A}$. Der Beweis für $\bigcap\limits_{j=1}^{n} A_j$ sei dem Leser als Übung überlassen.

Für alle Elemente einer Ereignisalgebra gilt somit, dass ihre mengentheoretisch definierten Verknüpfungen zu Vereinigungs-, Durchschnitts- und Differenzereignisssen wiederum Ereignisse der Ereignisalgebra sind. Man bezeichnet daher eine Ereignisalgebra als abgeschlossen.

Eine Ereignisalgebra heißt **Sigma–Algebra**, in Kurzschreibweise **$\boldsymbol{\sigma}$–Algebra**, wenn Bedingung (3) der Definition 2.1 für abzählbar unendlich viele Ereignisse verallgemeinert wird. Sie geht dann über in:

(3a) Ist $A_j \in \mathcal{A}$ für $j = 1, 2, \ldots$, dann gilt auch $\bigcup\limits_{j=1}^{\infty} A_j \in \mathcal{A}$.

Jede σ–Algebra ist auch eine Ereignisalgebra,[8] da mit Gültigkeit der Bedingung (3a) auch Bedingung (3) erfüllt ist. Satz 2.1 gilt daher auch für σ–Algebren.

Für jeden Zufallsvorgang lassen sich stets mehrere Algebren angeben. Aus Definition 2.1 folgt, dass die Potenzmenge auch immer eine Algebra ist. Sie

[7] Die Regeln von de Morgan lauten: (1) $\overline{A_1 \cap A_2} = \bar{A}_1 \cup \bar{A}_2$ und (2) $\overline{A_1 \cup A_2} = \bar{A}_1 \cap \bar{A}_2$.
[8] Die Umkehrung gilt jedoch nicht.

stellt jedoch eine Algebra mit der größten Anzahl an Elementen, d.h. Ereignissen dar. Es kann leicht selbst gezeigt werden, dass für jeden Zufallsvorgang das Teilmengensystem $\{\emptyset, \Omega\}$ ebenfalls eine Algebra ist, jetzt jedoch mit der geringsten Anzahl an Elementen. Sollen in einer Algebra aber bestimmte Ereignisse, z.B. A_1 und A_2 vorkommen, und sind diese nicht das sichere oder unmögliche Ereignis, scheidet die Algebra $\{\emptyset, \Omega\}$ aus. Eine Algebra, die A_1 und A_2 enthält und zudem noch die geringste Anzahl weiterer Ereignisse aufweist, heißt **Borel'sche Algebra**. Um für bestimmte, vorgegebene Ereignisse eine Borelsche Algebra zu finden, müssen die Komplementärereignisse hinzugefügt und dann mit allen Ereignissen die möglichen Vereinigungs- und Durchschnittsereignisse gebildet werden. Nur bei bestimmten Zufallsvorgängen ist die Potenzmenge, obwohl sie die vorgegebenen Ereignisse als Elemente enthält, auch eine Borel'sche Algebra.

Beispiel 2.2 *Das Zufallsexperiment „Werfen einer Münze" hat die beiden Ausgänge Kopf* (K) *oder Zahl* (Z)*. Der Stichprobenraum wird gegeben durch:* $\Omega = \{K, Z\}$*. Die Potenzmenge* $PM(\Omega)$ *enthält* $2^2 = 4$ *Ereignisse; diese sind:* $\emptyset, \Omega, \{K\}, \{Z\}$*.* $PM(\Omega) = \{\emptyset, \Omega, \{K\}, \{Z\}\}$ *stellt genau wie* $\{\emptyset, \Omega\}$ *eine Algebra dar. Es soll nun die Borel'sche Algebra erzeugt werden, die das Ereignis* $\{K\}$ *enthält. Hierzu muss das zu* $\{K\}$ *komplementäre Ereignis* $\{Z\}$ *aufgenommen werden. Mit den Ereignissen* $\{K\}$ *und* $\{Z\}$ *sind alle möglichen Vereinigungs- und Durchschnittsereignisse zu bilden. Man erhält:* $\{K\} \cup \{Z\} = \{K, Z\} = \Omega$ *und* $\{K\} \cap \{Z\} = \emptyset$*. Die Borel'sche Algebra lautet:* $\{\emptyset, \Omega, \{K\}, \{Z\}\}$ *und ist hier mit der Potenzmenge identisch.*

Die Ergebnisse des Beispiels 2.2 fasst Satz 2.2 zusammen:

Satz 2.2 *Enthält der Stichprobenraum* Ω *nur zwei Elemente* ω_1 *und* ω_2*, existieren für ihn auch nur die beiden Ereignisalgebren* $\{\emptyset, \Omega\}$ *und* $PM(\Omega)$*, wobei die Potenzmenge* $PM(\Omega)$ *eine Borel'sche Algebra für das Elementarereignis* $\{\omega_i\}$*,* $i = 1, 2$ *ist.*

Beispiel 2.3 *Ein Zufallsexperiment lautet: „Zweimaliges Werfen einer Münze". Als Stichprobenraum erhält man:* $\Omega = \{(K, K), (K, Z), (Z, K), (Z, Z)\}$*, wobei das Elementarereignis z.B.* (K, K) *bedeutet, dass sowohl beim ersten als auch beim zweiten Wurf der Münze die Seite mit dem Kopf oben liegt. Die Potenzmenge enthält bereits* $2^4 = 16$ *Ereignisse, die aufzuzählen dem Leser als Übung überlassen bleibt. Es seien* $A_1 = \{(K, K)\}$*,* $A_2 =$

$\{(K,Z),(Z,K),\ (Z,Z)\}$, $A_3 = \{(K,Z),(Z,K)\}$, $A_4 = \{(K,K),(Z,Z)\}$, $A_5 = \{(K,K),(K,Z),(Z,K)\}$ und $A_6 = \{(Z,Z)\}$. Die Menge $\mathcal{A} = \{\emptyset, \Omega, A_1, A_2, A_3, A_4, A_5, A_6\}$ stellt eine Algebra dar. Die Borel'sche Algebra für A_1 jedoch wird gegeben durch: $\{\emptyset, \Omega, A_1, A_2\}$ und enthält weniger Elemente als die Menge $\mathcal{A}$.

Ein Zufallsvorgang kann jetzt durch den Stichprobenraum Ω und durch eine geeignete Algebra $\mathcal{A}$ beschrieben werden. Das Paar $(\Omega, \mathcal{A})$ heißt **Messraum**. Für abzählbare Stichprobenräume wird die umfassendste Ereignisalgebra durch die Potenzmenge $PM(\Omega)$ gegeben. Je nachdem, an welchen Ereignissen eines Zufallsvorgangs man interessiert ist, lässt sich $PM(\Omega)$ durch eine Algebra mit weniger Elementen, z.B. durch die Borel'sche Algebra ersetzen. Bei überabzählbaren Stichprobenräumen, wie das z.B. bei $\Omega = \mathbb{R}$ der Fall ist, muss eine σ-Algebra aus offenen, halboffenen und geschlossenen Intervallen der reellen Zahlen bestehen.[9] Im Allgemeinen finden bei überabzählbar unendlichen Stichprobenräumen Borel'sche Algebren Verwendung.

Übungsaufgaben zu 2.2

2.2.1 Der Stichprobenraum eines Zufallsvorgangs wird gegeben durch $\Omega = \{1,2,3,4\}$.

a) Geben Sie für diesen Zufallsvorgang eine Ereignisalgebra an !

b) Stellen Sie die Borel'sche Algebra auf, die das Ereignis $B = \{1,2\}$ enthält !

c) Warum ist das vollständige Ereignissystem $\{A_1, A_2, A_3, A_4\}$ mit $A_j = \{j\}$, $j = 1, \ldots, 4$ keine Ereignisalgebra ?

d) Begründen Sie die Verallgemeinerung der Aussage unter (c), dass ein vollständiges Ereignissystem definitionsgemäß nie eine Ereignisalgebra sein kann !

[9] Ein Intervall der reellen Zahlen mit den Grenzen $a < b$ heißt offen, wenn gilt: $a < x < b$, halboffen bei $a \leq x < b$ bzw. $a < x \leq b$ und geschlossen für $a \leq x \leq b$.

2.3 Der Wahrscheinlichkeitsbegriff

Da es bei Zufallsvorgängen immer ungewiss ist, welches ihrer möglichen Ereignisse tatsächlich eintritt, wäre die Angabe von Zahlen nützlich, um die Chance des Eintretens zu quantifizieren. Diese Zahlen heißen Wahrscheinlichkeitsmaße oder kurz **Wahrscheinlichkeiten** und werden mit P symbolisiert. Stellt $A \subset \Omega$ ein Ereignis dar, gibt $P(A)$ jetzt die Wahrscheinlichkeit für das Eintreten des Ereignisses A an. Es handelt sich bei P also um eine Funktion, die den Ereignissen reelle Zahlen zuordnet,[10] die Wahrscheinlichkeiten heißen. Diese Zahlenzuordnung zur Quantifizierung der Chance kann natürlich nicht willkürlich geschehen, sondern muss bestimmten Grundanforderungen genügen. Besteht ein Zufallsvorgang z.B. darin, aus einem Topf mit 6 weißen und 4 roten Kugeln zufällig eine Kugel zu entnehmen, so ist die Chance, eine weiße Kugel zu erhalten, größer als die Chance, eine rote Kugel zu ziehen. Es leuchtet in diesem Fall unmittelbar ein, sich bei der Quantifizierung der Chancen auf die relativen Häufigkeiten h zu beziehen, die hier $h(\text{weiß}) = 0,6$ und $h(\text{rot}) = 0,4$ betragen. Diese beiden Zahlen vermitteln eine bessere intuitive Vorstellung von den Chancen der beiden Ereignisse als z.B. die Zahlenzuordnung 30 und 20.

Die axiomatische Grundlage für Wahrscheinlichkeiten wurde von Kolmogoroff entwickelt. Auf der Basis dieser Axiomatik lässt sich die Wahrscheinlichkeitsfunktion wie folgt definieren.

Definition 2.2 *Eine auf einer Algebra $\mathcal{A}$ definierte Funktion $P\colon \mathcal{A} \longrightarrow \mathbb{R}$ heißt* **Wahrscheinlichkeitsfunktion***, wenn sie die folgenden drei Axiome, die* **Kolmogoroff–Axiome** *heißen, erfüllt:*

(1) $P(A) \geq 0$ *für alle* $A \in \mathcal{A}$,

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

(3) $P(\bigcup_{j=1}^{\infty} A_j) = P(A_1) + P(A_2) + \ldots = \sum_{j=1}^{\infty} P(A_j)$
für alle A_i und A_j, die paarweise disjunkt sind: $A_i \cap A_j = \emptyset$, $i \neq j$.

