Kapitel I
Grundbegriffe der stochastischen Modellierung

Unter einem *Zufallsexperiment (ZE)* verstehen wir einen Vorgang, dessen Ausgang aus Sicht des Beobachters nicht exakt determiniert werden kann. Der Grund für diese Ungewißheit besteht entweder in der faktischen Unmöglichkeit einer präzisen Bestimmung, z.B. bei der Prognose des Wetters von morgen oder der des deutschen Fußballmeisters im nächsten Jahr, oder in einem Fehlen vollständiger Information über ein de facto vorliegendes Ergebnis, z.B. die aktuelle Zahl von Kunden in einem großen Warenhaus, der insgesamt angefallene Materialausschuß während eines Tages in einem Produktionsbetrieb oder auch der aktuelle Bestand einer bestimmten Baumart in einem Forstgebiet. Um ein ZE hinsichtlich mathematischer Gesetzmäßigkeiten analysieren zu können, bedarf es natürlicherweise als erstes seiner Übersetzung in mathematische Sprache. Dies geschieht in Form eines *stochastischen Modells*, das eine i.a. idealisierte, jedoch hinreichend genaue Abbildung der Realität liefern und folgende Eigenschaften besitzen soll:

- Die *Versuchsbedingungen* sind exakt festgelegt.
- Die Menge der möglichen Ausgänge (Ergebnisse) ist vorab bekannt.
- Das Experiment ist zumindest prinzipiell beliebig oft unter gleichen Bedingungen wiederholbar.

Mit zunehmender Komplexität des betrachteten ZE's liegt die Auswahl eines adäquaten Modells allerdings keineswegs auf der Hand und mündet in der Frage, was denn unter einer hinreichend genauen Abbildung der Realität zu verstehen ist. Der Mathematiker tut sich hier bisweilen schwer, denn eine solche Auswahl entzieht sich dem gewohnten Rückgriff auf ein formales Regelwerk, mit dem man, notfalls schematisch vorgehend, zum Ziel gelangt. Zwei in aller Regel konkurrierende Zielsetzungen sind in Einklang zu bringen:

- Das Modell muß *mathematisch handhabbar* sein, d.h. von einer Komplexität, die das Erkennen mathematischer Gesetzmäßigkeiten und Berechnen von Wahrscheinlichkeiten überhaupt erlaubt.
- Das Modell muß *die wesentlichen Einflußgrößen* des realen ZE's *berücksichtigen*, damit die Ergebnisse der Modellanalyse in Bezug auf die Realität tatsächlich Aussagekraft besitzen und einer Überprüfung standhalten.

Daß ein solcher Kompromiß Grenzen hat, versteht sich von selbst. Es kann und wird aber nicht Ziel dieser Einführungsvorlesung sein, diese Grenzen auszuloten. Vielmehr sollen dem Neuling im "Reich des Zufalls" fundamentale Begriffsbildungen und mathematische Gesetzmäßigkeiten anhand von Modellen, deren Komplexität von einfach bis moderat reicht, vermittelt werden.

Wie bereits in der Einführung bemerkt, zwingt uns der gänzliche Verzicht auf Kenntnisse aus der Maß- und Integrationstheorie zudem zur Beschränkung auf solche ZE, in denen sich der allgemein nicht verfügbare Begriff des Maßintegrals zu einer Summe oder einem Riemann-Integral vereinfacht. Betrachtet werden deshalb in erster Linie

- *Diskrete ZE*, deren Ergebnisraum abzählbar ist

sowie in begrenztem Umfang auch

- *Regulär stetige ZE*, deren Ergebnisraum ein Intervall im $\mathbb{R}^d$ und somit überabzählbar ist, und die eine gewisse hier noch nicht spezifizierte Regularitätsbedingung erfüllen.

Ein wesentlicher Grund, hier nicht ganz auf ZE mit überabzählbarem Ergebnisraum zu verzichten, liegt darin, daß einige Verteilungen von herausragender Bedeutung, die auch im Schulunterricht nicht vollständig fehlen sollten, nur in diesem Rahmen definiert werden können. Hierzu zählen besonders die Normal-, die Gleich- sowie die Exponentialverteilung.

1. Diskrete Zufallsexperimente

Viele denken bei ZE zuerst an einfache Glücksspiele wie das Werfen einer Münze, das Werfen eines oder mehrerer Würfel oder auch das Ziehen von numerierten oder gefärbten Kugeln aus einer Urne (Lotto). Jedes dieser Beispiele gehört in die Klasse der *endlichen ZE*, da die jeweilige Menge aller möglichen *Ergebnisse*, genannt *Ergebnisraum*, endlich ist. Die zufällige Anzahl von Zerfällen einer radioaktiven Substanz in einem Zeitintervall $[0,t]$ bildet ein Beispiel für ein ZE mit abzählbar unendlichem Ergebnisraum. Aber auch ZE mit überabzählbarem Ergebnisraum gibt es zuhauf. Brechen wir z.B. einen Stab der Länge L in zwei Stücke zufälliger Länge und betrachten als Ergebnis die Länge des kürzeren Stücks, so ist der Ergebnisraum offenbar das Intervall $(0,L/2]$.

Im folgenden konzentrieren wir uns aber zunächst auf *diskrete ZE*, in denen der Ergebnisraum abzählbar ist. Ihre axiomatische Einführung ist sowohl aus intuitiver als auch aus formaler Sicht einfach. Das Grundaxiom besteht darin, daß jedem Ergebnis eine Zahl zwischen 0 und 1, genannt *Wahrscheinlichkeit*, zugeordnet ist und daß die Gesamtsumme der Ergebniswahrscheinlichkeiten 1 beträgt.

1.1. Definition. Ein *diskretes Zufallsexperiment* ist ein Paar (Ω,p), bestehend aus einer abzählbaren, nichtleeren Menge Ω, genannt *Ergebnisraum*, und einer Funktion $p:\Omega\to[0,1]$ mit der Eigenschaft

$$\sum_{\omega\in\Omega} p(\omega) \;=\; 1.$$

Wir interpretieren $p(\omega)$ als die Wahrscheinlichkeit für das Auftreten von ω. Ist Ω endlich, heißt (Ω,p) auch *endliches Zufallsexperiment*.

Basierend auf dieser Definition, ordnen wir als nächstes jeder Teilmenge E von Ω, genannt *Ereignis*, eine Wahrscheinlichkeit zu. Es liegt auf der Hand zu sagen, daß E eingetreten bzw. nicht eingetreten ist, wenn das beobachtete Ergebnis ω aus E bzw. E^c stammt. Die Wahrscheinlichkeit von E definiert man demzufolge durch die Summe der Wahrscheinlichkeiten aller Elemente von E:

1.2. Definition. Gegeben ein diskretes ZE (Ω, p), sei $P : \mathfrak{P}(\Omega) \to [0,1]$ durch

$$P(E) \stackrel{\text{def}}{=} \sum_{\omega \in E} p(\omega) \tag{1.1}$$

definiert, wobei $P(\emptyset) \stackrel{\text{def}}{=} 0$ (leere Summe). P heißt die zu p gehörende *diskrete Wahrscheinlichkeitsverteilung* (kurz W-Verteilung) *über* Ω und $P(E)$ *die Wahrscheinlichkeit von* E.

Die folgenden Eigenschaften von P lassen sich direkt aus der Definition folgern:

1.3. Korollar. *Gegeben ein diskretes ZE (Ω, p) mit zugehöriger W-Verteilung P, gilt:*

(a) P ist normiert, d.h. $P(\Omega) = 1$.

(b) P ist σ-additiv, d.h. $P(\cup_{i=1}^\infty E_i) = \sum_{i=1}^\infty P(E_i)$ für jede Auswahl paarweise disjunkter (p.d.) Ereignisse $E_1, E_2, ... \subset \Omega$.

Im folgenden benutzen wir im Fall der Vereinigung p.d. Ereignisse zur besseren Hervorhebung auch das Summenzeichen "Σ" anstelle von "$\cup$". Für P gilt also

$$P\left(\sum_{i=1}^\infty E_i\right) = \sum_{i=1}^\infty P(E_i). \tag{1.2}$$

Beweis: Da (a) trivial ist, kommen wir gleich zu Aussage (b) und betrachten eine beliebige Folge $(E_i)_{i\ge1}$ p.d. Teilmengen von Ω. Dann gilt unter Benutzung von (1.2) und der Tatsache, daß für absolut konvergente Reihen die Summationsreihenfolge beliebig gewählt werden kann,

$$P\left(\sum_{i=1}^\infty E_i\right) = \sum_{\omega \in \Sigma_{i\ge1} E_i} p(\omega) = \sum_{i\ge1}\sum_{\omega\in E_i} p(\omega) = \sum_{i\ge1} P(E_i),$$

also das Gewünschte. $\diamondsuit$

1.4. Bemerkungen. (a) Aus der σ-Additivität von P folgt insbesondere dessen *endliche Additivität*, gegeben furch

$$P\left(\sum_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i) \tag{1.3}$$

für jede endliche Auswahl $E_1, ..., E_n$ p.d. Teilmengen von Ω. Setzt man nämlich $E_{n+1} = E_{n+2} = ... = \emptyset$, so ist $(E_i)_{i\geq 1}$ eine Folge p.d. Teilmengen von Ω mit $P(E_i) = 0$ für $i > n$, was vermöge der σ-Additivität

$$P\left(\sum_{i=1}^{n} E_i\right) = P\left(\sum_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i) = \sum_{i=1}^{n} P(E_i),$$

also (1.3) liefert.

(b) Ist Ω und damit auch $\mathfrak{P}(\Omega)$ endlich, so kann eine Folge $(E_i)_{i\geq 1}$ p.d. Teilmengen von Ω offenkundig höchstens endlich viele nichtleere Elemente enthalten. Jede endlich additive, normierte Mengenfunktion $P : \mathfrak{P}(\Omega) \to [0, 1]$ ist folglich schon σ-additiv. Ohne Beweis notieren wir, daß dies für abzählbar unendliches Ω nicht stimmt.

Während zu jedem diskreten ZE (Ω, p) vermöge (1.1) offenkundig eine eindeutig festgelegte diskrete W-Verteilung P gehört, zeigt der anschließende einfache Satz, daß umgekehrt jede normierte und σ-additive Mengenfunktion $P : \mathfrak{P}(\Omega) \to [0, 1]$ in eindeutiger Weise ein diskretes ZE (Ω, p) beschreibt.

