1 Wahrscheinlichkeitslehre

1.1 Grundlagen der Wahrscheinlichkeitsrechnung

Die Wahrscheinlichkeitslehre ist ein elementarer Bestandteil der Statistik. Die mathematische Wahrscheinlichkeitslehre umfasst ein kompliziertes System unterschiedlicher Regeln und Gesetzmäßigkeiten, die hier nur insoweit dargestellt werden, als es für das Verständnis der verteilungsfreien Methoden erforderlich ist. Wir behandeln zunächst die wichtigsten Grundlagen der Wahrscheinlichkeitsrechnung und gehen anschließend auf die Darstellung einiger ausgewählter Wahrscheinlichkeitsverteilungen über.

1.1.1 Vorbemerkungen

Wir alle kennen das auf die beschreibende Statistik gemünzte Wort: „Mit Statistik kann man alles beweisen!“ Richtiger müsste es – wenn man schon mit Aphorismen operiert – auf die Inferenzstatistik bezogen heißen: *Mit Statistik kann man gar nichts beweisen*, keinen Unterschied, keinen Zusammenhang, keine Gesetzmäßigkeit, sofern man von einem Beweis fordert, dass er logisch und sachlich unwidersprochen bleiben soll.

Was kann die moderne Statistik als wissenschaftliche Methode wirklich leisten? Sie gibt Auskunft darüber, mit welcher Wahrscheinlichkeit Unterschiede, Zusammenhänge und Gesetzmäßigkeiten, die wir in Stichprobenerhebungen gefunden haben, rein zufällig entstanden sein können und inwieweit sie als allgemein, auch für ein größeres Kollektiv gültig, anzusehen sind. Absolut sichere Aussagen und Voraussagen sind mit Hilfe der Statistik unmöglich. Jedoch liegt es an uns, das Risiko, das wir für unsere Aussage zulassen wollen, gemäß der wissenschaftlichen Fragestellung höher oder niedriger anzusetzen.

Naturwissenschaftliche Aussagen und Voraussagen gründen sich auf Messungen: Die (klassische) Physik kennt kein Problem der Messung, höchstens eines der Messgenauigkeit, sie hat ihr Zentimeter-Gramm-Sekunden-System und kann damit die anstehenden Probleme adäquat lösen. Die biologischen wie auch die Sozialwissenschaften haben es nicht so leicht. In ihren empirisch ausgerichteten Bereichen sind sie unentwegt auf der Suche nach optimalen Dimensionen einer gültigen Messung, sind auf der Suche nach immer raffinierteren Methoden der Versuchsplanung zur Kontrolle des meist bedeutsamen Fehlers der individuellen Messung eines Merkmals. Ein ganzer Wissenschaftszweig, die Biometrie, beschäftigt sich mit den Voraussetzungen für objektive, zuverlässige und gültige Ausgangsdaten. Auf diesen Voraussetzungen erst baut die Statistik auf.

Die statistische Methode und ihre praktische Anwendung setzen eine eigene, dem Anfänger ungewohnte Art des induktiven Denkens voraus. Im logischen Denkakt folgt jeder Schluss stets zwingend und für den Einzelfall gültig aus seinen Prämissen; der statistische Denkakt dagegen führt zu Schlüssen, die nur für ein (theoretisch unendlich großes) Kollektiv von Ereignissen gelten und *für den Einzelfall nicht zwingend zutreffen.* Er orientiert sich an einem Beweissystem, das in der mathematischen Theorie der Statistik formal exakt begründet ist und das erst in dem Grad, in dem man seiner inne wird, ein Evidenz- und Stringenzerlebnis von ähnlicher Art vermittelt wie das Begriffssystem der Elementarlogik.

Grundlage allen statistischen Denkens ist der Wahrscheinlichkeitsbegriff. Beginnen wir deshalb mit einer kurzen Einführung zu diesem Begriff.

1.1.2 Begriff der Wahrscheinlichkeit

Die Wahrscheinlichkeit kann in verschiedener Weise eingeführt werden. Eine inzwischen klassische Einführung in Form von anschaulichen Vorlesungen mit engem Bezug zur Anwendung findet sich bei Mises (1931).