Das Axiom (1) besagt, dass die Wahrscheinlichkeit nie negativ werden kann; Axiom (2) normiert die Wahrscheinlichkeit. Axiom (3), auch **Voll-**

[10] Da Ereignisse immer Teilmengen des Stichprobenraumes Ω sind, bezeichnet man P auch als **Mengenfunktion** da formal Mengen in die reellen Zahlen abgebildet werden.

bzw. $\boldsymbol{\sigma}$**–Additivität** genannt, gibt die Berechnung der Wahrscheinlichkeit für die Vereinigung paarweise disjunkter Ereignisse als Summe der Einzelwahrscheinlichkeiten an. Jedes Ereignis $A \subset \Omega$ wird somit durch P in das geschlossene Intervall $[0,1] \subset \mathbb{R}$ abgebildet. Ein Zufallsvorgang lässt sich nun durch das Tripel $(\Omega, \mathcal{A}, P)$ charakterisieren, das **Wahrscheinlichkeitsraum** heißt. Definition 2.2 bleibt auch dann gültig, wenn Bedingung (3) nur in schwächerer Form als $P(\bigcup\limits_{j=1}^{n} A_j) = \sum\limits_{j=1}^{n} P(A_j)$ eingeführt wird.

Obwohl aus den Axiomen von Kolmogoroff nicht folgt, wie groß die Wahrscheinlichkeit für das Eintreten bestimmter Ereignisse eines Zufallsvorgangs ist, lassen sich dennoch aus ihnen bestimmte Rechenregeln für Wahrscheinlichkeiten ableiten, die im Satz 2.3 zusammengefasst sind.

Satz 2.3 *Es sei P eine Wahrscheinlichkeitsfunktion für eine Algebra $\mathcal{A}$. Dann gilt:*

(a) $P(A) + P(\bar{A}) = 1,$

(b) $P(A) \leq P(B)$ *für* $A \subset B,$

(c) $P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum\limits_{j=1}^{n} P(A_j)$ *für paarweise disjunkte Ereignisse,*

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

Die Rechenregeln 2.3.a bis 2.3.d lassen sich mit den Kolmogoroff–Axiomen beweisen, die hier mit $(K1)$, $(K2)$ und $(K3)$ abgekürzt werden.

Regel (a): Aus $A \cup \bar{A} = \Omega$ und $A \cap \bar{A} = \emptyset$ gilt wegen (K2) und (K3): $P(A \cup \bar{A}) = P(\Omega) = 1 = P(A) + P(\bar{A})$.

Regel (b): Ist A ein Teilereignis von B, dann gilt: $\bar{A} \cap B \neq \emptyset$. Das Ereignis B kann daher zerlegt werden in: $B = A \cup (\bar{A} \cap B)$. Da A und $(\bar{A} \cap B)$ disjunkt sind, gilt wegen (K3): $P(B) = P(A) + P(\bar{A} \cap B)$ oder: $P(A) = P(B) - P(\bar{A} \cap B)$ bzw. $P(A) \leq P(B)$, da $P(\bar{A} \cap B) \geq 0$.

Regel (c): Diese Regel folgt direkt aus (K3) für eine endliche Anzahl paarweise disjunkter Ereignisse.

Regel (d): Das Ereignis $(A \cup B)$ lässt sich in drei paarweise disjunkte Ereignisse zerlegen:

$$(A \cup B) = (A \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap B).$$

Wegen (K3) gilt:
($\star$) $P(A \cup B) = P(A \cap B) + P(A \cap \bar{B}) + P(\bar{A} \cap B)$.
Die Ereignisse A und B lassen sich ebenfalls in paarweise disjunkte Ereignisse zerlegen:
$A = (A \cap B) \cup (A \cap \bar{B})$ und $B = (A \cap B) \cup (\bar{A} \cap B)$. Somit gilt für $P(A \cap \bar{B})$ bzw. $P(\bar{A} \cap B)$:

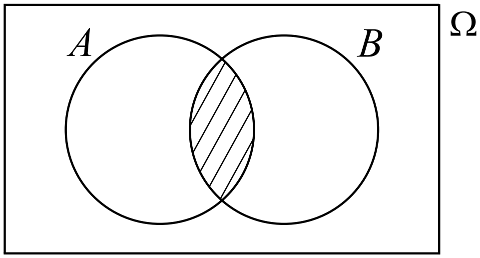
$$\begin{aligned} P(A \cap \bar{B}) &= P(A) - P(A \cap B) \quad \text{und} \\ P(\bar{A} \cap B) &= P(B) - P(A \cap B). \end{aligned}$$

Ersetzt man in Gleichung ($\star$) die entsprechenden Terme durch diese Ergebnisse, folgt:

$$\begin{aligned} P(A \cup B) &= P(A \cap B) + P(A) - P(A \cap B) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

Regel 2.3.d, die **Additionssatz** der Wahrscheinlichkeitsrechnung heißt, kann anhand eines Venn–Diagramms veranschaulicht werden.

Abb. 2.7: Veranschaulichung des Additionssatzes



Die Ereignisse A und B sind nicht disjunkt, da $A \cap B \neq \emptyset$; das Durchschnittsereignis $A \cap B$ entspricht der schraffierten Fläche in Abbildung 2.7. Die Berechnung der Wahrscheinlichkeit für das Vereinigungsereignis $(A \cup B)$ mit $P(A)+P(B)$ würde die Wahrscheinlichkeit für das Ereignis $A \cap B$ doppelt erfassen; folglich muss $P(A \cap B)$ subtrahiert werden. Analog hierzu berechnet man die Wahrscheinlichkeit für das Vereinigungsereignis $(A \cup B \cup C)$ dreier nicht disjunkter Ereignisse A, B und C als:

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

Bereits diese Erweiterung zeigt, dass der Additionssatz recht aufwendig wird. Seine Verallgemeinerung für abzählbar viele Ereignisse A_j, $j = 1, \ldots, n$ (Formel von Sylvester) soll daher hier entfallen.[11] Bei paarweise disjunkten Ereignissen A_j, $j = 1, \ldots, n$ entspricht der Additionssatz der Regel (c) des Satzes 2.3.

Aus dem **Additionssatz** 2.3.c **für disjunkte Ereignisse** folgt eine einfache Vorschrift für die Berechnung der Wahrscheinlichkeit eines Ereignisses $A = \{\omega_1, \omega_2, \ldots, \omega_n\}$ eines abzählbar unendlichen Stichprobenraumes Ω. In kanonischer Darstellung lässt sich A schreiben als Vereinigung von Elementarereignissen: $A = \{\omega_1\} \cup \{\omega_2\} \cup \ldots \cup \{\omega_n\}$. Da die Elementarereignisse $\{\omega_j\}$, $j = 1, \ldots, n$, paarweise disjunkt sind, folgt:

Satz 2.4 $P(A) = P(\{\omega_1\}) + P(\{\omega_2\}) + \ldots + P(\{\omega_n\}) = \sum\limits_{\omega_j \in A} P(\{\omega_j\}).$

Mit den Axiomen von Kolmogoroff und den Regeln (a) bis (d) des Satzes 2.3 sind nur die allgemeinen, formalen Eigenschaften der Wahrscheinlichkeiten festgelegt, nicht jedoch, welche Werte sie bei bestimmten Ereignissen eines Zufallsvorgangs annehmen. Hierzu muss erst eine Entscheidung über die dem Zufallsvorgang adäquate Wahrscheinlichkeitsinterpretation getroffen werden. Wahrscheinlichkeiten lassen sich subjektiv oder objektiv begründen. Die **subjektive Wahrscheinlichkeit** ist ein Maß für den Grad der Überzeugtheit einer Person, dass ein Ereignis A eintritt. Sie heißt daher auch **assertorische Wahrscheinlichkeit**. Da die subjektive Wahrscheinlichkeit personenabhängig ist, können demselben Ereignis verschiedene (subjektive) Wahrscheinlichkeiten zugeordnet werden. Dies ist auch der Haupteinwand gegen die Verwendung subjektiver Wahrscheinlichkeiten. Dennoch kommt diese Interpretation in der Realität oft vor. Subjektive Wahrscheinlichkeiten liegen stets Wetten zugrunde. Ist z.B. jemand bereit, auf das Eintreten des Ereignisses A im Verhältnis von $a : 1$ zu wetten, beträgt die subjektive

[11] Der daran interessierte Leser sei auf Schlittgen (1996), S. 8 ff. verwiesen.

Wahrscheinlichkeit $P_S(A)$ mindestens $\frac{a}{1+a}$. Alternativ kann auch der Einsatz einer Wette zur Ermittlung der subjektiven Wahrscheinlichkeiten dienen. Erhält man bei Eintritt des Ereignisses A eine Auszahlung in Höhe von G und bei Eintreten des Komplementärereignisses $\bar{A}$ nichts, gilt für den maximalen Einsatz E, den man für dieses Spiel zu zahlen bereit ist: $E = P_S(A)G + [1 - P_S(A)] \cdot 0 = P_S(A)G$ oder $P_S(A) = \frac{E}{G}$.

Auch bei der Beurteilung ökonomischer Entwicklungen werden subjektive Wahrscheinlichkeiten vereinzelt herangezogen. So verwendet das Ifo–Institut in München einen Konjunkturtest, der auf der subjektiven Einschätzung befragter Unternehmen basiert.[12]

Bei der **objektiven Wahrscheinlichkeitsinterpretation** können zwei Richtungen unterschieden werden: die a–priori und die statistische bzw. frequentistische Interpretation der Wahrscheinlichkeit. Die a–priori Wahrscheinlichkeit kann weiter in die klassische bzw. Laplace–Wahrscheinlichkeit und in die geometrische Wahrscheinlichkeit unterteilt werden. Die verschiedenen möglichen Unterteilungen der objektiven Wahrscheinlichkeit sind in Abbildung 2.8 zusammengestellt.