1.5. Satz. *Sei Ω eine abzählbare, nichtleere Menge und $P : \mathfrak{P}(\Omega) \to [0, 1]$ eine normierte, σ-additive Mengenfunktion. Dann gibt es genau ein diskretes ZE (Ω, p), so daß P die zu p gehörende diskrete W-Verteilung über Ω ist.*

Beweis: Wegen $\{\omega\} \subset \Omega$ muß $p(\omega) \stackrel{\text{def}}{=} P(\{\omega\})$ gelten, und wir erhalten dann vermöge der vorausgesetzten Eigenschaften von P

$$P(E) = P\left(\sum_{\omega\in E} \{\omega\}\right) = \sum_{\omega\in E} p(\omega)$$

sowie

$$\sum_{\omega\in\Omega} p(\omega) = \sum_{\omega\in\Omega} P(\{\omega\}) = P(\Omega) = 1. \qquad \diamond$$

Wenden wir uns nun einer Reihe von grundlegenden Beispielen zu:

1.6. Beispiel. ***(Einfacher Münzwurf, Bernoulli-Verteilung)*** Das einmalige Werfen einer Münze wird durch das einfache ZE (Ω, p) mit $\Omega = \{\text{Kopf, Zahl}\}$ und

$$p(\text{Zahl}) = 1 - p(\text{Kopf}) = \theta$$

für ein $\theta \in [0, 1]$ beschrieben. Da es für eine Analyse dieses Experiments offenbar keine Rolle spielt, wie die Elemente von Ω bezeichnet werden, identifiziert man aus Zweckmäßigkeitsgründen meistens "Kopf" mit 0 und "Zahl" mit 1. Dann gilt $\Omega = \{0, 1\}$ und $p(1) = 1 - p(0) = \theta$. (Ω, p) heißt *Bernoulli-Experiment*, die zu p gehörende W-Verteilung P *Bernoulli-Verteilung mit Parameter θ*, kurz *$B(1, \theta)$-Verteilung*. Der auftretende Parameter 1 bedeutet, daß die Münze

$n = 1$ mal geworfen wird (☞ Beispiel 1.8 für den Fall $n \geq 2$). Wir sprechen von einer *fairen* Münze, wenn beide Seiten mit gleicher Wahrscheinlichkeit $\theta = \frac{1}{2}$ auftreten.

1.7. Beispiel. ***(Würfelwurf, Laplace-Verteilung)*** Das einmalige Werfen eines fairen Würfels wird durch das ZE (Ω, p) mit Ergebnisraum $\Omega = \{1, 2, 3, 4, 5, 6\}$ und

$$p(1) \;=\; ... \;=\; p(6) \;=\; \frac{1}{|\Omega|} \;=\; \frac{1}{6}$$

beschrieben. Entsprechend ergibt sich bei n-maligem Werfen $\Omega = \{1, 2, 3, 4, 5, 6\}^n$ und

$$p(\omega) \;=\; \frac{1}{|\Omega|} \;=\; \frac{1}{6^n}$$

für alle $\omega = (\omega_1, ..., \omega_n) \in \Omega$. Man nennt allgemein jedes endliche ZE (Ω, p), dessen Ergebnisse alle gleichwahrscheinlich sind, *Laplace-Experiment* und die zugehörige W-Verteilung *Laplace-Verteilung über* Ω. Laplace-Experimenten werden wir aufgrund ihrer besonders einfachen Struktur in Abschnitt 3 größere Aufmerksamkeit schenken. Beachte, daß die $B(1, \frac{1}{2})$-Verteilung eine spezielle Laplace-Verteilung ist.

1.8. Beispiel. ***(Mehrfacher Münzwurf, Binomialverteilung)*** Wird eine Münze $n \geq 2$ mal geworfen und am Ende als Ergebnis notiert, wie oft "Zahl" aufgetreten ist, so erhält man als Ergebnisraum offenkundig $\Omega = \{0, ..., n\}$. Sei $\theta \in (0, 1)$ die Wahrscheinlichkeit für "Zahl" bei einmaligem Werfen der Münze. Wir werden später zeigen (☞ 4.1 und 10.8), daß

$$p(k) \;=\; \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

für $k = 0, ..., n$ gilt, sofern die einzelnen Würfe sich nicht gegenseitig beeinflussen und unter stets gleichen Bedingungen stattfinden. Die notwendige Eigenschaft

$$\sum_{k \in \Omega} p(k) \;=\; \sum_{k=0}^{n} \binom{n}{k} \theta^k (1-\theta)^{n-k} \;=\; 1$$

ergibt sich vermöge des binomischen Lehrsatzes. Die zu p gehörende W-Verteilung P heißt deshalb *Binomialverteilung mit Parametern n und θ* oder kurz *$B(n, \theta)$-Verteilung.* Die Bernoulli-Verteilung ist die spezielle Binomialverteilung mit $n = 1$.

1.9. Beispiel. ***(Warten auf die Sechs, geometrische Verteilung)*** Unser letztes Beispiel zeigt, daß auch ZE mit abzählbar unendlichem Ergebnisraum in vertrauten Situationen auftreten. Jeder weiß, daß bei dem Gesellschaftsspiel "Mensch ärgere dich nicht" ein Spieler seine Figur erst dann ziehen darf, wenn er zum ersten Mal eine Sechs wirft. Zugrundeliegt hier das diskrete ZE des sukzessiven Werfens eines fairen Würfels bis zur ersten Sechs, wobei wir als Ergebnis die Anzahl der notwendigen Fehlversuche ansehen wollen. Dann erhalten wir als

Ergebnisraum Ω die Menge $\mathbb{N}_0 \cup \{\infty\}$, wobei ∞ den Fall beschreibt, daß niemals eine Sechs erscheint, was ja denkbar ist. Anschauliche Überlegungen zeigen, daß

$$\begin{aligned} p(0) &= \frac{1}{6}, \\ p(1) &= \frac{5}{6} \cdot \frac{1}{6}, \\ p(2) &= \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \;=\; \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6} \end{aligned}$$

und allgemein

$$p(n) = \left(\frac{5}{6}\right)^n \cdot \frac{1}{6}$$

für alle $n \in \mathbb{N}_0$ gilt, wiederum unter der Voraussetzung, daß sich die Würfe nicht gegenseitig beeinflussen und unter stets gleichen Bedingungen stattfinden. Wegen

$$\sum_{n \in \mathbb{N}_0} p(n) = \frac{1}{6} \sum_{n \geq 0} \left(\frac{5}{6}\right)^n = 1$$

folgt $p(\infty) = 0$, was zweifellos unserer Anschauung entspricht. Wir sehen andererseits, daß ein Ergebnis durchaus Wahrscheinlichkeit 0 besitzen kann, obwohl es faktisch möglich ist. Aufgrund der auftretenden geometrischen Reihe nennt man die zu (Ω, p) gehörende W-Verteilung *geometrische Verteilung mit Parameter* $\frac{1}{6}$. Ersetzt man $\frac{1}{6}$ durch ein beliebiges $\theta \in (0,1)$ (folglich $\frac{5}{6}$ durch $1-\theta$), so erhält man entsprechend eine geometrische Verteilung mit Parameter θ (z.B. beim Werfen einer Münze so lange, bis zum ersten Mal "Zahl" erscheint, wenn θ die Wahrscheinlichkeit für "Zahl" in einem Wurf angibt).

Nachdem wir nunmehr diskrete ZE formal eingeführt haben, lohnt es, kurz innezuhalten, um in Form einer Übersicht die im weiteren Verlauf verwendeten mathematischen Objekte und Operationen ihrer Übersetzung in die Realität gegenüberzustellen (☞ nächste Seite). Dabei werden die folgenden vielleicht noch nicht geläufigen Mengenoperationen verwendet:

- Die *symmetrische Differenz* zweier Mengen:

$$E_1 \Delta E_2 \stackrel{\text{def}}{=} E_1 \cap E_2^c + E_1^c \cap E_2.$$

- Der *Limes superior* einer Mengenfolge:

$$\limsup_{i \to \infty} E_i \stackrel{\text{def}}{=} \bigcap_{k \geq 1} \bigcup_{i \geq k} E_i = \{\omega \in \Omega : \omega \text{ liegt in unendlich vielen } E_i\}.$$

- Der *Limes inferior* einer Mengenfolge:

$$\liminf_{i \to \infty} E_i \stackrel{\text{def}}{=} \bigcup_{k \geq 1} \bigcap_{i \geq k} E_i = \{\omega \in \Omega : \omega \text{ liegt in fast allen } E_i\}.$$

Realität/Interpretation	Math. Modell
Ergebnisraum	Ω
mögliches Ergebnis, ω tritt ein	$\omega(\in \Omega)$
Ereignis	$E(\subset \Omega)$
sicheres Ereignis	Ω
unmögliches Ereignis	$\emptyset$
Ereignis E_1 *oder* Ereignis E_2 tritt ein	$E_1 \cup E_2$
Ereignis E_1 *und* Ereignis E_2 tritt ein	$E_1 \cap E_2$
Ereignis E tritt *nicht* ein	E^c
das Eintreten von E_1 impliziert das Eintreten von E_2	$E_1 \subset E_2$
die Ereignisse E_1 und E_2 sind unverträglich	$E_1 \cap E_2 = \emptyset$
eines der beiden unverträglichen Ereignisse E_1 und E_2 tritt ein	$E_1 + E_2$
genau eines der beiden Ereignisse E_1 und E_2 tritt ein	$E_1 \Delta E_2$
mindestens eines der Ereignisse $E_i, i \geq 1$, tritt ein	$\cup_{i\geq 1} E_i$
alle Ereignisse $E_i, i \geq 1$, treten ein	$\cap_{i\geq 1} E_i$
unendlich viele der Ereignisse $E_i, i \geq 1$, treten ein	$\limsup_{i\to\infty} E_i$
fast alle der der Ereignisse $E_i, i \geq 1$, treten ein	$\liminf_{i\to\infty} E_i$

Eine weitere Liste gibt entsprechende "Sprachregelungen" für Wahrscheinlichkeiten:

Realität/Interpretation	Math. Modell
Wahrscheinlichkeit für das Eintreten von E	$P(E)$
E tritt (P-)fast sicher ein	$P(E) = 1$ $[\not\Rightarrow E = \Omega]$
E ist (P)-fast sicher unmöglich	$P(E) = 0$ $[\not\Rightarrow E = \emptyset]$
E tritt mit mindestens α-prozentiger Sicherheit ein	$P(E) \geq \alpha/100$
E_1 ist wahrscheinlicher als E_2	$P(E_1) > P(E_2)$
E_1 und E_2 sind gleichwahrscheinlich	$P(E_1) = P(E_2)$

Zur Veranschaulichung von diskreten W-Verteilungen stellen wir uns den Ergebnisraum Ω als Gesamtheit von Massepunkten vor, wobei sich in jedem $\omega \in \Omega$ die Masse $p(\omega)$ befindet. Für $E \subset \Omega$ gibt dann $P(E)$ die Masse in E an. Bild 1.1 illustriert diese Vorstellung für Ergebnisräume $\Omega \subset \mathbb{R}$ und $\Omega \subset \mathbb{R}^2$. Da vor allem Massepunkte mit positiver Masse von Bedeutung sind, führt man für deren Gesamtheit eine eigene Bezeichnung ein:

1.10. Definition. Gegeben ein diskretes ZE (Ω, p), heißt die Menge

$$\Omega_0 \stackrel{\text{def}}{=} \{\omega \in \Omega : p(\omega) > 0\}$$

der *Träger* der zu p gehörenden diskreten W-Verteilung P.