Für unsere Zwecke soll genügen: *Wenn unter* n *möglichen, einander ausschließenden Ereignissen, von denen eines mit Sicherheit eintritt,* g *von bestimmter Art sind, dann ist die Wahrscheinlichkeit, dass eines dieser* g *Ereignisse eintritt, gleich dem Bruch* g/n (g günstige unter n möglichen Ereignissen). Diese Wahrscheinlichkeit wird mit p bezeichnet. Dazu einige Beispiele: (A) Die Wahrscheinlichkeit, mit einem Würfel irgendeine Zahl von 1 bis 6 zu werfen, beträgt ohne Zweifel $p=1$. (B) Die Wahrscheinlichkeit, aus einer Urne mit Losen von 1 bis 10 das Los 7 oder ein Los mit kleinerer Nummer herauszuziehen, beträgt entsprechend der obigen Definition $p=0{,}7$. (C) Die Wahrscheinlichkeit, aus einem verdeckten Bridgespiel gerade das Herzass zu ziehen, beträgt analog $p=1/52\approx 0{,}02$. (D) Die Wahrscheinlichkeit, aus demselben Kartenspiel mehr als 4 Asse zu ziehen, ist naturgemäß $p=0$. (E) Die Wahrscheinlichkeit, mit einer Münze „Zahl" zu werfen, beträgt $p=0{,}5$. (F) Die Wahrscheinlichkeit, mit einem Würfel eine Sechs zu erzielen, ergibt $p=1/6=0{,}167$.

Jede Wahrscheinlichkeit hat einen Wert, der nicht negativ und nicht größer als 1 ist. Die Gesamtheit aller Wahrscheinlichkeitswerte konstituiert die im folgenden dargestellte Wahrscheinlichkeitsskala, die sich von 0 bis 1 erstreckt; sie enthält die oben herausgehobenen Wahrscheinlichkeitswerte an den entsprechenden Stellen markiert.

Die im Beispiel genannten Ereignisse A, B, C, D, E und F besitzen eine ihrer Skalenposition entsprechende Wahrscheinlichkeit.

Wir haben den Begriff der Wahrscheinlichkeit noch etwas näher zu erläutern. Halten wir uns dabei an das Würfelbeispiel (F): Von den 6 möglichen Augenzahlen

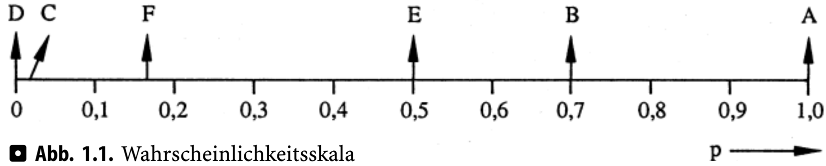

Abb. 1.1. Wahrscheinlichkeitsskala

tritt eine mit Sicherheit ein. Diese 6 Ereignisse sind auch gleichwertig, denn jedes Ereignis hat die gleiche Chance aufzutreten, wenn der Würfel nicht gefälscht ist.

Nehmen wir an, als günstiges Ereignis werde ein Wurf mit gerader Augenzahl angesehen. Drei Würfelflächen enthalten gerade Augenzahlen, daher beträgt die Wahrscheinlichkeit des Auftretens einer geraden Augenzahl 3/6 = 0,5.

Die beiden Begriffe *„gleichwertig“* und *„einander ausschließend“* sollen noch an 2 Beispielen illustriert werden.

Beispiel 1: Jemand möchte die Wahrscheinlichkeit, aus einem Skatspiel entweder ein As oder eine Herzkarte zu ziehen, ermitteln. Das Kartenspiel enthält 32 Karten, darin befinden sich 4 Asse und 8 Herzkarten. Folglich stehen – so möchte man meinen – die günstigen Ereignisse im Verhältnis zu den möglichen Ereignissen wie 12 : 32, also ist p = 0,375. Diese Schlussfolgerung ist aber unrichtig, denn ein Ass (das Herzass) gilt zugleich auch als Herzkarte. Das Auftreten eines Asses schließt also das Auftreten einer Herzkarte nicht aus. Die Bedingung, dass die Ereignisse einander ausschließen sollen, ist nicht erfüllt. Daher sind wir zu einem unrichtigen Wahrscheinlichkeitswert gekommen. Der richtige beträgt p = 11/32 = 0,344.