Können bei Zufallsvorgängen die Wahrscheinlichkeiten für das Eintreten der Ereignisse allein durch logische Schlüsse berechnet werden, liegt eine **a–priori Interpretation des Wahrscheinlichkeitsbegriffs** vor. Die Wahrscheinlichkeiten lassen sich angeben, ohne dass der Zufallsvorgang jemals durchgeführt werden müsste. Die **geometrische Konzeption der Wahrscheinlichkeit** basiert auf dieser Interpretation. Hierbei sind die Wahrscheinlichkeiten durch das Verhältnis geometrischer Figuren definiert, z.B. durch Flächen-, Winkel- oder Streckenverhältnisse. So könnte man in Abbildung 2.7 die Wahrscheinlichkeit für das Ereignis A als das Verhältnis der Kreis- zur Rechteckfläche berechnen. Analog hierzu ergeben sich die Wahrscheinlichkeiten für die Ereignisse A, B und C, dargestellt als Kreissektoren eines Glücksrades (siehe Abbildung 2.9), als Winkelverhältnisse. Da dem Ereignis C ein Winkel von 90° entspricht und die Winkel der Ereignisse A und B gleich groß sind, betragen die geometrischen Wahrscheinlichkeiten: $P(A) = P(B) = \frac{3}{8}$ und $P(C) = \frac{1}{4}$.

Obwohl das geometrische Wahrscheinlichkeitskonzept auch für überabzählbar unendliche Stichprobenräume geeignet ist, existieren nur wenige praktische

[12] Zum Aufbau dieses Tests siehe Assenmacher (1998b), S. 27 ff.

Abb. 2.8: Wahrscheinlichkeitsinterpretationen

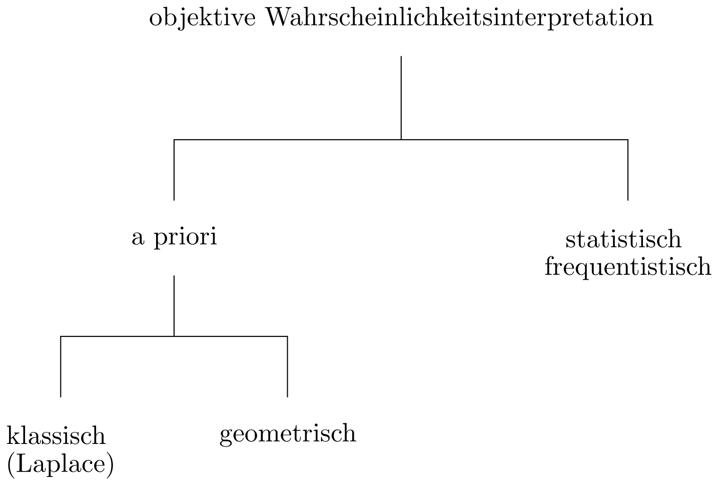
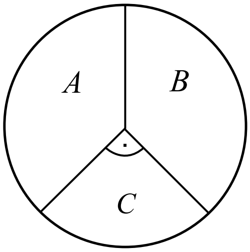


Abb. 2.9: Glücksrad



Zufallsvorgänge, bei denen die Überführung der Ereignisse in geometrische Figuren gelingt.

Die klassische Interpretation der Wahrscheinlichkeit, zu Ehren ihres Begründers auch **Laplace–Wahrscheinlichkeit** genannt, findet bei Zufallsvorgängen Anwendung, deren Stichprobenraum endlich ist und deren Elementarereignisse gleiche Chancen besitzen, einzutreten. Zufallsvorgänge mit dieser Eigenschaft heißen **Laplace–Experimente**. Die Wahrscheinlichkeit für ein beliebiges Elementarereignis $\{\omega_i\}$, $i = 1, \ldots, m$ beträgt dann $P(\{\omega_i\}) = \frac{1}{m}$. Nach Satz 2.4 erhält man die Wahrscheinlichkeit $P(A)$ für ein Ereignis A als:

$$P(A) = \frac{\text{Anzahl der für } A \text{ günstigen Ausgänge}}{\text{Anzahl der möglichen Ausgänge}}.$$

Die Anzahl der Elemente einer Menge M heißt **Mächtigkeit** und wird mit $|M|$ symbolisiert. Die Formel der klassischen bzw. Laplace–Wahrscheinlichkeit lautet dann:

$$P(A) = \frac{|A|}{|\Omega|}.$$

Laplace–Experimente können als die zufällige Entnahme aus einer endlichen Menge von Objekten charakterisiert werden. Die Vorschrift „zufällig" sichert, dass alle Objekte die gleiche Chance haben, gezogen zu werden. Eine solche Vorgehensweise wird als „**zufälliges Ziehen aus einer Urne**" oder als „**Laplace–Urnen–Modell**" bezeichnet. Geschieht das Ziehen mehrmals hintereinander, muss noch zwischen „Ziehen mit Zurücklegen" oder „Ziehen ohne Zurücklegen" unterschieden werden.

Beispiel 2.4 *(a) In einer Urne befinden sich 20 Kugeln, von denen 8 rot sind. Die Wahrscheinlichkeit, eine rote Kugel bei einer zufälligen Entnahme zu erhalten, beträgt* $P(rot) = \frac{8}{20} = 0,4$. *Beim „Ziehen mit Zurücklegen" bleibt diese Wahrscheinlichkeit von Zug zu Zug gleich. Wird die entnommene Kugel nach dem ersten Zug nicht wieder in die Urne zurückgelegt, befinden sich nur noch 19 Kugeln in der Urne, die aber bei zufälligem Ziehen wieder die gleiche Chance haben, entnommen zu werden. Wurde im ersten Zug eine rote Kugel gezogen, sind von den 19 Kugeln 7 rot, und die klassische Wahrscheinlichkeit beträgt jetzt* $P(rot) = \frac{7}{19}$.

(b) Ein **Laplace–Würfel**[13] *wird geworfen. Die Wahrscheinlichkeit, dass eine Augenzahl größer als 2 oben liegt, beträgt:*

$$P(Augenzahl > 2) = \frac{|\{3,4,5,6\}|}{|\{1,2,3,4,5,6\}|} = \frac{4}{6} = \frac{2}{3}.$$

Obwohl die klassische Definition der Wahrscheinlichkeit intuitiv überzeugt, bleibt sie doch logisch unbefriedigend. Die Definition basiert auf „gleichmöglichen" Elementarereignissen, also auf dem zu definierenden Begriff selbst

[13] Ein Laplace–Würfel ist so gefertigt, dass alle Seiten die gleiche Chancen haben, oben zu liegen. Man bezeichnet ihn auch als „**idealen**" bzw. „**fairen**" **Würfel**.

und ist daher zirkulär. Auch muss angemerkt werden, dass die Voraussetzungen zur Berechnung der klassischen Wahrscheinlichkeit bei vielen praktischen Problemen nicht vorliegen: Entweder lassen sich die Wahrscheinlichkeiten nicht durch logische Schlüsse aus einem konkreten Sachverhalt ableiten oder die Elementarereignisse besitzen keine gleichen Eintrittswahrscheinlichkeiten.

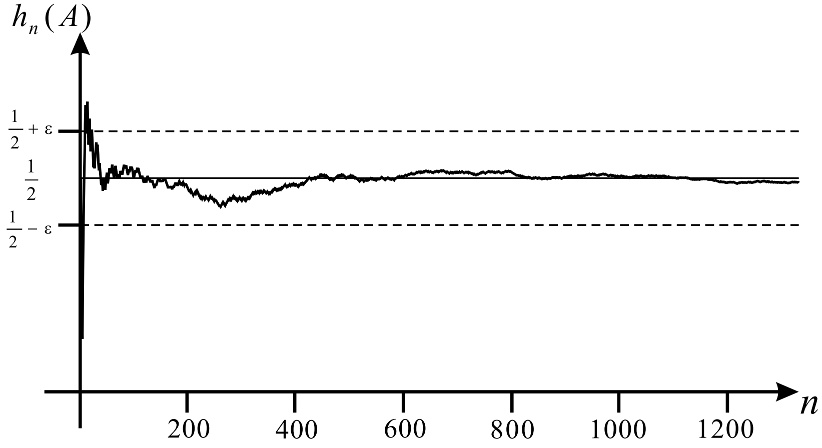
Bei der **statistischen** bzw. **frequentistischen Interpretation** werden die Wahrscheinlichkeiten nicht aufgrund logischer Schlüsse, sondern erst nach einer, in der Regel großen Anzahl von Durchführungen eines Zufallsexperiments quantifiziert. Es handelt sich somit um eine **a–posteriori** Bestimmung der Wahrscheinlichkeit. Diese Interpretation der Wahrscheinlichkeit verdeutlicht das folgende Experiment. Eine Laplace–Münze wird n-mal geworfen. Für jeden Wurf lautet der Stichprobenraum $\Omega = \{Z, K\}$ mit Z: Zahl und K: Kopf. Das Ereignis A sei definiert als $A = \{Z\}$; seine Wahrscheinlichkeit lässt sich a–priori mit $\frac{1}{2}$ angeben. Die Anzahl der Würfe, für die das Ereignis A eintritt, wird mit $n(A)$ bezeichnet; die relative Häufigkeit für A beträgt dann: $h_n(A) = \dfrac{n(A)}{n}$. Wird die Münze sehr oft geworfen, schwankt die relative Häufigkeit nur noch geringfügig um den Wert 0,5 (siehe Abbildung 2.10).

Dieses empirisch bestätigte stabile Verhalten der relativen Häufigkeiten nutzte Richard von Mises (1883 – 1953) zur Definition der **statistischen Wahrscheinlichkeit**: Strebt die Anzahl n der Durchführungen eines Zufallsexperimentes gegen unendlich und konvergiert die relative Häufigkeit $h_n(A)$ eines Ereignisses A gegen die Zahl p_A, so stellt p_A die statistische Wahrscheinlichkeit des Ereignisses A dar. Formal schreibt man diese Definition als:

$$P(A) = \lim_{n\to\infty} h_n(A) = p_A. \tag{2.2}$$

Lassen sich bei Zufallsexperimenten den Ereignissen nicht wie beim obigen Beispiel a–priori Wahrscheinlichkeiten zuordnen, ist p_A unbekannt. Da unendlich viele Wiederholungen des Zufallsexperimentes unmöglich sind, dient die für endlich viele Durchführungen berechnete relative Häufigkeit zur Schätzung der unbekannten Wahrscheinlichkeit. Dass dies auch theoretisch die richtige Vorgehensweise ist, wird in Kapitel 6.1 gezeigt.

Abb. 2.10: Relative Häufigkeiten beim Werfen einer fairen Münze



Übungsaufgaben zu 2.3

2.3.1 Die Ereignisse A, B und $A \cap B$ haben folgende Wahrscheinlichkeiten:

$$P(A) = 0,7;\ P(B) = 0,5 \text{ und } P(A \cap B) = 0,4.$$

Berechnen Sie:

a) $P(A \cup B)$, *b*) $P(\bar{A} \cup \bar{B})$, *c*) $P(\bar{A} \cap \bar{B})$,

d) $P(\bar{A} \cup B)$, *e*) $P(A \cap \bar{B})$, *f*) $P(\overline{A \cup B})$!