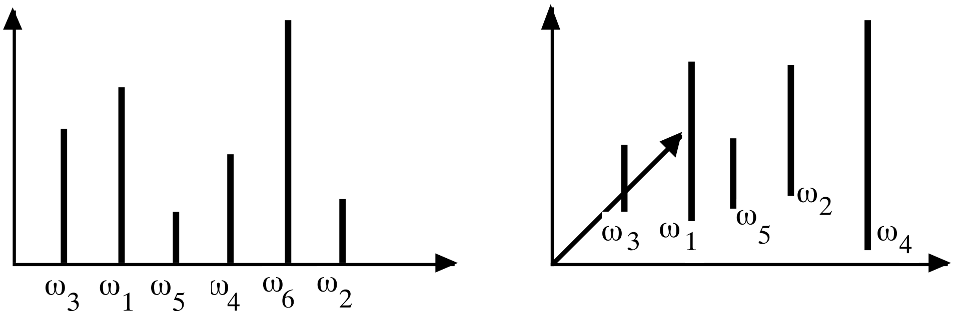

Bild 1.1. Illustration diskreter W-Verteilungen im Fall $\Omega \subset \mathbb{R}$ und $\Omega \subset \mathbb{R}^2$.

Zur Berechnung von Wahrscheinlichkeiten spielen wegen

$$P(E) \;=\; \sum_{\omega \in E} p(\omega) \;=\; \sum_{\omega \in E \cap \Omega_0} p(\omega) \;=\; P(E \cap \Omega_0)$$

Punkte außerhalb von Ω_0 keine Rolle. Dann ist es aber auch egal, wieviele das sind. Wir können deshalb folgende Erweiterung unserer ursprünglichen Definition diskreter ZE geben:

1.11. Definition (Verallgemeinerung von 1.1 und 1.2). Ein *diskretes ZE* ist ein Paar (Ω, p), bestehend aus einer beliebigen nichtleeren Menge Ω und einer Funktion $p : \Omega \to [0,1]$, für die $\Omega_0 = \{\omega \in \Omega : p(\omega) > 0\}$ abzählbar ist und $\sum_{\omega \in \Omega_0} p(\omega) = 1$ gilt. Die durch

$$P(E) \stackrel{\text{def}}{=} \sum_{\omega \in E \cap \Omega_0} p(\omega) \tag{1.4}$$

definierte Abbildung $P : \mathfrak{P}(\Omega) \to [0,1]$ heißt die zu p gehörende *diskrete W-Verteilung* über Ω. Ferner heißt p die *Zähldichte* von P und Ω_0 der *Träger* von P.

1.12. Bemerkung. In Analogie zu Satz 1.5 halten wir fest: Ist $\Omega \neq \emptyset$ eine beliebige Menge und $P : \mathfrak{P}(\Omega) \to [0,1]$ eine σ-additive Mengenfunktion, zu der eine abzählbare Menge $\Omega_0 \subset \Omega$ existiert mit $P(\Omega_0) = 1$, so gibt es genau ein diskretes ZE (Ω, p) mit zugehöriger diskreter W-Verteilung P, nämlich das durch $p(\omega) \stackrel{\text{def}}{=} P(\{\omega\})$ definierte (Ω, p).

2. Eigenschaften diskreter Wahrscheinlichkeitsverteilungen

In diesem Abschnitt wollen wir die wichtigsten Eigenschaften diskreter W-Verteilungen zusammentragen.

2.1. Satz. *Sei P eine diskrete W-Verteilung über Ω. Dann gilt:*

(a) $P(E^c) = 1 - P(E)$ für alle $E \subset \Omega$.

(b) P ist monoton, d.h. $P(E_1) \leq P(E_2)$ für alle $E_1 \subset E_2 \subset \Omega$.

(c) P ist subadditiv, d.h. $P(\cup_{i\geq 1}E_i) \leq \sum_{i\geq 1} P(E_i)$ für jede Folge $E_1, E_2, ...$ von Teilmengen von Ω.

(d) $P(\cup_{i=1}^n E_i) = \sum_{k=1}^n (-1)^{k+1} \sum_{1\leq i_1<...<i_k\leq n} P(E_{i_1}\cap ... \cap E_{i_k})$ für jedes $n \geq 1$ und jede Auswahl $E_1, ..., E_n$ von Teilmengen von Ω (Siebformel).

Beweis: (a) folgt aus $1 = P(\Omega) = P(E)+P(E^c)$ und (b) aus $P(E_2) = P(E_1+E_2\cap E_1^c) = P(E_1) + P(E_2 \cap E_1^c) \geq P(E_1)$.

(c) Sei Ω_0 der abzählbare Träger von P. Dann gilt unter Hinweis auf (1.4)

$$\begin{aligned} P\left(\bigcup_{i\geq 1} E_i\right) &= \sum_{\omega\in\Omega_0\cap(\cup_{i\geq 1}E_i)} P(\{\omega\}) \\ &\leq \sum_{\omega\in\Omega_0\cap E_1} P(\{\omega\}) \;+\; \sum_{\omega\in\Omega_0\cap E_2} P(\{\omega\}) \;+ ... \\ &= \sum_{i\geq 1}\sum_{\omega\in\Omega_0\cap E_i} P(\{\omega\}) \;=\; \sum_{i\geq 1} P(E_i). \end{aligned}$$

(d) Wir benutzen vollständige Induktion: Für $n = 1$ ist die Formel offensichtlich wahr. Für $n = 2$ (wird für den Induktionsschritt benötigt) müssen wir

$$P(E_1 \cup E_2) \;=\; P(E_1) + P(E_2) - P(E_1 \cap E_2) \tag{2.1}$$

nachweisen. Aus $E_1 \cup E_2 = E_2 + E_1 \cap E_2^c$ und $E_1 = E_1 \cap E_2 + E_1 \cap E_2^c$ folgt

$$P(E_1 \cup E_2) \;=\; P(E_2) + P(E_1 \cap E_2^c) \quad \text{bzw.} \quad P(E_1 \cap E_2^c) \;=\; P(E_1) - P(E_1 \cap E_2),$$

was insgesamt (2.1) ergibt.

Induktionsschritt: Aus $P(\cup_{i=1}^n E_i) = \sum_{k=1}^n (-1)^{k+1} \sum_{1\leq i_1<...<i_k\leq n} P(E_{i_1}\cap ... \cap E_{i_k})$ für $n \geq 2$ folgt mit $\hat{E}_i \overset{\text{def}}{=} E_i \cap E_{n+1}$

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} E_i\right) &\overset{(2.1)}{=} P\left(\bigcup_{i=1}^{n} E_i\right) \;+\; P(E_{n+1}) \;-\; P\left(E_{n+1}\cap\bigcup_{i=1}^{n} E_i\right) \\ &= P\left(\bigcup_{i=1}^{n} E_i\right) \;+\; P(E_{n+1}) \;-\; P\left(\bigcup_{i=1}^{n} E_i \cap E_{n+1}\right) \\ &= \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\leq i_1<...<i_k\leq n} P(E_{i_1}\cap ... \cap E_{i_k}) \;+\; P(E_{n+1}) \\ &\quad - \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\leq i_1<...<i_k\leq n} P(\underbrace{\hat{E}_{i_1}\cap ... \cap \hat{E}_{i_k}}_{=E_{i_1}\cap ... \cap E_{i_k}\cap E_{n+1}}) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\leq i_1<\ldots<i_k\leq n} P(E_{i_1}\cap\ldots\cap E_{i_k}) \;+\; P(E_{n+1}) \\
&+ \sum_{j=2}^{n+1}(-1)^{j+1} \sum_{1\leq i_1<\ldots<i_{j-1}\leq n, i_j=n+1} P(E_{i_1}\cap\ldots\cap E_{i_{j-1}}\cap E_{i_j}) \\
&= \sum_{k=1}^{n+1}(-1)^{k+1} \sum_{1\leq i_1<\ldots<i_k\leq n+1} P(E_{i_1}\cap\ldots\cap E_{i_k}).
\end{aligned}$$

◊

Um als nächstes die Stetigkeit diskreter W-Verteilungen zu zeigen, benötigen wir zuerst folgende Definition:

2.2. Definition. Eine Folge $(A_n)_{n\geq 1}$ von Teilmengen einer beliebigen Menge Ω heißt *monoton wachsend* oder *isoton* mit Limes $A \stackrel{\text{def}}{=} \cup_{n\geq 1}A_n$, wenn $A_1 \subset A_2 \subset \ldots$. Entsprechend heißt $(A_n)_{n\geq 1}$ *monoton fallend* oder *antiton* mit Limes $A \stackrel{\text{def}}{=} \cap_{n\geq 1}A_n$, wenn $A_1 \supset A_2 \supset \ldots$. Für die Konvergenz schreiben wir kurz $A_n \uparrow A$ bzw. $A_n \downarrow A$.

In Analogie zum Supremum bzw. Infimum reeller Zahlenfolgen entspricht also der Limes einer isotonen bzw. antitonen Mengenfolge $(A_n)_{n\geq 1}$ der "kleinsten oberen" bzw. "größten unteren Schranke" der A_n, d.h. der kleinsten Menge, die alle A_n enthält, bzw. größten Menge, die in allen A_n enthalten ist.

2.3. Satz. *Sei Ω eine nichtleere Menge und $P : \mathfrak{P}(\Omega) \to [0,1]$ eine normierte, endlich additive Mengenfunktion. Dann sind folgende Aussagen äquivalent:*

(a) P ist σ-additiv.

(b) P ist stetig von unten, d.h. $\lim_{n\to\infty} P(A_n) = P(A)$ für jede isotone Mengenfolge $(A_n)_{n\geq 1}$ mit Limes A.

(c) P ist stetig von oben, d.h. $\lim_{n\to\infty} P(A_n) = P(A)$ für jede antitone Mengenfolge $(A_n)_{n\geq 1}$ mit Limes A.

(d) P ist stetig in der leeren Menge, d.h. $\lim_{n\to\infty} P(A_n) = 0$ für jede antitone Mengenfolge $(A_n)_{n\geq 1}$ mit Limes $\emptyset$.

Beweis: "(a)⇒(b)" Sei $(A_n)_{n\geq 1}$ eine isotone Mengenfolge mit $A_n \uparrow A$. Mit $A_0 \stackrel{\text{def}}{=} \emptyset$ gilt

$$A \;=\; \bigcup_{n\geq 1} A_n \;=\; \sum_{k\geq 1}(A_k \cap A_{k-1}^c)$$

und folglich vermöge der σ-Additivität

$$P(A) \;=\; P\left(\sum_{k\geq 1}(A_k\cap A_{k-1}^c)\right) \;=\; \sum_{n\geq 1} P(A_n\cap A_{n-1}^c)$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} P(A_k \cap A_{k-1}^c) \;=\; \lim_{n\to\infty} \sum_{k=1}^{n} \Big(P(A_k) - P(A_{k-1})\Big)$$
$$= \lim_{n\to\infty} P(A_n).$$

"(b)⇒(c)" Aus $A_n \downarrow A$ folgt offenkundig $A_n^c \uparrow A^c = (\cap_{n\geq 1} A_n)^c = \cup_{n\geq 1} A_n^c$. Nach Voraussetzung gilt also $\lim_{n\to\infty} P(A_n^c) = P(A^c)$ und damit

$$P(A) \;=\; 1 - P(A^c) \;=\; 1 - \lim_{n\to\infty} P(A_n^c) \;=\; \lim_{n\to\infty} \big(1 - P(A_n^c)\big) \;=\; \lim_{n\to\infty} P(A_n).$$

"(c)⇒(d)" ist klar.