Beispiel 2: Angenommen, jemand möchte die Wahrscheinlichkeit, bei 2 hintereinander durchgeführten Würfen mit einer Münze 2-mal Zahl zu erhalten, ermitteln. Die 3 möglichen Ergebnisse, 2-mal Zahl, 2-mal Adler sowie 1-mal Zahl und 1-mal Adler, schließen sich gegenseitig aus. Man könnte schlussfolgern, die Wahrscheinlichkeit, 2-mal Zahl zu werfen, betrage 1/3. Diese Überlegung ist falsch, denn die 3 Ereignisse sind nicht gleichwertig. Das 3. Ereignis (Zahl–Adler) kann nämlich in 2facher Weise zustandekommen: das erste Mal Zahl und das zweite Mal Adler oder umgekehrt das erste Mal Adler und das zweite Mal Zahl. Richtig wäre folgende Überlegung gewesen: Es resultieren 4 gleichwertige Ereignisse: Zahl–Zahl, Adler–Adler, Zahl–Adler und Adler–Zahl. Daraus ersehen wir, dass die Wahrscheinlichkeit, 2-mal Zahl zu werfen, nicht p = 1/3, sondern p = 1/4 ausmacht. Dadurch, dass wir die Aufeinanderfolge von Zahl und Adler außer Acht gelassen haben, sind die Ereignisse nicht mehr gleich wahrscheinlich.

1.1.3 Theoretische und empirische Wahrscheinlichkeit

Wenn wir eine Münze werfen, so erwarten wir das Resultat „Zahl“ mit einer Wahrscheinlichkeit von p = 1/2. Wir folgern nämlich: Es gibt nur 2 mögliche Resultate, von denen eines im gegebenen Fall mit Sicherheit eintreten muss, so dass – wenn die Münze nicht verfälscht ist – jedes der beiden Resultate die gleiche Wahrscheinlichkeit für sich hat. Da wir dieses Resultat allein auf logischem Weg erzielt haben, sprechen wir von einer *theoretischen*, einer erwarteten oder einer **A-priori-Wahrscheinlichkeit.**

Werfen wir dagegen eine Münze, deren eine Kante stark abgenutzt wurde, so dürfen wir nicht mehr erwarten, dass bei einem beliebigen Wurf das Symbol „Zahl“ mit der Wahrscheinlichkeit p = 1/2 nach oben zu liegen kommen wird. Auf die Größe der Wahrscheinlichkeit, in diesem Fall Zahl zu werfen, kann uns nur ein Experiment einen Hinweis geben: Wir werfen die Münze einige hundert Male und zählen aus, wie oft wir das Resultat Zahl erhalten. Bilden wir Quotienten aus der Anzahl der „Zahlen" und der Anzahl der Würfe, so erhalten wir eine relative Häufigkeit, die wir als *empirische* oder als **A-posteriori-Wahrscheinlichkeit** bezeichnen. Mit einer zunehmenden Anzahl von Versuchen konvergiert die relative Häufigkeit auf einen konstanten Wert p (A): Bezeichnen wir die Häufigkeit eines Ereignisses A mit f (A) und die Anzahl aller Ereignisse einer Versuchsreihe mit n, so ergibt sich als Formel für die A-posteriori-Wahrscheinlichkeit:

$$p(A) = \lim_{n \to \infty} \frac{f(A)}{n}. \tag{1.1}$$

Im Folgenden wenden wir uns den wichtigsten Gesetzen der Wahrscheinlichkeitsrechnung zu, dem Additions- und Multiplikationssatz für Wahrscheinlichkeiten.