2.3.2 Prüfen Sie, ob die Wahrscheinlichkeitsangaben $P(A) = 0,5$, $P(B) = 0,2$ und $P(A \cap B) = 0,3$ richtig sind!

2.3.3 a) Bestimmen Sie die Wahrscheinlichkeit $P(A \cup B)$ mit $A = \{1\}$ und $B = \{3,4\}$ für das Zufallsexperiment „Werfen eines idealen Würfels“!

b) Bestimmen Sie die Wahrscheinlichkeit $P(A \cup B)$ mit $A = \{1\}$ und $B = \{3,4\}$ für das Zufallsexperiment „Zweimaliges Werfen eines idealen Würfels“, wobei sich das Ereignis A auf den ersten Wurf und Ereignis B auf den zweiten Wurf bezieht!

2.3.4 Ein idealer Würfel wird zweimal geworfen. Die Ereignisse A, B und C sind definiert als:

A: Augenzahl 6 im ersten Wurf, B: gerade Augenzahl im zweiten Wurf, C: Summe aus Augenzahl im ersten und zweiten Wurf ist größer oder gleich 10.

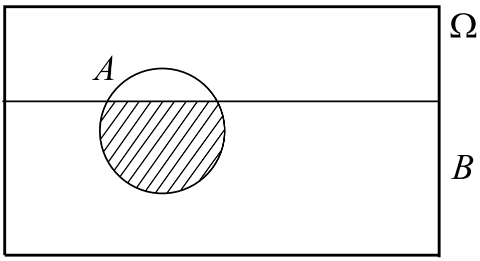
a) Berechnen Sie $P(A), P(B)$ und $P(C)$!

b) Wie groß ist die Wahrscheinlichkeit, dass wenigstens eins der drei Ereignisse eintritt?

2.4 Bedingte Wahrscheinlichkeit, stochastische Unabhängigkeit und Multiplikationssätze

Die Berechnung von Wahrscheinlichkeiten erfolgt bislang unter Bezug auf den ganzen Stichprobenraum Ω. Es lassen sich aber auch dann Wahrscheinlichkeiten für ein Ereignis A berechnen, wenn nicht mehr der gesamte Stichprobenraum, sondern nur noch ein Teil davon relevant ist. Abbildung 2.11 verdeutlicht die Veränderung des Bezugssystems. Das Ereignis A wird durch den

Abb. 2.11: Änderung des Stichprobenraumes



Kreis, das Ereignis B durch das untere Rechteck[14] und das Durchschnittsereignis $A \cap B$ durch das schraffierte Kreissegment wiedergegeben. $P(A)$, $P(B)$ und $P(A \cap B)$ sind die Wahrscheinlichkeiten für die Ereignisse A, B und $A \cap B$, wenn der Stichprobenraum Ω zugrunde liegt. Es kann aber auch die Wahrscheinlichkeit für das Eintreten des Ereignisses A unter der Bedingung

[14] Da B offensichtlich Teilmenge von Ω ist, stellt B auch ein zu Ω gehörendes Ereignis dar.

berechnet werden, dass nur noch die Ausgänge von Interesse sind, für die auch das Ereignis B eintritt. Diese Wahrscheinlichkeit wird mit $P(A|B)$ bezeichnet und heißt **bedingte Wahrscheinlichkeit**.

Bedingten Wahrscheinlichkeiten liegt nicht mehr der Stichprobenraum Ω, sondern ein neuer Stichprobenraum zugrunde, der durch die Bedingung, hier durch das Ereignis B, gegeben wird. Die Auswirkung des Stichprobenraumwechsels lässt sich an Abbildung 2.11 leicht nachvollziehen. Verwendet man das geometrische Wahrscheinlichkeitskonzept, erhält man $P(A)$ als Verhältnis der Kreisfläche zur Fläche des Rechtecks Ω; die bedingte Wahrscheinlichkeit $P(A|B)$ jedoch als Verhältnis der schraffierten Fläche des Kreissegmentes zur Fläche des Rechtecks B. Da die bedingte Wahrscheinlichkeit die Bedingung vorraussetzt, nennt man sie auch a–posteriori Wahrscheinlichkeit.

Beim Stichprobenraum B stellt das Ereignis B das sichere Ereignis dar, für dessen Wahrscheinlichkeit gelten muss: $P(B|B) = 1$. Um den Wert 1 zu erhalten, dividiert man $P(B)$ einfach durch $P(B)$. Analog hierzu berechnet man die Wahrscheinlichkeit für das Ereignis $A \cap B$, unter der Bedingung, dass B der Stichprobenraum ist, als: $P(A \cap B|B) = P(A \cap B)/P(B)$. Da $[(A \cap B)|B]$ und $(A|B)$ dieselben Ereignisse sind, ist eine einfache Formel für das Berechnen bedingter Wahrscheinlichkeiten gefunden:

Satz 2.5 *Es seien A, $B \subset \Omega$ Ereignisse und $(\Omega, \mathcal{A}, P)$ der Wahrscheinlichkeitsraum, $P(B) > 0$. Für die* **bedingte Wahrscheinlichkeit** *$P(A|B)$ gilt:*

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

Ist $P(B) = 0$, ist B das unmögliche Ereignis; dann kann aber auch nicht das Ereignis $A|B$ eintreten. Seine Wahrscheinlichkeit wird daher festgelegt als: $P(A|B) = 0$ für $P(B) = 0$. Der Wechsel des Bezugssystems von Ω zu B bei der Wahrscheinlichkeitsberechnung kommt auch bei der Angabe des Wahrscheinlichkeitsraumes für bedingte Ereignisse zum Ausdruck. Dieser lautet jetzt $\left(B, \mathcal{A}_B, \dfrac{P(\cdot \cap B)}{P(B)}\right)$, wobei $\mathcal{A}_B$ eine für den Stichprobenraum B definierte Algebra mit der hierauf definierten bedingten Wahrscheinlichkeit $P(\cdot \cap B)/P(B)$ ist.

Die Axiome von Kolmogoroff gelten auch bei bedingten Wahrscheinlichkeiten. Während bereits bei der Herleitung der bedingten Wahrscheinlichkeit von den Axiomen (1) und (2) der Definition 2.2 Gebrauch gemacht wurde, lautet das 3. Axiom für $A_1, A_2, B \subset \Omega$ mit $A_1 \cap A_2 = \emptyset$ und $P(B) > 0$:

$$P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B).$$

Für $A_2 = \bar{A}_1$ erhält man:

$$P(A_1 \cup \bar{A}_1|B) = P(A_1|B) + P(\bar{A}_1|B) = 1.$$

Beispiel 2.5 *a) Ein idealer Würfel wird geworfen. Stichprobenraum und Wahrscheinlichkeiten lauten:* $\Omega = \{1, 2, \ldots, 6\}$ *und* $P(\{\omega_i\}) = \frac{1}{6}$ *für* $i = 1, \ldots, 6$. *A ist das Ereignis, eine „1“ zu würfeln; B das Ereignis, eine ungerade Augenzahl zu erhalten. Das Ereignis $A \cap B$ tritt ein, wenn nach dem Wurf die „1“ oben liegt. Die Wahrscheinlichkeiten für diese Ereignisse betragen:* $P(A) = \frac{1}{6}$, $P(B) = \frac{1}{2}$ *und* $P(A \cap B) = \frac{1}{6}$. *Die Wahrscheinlichkeit, eine „1“ zu würfeln unter der Bedingung, dass eine ungerade Augenzahl eintritt, lässt sich direkt ermitteln. Der durch die Bedingung gegebene Stichprobenraum lautet:* $B = \{1, 3, 5\}$, *also gilt:* $P(A|B) = \frac{1}{3}$. *Denselben Wert erhält man nach Satz 2.5:*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$

b) In einer Urne liegen n Kugeln, von denen $k < n$ die Eigenschaft A und $m < n$ die Eigenschaft B aufweisen; i Kugeln besitzen beide Eigenschaften. Die Wahrscheinlichkeiten, eine Kugel mit der Eigenschaft A, B oder $A \cap B$ zu ziehen, betragen: $P(A) = \frac{k}{n}$, $P(B) = \frac{m}{n}$ *und* $P(A \cap B) = \frac{i}{n}$. *Die Wahrscheinlichkeit, eine Kugel mit der Eigenschaft A zu erhalten unter der Bedingung, dass sie auch die Eigenschaft B besitzt, beträgt* $P(A|B) = \frac{i}{m}$, *da m Kugeln die Eigenschaft B besitzen, von denen i Kugeln noch die Eigenschaft A aufweisen. Erweitert man den Bruch mit $\frac{1}{n}$, folgt:*

$$P(A|B) = \frac{i}{m} = \frac{\frac{i}{n}}{\frac{m}{n}} = \frac{P(A \cap B)}{P(B)} \quad (\textit{Satz } 2.5).$$

An dieser Berechnung erkennt man auch, dass für $P(B|A)$ gelten muss:

$$P(B|A) = \frac{i}{k} = \frac{\frac{i}{n}}{\frac{k}{n}} = \frac{P(A \cap B)}{P(A)}.$$

Die bedingte Wahrscheinlichkeit $P(A|B)$ lässt sich als ein Maß für die Wahrscheinlichkeit des Eintretens von A unter der zusätzlichen Information interpretieren, dass B bereits eingetreten ist. Übt diese Zusatzinformation keinen Einfluß auf die Eintrittswahrscheinlichkeit von A aus, gilt: $P(A|B) = P(A)$; A ist unabhängig von B. Dann ist aber auch B unabhängig von A, wie folgende Umformung zeigt:

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} \quad \text{(wegen Satz 2.5)} \\ &= \frac{P(A)P(B)}{P(A)} = P(B) \quad \text{(wegen } P(A|B) = P(A)). \end{aligned}$$

Die Wahrscheinlichkeit für B ist von der Bedingung A unabhängig. Diese, über die Wahrscheinlichkeiten festgelegte Eigenschaft zweier Ereignisse heißt paarweise stochastische Unabhängigkeit. Hieraus lässt sich eine einfache Überprüfungsmöglichkeit auf stochastische Unabhängigkeit gewinnen, die in der Definition 2.3 festgehalten ist.

Definition 2.3 *Zwei Ereignisse A und B heißen* **paarweise stochastisch unabhängig***, wenn eine der drei folgenden Bedingungen erfüllt ist:*

$$\begin{aligned} &(1) \quad P(A \cap B) = P(A)P(B), \\ &(2) \quad P(A|B) = P(A) \quad \textit{falls } P(B) > 0, \\ &(3) \quad P(B|A) = P(B) \quad \textit{falls } P(A) > 0; \end{aligned}$$

andernfalls bezeichnet man sie als **stochastisch abhängig**.