"(d)⇒(a)" Sei $(E_n)_{n\geq 1}$ eine Folge p.d. Teilmengen von Ω und $E = \sum_{n\geq 1} E_n$. Dann folgt $A_n \stackrel{\text{def}}{=} E \cap (\sum_{k=1}^n E_k)^c \downarrow \emptyset$ und somit unter Verwendung der endlichen Additivität von P und der Voraussetzung (d)

$$0 \;=\; \lim_{n\to\infty} P(A_n) \;=\; \lim_{n\to\infty} \left(P(E) - P\left(\sum_{k=1}^{n} E_k\right)\right) \;=\; \lim_{n\to\infty} \left(P(E) - \sum_{k=1}^{n} P(E_k)\right)$$
$$= \; P(E) - \lim_{n\to\infty} \sum_{k=1}^{n} P(E_k) \;=\; P(E) - \sum_{k\geq 1} P(E_k),$$

d.h. $P(E) = \sum_{k\geq 1} P(E_k)$. ◊.

Da diskrete W-Verteilungen σ-additiv sind, können wir als direkte Folgerung festhalten:

2.4. Korollar. *Jede diskrete W-Verteilung ist stetig von oben und unten.*

Die Menge der diskreten W-Verteilungen über einer Menge Ω ist abgeschlossen unter der Bildung konvexer Linearkombinationen:

2.5. Lemma. *Für $n \in \mathbb{N}$ seien $\lambda_n \geq 0$ mit $\sum_{n\geq 1} \lambda_n = 1$ und P_n diskrete W-Verteilungen über Ω. Dann ist auch*

$$P \stackrel{\text{def}}{=} \sum_{n\geq 1} \lambda_n P_n \tag{2.2}$$

eine diskrete W-Verteilung über Ω.

Beweis: Offensichtlich gilt für jedes $E \subset \Omega$

$$0 \;\leq\; P(E) \;=\; \sum_{n\geq 1} \lambda_n P_n(E) \;\leq\; \sum_{n\geq 1} \lambda_n \;=\; 1,$$

d.h. $P : \mathfrak{P}(\Omega) \to [0,1]$. Gemäß Satz 1.5 müssen wir noch zeigen, daß P wiederum normiert und σ-additiv ist. Die Normiertheit ergibt sich sofort, wenn wir in der obigen Ungleichung $E =$

Ω wählen. Betreffend der σ-Additivität liefert der Umordnungssatz für absolut konvergente Reihen zusammen mit der σ-Additivität der P_n für jede Folge $(E_k)_{k\geq 1}$ p.d. Teilmengen von Ω

$$\begin{aligned} P\left(\sum_{k\geq 1} E_k\right) &= \sum_{n\geq 1} \lambda_n P_n\left(\sum_{k\geq 1} E_k\right) = \sum_{n\geq 1} \lambda_n \sum_{k\geq 1} P_n(E_k) \\ &= \sum_{k\geq 1}\sum_{n\geq 1} \lambda_n P_n(E_k) = \sum_{k\geq 1} P(E_k). \end{aligned}$$ ◊

Mittels dieses Satzes gelangen wir nun leicht zu einer "standardisierten" Darstellungsform allgemeiner diskreter W-Verteilungen über einer Menge Ω unter Verwendung besonders einfacher diskreter W-Verteilungen, die wir zunächst definieren müssen:

2.6. Definition. Es seien Ω eine beliebige nichtleere Menge und $\omega_0 \in \Omega$. Dann heißt die durch $\delta_{\omega_0} : \mathfrak{P}(\Omega) \to [0,1]$,

$$\delta_{\omega_0}(E) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{falls } \omega_0 \in E \\ 0, & \text{falls } \omega_0 \notin E \end{cases},$$

definierte diskrete W-Verteilung über Ω *Dirac-Verteilung im Punkt* ω_0.

Daß δ_{ω_0} wirklich eine diskrete W-Verteilung über Ω definiert, folgt aus $\delta_{\omega_0}(\Omega) = 1$ und

$$\delta_{\omega_0}\left(\sum_{i\geq 1} E_i\right) = \left\{\begin{matrix} 1, & \text{falls } \omega_0 \in E_{i_0} \text{ für genau ein } i_0 \in \mathbb{N} \\ 0, & \text{sonst} \end{matrix}\right\} = \sum_{i\geq 1} \delta_{\omega_0}(E_i)$$

für jede Folge $(E_i)_{i\geq 1}$ p.d. Teilmengen von Ω. δ_{ω_0} bildet das stochastische Analogon zu dem *deterministischen Experiment* mit Ausgang ω_0 ($\{\omega_0\}$ ist ein fast sicheres Ereignis).

2.7. Satz. *Sei Ω eine beliebige nichtleere Menge und P eine diskrete W-Verteilung über Ω mit Träger Ω_0. Dann besitzt P die Darstellung*

(2.3) $$P = \sum_{\omega\in\Omega_0} P(\{\omega\})\delta_\omega.$$

Beweis: Für jedes $E \subset \Omega$ gilt

$$\begin{aligned} P(E) &= P(E\cap\Omega_0) = \sum_{\omega\in E\cap\Omega_0} P(\{\omega\}) \\ &= \sum_{\omega\in\Omega_0} P(\{\omega\})\delta_\omega(E) = \left(\sum_{\omega\in\Omega_0} P(\{\omega\})\delta_\omega\right)(E). \end{aligned}$$ ◊

Wie der Leser sofort einsieht, kann man in (2.3) den Träger Ω_0 durch jede abzählbare Obermenge Ω^* ersetzen.

3. Laplace-Experimente und Kombinatorik

Obgleich wir Laplace-Experimente schon in Beispiel 1.6(b) kurz eingeführt hatten, hier noch einmal die genaue Definition:

3.1. Definition. Ein endliches ZE (Ω, p) heißt *Laplace-Experiment* und die zugehörige diskrete W-Verteilung P *Laplace-Verteilung über* Ω, wenn $p(\omega) = 1/|\Omega|$ für alle $\omega \in \Omega$ und damit

$$P(E) = \frac{|E|}{|\Omega|} \tag{3.1}$$

für alle $E \subset \Omega$ gilt.

Laplace-Experimente stehen aufgrund ihrer besonders einfachen Struktur sehr häufig am Anfang einer Einführung in die Wahrscheinlichkeitsrechnung. Dies sollte aber nicht darüber hinwegtäuschen, daß schon unter ihnen zahlreiche interessante Beispiele existieren, für die die Berechnung konkreter Wahrscheinlichkeiten alles andere als einfach ist. Auf einige derartige Beispiele kommen wir am Ende des Abschnitts zurück.

Formel (3.1) für "Laplace-Wahrscheinlichkeiten" liest man - besonders in Schulbüchern - oft in der verbalisierten Form

$$P(E) = \frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}},$$

wobei natürlich die Elemente von E als die günstigen, die Elemente des Ergebnisraums Ω als die möglichen Fälle angesehen werden. Die Berechnung von Wahrscheinlichkeiten führt hier also auf das Abzählen von Elementen einer gegebenen Menge, was sich bei großen Grundgesamtheiten durchaus als äußerst schwierig entpuppen kann. Der Teilbereich der Mathematik, der sich mit der "Kunst des Abzählens" in systematischer Weise befaßt, heißt *Kombinatorik*. Deren vier wichtigsten Grundregeln wollen wir im folgenden kurz vorstellen. Sie lassen sich besonders gut anhand von zwei Vorgängen erläutern: dem Ziehen numerierter Kugeln aus einer Urne *(Urnenmodell)* sowie dem Verteilen von Teilchen auf eine gegebene Menge von Fächern *(Teilchen-Fächer-Modell)*.

1. Permutationen mit Wiederholung. Betrachten wir eine Urne mit $n \geq 1$ von 1 bis n durchnumerierten Kugeln. Zieht man aus dieser Urne sukzessiv k Kugeln mit Zurücklegen und notiert jeweils die Nummer der gezogenen Kugel, so erhält man als Ergebnis irgendein k-Tupel $(\omega_1, ..., \omega_k) \in \{1, ..., n\}^k$ und nennt dieses *k-Permutation mit Wiederholung* oder auch *geordnete k-Stichprobe mit Zurücklegen* der Menge $\{1, ..., n\}$. Das Attribut "geordnet" soll auf die Tatsache hinweisen, daß die Reihenfolge der gezogenen Nummern berücksichtigt wird. Da in jeder Komponente einer solchen k-Permutation eines der Elemente 1 bis n auftreten kann, gilt:

3.2. Satz. *Die Anzahl der k-Permutationen mit Wiederholung einer n-elementigen Menge entspricht der Mächtigkeit der Menge $\{1, ..., n\}^k$ und ist durch n^k gegeben.*

Dasselbe Ergebnis ergibt sich, wenn wir k unterscheidbare Teilchen auf n Fächer verteilen und Mehrfachbesetzungen zulassen. In dieser Situation bezeichnet ω_j die Nummer des Faches, in das das j-te Teilchen gelegt wurde ($1 \leq j \leq k$).

2. Permutationen ohne Wiederholung. Modifiziert man den obigen Ziehungsmechanismus dahingehend, daß gezogene Kugeln nicht zurückgelegt werden, so erhält man als Ergebnis ein k-Tupel aus der reduzierten Menge

$$[1, ..., n]^k \stackrel{\text{def}}{=} \{(\omega_1, ..., \omega_k) \in \{1, ..., n\}^k : \omega_i \neq \omega_j, \text{ falls } i \neq j\},$$

genannt *k-Permutation ohne Wiederholung* oder auch *geordnete k-Stichprobe ohne Zurücklegen* der Menge $\{1, ..., n\}$. In diesem Fall muß natürlich $k \leq n$ sein, weil nicht mehr Kugeln als sich in der Urne befinden gezogen werden können. Zur Berechnung der Anzahl der verschiedenen derartigen k-Permutationen überlegt man sich, daß die erste Ziehung n mögliche Ausgänge hat, die zweite dann noch $n-1$, etc. Es folgt:

3.3. Satz. *Die Anzahl der k-Permutationen ohne Wiederholung einer n-elementigen Menge entspricht der Mächtigkeit der Menge $[1, ..., n]^k$ und ist durch*

$$\frac{n!}{(n-k)!} = n \cdot (n-1) \cdot ... \cdot (n-k+1)$$

gegeben.

Im Spezialfall $k = n$ liefert der obige Ziehungsmechanismus offenkundig irgendeine Anordnung der Zahlen $1, ..., n$, die einfach als *Permutation der Zahlen* $1, ..., n$ bezeichnet wird. Davon gibt es nach Satz 3.3 $\frac{n!}{0!} = n!$ verschiedene.

In einem Teilchen-Fächer-Modell erhalten wir dasselbe Ergebnis, wenn wir k unterscheidbare Teilchen auf n Fächer verteilen, aber keine Mehrfachbesetzungen erlauben. ω_j bezeichnet wiederum die Nummer des Faches, in das das j-te Teilchen gelegt wurde ($1 \leq j \leq k$).