1.1.4 Additions- und Multiplikationssatz

Beispiel 1: Beim Würfelspiel können wir uns fragen, wie groß die Wahrscheinlichkeit ist, eine Sechs *oder* eine Fünf zu werfen. Da wir es hier mit 2 günstigen unter 6 möglichen Fällen zu tun haben, ist $p = 2/6 \cong 0{,}33$. Die Wahrscheinlichkeit, eine Sechs, eine Fünf oder eine Zwei zu werfen, ist entsprechend durch $1/6 + 1/6 + 1/6 = 0{,}5$ gegeben. Sie ist also die Summe der Wahrscheinlichkeiten, eine Sechs, eine Fünf oder eine Zwei zu werfen.

Die Verallgemeinerung dieser Überlegung führt zum **Additionssatz** der Wahrscheinlichkeit. Er lautet: Die Wahrscheinlichkeit p, dass von n einander ausschließenden Ereignissen das erste oder das zweite oder das dritte oder das n-te eintritt, ist gleich der Summe der Wahrscheinlichkeiten für das Auftreten der Einzelereignisse. Bezeichnen wir allgemein mit p_i die Wahrscheinlichkeit des i-ten Ereignisses, so beträgt die zusammengesetzte Wahrscheinlichkeit

$$p = p_1 + p_2 + p_3 + \ldots + p_i + \ldots + p_n = \sum_{i=1}^{n} p_i. \tag{1.2}$$

Beispiel 2: Wenn wir einen Würfel 2-mal hintereinander werfen, so können wir uns fragen: Wie groß ist die Wahrscheinlichkeit p, dass wir 2-mal eine Sechs werfen? Dieselbe Frage wäre auch für den gleichzeitigen Wurf zweier Würfel zu stellen. Die theoretische Wahrscheinlichkeit könnten wir genauso wie im Beispiel 1 bestimmen; sie leitet sich aus folgender Überlegung her: Die Wahrscheinlichkeit, dass der 1. Wurf eine Sechs ist, beträgt $p_1 = 1/6$. Angenommen, wir hätten geworfen und wirklich eine Sechs erhalten. In diesem Fall besteht wiederum eine Wahrscheinlichkeit von $p_2 = 1/6$, dass auch der 2. Wurf eine Sechs ergibt. Dieselbe Wahrscheinlichkeit $p_2 = 1/6$ hätte auch in jenen 5 Fällen Geltung, in denen wir beim 1. Wurf keine Sechs erhalten hätten. Die beiden Würfe sind nämlich *voneinander unabhängig*.

Die Wahrscheinlichkeit p, 2-mal nacheinander eine Sechs zu werfen, beträgt demgemäß nur 1/6 der Wahrscheinlichkeit, überhaupt eine Sechs zu werfen. Folglich ist $p = p_1 \cdot p_2 = 1/6 \cdot 1/6 = 1/36$. Entsprechend ist die Wahrscheinlichkeit, mit einer Münze 2-mal „Zahl" zu werfen: $p = p_1 \cdot p_2 = 1/2 \cdot 1/2 = 1/4$. Wir können diesen als **Multiplikationssatz** der Wahrscheinlichkeit bekannten Tatbestand allgemein so formulieren: Die Wahrscheinlichkeit p, dass n voneinander unabhängige Ereignisse gemeinsam auftreten, ist gleich dem Produkt der Einzelwahrscheinlichkeiten p_i dieser Ereignisse.

$$p = p_1 \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_i \cdot \ldots \cdot p_n = \prod_{i=1}^{n} p_i \, . \tag{1.3}$$

Additions- und Multiplikationssatz sind wichtige Ausgangspunkte der folgenden Ausführungen über die Kombinatorik und der späteren über die statistische Entscheidung.

1.1.5 Punktwahrscheinlichkeiten

Wenden wir uns von den Würfelversuchen, die 6 mögliche Resultate ergeben, wieder dem einfacheren Münzenversuch mit 2 Alternativen zu. Fragen wir uns, welche Kombinationen von „Zahl" (Z) und „Adler" (A) wir bei gleichzeitigem Wurf mit 3 Münzen theoretisch erhalten könnten. Im Folgenden sind die Möglichkeiten vollzählig zusammengestellt: ZZZ, ZZA, ZAZ, ZAA, AAA, AAZ, AZA, AZZ.