Beispiel 2.6 *Die Ereignisse des Beispiels 2.5.a sind stochastisch abhängig, da $P(A \cap B) = \frac{1}{6} \neq P(A)P(B) = \frac{1}{6} \cdot \frac{1}{2}$.*

Beispiel 2.7 *Ein idealer Würfel wird zweimal geworfen. Das Ereignis A ist festgelegt als „Augenzahl 1 im ersten Wurf“; das Ereignis B lautet „Augenzahl 2 im zweiten Wurf“. Der Stichprobenraum umfasst 36 geordnete Zahlenpaare, weil zu unterscheiden ist, wann welche Augenzahl eintritt:*

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

Für die Ereignisse erhält man:

$$\begin{aligned}
A &= \{(1,1)\dots(1,6)\},\ P(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6},\\
B &= \{(1,2)\dots(6,2)\},\ P(B) = \frac{|B|}{|\Omega|} = \frac{6}{36} = \frac{1}{6},\\
A \cap B &= \{(1,2)\} \qquad ,\ P(A\cap B) = \frac{|A\cap B|}{|\Omega|} = \frac{1}{36}.
\end{aligned}$$

Da gilt: $P(A \cap B) = \frac{1}{36} = P(A)P(B) = \frac{1}{6}\cdot\frac{1}{6}$, folgt paarweise stochastische Unabhängigkeit.

Die Verallgemeinerung der stochastischen Unabhängigkeit für mehr als zwei Ereignisse gibt Definition 2.4 wieder.

Definition 2.4 *Es seien $A_j \in \mathcal{A}$, $j = 1, 2, \dots n$ Ereignisse und $(\Omega, \mathcal{A}, P)$ der Wahrscheinlichkeitsraum. Die Ereignisse A_j, $j = 1, 2, \dots n$ heißen* **gemeinsam bzw. vollständig stochastisch unabhängig***, wenn für jede Teilmenge $\{j_1, j_2, \dots, j_m\} \subset \{1, 2, \dots, n\}$ gilt:*

$$P(A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_m}) = P(A_{j_1})P(A_{j_2}) \cdot \ldots \cdot P(A_{j_m}).$$

Diese formale Definition besagt anschaulich, dass alle Schnittereignisse, die mit den Ereignissen A_j, $j = 1, \dots, n$ gebildet werden können, auf stochastische Unabhängigkeit zu untersuchen sind. Das folgende Beispiel verdeutlicht diese Definition und zeigt, dass aus paarweiser stochastischer Unabhängigkeit nicht immer auch gemeinsame stochastische Unabhängigkeit folgt.[15]

Beispiel 2.8 *Ein idealer Würfel wird zweimal geworfen. Das Ereignis A liegt vor, wenn beim ersten Wurf eine ungerade Augenzahl realisiert wird; das Ereignis B tritt bei einer ungeraden Augenzahl im zweiten Wurf ein. Das Ereignis C bedeutet eine ungerade Summe der Augenzahl beider Würfe. Der Stichprobenraum ist hier derselbe wie beim Beispiel 2.7; die Ereignisse lassen sich daher angeben als:*

$$\begin{aligned}
A &= \{(1,j),(3,j),(5,j)|j = 1,2,\dots,6\},\\
B &= \{(i,1),(i,3),(i,5)|i = 1,2,\dots,6\},\\
C &= \{(i,j)|i+j:\ \textit{ungerade}\}.
\end{aligned}$$

[15] Paarweise stochastische Unabhängigkeit ist somit eine notwendige, aber keine hinreichende Bedingung für gemeinsame bzw. vollständige stochastische Unabhängigkeit. Sind Ereignisse gemeinsam stochastisch unabhängig, dann sind sie es auch paarweise; die Umkehr jedoch gilt nicht.

Die Wahrscheinlichkeiten für A und B sind leicht zu berechnen als: $P(A) = P(B) = \frac{18}{36} = \frac{1}{2}$. *Eine ungerade Summe* $i + j$ *tritt nur dann ein, wenn* i *oder* j, *aber nicht beide, ungerade sind. Insgesamt gibt es 18 solcher Summen, also:* $P(C) = \frac{18}{36} = \frac{1}{2}$. *Die Ereignisse* A, B *und* C *wären gemeinsam bzw. vollständig stochastisch unabhängig, wenn gemäß der Definition 2.4 neben der paarweisen stochastischen Unabhängigkeit noch gilt:* $P(A \cap B \cap C) = P(A)P(B)P(C)$. *Die Schnittereignisse* $A \cap B$, $A \cap C$ *und* $B \cap C$ *kommen im Stichprobenraum jeweils 9-mal vor; ihre Wahrscheinlichkeiten stimmen daher überein und betragen* $\frac{9}{36} = \frac{1}{4}$. *Die drei Ereignisse sind paarweise stochastisch unabhängig, denn es gilt z.B. für* $P(A \cap B) : P(A \cap B) = \frac{1}{4} = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2}$. *Das Ereignis* $A \cap B \cap C$ *ist das unmögliche Ereignis, da eine ungerade Augenzahl im ersten und zweiten Wurf immer zu einer geraden Summe führt. Daher folgt:*

$$P(A \cap B \cap C) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2};$$

die Ereignisse A, B *und* C *sind nicht vollständig stochastisch unabhängig.*

Mit dem Konzept der bedingten Wahrscheinlichkeiten lassen sich nützliche Sätze der Wahrscheinlichkeitsalgebra gewinnen. Die Auflösung von Satz 2.5 nach $P(A \cap B)$ heißt Multiplikationssatz für zwei Ereignisse, wobei noch zu unterscheiden ist, ob sie stochastisch abhängig sind oder nicht. Satz 2.6 gibt die Regeln an.

Satz 2.6 Multiplikationssatz für zwei Ereignisse
Es seien $A, B \subset \Omega$ *Ereignisse mit* $P(A) > 0$ *und* $P(B) > 0$. *Dann gilt*

a) bei stochastischer Abhängigkeit:

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

b) bei stochastischer Unabhängigkeit:

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

Die Verallgemeinerung des Satzes 2.6 für n Ereignisse $A_j \subset \Omega$, $j = 1, \ldots, n$ führt zu Satz 2.7.

Satz 2.7 Multiplikationssatz für n Ereignisse
Es seien $A_j \subset \Omega$, $j = 1, \ldots, n$ *Ereignisse mit* $A_1 \cap A_2 \cap \ldots \cap A_{n-1} \neq \emptyset$. *Dann gilt*

a) bei stochastischer Abhängigkeit:

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

b) bei gemeinsamer stochastischer Unabhängigkeit:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2) \cdot \ldots \cdot P(A_n).$$

Jede Wahrscheinlichkeit $P(B) > 0$ kann auf bedingte Wahrscheinlichkeiten zurückgeführt werden. Wie dies geschieht, zeigt der Satz von der totalen Wahrscheinlichkeit.

Satz 2.8 Satz von der totalen Wahrscheinlichkeit
Es seien $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $\{A_1, \ldots, A_n\}$, mit $A_j \in \mathcal{A}$ für $j = 1, \ldots, n$ ein vollständiges System von Ereignissen. Für jedes Ereignis $B \in \mathcal{A}$ gilt dann:

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \ldots + P(B|A_n)P(A_n) \\ &= \sum_{j=1}^{n} P(B|A_j)P(A_j). \end{aligned}$$

Der Beweis dieses Satzes verdeutlicht seine Anwendungsmöglichkeiten. Das Ereignis B lässt sich disjunkt zerlegen in:

$$\begin{aligned} &B = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n) \text{ mit} \\ &(B \cap A_i) \cap (B \cap A_j) = \emptyset \text{ für } i \neq j \text{ und } i, j = 1, \ldots, n. \end{aligned}$$

Abbildung 2.12 zeigt eine solche Zerlegung für das Ereignissystem $\{A_1, \ldots, A_4\}$.[16]

Nach dem Additionssatz für disjunkte Ereignisse erhält man:

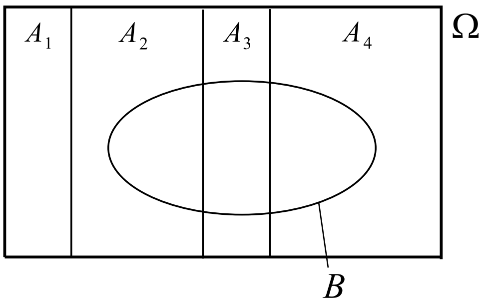
$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_n);$$

wegen Satz 2.5 gilt: $P(B \cap A_j) = P(B|A_j)P(A_j), j = 1, \ldots, n$ und daher:

$$P(B) = \sum_{j=1}^{n} P(B|A_j)P(A_j).$$

[16] Bei einer Zerlegung kann durchaus vorkommen: $B \cap A_1 = \emptyset$, wie dies in Abbildung 2.12 bei $A_1 \cap B$ der Fall ist.

Abb. 2.12: Zerlegung eines Ereignisses B



Die „totale" Wahrscheinlichkeit $P(B)$ ist gleich dem gewogenen arithmetischen Mittel der bedingten Wahrscheinlichkeiten $P(B|A_j)$ und den Gewichten $P(A_j)$.

Mit den Sätzen 2.5 und 2.6 lässt sich eine weitere Regel gewinnen, vorausgesetzt es gilt $P(B) > 0$:

$$\begin{aligned} P(A_i|B) &= \frac{P(A_i \cap B)}{P(B)} \quad \text{(nach Satz 2.5)} \\ &= \frac{P(B|A_i)P(A_i)}{P(B)} \quad \text{(nach Satz 2.6)} \end{aligned}$$

Ersetzt man im letzten Bruch $P(B)$ durch Satz 2.8, folgt:

Satz 2.9 Bayes'sches Theorem

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum\limits_{j=1}^{n} P(B|A_j)P(A_j)} \quad \textit{für alle } i = 1, \ldots, n.$$

Im nächsten Beispiel kommen die Sätze zur Anwendung.