3. Kombinationen ohne Wiederholung. Betrachten wir nun die Situation, daß $k \leq n$ Kugeln *gleichzeitig* gezogen werden (einmaliges Greifen in die Urne). In diesem Fall läßt sich nur schwerlich eine ausgezeichnete Reihenfolge der gezogenen Nummern angeben, d.h., das Ergebnis besteht diesmal in einer Teilmenge $\{\omega_1, ..., \omega_k\}$ von $\{1, ..., n\}$, auch *k-Kombination ohne Wiederholung* oder *ungeordnete k-Stichprobe ohne Zurücklegen* genannt. Wir überlegen uns, daß jede solche ungeordnete k-Stichprobe auf genau $k!$ verschiedene Weise in eine geordnete k-Stichprobe überführt werden kann. Es existieren folglich $k!$ mal so viele geordnete wie ungeordnete k-Stichproben ohne Zurücklegen der Menge $\{1, ..., n\}$. Unter Benutzung von Satz 3.3 liefert dies:

3.4. Satz. *Die Anzahl der k-Kombinationen ohne Wiederholung einer n-elementigen Menge entspricht der Anzahl der k-elementigen Teilmengen einer n-elementigen Menge und ist durch*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{k!}$$

gegeben.

Das korrespondierende Teilchen-Fächer-Modell liegt vor, wenn man k ununterscheidbare Kugeln auf n Fächer ohne Mehrfachbesetzung verteilt. Das Ergebnis $\{\omega_1,...,\omega_k\}$ beschreibt in diesem Fall die Nummern derjenigen Fächer, die (genau) eine Kugel enthalten.

4. Kombinationen mit Wiederholung. Wenden wir uns als letztes dem kniffeligsten Fall zu, daß aus der Urne sukzessiv k Kugeln mit Zurücklegen gezogen werden, jedoch ohne am Ende die Reihenfolge zu berücksichtigen. Unser Ziehungsergebnis, genannt *k-Kombination mit Wiederholung* oder *ungeordnete k-Stichprobe mit Zurücklegen*, besteht dann nur noch aus dem Vektor der Häufigkeiten, mit denen die einzelnen Kugeln aufgetreten sind, d.h. aus einem Element der Menge

$$\mathcal{S}_k^n \stackrel{\text{def}}{=} \left\{(\omega_1,...,\omega_n) \in \{0,...,k\}^n : \sum_{j=1}^n \omega_j = k\right\}.$$

Um die Mächtigkeit von $\mathcal{S}_k^n$ zu bestimmen, betrachten wir ein äquivalentes Abzählungsproblem und codieren zunächst $(\omega_1,...,\omega_k) \in \mathcal{S}_k^n$ in eineindeutiger Weise durch ein $(n+k-1)$-Tupel, bestehend aus k Sternchen "$*$" und $n-1$ Trennstrichen "$|$": Für jedes ω_j, $1 \le j \le n$, besetzen wir entsprechend viele aufeinanderfolgende Positionen mit Sternchen und trennen die so entstehenden n Blöcke durch $n-1$ Trennstriche.

$$(\omega_1,...,\omega_n) \quad \leftrightarrow \quad (\underbrace{*,...,*}_{\omega_1\text{-mal}}|\underbrace{*,...,*}_{\omega_2\text{-mal}}|...|\underbrace{*,...,*}_{\omega_n\text{-mal}})$$

Beachte, daß auch Blöcke der Länge 0 zugelassen sind. Die Anzahl der Elemente von $\mathcal{S}_k^n$ entspricht damit der Anzahl der Möglichkeiten, $(n+k-1)$-Tupel aus k Sternchen und $n-1$ Trennstrichen zu bilden. Da jedes solche Tupel durch die Positionen, an denen die Trennstriche stehen, vollständig festgelegt ist (an den übrigen Positionen stehen gerade die Sternchen), also durch genau eine $(n-1)$-elementige Teilmenge von $\{1,...,n+k-1\}$, folgern wir mit Satz 3.4:

3.5. Satz. *Die Anzahl der k-Kombinationen mit Wiederholung einer n-elementigen Menge entspricht der Anzahl der $(n-1)$-elementigen Teilmengen einer $(n+k-1)$-elementigen Menge und ist durch*

$$\binom{n+k-1}{n-1} = \binom{n+k-1}{k} = \frac{(n+k-1)\cdot\ldots\cdot(n+1)n}{k!}$$

gegeben.

Auch hier hat man natürlich ein Analogon unter den Teilchen-Fächer-Modellen, und zwar, wenn k ununterscheidbare Kugeln auf n Fächer mit Mehrfachbesetzung verteilt werden. In diesem Fall gibt die j-te Komponente ω_j des Ergebnisses $(\omega_1, ..., \omega_n) \in \mathcal{S}_k^n$ die Anzahl der Teilchen, die im j-ten Fach liegen, an.

Mittels der soeben vorgestellten vier Grundregeln der Kombinatorik lassen sich eine große Zahl von Abzählaufgaben erfolgreich behandeln. Dabei hat man jedoch i.a. gleich mehrere dieser Regeln in geeigneter Weise zu kombinieren, denn nur in seltenen Fällen entspricht ein gegebenes Problem einem der obigen "Grundtypen". Zur besseren Übersicht fassen wir die Ergebnisse in einer Tabelle zusammen:

Ziehen von k numerierten Kugeln aus einer Urne mit n Kugeln *Verteilen von k Teilchen auf n Fächer*			
Reihenfolge wird berücksichtigt *Teilchen sind unterscheidbar*	Wiederholungen *Mehrfachbesetzungen*	Modell	Anzahl
ja	ja	1	n^k
ja	nein	2	$\frac{n!}{(n-k)!}$
nein	nein	3	$\binom{n}{k}$
nein	ja	4	$\binom{n+k-1}{n-1}$

5. Binomialkoeffizienten. Aufgrund ihres häufigen Auftretens in kombinatorischen Formeln wollen wir in Kürze einige wichtige Eigenschaften der Binomialkoeffizienten $\binom{n}{k}$ zusammentragen. Die Identität, nach der sie benannt wurden, ist der *binomische Lehrsatz*:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \tag{3.2}$$

für alle $a, b \in \mathbb{R}$ und $n \in \mathbb{N}_0$. Er liefert insbesondere

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n, \tag{3.3}$$

wenn man $a = b = 1$ wählt, sowie

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0 \tag{3.4}$$

im Fall $a = -1, b = 1$. Differenziert man (3.2) nach a und setzt anschließend $a = b = 1$, folgt außerdem

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}. \tag{3.5}$$

Für $k \in \mathbb{Z} - \{0, ..., n\}$ wird aus Zweckmäßigkeitsgründen $\binom{n}{k} \overset{\text{def}}{=} 0$ vereinbart.

Nach Satz 3.4 gibt $\binom{n}{k}$ die Anzahl der k-elementigen Teilmengen von $\{1, ..., n\}$ an. Von diesen enthalten $\binom{n-1}{k-1}$ das Element 1 und $\binom{n-1}{k}$ nicht, was

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \tag{3.6}$$

impliziert. Erwähnen wollen wir auch die sofort aus der Definition folgende Symmetrieeigenschaft

$$\binom{n}{k} = \binom{n}{n-k}. \tag{3.7}$$

Schreibt man die Binomialkoeffizienten im sogenannten *Pascalschen Dreieck*

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
 & & \vdots & & & & & & \vdots & &
\end{array}$$

nieder, wobei die $(n+1)$-te Reihe $\binom{n}{k}$ für $k = 0, ..., n$ enthält, so besagt (3.6), daß sich jeder Wert im Inneren des Dreiecks als Summe der beiden schräg über ihm stehenden Werte berechnet. Man sieht ferner, daß die Binomialkoeffizienten bis zur Mitte hin ansteigen, d.h.

$$\binom{n}{k} < \binom{n}{k+1}, \quad \text{falls} \quad k+1 < \frac{n}{2}. \tag{3.8}$$

Für große n und k macht die Berechnung von $\binom{n}{k}$ i.a. beträchtliche Mühe. In diesen Fällen erweist sich die sogenannte *Stirlingsche Formel* als "handliche" Approximation für Fakultäten. Sie lautet:

$$(2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n+\frac{1}{12n+1}} < n! < (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n+\frac{1}{12n}} \tag{3.9}$$

für alle $n \in \mathbb{N}$, was insbesondere

$$n! \simeq (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n} \tag{3.10}$$

für $n \to \infty$ liefert. Einen Beweis findet der Leser z.B. in [1], S. 52-54.

6. Anwendungsbeispiele. Es ist Zeit für eine Reihe von Beispielen, in denen das zuvor Hergeleitete Anwendung findet. Wir beginnen mit einem kleinen Exkurs in die statistische Physik.

3.6. Teilchen-Fächer-Modelle in der statistischen Physik. [nach [3], S. 18-20] Da es für große Systeme von Teilchen aussichtslos ist, deren Bewegungen nach den Gesetzen der Newtonschen Mechanik zu beschreiben — die zugehörigen Differentialgleichungssysteme sind einfach zu groß — wählt man stattdessen einen stochastischen Modellansatz. Der Zustand eines Teilchens zu einem festen Zeitpunkt ist durch seine drei Orts- und seine drei Geschwindigkeitskoordinaten gegeben. Für ein geschlossenes System dürfen wir annehmen, daß die Menge der möglichen Zustände, genannt *Phasenraum*, eine beschränkte Teilmenge des $\mathbb{R}^6$ bildet, die in 6-dimensionale gleichgroße Quader, genannt Zellen oder Fächer, eingeteilt wird, welche so klein sind, daß die relevanten Größen wie z.B. die Energie in jeder Zelle als konstant angesehen werden können. Sei n die Anzahl der Zellen und k die Anzahl der Teilchen.

Interessiert man sich für die Energie der Teilchen, so teilt man die Zellen in Gruppen gleichen Energieniveaus ein. Es gebe m Energieniveaus $\mathfrak{E}_1, ..., \mathfrak{E}_m$ und n_j Zellen vom Energieniveau $\mathfrak{E}_j$, $j = 1, ..., m$. Wir dürfen die Zellen o.E. so numerieren, daß die ersten n_1 das Energieniveau $\mathfrak{E}_1$, die folgenden n_2 das Energieniveau $\mathfrak{E}_2$, etc. besitzen. Bezeichnet ω_j die Nummer der Zelle, in der sich das j-te Teilchen gerade aufhält, so beschreibt $(\omega_1, ..., \omega_k)$ den *Mikrozustand* des Systems (zu einem festen Zeitpunkt). Sei $N_j(\omega)$ die *Besetzungszahl* des j-ten Energieniveaus, d.h. die Anzahl der Teilchen, die sich in Zellen vom Energieniveau $\mathfrak{E}_j$ aufhalten. Wir interessieren uns für die Wahrscheinlichkeit des Ereignisses

$$E_{k_1,...,k_m} \stackrel{\text{def}}{=} \{\omega : N_1(\omega) = k_1, ..., N_m(\omega) = k_m\},$$

daß sich für $j = 1, ..., m$ gerade k_j Teilchen in Zellen vom Energieniveau $\mathfrak{E}_j$ aufhalten, wobei natürlich $k_1 + ... + k_m = k$ gelten muß. Eine angemessene diskrete W-Verteilung über dem Grundraum Ω wird nun in Abhängigkeit von der Art der betrachteten Teilchen vorgenommen.