Unter den $2^3 = 8$ möglichen Resultaten finden wir nur eins, bei dem alle Münzen auf „Zahl" fallen. Die Wahrscheinlichkeit, 3-mal „Zahl" zu erhalten, ist demnach $p = 1/8$. Die Wahrscheinlichkeit, dass wir bei einem Wurf die Kombination 2-mal „Zahl", 1-mal „Adler" antreffen werden, beträgt 3/8 wie auch für die Kombination 1-mal „Zahl" und 2-mal „Adler". Die Wahrscheinlichkeit, 3-mal „Adler" zu werfen, ergibt sich wiederum zu 1/8.

Die Wahrscheinlichkeit, *ein bestimmtes Ereignis* (z. B. $2 \times Z$, $1 \times A$) zu erzielen, nennt man *Punktwahrscheinlichkeit*. Man erhält die Punktwahrscheinlichkeit – mit (klein) p bezeichnet –, indem man die Häufigkeit der 4 im vorigen Absatz als Beispiel genannten Kombinationen von $n = 3$ Elementen durch 8 als Anzahl aller möglichen Kombinationen dividiert. Diese p-Werte erhalten wir auch, wenn wir die Zahlen der 4. Zeile aus ◘ Tab. 1.1, dem sog. Pascalschen Dreieck, durch $2^3 = 8$ dividieren.

Das Pascalsche Dreieck in ◘ Tab. 1.1 wurde für $n = 1$ bis $n = 5$ in Einserschritten oder kurz für $n = 1(1)5$ entwickelt. (Die in Klammern gesetzte Zeile $n = 0$ wurde

◘ **Tabelle 1.1**

Überwiegen von „Zahl"	Überwiegen von „Adler"	n	2^n
(1)		(0)	(1)
1 1 Z A		1	2
1 2 1 ZZ ZA AA		2	4
1 3 3 1 ZZZ ZZA ZAA AAA		3	8
1 4 6 4 1 ZZZZ ZZZA ZZAA ZAAA AAAA		4	16
1 5 10 10 5 1 ZZZZZ ZZZZA ZZZAA ZZAAA ZAAAA AAAAA		5	32

der Vollständigkeit halber mit aufgenommen.) In allgemeiner Schreibweise kennzeichnet a(i)n, dass von a bis n in Intervallen der Größe i gezählt wird.

Wie man leicht erkennt, ergeben sich die Werte einer Zeile als Summen von jeweils 2 benachbarten Werten der vorangehenden Zeile, ergänzt durch die Zahl 1 am Anfang und am Ende der Zeile. Diesem Prinzip folgend, lässt sich das Pascalsche Dreieck in ◘ Tab. 1.1 beliebig ergänzen.

Aus dieser Tabelle entnehmen wir weiter, dass bei einem Wurf mit $n = 4$ Münzen $p(4 \times Z) = 1/16$, $p(3 \times Z, 1 \times A) = 4/16$, $p(2 \times Z, 2 \times A) = 6/16$, $p(1 \times Z, 3 \times A) = 4/16$ und $p(4 \times A) = 1/16$ resultieren.

Entsprechend sind die Punktwahrscheinlichkeiten für bestimmte Adler-Zahl-Kombinationen für mehr als 4 Münzen zu berechnen (zur Herleitung des Pascalschen Dreiecks ► S. 14).

Die Berechnung von Punktwahrscheinlichkeiten ist essenziell für viele verteilungsfreie Verfahren. Allerdings werden wir dazu wie auch für die im folgenden zu besprechenden Überschreitungswahrscheinlichkeiten in der Regel kompliziertere Wahrscheinlichkeitsmodelle benötigen als das beispielhaft verwendete Wahrscheinlichkeitsmodell des Münzwurfes (► Abschn. 1.1.7).