Beispiel 2.9 *Bei der Fertigung von Sicherungen sind 10% der Produkte defekt. Ein zur Qualitätskontrolle durchgeführter Test zeigt bei fehlerhaften Sicherungen in 95%, bei fehlerfreien Sicherungen in 15% der Fälle einen Defekt an. Wie groß ist die Wahrscheinlichkeit, dass*

a) eine Sicherung, für die der Test einen Defekt anzeigt, auch tatsächlich defekt ist, *b) eine Sicherung fehlerfrei ist, wenn dies der Test anzeigt?*

Der Stichprobenraum Ω *besteht hier aus den Ausgängen* ω_1 *= fehlerhafte Sicherung und* ω_2 *= fehlerfreie Sicherung; die Ereignisse* $A_1 = \{\omega_1\}$ *und* $A_2 = \{\omega_2\}$ *bilden ein vollständiges System von Ereignissen. Das Ereignis* B *lautet: Test zeigt Fehler an. Damit sind folgende Wahrscheinlichkeiten bekannt:*

$$P(A_1) = 0,1,\ P(A_2) = 0,9,\ P(B|A_1) = 0,95,\ P(\bar{B}|A_1) = 0,05$$
$$P(B|A_2) = 0,15 \text{ und } P(\bar{B}|A_2) = 0,85.$$

Die (unbedingte=totale) Wahrscheinlichkeit, dass der Test einen Fehler anzeigt, erhält man mit dem Satz von der totalen Wahrscheinlichkeit als:

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = 0,95 \cdot 0,1 + 0,15 \cdot 0,9 \\ &= 0,23. \end{aligned}$$

Die unter a) gesuchte Wahrscheinlichkeit für das Ereignis $A_1|B$ *beträgt nach dem Theorem von Bayes:*

$$\begin{aligned} P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} = \frac{P(B|A_1)P(A_1)}{P(B)} \\ &= \frac{0,95 \cdot 0,1}{0,23} = 0,413. \end{aligned}$$

Für die unter b) gesuchte Wahrscheinlichkeit $P(A_2|\bar{B})$ *benötigt man die Wahrscheinlichkeit des Komplementärereignisses* $\bar{B}$*. Diese beträgt:* $P(\bar{B}) = 1 - P(B) = 0,77$*. Mithin gilt nach dem Bayesschen Theorem:*

$$P(A_2|\bar{B}) = \frac{P(\bar{B}|A_2)P(A_2)}{P(\bar{B})} = \frac{0,85 \cdot 0,9}{0,77} = 0,994.$$

$P(\bar{B})$ *erhält man auch über den Satz von der totalen Wahrscheinlichkeit:*

$$P(\bar{B}) = P(\bar{B}|A_1)P(A_1) + P(\bar{B}|A_2)P(A_2) = 0,05 \cdot 0,1 + 0,85 \cdot 0,9 = 0,77.$$

Das Ereignis $A_1 \cap B$ *liegt vor, wenn eine fehlerhafte Sicherung eintritt* <u>*und*</u> *der Test dieses anzeigt;*[17] *die Wahrscheinlichkeit hierfür berechnet man nach dem Multiplikationssatz 2.6.a als:*

$$P(A_1 \cap B) = P(B|A_1)P(A_1) = 0,95 \cdot 0,1 = 0,095, \text{ oder:}$$
$$P(A_1 \cap B) = P(A_1|B)P(B) = 0,413 \cdot 0,23 = 0,095.$$

[17] Man mache sich die Unterschiede der Ereignisse $A_1 \cap B$, $A_1|B$ und $B|A_1$ ganz klar!

Die auf den ersten Blick geringe Wahrscheinlichkeit ist dadurch zu erklären, dass fehlerhafte Sicherungen nur bei 10% der Produktion vorliegen. Schließlich kann noch gezeigt werden, dass die Ereignisse B und $B|A_1$ stochastisch abhängig sind, denn: $P(B) = 0,23 \neq P(B|A_1) = 0,95$.

Das Theorem von Bayes gibt an, wie die Wahrscheinlichkeit für ein Ereignis A nach Vorliegen des Ergebnisses eines Experimentes oder einer Beobachtung neu berechnet werden kann. Hierzu wird das vollständige Ereignissystem A_j, $j = 1, ..., n$ als Menge sich ausschließender Hypothesen oder disjunkter Zustände interpretiert. Die Wahrscheinlichkeit $P(A_j)$ der Zustände ist aufgrund von Ausgangsinformationen bestimmt; sie stellt daher die a–priori Wahrscheinlichkeit für jedes A_j dar. Mit dem Bayes'schen Theorem erhält man jetzt die a–posteriori Wahrscheinlichkeiten für A_j, wenn die Beobachtung B bzw. das experimentelle Ergebnis B vorliegt. Das Eintereten des Ereignisses B bedeutet eine zusätzliche Information, die genutzt werden kann.

Wird ein Stichprobenraum Ω in zwei vollständige Ereignissysteme A_i, $i = 1, \ldots, m$ und B_j, $j = 1, \ldots, n$ zerlegt, können insgesamt mn Durchschnittsereignisse $(A_i \cap B_j)$ gebildet werden. Ihre Wahrscheinlichkeiten lassen sich übersichtlich in einer Wahrscheinlichkeitstabelle zusammenfassen (vgl. Tabelle 2.1). In der Vorspalte steht die Zerlegung A_i, in der Kopfzeile die Zerlegung B_j; die Wahrscheinlichkeiten $P(A_i \cap B_j)$ der Durchschnittsereignisse stellen die Feldwerte der Tabelle dar.

Weil A_i, $i = 1, \ldots, m$ und B_j, $j = 1, \ldots, n$ jeweils vollständige Ereignissysteme sind, muss jede Zeilensumme die Wahrscheinlichkeit für A_i, $P(A_i) = \sum\limits_{j=1}^{n} P(A_i \cap B_j)$ und jede Spaltensumme die Wahrscheinlichkeit für B_j, $P(B_j) = \sum\limits_{i=1}^{m} P(A_i \cap B_j)$ sein. Im Zusammenhang mit Wahrscheinlichkeitstabellen bezeichnet man diese Wahrscheinlichkeiten als Randwahrscheinlichkeiten.[18]

[18] In Kapitel 5 wird auf Randwahrscheinlichkeiten genauer eingegangen.

Tab. 2.1: Wahrscheinlichkeitstabelle

	B_1	B_2	$\ldots$	B_j	$\ldots$	B_n	$P(A_i)$
A_1							$P(A_1)$
A_2							$P(A_2)$
$\vdots$							$\vdots$
A_i				$P(A_i \cap B_j)$			$P(A_i)$
$\vdots$							$\vdots$
A_m							$P(A_m)$
$P(B_j)$	$P(B_1)$	$P(B_2)$	$\ldots$	$P(B_j)$	$\ldots$	$P(B_n)$	

Übungsaufgaben zu 2.4

2.4.1 a) Sind disjunkte Ereignisse A und B mit $P(A) > 0$ und $P(B) > 0$ auch stochastisch unabhängig?

b) Es seien A und B zwei stochastisch unabhängige Ereignisse. Zeigen Sie, dass dann auch die Ereignispaare $(A, \bar{B}), (\bar{A}, B)$ und $(\bar{A}, \bar{B})$ stochastisch unabhängig sind!

2.4.2 Beweisen Sie mit Satz 2.6, dass für drei Ereignisse A, B und C gilt: $P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C)$!

2.4.3 Aus einem Satz von 32 Spielkarten wird zufällig eine Karte gezogen.

a) Wie groß ist die Wahrscheinlichkeit für das Eintreten der folgenden Ereignisse:

A: es wird eine Dame gezogen,

B: es wird eine rote Karte gezogen,

C: es wird kein Kreuz gezogen,

D: es wird Herz König gezogen?

b) Untersuchen Sie die Ereignisse auf paarweise und gemeinsame stochastische Unabhängigkeit!

2.4.4 Drei Urnen sind mit Kugeln gefüllt. In der Urne A befinden sich 3 rote und 5 grüne Kugeln. Die Urne B enthält 5 rote und 3 grüne und in der Urne C sind 2 rote und 6 grüne Kugeln. Um zu entscheiden, aus welcher Urne Kugeln entnommen werden, wird eine Kugel aus einer vierten Urne gezogen, die 5 Kugeln für A, 4 für B und 3 für C enthält. Wie groß ist die Wahrscheinlichkeit, dass

a) man aus der Urne B zieht?

b) wenn man aus der Trommel A zieht, eine rote Kugel erhält?

c) man eine grüne Kugel zieht?

d) man aus der Urne C zieht und eine rote Kugel erhält?

e) wenn man eine rote Kugel zieht, sie aus Urne A kommt?

2.4.5 Am Düsseldorfer Flughafen kommen 50% der landenden Maschinen aus Hamburg, 30% aus Berlin und 20% aus München. Man weiß, dass 10% der Maschinen aus Hamburg, 12% der aus Berlin und 8% der aus München Verspätung haben. Wie groß ist die Wahrscheinlichkeit, dass

a) eine landende Maschine verspätet eintrifft,

b) ein Flugzeug zu spät ankommt, wenn es in Berlin abgeflogen ist,

c) eine Maschine aus München kommt und keine Verspätung hat,

d) wenn eine Maschine Verspätung hat, sie aus Hamburg kommt?

2.4.6 In einer Urne befinden sich 3 schwarze und 2 weiße Kugeln. Wie groß ist die Wahrscheinlichkeit, nach dreimaligem Ziehen ohne Zurücklegen

a) eine weiße Kugel im zweiten Zug,

b) im dritten Zug die erste weiße Kugel,

c) alle weißen Kugeln

zu ziehen?

2.4.7 In einem Karton befinden sich fünf Glühbirnen, von denen drei defekt sind. Wenn diese Glühbirnen nacheinander getestet werden, wie groß ist die Wahrscheinlichkeit, dass die zweite defekte Glühbirne

a) die zweite Glühbirne, b) die dritte Glühbirne ist,

die getestet wurde?

2.4.8 Zwei Angestellte fehlen jeweils 2 von 10 Tagen unabhängig voneinander. Wie groß ist die Wahrscheinlichkeit, dass mindestens einer von ihnen anwesend ist?

2.4.9 Eine Werbeaktion für ein bestimmtes Produkt erreicht 50% aller Konsumenten, von denen sich dann 60% für einen Kauf entscheiden. Von denjenigen ohne Kenntnis der Werbeaktion kaufen 30% das Produkt.

a) Wie groß ist die Wahrscheinlichkeit, dass das Produkt gekauft wird?

b) Berechnen Sie die Wahrscheinlichkeit, dass ein Käufer die Werbeaktion kennt!

2.5 Grundlagen der Kombinatorik

Zur Berechnung von Wahrscheinlichkeiten nach dem klassischen Konzept müssen die hinsichtlich einer Fragestellung günstigen Ausgänge zu den möglichen Ausgängen ins Verhältnis gesetzt werden. Oft bereitet aber gerade die Ermittlung der günstigen Ausgänge erhebliche Schwierigkeiten.[19] Es ist daher vorteilhaft, Formeln für ihre Berechnung zu entwickeln. Der Zweig der Mathematik, der sich hiermit beschäftigt, ist die **Kombinatorik**. Sie untersucht die verschiedenen Möglichkeiten der Anordnung oder Auswahl von Objekten, die im Folgenden Elemente heißen.