(a) Maxwell-Boltzmann-Statistik. Können sich beliebig viele Teilchen gleichzeitig in einer Zelle aufhalten und sind diese unterscheidbar, so ist $\Omega = \{1, ..., n\}^k$ die Menge der Mikrozustände. Man spricht von der Maxwell-Boltzmann-Statistik, wenn über diesem Grundraum eine Laplace-Verteilung zugrundegelegt wird, d.h. $P(\{\omega\}) = n^{-k}$ für alle $\omega \in \Omega$ gilt. Es gibt

$$\binom{k}{k_1}\binom{k-k_1}{k_2} \cdot ... \cdot \binom{k-k_1-...-k_{m-1}}{k_m} = \frac{k!}{k_1! \cdot ... \cdot k_m!}$$

Aufteilungen der k Teilchen auf die m Energieniveaus und ferner $n_1^{k_1} n_2^{k_2} \cdot ... \cdot n_m^{k_m}$ Aufteilungen der Teilchen auf die einzelnen Zellen mit jeweiligem Energieniveau, was

$$|E_{k_1,...,k_m}| = \frac{k!}{k_1! \cdot ... \cdot k_m!} n_1^{k_1} n_2^{k_2} \cdot ... \cdot n_m^{k_m}$$

und damit

$$P(E_{k_1,\ldots,k_m}) = \frac{k!}{k_1! \cdot \ldots \cdot k_m!} \left(\frac{n_1}{n}\right)^{k_1} \left(\frac{n_2}{n}\right)^{k_2} \cdot \ldots \cdot \left(\frac{n_m}{n}\right)^{k_m} \tag{3.11}$$

impliziert. Eine gute Übereinstimmung dieser an sich aus einem willkürlichen Modellansatz gewonnenen Energieverteilung läßt sich etwa für Gasmoleküle bei mittleren und hohen Temperaturen nachweisen.

(b) Fermi-Dirac-Statistik. Ein anderes Bild ergibt sich bei Betrachtung von Elementarteilchen. Hier stimmen die Resultate, die man für das obige Modell ableiten kann, nicht mit den Beobachtungen überein. Für Elementarteilchen mit halbzahligem Spin, genannt Fermionen (z.B. Elektronen, Protonen, Neutronen), gelangt man aber zu guten Übereinstimmungen, wenn man eine Laplace-Verteilung auf dem Raum $\Omega = [1, ..., n]^k$ zugrundelegt und somit Mehrfachbesetzungen ausschließt. Für Fermionen gilt demnach das sogenannte *Pauli-Verbot*: Es ist verboten, daß sich in einer Zelle mehr als ein Teilchen aufhält. Zur Berechnung von $P(E_{k_1,\ldots,k_m})$ notieren wir, daß es nach wie vor $\frac{k!}{k_1! \cdot \ldots \cdot k_m!}$ Aufteilungen der k Teilchen auf die m Energieniveaus gibt, daß aber aufgrund des Pauli-Verbots nur $\frac{n_1!}{(n_1-k_1)!} \cdot \ldots \cdot \frac{n_m!}{(n_m-k_m)!}$ Aufteilungen der Teilchen auf die Zellen mit jeweiligem Energieniveau existieren. Es folgt

$$|E_{k_1,\ldots,k_m}| = \frac{k!}{k_1! \cdot \ldots \cdot k_m!} \cdot \frac{n_1!}{(n_1 - k_1)!} \cdot \ldots \cdot \frac{n_m!}{(n_m - k_m)!}$$

und weiter unter Beachtung von $|\Omega| = \frac{n!}{(n-k)!}$ (Satz 3.3)

$$\begin{aligned} P(E_{k_1,\ldots,k_m}) &= \frac{k!}{k_1! \cdot \ldots \cdot k_m!} \cdot \frac{n_1!}{(n_1 - k_1)!} \cdot \ldots \cdot \frac{n_m!}{(n_m - k_m)!} \cdot \frac{(n-k)!}{n!} \\ &= \frac{\binom{n_1}{k_1} \cdot \ldots \cdot \binom{n_m}{k_m}}{\binom{n}{k}}. \end{aligned} \tag{3.12}$$

Da jedes $(\omega_1, ..., \omega_k) \in [1, ..., n]^k$ auf $k!$ verschiedene Weisen permutierbar ist und dabei die Energieverteilung unverändert bleibt, können wir ebensogut zu $\{\omega_1, ..., \omega_k\}$ übergehen, d.h. die k-elementigen Teilmengen von $\{1, ..., n\}$ als gleichwahrscheinliche Mikrozustände des Systems auffassen. Dies bedeutet, daß wir nur noch angeben, welche der n Zellen mit einem Teilchen besetzt sind. Da es in diesem Fall $\binom{n}{k}$ Mikrozustände gibt (Satz 3.4) und $\binom{n_1}{k_1} \cdot \ldots \cdot \binom{n_m}{k_m}$ Besetzungskonfigurationen mit k_j Zellen zum Niveau $\mathfrak{E}_j$ für $j = 1, ..., m$, erhalten wir (natürlich) wiederum das obige Ergebnis für $P(E_{k_1,\ldots,k_m})$.

(c) Bose-Einstein-Statistik. Für Bosonen, d.h. Elementarteilchen mit ganzzahligem Spin (z.B. Photonen und Mesonen), kommt man zu Ergebnissen in Übereinstimmung mit physikalischen Beobachtungen, wenn man die Teilchen als nicht unterscheidbar ansieht, Mehrfachbesetzungen von Zellen erlaubt und als gleichwahrscheinliche Mikrozustände die möglichen Besetzungskonfigurationen $\omega = (\omega_1, ..., \omega_n)$ wählt, in denen die j-te Komponente die Anzahl der Teilchen in der j-ten Zelle angibt. Zugrundegelegt wird demnach eine Laplace-Verteilung auf dem

Raum $\Omega = \mathcal{S}_k^n$, der nach Satz 3.5 aus $\binom{n+k-1}{n-1}$ Elementen besteht. Man überlegt sich leicht, daß in diesem Modell

$$E_{k_1,\dots,k_m} \;=\; \mathcal{S}_{k_1}^{n_1} \times \dots \times \mathcal{S}_{k_m}^{n_m}$$

und folglich

$$P(E_{k_1,\dots,k_m}) \;=\; \frac{|\mathcal{S}_{k_1}^{n_1}| \cdot \dots \cdot |\mathcal{S}_{k_m}|}{|\mathcal{S}_k^n|} \;=\; \frac{\binom{n_1+k_1-1}{n_1-1} \cdot \dots \cdot \binom{n_m+k_m-1}{n_m-1}}{\binom{n+k-1}{n-1}} \tag{3.13}$$

gilt.

3.7. Das Paradoxon der ersten Kollision. [nach [2], S. 67-71] Am 29. 6. 1995 erfuhr man aus der Tagespresse die angebliche Sensation, daß erstmals in der 40jährigen Geschichte des deutschen Zahlenlottos zwei identische Gewinnreihen festgestellt wurden. Am 21. Juni des Jahres ergab die Ziehung A des Mittwochslottos die Zahlenreihe 15 – 25 – 27 – 30 – 42 – 48. Genau dieselben Zahlen wurden bei der 1628. Ausspielung des Samstaglottos schon einmal gezogen, nämlich am 26. 12. 1986. Welch ein Zufall, so scheint es, wenn man bedenkt, daß es fast 14 Millionen verschiedene Sechserreihen gibt.

Im folgenden wollen wir dieser angeblichen Sensation auf den Grund gehen und zunächst einmal klären, nach welchem Ereignis eigentlich gesucht wird. Als Sensation gilt offenbar, daß *irgendeine* Gewinnreihe *irgendeines* Zahlenlottos (Mittwochslotto A, Mittwochslotto B, Samstaglotto) schon in *irgendeiner* früheren Ziehung auftrat. Wir müssen daher die Ausspielungen aller drei wöchentlich stattfindenden Ziehungen zusammenfassen. Da bis zum 21. 6. 1995 2071 Ausspielungen des Samstaglottos und jeweils 472 Ausspielungen des Mittwochslottos A bzw. B stattgefunden hatten, scheint das sensationelle Ereignis darin zu bestehen, daß *schon* in 3016 Ausspielungen eine Gewinnreihe wiederholt auftrat.

Für unsere weiteren Überlegungen müssen wir als erstes ein geeignetes Laplace-Modell spezifizieren. Die Menge der möglichen Ziehungsergebnisse ist offenkundig durch $\mathfrak{P}_6(\{1,\dots,49\})$, die Menge der 6-elementigen Teilmengen der Menge $\{1,\dots,49\}$, gegeben, die nach Satz 3.4 aus $n \overset{\text{def}}{=} \binom{49}{6} = 13\,983\,316$ Elementen besteht. Identifizieren wir jede solche Teilmenge mit dem entsprechenden 6-Tupel der aufsteigend angeordneten Elemente, so können wir diese lexikographisch anordnen (☞ umseitige Tabelle). Um das gesuchte Ereignis in einem Laplace-Experiment zu beschreiben, denken wir uns die möglichen Gewinnreihen als Fächer und die Ausspielungen als Plazierungen von Teilchen in die entsprechenden Fächer. Es ist klar, daß nach spätestens $n+1$ Plazierungen mindestens ein Fach mehrfach besetzt ist, d.h. im realen Modell mindestens eine Gewinnreihe mehrfach aufgetreten ist. Aus diesem Grund betrachten wir nun das Laplace-Experiment, $n+1$ Teilchen sukzessiv auf n Fächer unter Zulassung von Mehrfachbesetzungen zu verteilen, und erhalten damit als Ergebnisraum

$$\Omega \;=\; \{1,\dots,n\}^{n+1}.$$

Nr.	Ziehung					
1	1	2	3	4	5	6
2	1	2	3	4	5	7
3	1	2	3	4	5	8
$\vdots$						
49	1	2	3	4	5	49
50	1	2	3	4	6	7
51	1	2	3	4	6	8
$\vdots$						
n	44	45	46	47	48	49

Lexikographische Anordnung aller möglichen Lottoergebnisse

Die j-te Komponente eines Ergebnisses $\omega = (\omega_1, ..., \omega_{n+1})$ gibt an, in welches Fach die j-te Kugel gelegt wird. Im Rahmen dieses Teilchen-Fächer-Modells besteht das gesuchte Ereignis E_{3016} darin, daß nach 3016 Plazierungen mindestens ein Fach mehrfach besetzt ist (eine Kollision auftritt). Eine formale Definition lautet

$$E_{3016} \;=\; \{\omega \in \Omega : |\{\omega_1, ..., \omega_{3016}\}| < 3016\}.$$