1.1.6 Überschreitungswahrscheinlichkeiten

Wir wollen im Folgenden noch eine andere Wahrscheinlichkeit kennenlernen, die sich am besten anhand eines Wettbeispiels einführen lässt: Angenommen, wir haben gewettet, mit $n = 4$ Münzen *mindestens* $x = 3$-mal „Zahl" zu werfen. Wie groß ist die Wahrscheinlichkeit, diese Wette zu gewinnen? Die Antwort ist einfach: „Mindestens 3-mal" bedeutet 3-mal oder 4-mal „Zahl" zu werfen; also ist die gesuchte Wahrscheinlichkeit – wir bezeichnen sie mit (groß) P und nennen sie **Überschreitungswahrscheinlichkeit** – nach dem Additionssatz gleich der Punktwahrscheinlichkeit, 3-mal „Zahl" zu werfen: $p(x = 3) = 4/16$ plus der Punktwahrscheinlichkeit, 4-mal „Zahl" zu werfen: $p(x = 4) = 1/16$; also ist $P = 4/16 + 1/16 = 5/16$. In gleicher Weise könnten wir nach der Wahrscheinlichkeit, mindestens 2-mal „Zahl" zu werfen, fragen: Sie beträgt analog $P = 6/16 + 4/16 + 1/16 = 11/16$. *Wir können die Überschreitungswahrscheinlichkeit definieren als die Wahrscheinlichkeit des Auftretens eines bestimmten Ereignisses, vermehrt um die Wahrscheinlichkeiten aller „extremeren" Ereignisse.*

Statt nach der Wahrscheinlichkeit für „mindestens 3-mal Zahl" hätten wir auch nach der Wahrscheinlichkeit für „*höchstens* 1-mal Adler" fragen können. Für beide Fälle ist die Überschreitungswahrscheinlichkeit natürlich identisch. Allgemein: Die Wahrscheinlichkeit, dass ein Ereignis A bei n Versuchen mindestens x-mal auftritt, entspricht der Wahrscheinlichkeit, dass das zu A komplementäre Ereignis $\bar{A}$ (lies: non-A) höchstens $(n - x)$-mal auftritt.

In dem obigen Beispiel haben wir sozusagen *einseitig* gewettet. Was unter einer einseitigen Wette zu verstehen ist, wollen wir gleich am entgegengesetzten Fall einer zweiseitigen Wette illustrieren: Wir wetten, bei 4 Würfen entweder 4-mal oder 0-mal „Zahl" zu werfen. Wie groß ist die Chance, diese Wette zu gewinnen? Die Punktwahrscheinlichkeit für $x = 4$ beträgt $p(x = 4) = 1/16$, und die Punktwahrscheinlichkeit für $x = 0$ ist $p(x = 0) = 1/16$, so dass die zweiseitige Überschreitungswahrscheinlichkeit, die wir durch P′ kennzeichnen, mit $P' = 2/16$ der doppelten einseitigen Überschreitungswahrscheinlichkeit entspricht. Hätten wir gewettet, mindestens 3-mal „Zahl" oder höchstens 1-mal „Zahl" zu werfen, so wäre dies ebenfalls eine zweiseitige Wette, deren Gewinnchancen nach dem Pascalschen Dreieck wie folgt zu berechnen wären: Mindestens 3-mal „Zahl" heißt 3- oder 4-mal „Zahl", deren Punktwahrscheinlichkeiten 4/16 und 1/16 betragen. Hinzu kommen die Wahrscheinlichkeiten für 1-mal Zahl ($p = 4/16$) und für 0-mal Zahl ($p = 1/16$). Die gesamte zweiseitige Überschreitungswahrscheinlichkeit ist also $P' = 1/16 + 4/16 + 4/16 + 1/16 = 10/16$.

Die Frage, ob es sich um eine einseitige oder zweiseitige Wette oder – in der Terminologie der Statistik – um einen *einseitigen oder zweiseitigen Test* handelt, ist für die Entscheidung bestimmter empirischer Fragestellungen von großer Bedeutung. Wir werden darauf an späterer Stelle (► Abschn. 2.2.1) noch zurückkommen. Festzuhalten ist, dass die Wahrscheinlichkeit für die zweiseitige Frage durch Verdopplung der Wahrscheinlichkeit für die einseitige Frage zu ermitteln ist, sofern die Wahrscheinlichkeitsverteilung für x symmetrisch ist (► Abschn. 1.2).