Allgemein können bei der Anordnung oder Auswahl von Elementen vier Kriterien unterschieden werden:

(1) Sind die Elemente verschieden oder einige gleich,

(2) sollen nur einige Elemente ausgewählt oder alle angeordnet werden,

(3) spielt die Reihenfolge der Elemente eine Rolle oder ist dies unerheblich und schließlich

(4) soll eine Wiederholung der Elemente zulässig sein oder nicht?

Mit diesen Kriterien lassen sich alle Kombinatorikprobleme klassifizieren. Werden alle N Elemente einer Gesamtheit angeordnet, liegt eine **Permutation** vor. Sind n Elemente auszuwählen, ohne dabei auf die Reihenfolge ihrer Auswahl zu achten, spricht man von **Kombinationen**. **Variationen**

[19] Dasselbe gilt auch für die Berechnung der Möglichkeiten, aus einer Grundgesamtheit nach bestimmten Verfahren Teilerhebungen zu gewinnen. Diese, für die induktive Statistik charakteristische Vorgehensweise wird in Kapitel 7 behandelt.

ergeben sich, wenn bei der Auswahl der Reihenfolge Bedeutung zukommt. Variationen stellen daher geordnete Kombinationen dar und werden mitunter auch so bezeichnet. Bei jeder der drei Klassifikationen ist schließlich noch zu unterscheiden, ob eine Wiederholung der Elemente möglich sein soll oder nicht.

Die Entwicklung der Kombinatorikformeln lässt sich anschaulich an einer Gesamtheit mit N nummerierten Kugeln $1, 2, \ldots, N$ zeigen. Jede Zusammenstellung aller N Kugeln (Zahlen) in irgendeiner Anordnung ist eine Permutation. Die Anzahl unterscheidbarer Permutationen erhält man durch folgende Überlegung. Um die erste Stelle der Anordnung zu besetzen, stehen N Kugeln zur Auswahl. Gleichgültig, welche man wählt, bleiben zur Besetzung der zweiten Stelle $N-1$, zur Besetzung der dritten Stelle $N-2$ Kugeln und schließlich zur Besetzung der letzten (N-ten) Stelle nur noch eine Kugel übrig. Die Gesamtzahl unterscheidbarer Permutationen ist das Produkt aus all diesen Besetzungsmöglichkeiten, also:

$$N(N-1)(N-2) \cdot \ldots \cdot 1.$$

Ein solches Produkt heißt **N–Fakultät** und wird kompakt geschrieben als: $N!$, wobei N eine natürliche Zahl ist. Für $N = 4$ erhält man: $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, für $N = 1$: $1! = 1$. Damit die Rekursionsformel $N! = N(N-1)!$ uneingeschränkt auf der Menge der natürlichen Zahlen gilt, muss definiert werden $0! = 1$. Bezeichnet man die Permutationen mit P,[20] gibt Satz 2.10 die Anzahl der Anordnungen von N verschiedenen Elementen wieder.

Satz 2.10 Permutationen ohne Wiederholung

$$P(N) = N! = N(N-1)(N-2) \cdot \ldots \cdot 1.$$

Da die Werte von $N!$ mit zunehmendem N sehr schnell anwachsen, kann bei großem N die **Stirling'sche Nährungsformel $\boldsymbol{St(N)}$** benutzt werden:

$$St(N) = \sqrt{2\pi N}\left(\frac{N}{e}\right)^N \approx N!\,.$$

[20] Obwohl Permutationen dieselbe Abkürzung wie Wahrscheinlichkeiten erhalten, geht aus dem Zusammenhang stets klar hervor, was mit P gemeint ist. Verwechslungen sind daher (nahezu) ausgeschlossen.

Der Approximationswert $St(N)$ ist stets kleiner als $N!$, und der Fehler $N! - St(N)$ divergiert gegen unendlich. Jedoch konvergiert der relative Fehler $\frac{N!-St(N)}{N!}$ gegen null für $N \to \infty$: $\lim\limits_{N\to\infty} \frac{N!-St(N)}{N!} = 1 - \lim\limits_{N\to\infty} \frac{St(N)}{N!} = 0$.

Beispiel 2.10 *Es sollen drei Kugeln mit den Ziffern 1,2 und 3 unterschiedlich angeordnet werden. Die möglichen Permutationen lauten:*

$$\begin{array}{ccccccccccc} 1 & 2 & 3 & \quad & 1 & 3 & 2 & \quad & 3 & 1 & 2 \\ 2 & 1 & 3 & \quad & 2 & 3 & 1 & \quad & 3 & 2 & 1 \end{array},$$

ihre Anzahl erhält man nach Satz 2.10 als $3! = 6$.

Wären die Kugeln mit den Ziffern 1 und 2 weiß (w), die Kugel mit der Ziffer 3 rot (r) und sollen jetzt die Permutationen hinsichtlich der Farbe gebildet werden, lassen sich die untereinander stehenden Permutationen nicht mehr unterscheiden, weil sich bei ihnen die weißen Kugeln auf denselben Positionen befinden. Es können nun Klassen gebildet werden, die durch die Positionen der weißen Kugeln definiert sind. In der ersten Klasse nehmen die weißen Kugeln die Plätze 1 und 2, in der zweiten Klasse die Plätze 1 und 3 und in der dritten Klasse die Plätze 2 und 3 ein. Unterscheidbar bleiben daher nur noch pro Klasse die Permutationen wwr, wrw und rww. Diese Art der Klassierung lässt sich auch auf den allgemeinen Fall übertragen. Von den N Elementen einer Gesamtheit seien nur $m < N$ verschieden. Die Gesamtheit lässt sich dann in m Gruppen zerlegen, deren Elemente gleich sind und die n_1-mal, n_2-mal und schließlich n_m-mal vorkommen, mit $\sum\limits_{i=1}^{m} n_i = N$. Es sei zunächst angenommen, alle Elemente seien verschieden. Die $N!$ Permutationen werden jetzt hinsichtlich der Positionen, die von den n_1 Elementen der Gruppe 1 in einer Permutation eingenommen werden können, wie in Beispiel 2.10 klassiert. Permutationen innerhalb einer Klasse weisen immer dieselbe Besetzung auf den $N - n_1$ übrigen Permutationsplätzten auf. Jede Klasse umfasst dann genau $n_1!$ Permutationen. Sind die Elemente der Gruppe 1 jedoch gleich, lassen sich die Permutationen in den einzelnen Klassen nicht mehr unterscheiden. Anstelle von $n_1!$ Permutationen bleibt pro Klasse nur eine Permutation übrig. Die Gesamtzahl unterscheidbarer Permutationen beträgt dann aber nur $\frac{N!}{n_1!}$, sofern die Elemente in den übrigen Gruppen 2 bis m verschieden wären. Da dies nicht der Fall ist, werden nun die $\frac{N!}{n_1!}$ Permutationen hinsichtlich der Positionen der Elemente der Gruppe 2 auf die beschriebene Weise klassiert. Jede Klasse enthält jetzt $n_2!$ Permutationen, die bei Gleichheit der Elemente der Gruppe 2 wiederum identisch sind. Die Anzahl unterscheidbarer Permutationen reduziert sich erneut und beträgt: $\frac{N!}{n_1!n_2!}$. Verfährt man so mit allen Gruppen, erhält man die Anzahl unterscheidbarer Anordnungen, die wegen der Merkmalsgleichheit einiger Elemente als Permutationen mit Wiederholung $P_W(N)$ bezeichnet werden.

Satz 2.11 Permutationen mit Wiederholung

$$P_W(N) = \frac{N!}{\prod\limits_{i=1}^{m} n_i!}.$$

Beispiel 2.11 *Von 10 durchnummerierten Kugeln sind fünf blau, vier rot und eine weiß. Die Anzahl der Permutationen hinsichtlich der Ziffern erhält man nach Satz 2.10 als:*

$$P(10) = 10! = 3\,628\,800.$$

Sucht man die Anordnungen gemäß der Farbe, sind nur die Ausprägungen blau, rot und weiß relevant, die mit den Häufigkeiten $n_1 = 5$, $n_2 = 4$ *und* $n_3 = 1$ *vorkommen. Nach Satz 2.11 folgt:*

$$P_W(10) = \frac{10!}{5!4!1!} = \frac{3\,628\,800}{2880} = 1260.$$

Will man wissen, wie viele dieser Permutationen eine blaue, rote oder weiße Kugel an erster Stelle aufweisen, ist die erste Position der Permutation mit der vorgegebenen Kugel besetzt, und es können nur noch 9 Kugeln mit Wiederholung permutiert werden. Die entsprechenden Anzahlen betragen:

$$\begin{aligned} P_{blau}(9) &= \frac{9!}{4!4!} = 630, \\ P_{rot}(9) &= \frac{9!}{5!3!} = 504, \\ P_{weiß}(9) &= \frac{9!}{5!4!} = 126. \end{aligned}$$

Die Auswahl von n Elementen, ohne dabei auf die Reihenfolge der Entnahme zu achten, aus einer Gesamtheit von N Elementen wird als **Kombination** von N Elementen zur Klasse n bezeichnet und mit $K(N, n)$ abgekürzt. Um die Anzahl möglicher Kombinationen zu ermitteln, muss unterschieden werden, ob eine Wiederholung der Elemente in der Kombination zulässig sein soll oder nicht. Die Formel für die Anzahl an Kombinationen ohne Wiederholung zur Klasse n mit $n < N$ kann über Permutationen entwickelt werden. Da bei Kombinationen die Reihenfolge der Auswahl unerheblich ist, lassen sich die N Elemente der Gesamtheit in n Elemente, die in eine Kombination aufgenommen werden und in $N - n$ Elemente, die ausgeschlossen bleiben, unterteilen. Die N Elemente unterscheiden sich jetzt nur noch hinsichtlich dieser beiden Merkmale. Berechnet man hierfür die Anzahl der Permutationen mit Wiederholung für die so unterteilte Gesamtheit, entspricht diese der Anzahl der Kombinationen ohne Wiederholung von N Elementen zur Klasse n. Gemäß Satz 2.11 ergibt sich:

$$K(N,n) = \frac{N!}{n!(N-n)!}.$$

Durch Kürzen kann der Bruch in eine andere Form gebracht werden:

$$\begin{aligned}\frac{N!}{n!(N-n)!} &= \frac{N(N-1)\cdot\ldots\cdot(N-n+1)(N-n)(N-n-1)\cdot\ldots\cdot 1}{n(n-1)\cdot\ldots\cdot 1\cdot(N-n)(N-n-1)\cdot\ldots\cdot 1}\\ &= \frac{N(N-1)\cdot\ldots\cdot(N-n+1)}{n(n-1)\cdot\ldots\cdot 1} = \binom{N}{n}.\end{aligned}$$

Der letzte Term der Umformung ist der **Binomialkoeffizient**, der als „N über n“ gelesen wird. Somit kann die Anzahl der Kombinationen ohne Wiederholung berechnet werden mit:

Satz 2.12 Kombinationen ohne Wiederholung zur Klasse n

$$K(N,n) = \frac{N!}{n!(N-n)!} = \binom{N}{n}.$$

Mit der Definition des Binomialkoeffizienten $\binom{N}{n} = \dfrac{N!}{n!(N-n)!}$ für natürliche Zahlen N, ganze Zahlen $n \geq 0$ und $N \geq n$ lässt sich zeigen, dass gilt:

$$\text{(a)} \binom{N}{n} = \binom{N}{N-n} \text{ und, wegen } 0! = 1, \quad \text{(b)} \binom{N}{N} = \binom{N}{0} = 1;$$

für $N < n$ wird definiert: $\binom{N}{n} = 0$.