Als Mathematiker sind wir natürlich an einer allgemeinen Lösung des Problems interessiert und ersetzen 3016 durch ein beliebiges i zwischen 2 und $n+1$. Gesucht ist demnach die Wahrscheinlichkeit des Ereignisses

$$E_i \;=\; \{\omega \in \Omega : |\{\omega_1, ..., \omega_i\}| < i\},$$

daß nach i Plazierungen von Teilchen mindestens ein mehrfach besetztes Fach vorliegt. Zur Bestimmung von $|E_i|$ stellen wir fest, daß dessen Komplement

$$E_i^c \;=\; [1, ..., n]^i \times \{1, ..., n\}^{n+1-i}$$

das Ereignis, daß bis zur i-ten Plazierung keine Mehrfachbesetzung auftritt, bezeichnet. Da offenbar $|E_i^c| = |[1, ..., n]^i| \cdot |\{1, ..., n\}^{n+1-i}| = \frac{n!}{(n-i)!} \cdot n^{n+1-i}$, folgt schließlich

$$\begin{aligned} P(E_i) \;&=\; 1 - P(E_i^c) \;=\; 1 - \frac{|E_i^c|}{|\Omega|} \\ &=\; 1 - \frac{\frac{n!}{(n-i)!} \cdot n^{n+1-i}}{n^{n+1}} \\ &=\; 1 - \frac{n(n-1) \cdot \ldots \cdot (n-i+1)}{n^i} \\ &=\; 1 - \prod_{j=1}^{i-1} \left(1 - \frac{j}{n}\right), \end{aligned} \tag{3.14}$$

für alle $2 \leq i \leq n+1$. Eine Auswertung dieses Ergebnisses für $i = 3016$ liefert

$$P(E_{3016}) = 0.2775...,$$

was in etwa der Wahrscheinlichkeit entspricht, mit zwei fairen Würfeln eine Augensumme von mindestens 9 zu erzielen ($\frac{10}{36} = 0.2777...$). Von einem sensationellen Ereignis kann also keine Rede mehr sein. Eine Aufstellung weiterer $P(E_i)$ für ausgewählte i im Fall unseres Lottobeispiels ($n = 13\,983\,816$) findet der Leser in der unteren Tabelle.

i	$P(E_i)$	i	$P(E_i)$	i	$P(E_i)$
500	0.0089	4500	0.5152	8500	0.9245
1000	0.0351	5000	0.5909	9000	0.9448
1500	0.0773	5500	0.6609	9500	0.9603
2000	0.1332	6000	0.7240	10000	0.9720
2500	0.2002	6500	0.7792	10500	0.9806
3000	0.2751	7000	0.8266	11000	0.9868
3500	0.3546	7500	0.8662	11500	0.9912
4000	0.4356	8000	0.8986	12000	0.9942

Es mag überraschend erscheinen, daß bei fast 14 Millionen möglichen Tipreihen schon nach 4500 Ziehungen eine mehr als 50prozentige Chance besteht, mindestens eine Tipreihe zweimal zu erhalten. Der Grund hierfür ist, daß wir auf *irgendeine* und nicht auf eine bestimmte Kollision warten. Es folgt aus dem anschließenden Satz u.a., daß die benötigte Mindestanzahl von Teilchen, um bei rein zufälliger Besetzung von n Fächern eine Kollisionswahrscheinlichkeit $\geq 1/2$ zu sichern, von der Größenordnung $n^{1/2}$ ist.

3.7.1. Satz. *Gegeben die vorherige Situation, jedoch mit beliebigem $n \in \mathbb{N}$, gilt*

$$1 - \exp\left(-\frac{i(i-1)}{2n}\right) \;\leq\; P(E_i) \;\leq\; 1 - \exp\left(-\frac{i(i-1)}{2(n-i+1)}\right) \tag{3.15}$$

für alle $2 \leq i \leq n+1$ und ferner

$$\lim_{n\to\infty} P(E_{[tn^{1/2}]}) \;=\; 1 - e^{-t^2/2} \tag{3.16}$$

für alle $t > 0$, wobei $[x] \stackrel{\text{def}}{=} \sup\{k \in \mathbb{N} : k \leq x\}$.

Beweis: Unter Benutzung der Ungleichung $\log x \leq x - 1$ für alle $x > 0$ sowie $\sum_{j=1}^{k} j = \frac{k(k+1)}{2}$ für alle $k \in \mathbb{N}$ ergibt sich in (3.14)

$$P(E_i) \;=\; 1 - \prod_{j=1}^{i-1}\left(1 - \frac{j}{n}\right)$$

$$
\begin{aligned}
&= 1 - \exp\left(\sum_{j=1}^{i-1} \log\left(1 - \frac{j}{n}\right)\right) \\
&\geq 1 - \exp\left(- \sum_{j=1}^{i-1} \frac{j}{n}\right) \\
&= 1 - \exp\left(- \frac{i(i-1)}{2n}\right),
\end{aligned}
$$

also die untere Schranke in (3.15). Entsprechend liefert $\log x \geq 1 - \frac{1}{x}$ für alle $x > 0$ (ersetze x durch $1/x$ in der obigen Ungleichung für $\log x$)

$$
\begin{aligned}
P(E_i) &= 1 - \prod_{j=1}^{i-1} \left(1 - \frac{j}{n}\right) \\
&= 1 - \exp\left(\sum_{j=1}^{i-1} \log\left(\frac{n-j}{n}\right)\right) \\
&\leq 1 - \exp\left(\sum_{j=1}^{i-1} \left(1 - \frac{n}{n-j}\right)\right) \\
&= 1 - \exp\left(- \sum_{j=1}^{i-1} \frac{j}{n-j}\right) \\
&\leq 1 - \exp\left(- \sum_{j=1}^{i-1} \frac{j}{n-i+1}\right) \\
&= 1 - \exp\left(- \frac{i(i-1)}{2(n-i+1)}\right),
\end{aligned}
$$

d.h. die obere Schranke in (3.15).

Zum Beweis von (3.16) fixieren wir ein $t > 0$ und bemerken, daß für jedes hinreichend große n

$$2 \leq k_n \stackrel{\text{def}}{=} [tn^{1/2}] \leq tn^{1/2} < k_n + 1 \leq n + 1$$

gilt. Da außerdem

$$\lim_{n\to\infty} \exp\left(- \frac{k_n(k_n-1)}{2n}\right) \;=\; \lim_{n\to\infty} \exp\left(- \frac{(k_n+1)k_n}{2(n-k_n+2)}\right) \;=\; e^{-t^2/2},$$

ergibt sich vermöge (3.15) und $E_{k_n} \subset E_{k_n+1}$

$$1 - e^{-t^2/2} \;\leq\; \lim_{n\to\infty} P(E_{k_n}) \;\leq\; \lim_{n\to\infty} P(E_{k_n+1}) \;\leq\; 1 - e^{-t^2/2},$$

und somit die Behauptung. $\diamondsuit$

Wie man nun sofort sieht, liefert (3.16) insbesondere

$$\lim_{n\to\infty} P\left(E_{\left[\sqrt{2n\log 2}\right]}\right) \;=\; \frac{1}{2}, \tag{3.17}$$

wobei im obigen Lottobeispiel $\left[\sqrt{2n \log 2}\,\right] = 4402$ gilt.

Beim *"Geburtstagsproblem"*, einer weitaus bekannteren Variante des Kollisionsproblems, wird nach der Wahrscheinlichkeit gefragt, daß in einer Gruppe von i zufällig ausgewählten Personen mindestens zwei am gleichen Tag Geburtstag haben unter der Annahme, daß alle 365 Tage eines Jahres als Geburtstage gleichwahrscheinlich sind. Unterstellen wir also ein Laplace-Modell, so ist dieses offenkundig formal äquivalent zu dem zuletzt beschriebenen Lottomodell mit lediglich anderen Werten für n und i. Das Ereignis einer mehrfach auftretenden Gewinnreihe entspricht hier dem Auftreten eines Doppelgeburtstages. Da die n Fächer nun natürlich den möglichen Geburtstagen gleichzusetzen sind, gilt $n = 365$. Die Wahrscheinlichkeit für das Auftreten eines Doppelgeburtstages unter i Personen beträgt demnach $1 - \prod_{j=1}^{i-1} \left(1 - \frac{j}{n}\right)$. Sie ist ab $i = 23$ Personen größer als $\frac{1}{2}$, was wiederum viele überraschen wird, die dies erst für eine wesentlich größere Anzahl von Personen erwartet hätten. Für all jene Skeptiker, die die Annahme gleichwahrscheinlicher Geburtstage (zurecht) für unrealistisch halten, sei bemerkt, daß die Wahrscheinlichkeit für das Auftreten eines Doppelgeburtstages unter jeder anderen Verteilungsannahme noch größer ist.

3.8. Das Koinzidenz-Paradoxon. [nach [2], S. 74-77] Hierbei handelt es sich um ein klassisches Problem der W-Theorie, das seinen Ursprung in der Untersuchung des *"treize"-Spiels* durch de Montmort (1708) hat.

- Beim "treize"-Spiel werden dreizehn Karten vom Wert 1,...,13 gut gemischt und dann eine nach der anderen aufgedeckt. Stimmt kein Kartenwert mit der Nummer der Ziehung überein, gewinnt der Spieler, andernfalls die Bank.

Dieselbe Problemstellung findet man unter dem Sammelbegriff *Rencontre-Problem* in zahllosen anderen Verkleidungen, z.B. als *Lambertsches Problem der vertauschten Briefe* (Lambert 1771):

- n Briefe werden rein zufällig in n adressierte Umschläge gesteckt. Mit welcher Wahrscheinlichkeit gelangt mindestens ein Brief in den richtigen Umschlag?

Der gemeinsame mathematische Nenner dieser Verkleidungen ist, daß die Zahlen $1, ..., n$ rein zufällig permutiert werden und nach der Wahrscheinlichkeit gefragt wird, daß diese Permutation mindestens ein Element "fest läßt", d.h. auf sich selbst abbildet. Man nennt ein solches Element *Fixpunkt* der Permutation. Wir betrachten demnach als Ergebnisraum die Menge $\Omega = [1, ..., n]^n$ aller Permutationen der Zahlen $1, ..., n$ mit der Laplace-Verteilung. Bezeichnet für $i = 1, ..., n$

$$E_i \;=\; \{(\omega_1, ..., \omega_n) \in \Omega : \omega_i = i\}$$

die Teilmenge aller Permutationen mit Fixpunkt i, so ist das gesuchte Ereignis E aller Permutationen mit mindestens einem Fixpunkt offenkundig durch deren Vereinigung gegeben, und es folgt vermöge der Siebformel (Satz 2.1(d)) und wegen $|\Omega| = n!$

$$\begin{aligned} P(E) &= \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\le i_1<...<i_k\le n} P(E_{i_1}\cap ... \cap E_{i_k}) \\ &= \frac{1}{n!}\sum_{k=1}^{n}(-1)^{k+1} \sum_{1\le i_1<...<i_k\le n} |E_{i_1}\cap ... \cap E_{i_k}|. \end{aligned}$$