1.1.7 Elemente der Kombinatorik

Es wäre unökonomisch, wollten wir A-priori-Wahrscheinlichkeiten für das Auftreten bestimmter Ereignisse auf die beschriebene Art ermitteln; außerdem würden wir komplexere Probleme mit unseren bisherigen Mitteln gar nicht bewältigen. Zur Berechnung komplexer A-priori-Wahrscheinlichkeiten bedienen wir uns verschiedener Formeln eines Teilgebietes der Mathematik, der *Kombinatorik*. Diese Formeln gründen sich auf 2 Prinzipien, die wir sofort als Analoga des Additions- und Multiplikationssatzes der Wahrscheinlichkeitsrechnung erkennen werden:

Prinzip 1: Wenn ein Ereignis A auf m-fache und ein anderes Ereignis B auf n-fache Weise entstehen kann, so kann das Ereignis A *oder* B auf $(m + n)$-fache Weise entstehen, vorausgesetzt, dass A und B nicht gleichzeitig auftreten können.

Prinzip 2: Wenn ein Ereignis A auf m-fache und ein Ereignis B auf n-fache Weise entstehen kann, dann kann das Ereignis (A, B), d.h. dass zunächst A *und* dann B eintritt, auf $(m \cdot n)$-fache Weise entstehen, vorausgesetzt, dass alle Möglichkeiten auftreten können.

Was diese beiden Sätze beinhalten, wollen wir uns wieder an einem einfachen Beispiel überlegen: Das Ereignis A – eine Herzkarte aus einem Skatblatt zu entnehmen – kann auf 8 verschiedene Weisen verwirklicht werden; das Ereignis B – eine Kreuzkarte zu entnehmen – kann ebenfalls auf 8 verschiedene Weisen erfolgen. Es gibt also $8 + 8 = 16$ verschiedene Möglichkeiten, eine Herz- oder eine Kreuzkarte aus einem Skatblatt von 32 Karten zu ziehen, oder die Wahrscheinlichkeit, eine Herz- oder Kreuzkarte aus dem Skatblatt zu ziehen, beträgt $p = 16/32 = 0{,}5$. Dies war das 1. Prinzip.

Das 2. Prinzip können wir uns dadurch veranschaulichen, dass wir nacheinander 2 Karten aus dem Skatspiel entnehmen. Bleiben wir bei den Farben Herz und Kreuz. Eine Herzkarte konnten wir auf 8fache Weise entnehmen, ebenso eine

Kreuzkarte. Auf wievielfache Weise können wir nun eine Herzkarte und eine Kreuzkarte entnehmen? Wir können das Herzass mit einem Kreuzass, einen Kreuzkönig, einer Kreuzdame usw. paaren; es resultieren 8 Paarungsmöglichkeiten. Dieselbe Anzahl von Möglichkeiten ergibt sich für den Herzkönig, für die Herzdame, den Buben usw. bei der Paarung mit einer Kreuzkarte. Im Ganzen gibt es also $8 \cdot 8 = 64$ Möglichkeiten.

Die beiden Prinzipien können von 2 auf k Ereignisse verallgemeinert werden. Für drei einander ausschließende Ereignisse A, B, C, die auf m-, n-, o-fache Weise entstehen können, gilt: Das Ereignis A oder B oder C kann auf $(m+n+o)$-fache Weise zustande kommen; das Ereignis (A, B, C) kann in $(m \cdot n \cdot o)$-facher Weise zustande kommen.