Beispiel 2.12 *Gegeben ist der endliche Stichprobenraum* $\Omega = \{\omega_1, \ldots, \omega_m\}$ *mit der Potenzmenge* $PM(\Omega) = \{\emptyset, \{\omega_1\}, \ldots, \{\omega_m\}, \{\omega_1, \omega_2\}, \ldots, \Omega\}$. *Die*

Elemente der Potenzmenge ergeben sich als Kombinationen zu den Klassen $0, 1, \ldots, m$. *Die Kombination zur Klasse* 0 *stellt die leere Menge, das unmögliche Ereignis dar; die Kombination zur Klasse* m *liefert die Menge* Ω, *das sichere Ereignis. Die Anzahl aller Kombinationen, summiert über die Klassen, beträgt:*

$$\binom{m}{0} + \binom{m}{1} + \binom{m}{2} + \ldots + \binom{m}{m}.$$

Gemäß des **binomischen Satzes** *gilt:*

$$(a+b)^m = \binom{m}{0} a^m b^0 + \binom{m}{1} a^{m-1} b + \ldots + \binom{m}{m} a^0 b^m.$$

Setzt man $a = b = 1$, *führt dies zu:*

$$2^m = \binom{m}{0} + \binom{m}{1} + \ldots + \binom{m}{m}.$$

Das beweist, dass die Potenzmenge 2^m *Elemente enthält.*

Lässt man bei Kombinationen eine Wiederholung zu, steht jedes Element mehrmals zur Verfügung, ausgewählt zu werden. Da es jetzt vorkommen kann, dass eine Kombination n-mal dasselbe Element enthält, muss jedes Element $(n-1)$-mal zusätzlich zur Verfügung stehen. Die Anzahl der Kombinationen mit Wiederholung von N Elementen zur Klasse n, $K_W(N, n)$, beträgt daher:

Satz 2.13 Kombinationen mit Wiederholung

$$K_W(N, n) = \binom{N+n-1}{n}.$$

Die Wiederholungsmöglichkeit der Elemente erreicht man dadurch, dass die Kombinationen durch „Ziehen mit Zurücklegen“ gebildet werden. Bei dieser Vorgehensweise erschöpft sich die Gesamtheit durch Entnahme ihrer Elemente nicht, und die bei Kombinationen ohne Wiederholung notwendige Einschränkung $n < N$ entfällt.

Beispiel 2.13 *Mit den Buchstaben b und u sollen alle Kombinationen mit Wiederholung zur Klasse 4 gebildet werden. Es gilt:* $N = 2$ *und* $n = 4$; *nach Satz 2.13 beträgt die Anzahl:*

$$K_W(2,4) = \binom{2+4-1}{4} = \binom{5}{4} = \frac{5\cdot 4\cdot 3\cdot 2}{4\cdot 3\cdot 2} = 5.$$

Die fünf Kombinationen lauten: $(bbbb), (bbbu), (bbuu), (buuu)$ *und* $(uuuu)$.

Variationen ohne Wiederholung $V(N,n)$ sind definitionsgemäß geordnete Kombinationen ohne Wiederholung. Jede dieser Kombinationen zur Klasse n kann auf $n!$ verschiedene Arten angeordnet (permutiert) werden. Da es insgesamt $\binom{N}{n}$ Kombinationen ohne Wiederholung gibt, erhält man die Anzahl der Variationen ohne Wiederholung als $\binom{N}{n}n!$. Dieses Produkt lässt sich leicht in eine Form bringen, wie sie in Satz 2.14 verwendet wird.

Satz 2.14 Variationen ohne Wiederholung

$$V(N,n) = \binom{N}{n}n! = \frac{N!}{(N-n)!}.$$

Bei Variationen mit Wiederholung $V_W(N,n)$ hat man N verschiedene Besetzungsmöglichkeiten für jeden Platz in der Variation. Ihre Anzahl gibt Satz 2.15 wieder.

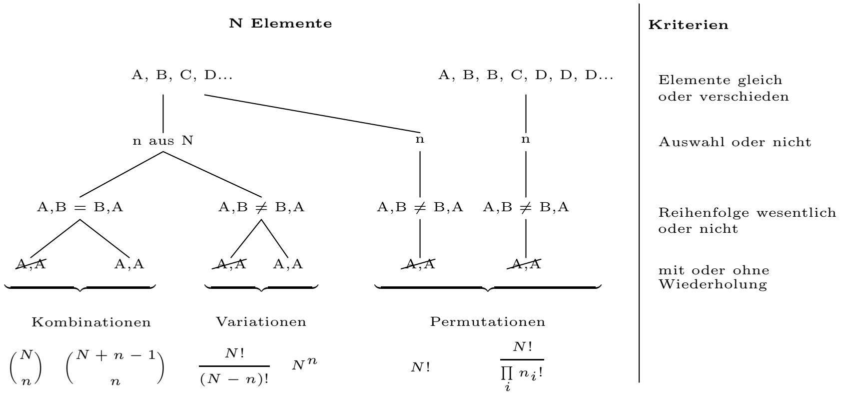
Satz 2.15 Variationen mit Wiederholung

$$V_W(N,n) = N^n.$$

Beispiel 2.14 *Für eine Gesamtheit mit* $N = 4$ *verschiedenen Elementen lassen sich* $\binom{4}{3} = 4$ *Kombinationen ohne Wiederholung zur Klasse* $n = 3$ *bilden. Jede Kombination kann* $3! = 6$ *mal permutiert werden. Die Anzahl der Variationen ohne Wiederholung beträgt 24. Dasselbe Ergebnis stellt sich nach Satz 2.14 als* $\frac{4!}{1!} = 4\cdot 3\cdot 2 = 24$ *ein. Sucht man die Anzahl der Variationen mit Wiederholung, erhält man nach Satz 2.15:* $V_W(4,3) = 4^3 = 64$.

In Abbildung 2.13 sind die Kriterien und entwickelten Regeln der Kombinatorik zusammengestellt. Identifiziert man bei einem kombinatorischen Problem die in Abbildung 2.13 rechts angeführten Kriterien, sollten die Pfade schnell zur entsprechenden Formel führen.

Abb. 2.13: Flussdiagramm zur Kombinatorik



Übungsaufgaben zu 2.5

2.5.1 Wie viele Möglichkeiten gibt es,

a) einen Lottoschein „6 aus 49“,

b) einen Totoschein für 12 Fußballspiele, bei denen die Tendenz getippt werden muss (0: Unentschieden, 1: Sieg der Heimmannschaft, 2: Sieg der Gastmannschaft),

c) auf der Pferderennbahn bei 18 teilnehmenden Pferden einen Wettschein, bei dem die drei bestplatzierten Pferde in richtiger Reihenfolge getippt werden sollen,

auszufüllen?

2.5.2 Zwölf verschiedene Bücher werden zufällig auf einem Regal angeordnet. Bestimmen Sie die Wahrscheinlichkeit dafür, dass vier bestimmte Bücher nebeneinander stehen!

2.5.3 Wie viele vierstellige Zahlen lassen sich mit den Ziffern 1 bis 5 bilden, wenn die Ziffer 3 immer enthalten sein soll und keine Ziffer mehrfach vorkommen darf?

2.5.4 In einem binären System werden die natürlichen Zahlen mit den Ziffern 0 und 1 dargestellt. Wie viele verschiedene natürliche Zahlen können mit

a) genau fünf Stellen, b) bis zu fünf Stellen

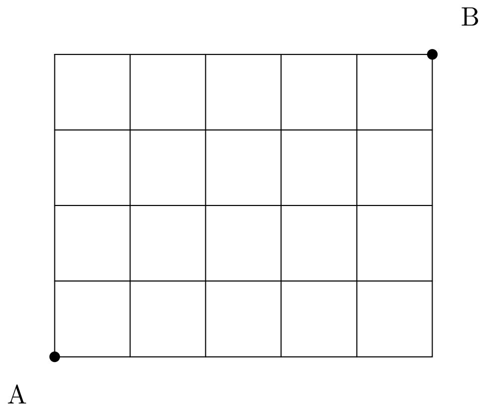
dargestellt werden?

2.5.5 Wie viele Anordnungen kann man aus den Buchstaben der Wörter:

a) „Symbol", b) „Bubu", c) „Statistik"

bilden, wenn jeweils alle Buchstaben verwendet werden?

2.5.6 Die folgende Abbildung stellt ein (stilisiertes) Straßennetz dar:



Auf wie vielen verschiedenen kürzesten Wegen kann man von A nach B gelangen?

2.5.7 Auf wie viele Arten lassen sich 10 Bücher auf 3 Personen verteilen, wenn immer Person A fünf, Person B drei und Person C zwei Bücher erhalten sollen?

2.5.8 Auf einer Statistikkonferenz waren 50 Teilnehmer. Wie oft wurden zur Begrüßung die Hände geschüttelt? Am darauf folgenden Weihnachtsfest schicken sich alle gegenseitig eine Karte. Wie viele Weihnachtsgrüße werden verschickt?

2.5.9 a) Wie viele Anordnungen der Personen gibt es für drei Damen und vier Herren, sich

(1) an einen runden Tisch, (2) auf eine Bank

zu setzen, wenn die drei Damen immer nebeneinander sitzen möchten?

b) Auf wie viele unterschiedliche Weisen lassen sich die 7 Personen auf die Plätze 1 bis 7 verteilen, wenn die Herren nur die ungeraden Plätze einnehmen?