Nun ist $E_{i_1}\cap...\cap E_{i_k}$ nicht anderes als die Menge aller Permutationen der Zahlen $1,..,n$, welche die Zahlen $i_1,...,i_k$ fest lassen. Da auf den übrigen $n-k$ Positionen dann noch $n-k$ Elemente frei permutierbar sind, gilt $|E_{i_1}\cap ... \cap E_{i_k}| = (n-k)!$, unabhängig von der Wahl von $i_1,...,i_k$. Da außerdem $\binom{n}{k}$ streng aufsteigende k-Tupel $(i_1,...,i_k) \in \{1,...,n\}^k$ existieren, wie man leicht einsieht, erhalten wir

$$P(E) = \frac{1}{n!}\sum_{k=1}^{n}(-1)^{k+1}\binom{n}{k}(n-k)! = \sum_{k=1}^{n}\frac{(-1)^{k+1}}{k!}. \tag{3.18}$$

Modell und Ergebnis hängen natürlich von der Anzahl n der permutierten Zahlen ab. Wenn wir dies dadurch zum Ausdruck bringen, daß wir die jeweilige Laplace-Verteilung mit n indizieren, also P_n für P schreiben, so folgt

$$\lim_{n\to\infty} P_n(E) = \sum_{k\ge 1}\frac{(-1)^{k+1}}{k!} = 1-\frac{1}{e} \approx 0.632. \tag{3.19}$$

Eine numerische Auswertung der in (3.18) auftretenden Summe liefert für $n = 2,3,4,5,6$ die bis auf vier Nachkommastellen gerundeten Werte 0.5, 0.6667, 0.6250, 0.6333 und 0.6319. Da die Summanden alternieren und $\frac{1}{n!} \le \frac{1}{7!} = 0.00019...$ für $n \ge 7$ gilt, erhalten wir das folgende überraschende Resultat:

3.8.1. Satz. *Die Wahrscheinlichkeit für das Auftreten mindestens eines Fixpunkts in einer rein zufälligen Permutation einer n-elementigen Menge ist für $n \ge 7$ bis auf eine Abweichung von höchstens* 0.0002 *konstant und bis auf drei Nachkommastellen durch 0.632 gegeben.*

3.9. Ein Auszählungproblem. Bei einer Wahl zwischen zwei Kandidaten A und B seien a Stimmen auf A und b Stimmen auf B gefallen, wobei $0 \le a < b$. Wir fragen nach der Wahrscheinlichkeit, daß bei einer zufälligen Auszählung der $a+b$ Stimmen Kandidat B stets vor Kandidat A liegt. Als Ergebnisraum wählen wir die Menge aller $(a+b)$-Tupel $(\omega_1,...,\omega_{a+b})$ bestehend aus a "-1"sen und b "1"sen, wobei -1 eine Stimme für A und 1 eine Stimmen für B bedeutet. Wie man sofort einsieht, gilt damit

$$\Omega = \mathcal{P}_{a,b} \stackrel{\text{def}}{=} \left\{(\omega_1,...,\omega_{a+b}) \in \{-1,1\}^{a+b} : \sum_{j=1}^{a+b}\omega_j = b-a\right\}.$$

Den Stimmenunterschied zwischen B und A bis zur k-ten Auszählung gibt $\sum_{j=1}^{k} \omega_j$ an. Das gesuchte Ereignis lautet deshalb

$$E = \left\{ (\omega_1, ..., \omega_{a+b}) \in \Omega : \sum_{j=1}^{k} \omega_j \geq 1 \text{ f.a. } 1 \leq k \leq a+b \right\}.$$

Da die Stimmenauszählung in zufälliger Reihenfolge erfolgt, legen wir eine Laplace-Verteilung über Ω zugrunde.

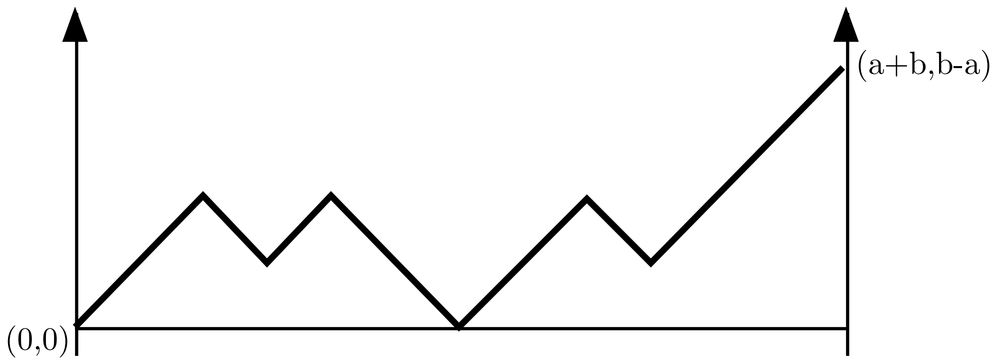

BILD 3.1. Der Pfad zu $(1, 1, -1, 1, -1, -1, 1, 1, -1, 1, 1, 1)$ $[m = 4, n = 8]$.

Zur Berechnung von $P(E)$ notieren wir als erstes, daß $|\Omega| = \binom{a+b}{a}$. Jedes $(a+b)$-Tupel aus Ω kann nämlich eineindeutig der a-elementigen Teilmenge von $\{1, ..., a+b\}$, deren Elemente die Positionen der "-1"sen angeben, zugeordnet werden. Die Bestimmung von $|E|$ ist nicht ganz so einfach und erfordert in der Tat einen hübschen "Abzähltrick", den man als *Spiegelungsprinzip* bezeichnet. Zu diesem Zweck identifizieren wir jedes $(\omega_1, ..., \omega_{a+b}) \in \Omega$ mit dem stückweise linearen Pfad von $(0, 0)$ nach $(a + b, b - a)$ durch die Punkte $(k, \sum_{j=1}^{k} \omega_j)$, $1 \leq k \leq a + b$ (☞ Bild 3.1). Auf der Basis dieser Identifikation entspricht E offenbar der Menge aller (derartigen) Pfade von (0,0) durch (1,1) nach $(a + b, b - a)$, die die x-Achse nicht berühren oder schneiden. Es folgt

$$\begin{aligned}
|E| &= \text{Anzahl der Pfade von } (1,1) \text{ nach } (a+b, b-a), \\
&\quad\ \text{die die } x\text{-Achse nicht berühren oder schneiden} \\
&= \text{Anzahl der Pfade von } (0,1) \text{ nach } (a+b-1, b-a), \\
&\quad\ \text{die die } x\text{-Achse nicht berühren oder schneiden} && \text{[Verschiebung]} \\
&= \text{Anzahl der Pfade von } (0,1) \text{ nach } (a+b-1, b-a) \\
&- \text{Anzahl der Pfade von } (0,1) \text{ nach } (a+b-1, b-a), \\
&\quad\ \text{die die } x\text{-Achse berühren oder schneiden} \\
&= \text{Anzahl der Pfade von } (0,0) \text{ nach } (a+b-1, b-a-1) && \text{[Verschiebung]} \\
&- \text{Anzahl der Pfade von } (0,1) \text{ nach } (a+b-1, b-a), \\
&\quad\ \text{die die } x\text{-Achse berühren oder schneiden.}
\end{aligned}$$

Da wir die Pfade von (0,0) nach $(a+b-1, b-a-1)$ mit der Menge $\mathcal{P}_{a,b-1}$ identifizieren können, entspricht deren Anzahl $\binom{a+b-1}{a}$. Um die Anzahl der Pfade von (0,1) nach $(a+b-1, b-a-1)$ zu bestimmen, die die x-Achse berühren oder schneiden, spiegeln wir dessen Anfangsstück bis zum ersten Berührungs- oder Schnittpunkt mit der x-Achse an dieser Achse (Spiegelungsprinzip, ☞ Bild 3.2) und erhalten:

$$\begin{aligned} & \text{Anzahl der Pfade von } (0,1) \text{ nach } (a+b-1, b-a), \\ & \quad \text{die die } x\text{-Achse berühren oder schneiden} \\ = \; & \text{Anzahl der Pfade von } (0,-1) \text{ nach } (a+b-1, b-a), \\ & \quad \text{die die } x\text{-Achse berühren oder schneiden} \end{aligned}$$

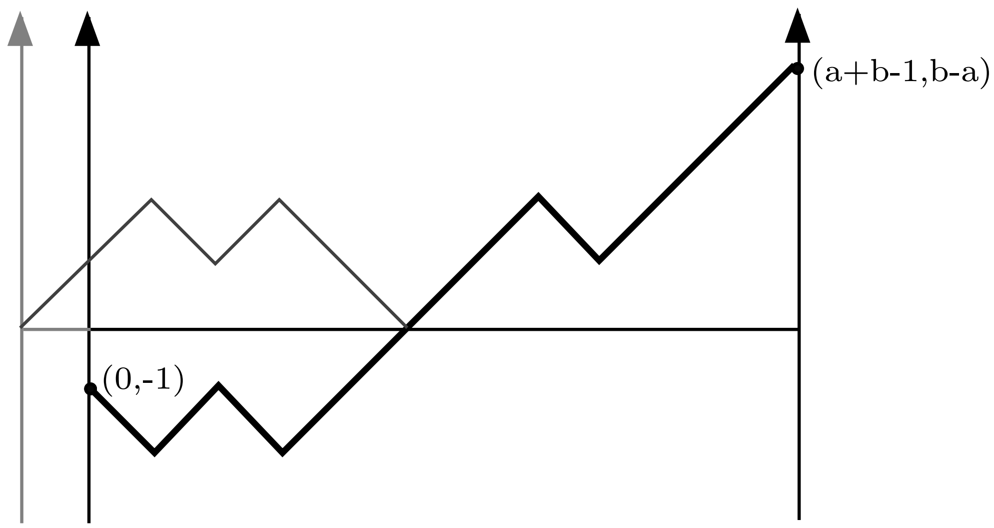

Bild 3.2. Gespiegelter Pfad.

Da aber die Pfade von $(0,-1)$ nach $(a+b-1, b-a)$ die x-Achse sowieso schneiden, gilt

$$\begin{aligned} & \text{Anzahl der Pfade von } (0,-1) \text{ nach } (a+b-1, b-a), \\ & \quad \text{die die } x\text{-Achse berühren oder schneiden} \\ = \; & \text{Anzahl der Pfade von } (0,-1) \text{ nach } (a+b-1, b-a), \\ = \; & \text{Anzahl der Pfade von } (0,0) \text{ nach } (a+b-1, b-a+1) \qquad \text{[Verschiebung]} \\ = \; & |\mathcal{P}_{a-1,b}| \; = \; \binom{a+b-1}{a-1}. \end{aligned}$$

Insgesamt erhalten wir

$$P(E) \;=\; \frac{|E|}{|\Omega|} \;=\; \frac{\binom{a+b-1}{a} - \binom{a+b-1}{a-1}}{\binom{a+b}{a}} \;=\; \frac{b-a}{a+b}.$$

Literatur

[1] FELLER, W. *Introduction to Probability Theory and its Applications. Volume I (3. Auflage).* Wiley, New York (1968).

[2] HENZE, N. *Stochastik für Einsteiger.* Vieweg, Braunschweig, Wiesbaden (1997).

[3] KRENGEL, U. *Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Vieweg, Braunschweig, Wiesbaden (1991).

SCHMITZ, N. *Stochastik für Lehramtsstudenten.* LIT-Verlag, Münster (1997).