Permutationen und Variationen

Überlegen wir uns einmal, auf wieviele Weisen wir die 3 Buchstaben des Wortes ROT anordnen können. Versuchen wir es erst durch Probieren:

ROT RTO OTR ORT TRO TOR

Es haben sich 6 Anordnungen ergeben. Wie ist die Entstehung dieser Anordnungen zu denken? Wir haben 3 Möglichkeiten, einen der 3 Buchstaben an die 1. Stelle zu setzen. Nach der Entscheidung für eine Möglichkeit, z. B. das R, haben wir nur mehr 2 Möglichkeiten, einen der verbleibenden Buchstaben an die 2. Stelle zu setzen; wir wählen z. B. das O. Für die Besetzung der 3. Stelle ergibt sich nur noch eine Möglichkeit, das restliche T. Die 1. Stelle kann also auf 3fache, die 2. auf 2fache und die 3. Stelle auf 1fache Weise besetzt werden. Betrachten wir die Besetzung der 3 Stellen, so ergibt sich unmittelbar, dass sie auf $3 \cdot 2 \cdot 1 = 6$fache Weise möglich ist.

Die 6 möglichen Anordnungen der 3 Buchstaben des Wortes ROT sind die **Permutationen** der Elemente R, T, O; die 4 Ziffern 3, 5, 6 und 9 ergeben nach derselben Regel $4 \cdot 3 \cdot 2 \cdot 1 = 24$ Permutationen, k Objekte liefern entsprechend $k\,(k-1)\,\ldots\,2 \cdot 1$ Permutationen. Schreiben wir das fortlaufende Produkt der natürlichen Zahlen von 1 bis k vereinfachend als k! (lies: k Fakultät), so ist die Zahl der Permutationen Pk von k Elementen durch die Gleichung

$$P_k = k! \tag{1.4}$$

gegeben.

Wie steht es nun mit der Permutationszahl von n Elementen, wenn jeweils nur ein Teil, also z. B. k Elemente in einer Anordnung benutzt werden? Wieviele Permutationen zu je $k = 4$ Buchstaben lassen sich beispielsweise aus dem Wort MORGEN mit $n = 6$ bilden?

Stellen wir analoge Überlegungen wie oben an. Die 1. Stelle der Anordnung kann auf n-fache Weise besetzt werden, die 2. kann in $(n-1)$-facher Weise besetzt werden, die 3. in $(n-2)$-facher Weise usw. bis zur k-ten Stelle. Ehe wir die k-te Stelle einsetzen, haben wir $(k-1)$ von den n Elementen in der Anordnung untergebracht, und es verbleiben noch $n-(k-1) = n-k+1$ Elemente zur Disposition für die Besetzung der letzten Stelle. Die Anzahl dieser möglichen Permutationen beträgt demnach $n(n-1)\,\ldots\,(n-k+1)$.

Dass auch hier wieder die Fakultätenschreibweise möglich ist, wird deutlich, wenn wir dieses Produkt erweitern, indem wir es mit der Anzahl aller Faktoren multiplizieren, die zwischen $(n-k+1)$ und 1 liegen, und es durch dieselben Faktoren dividieren:

$$\frac{n\cdot(n-1)\cdot(n-2)\ldots(n-k+1)\cdot(n-k)\cdot(n-k-1)\cdot(n-k-2)\cdot\ldots(2)\cdot(1)}{(n-k)\cdot(n-k-1)\cdot(n-k-2)\cdot\ldots(2)\cdot(1)}\ .$$

Wenden wir auf diesen Ausdruck die Fakultätenschreibweise an, so erhalten wir die Anzahl der Permutationen von n Elementen zu je k Elementen oder – wie man auch sagt – **Variationen** von n Elementen zur k-ten Klasse nach der Gleichung:

$$_{n}P_{k} = \frac{n!}{(n-k)!}\ . \qquad (1.5)$$

Aus dem Wort MORGEN lassen sich also $6!/(6-4)!=360$ Permutationen zu je 4 Buchstaben herstellen.

Kann ein Element mehrfach eingesetzt werden, so spricht man von **Variationen mit Wiederholungen.** Die Zahl der Variationen von n Elementen zur k-ten Klasse (in Kombination zu je k Elementen mit Wiederholungen) beträgt:

$$_{n}V_{k} = n^{k}\ . \qquad (1.6)$$

Demnach lassen sich z. B. aus dem Wort MOST $(n=4)$ $4^2=16$ verschiedene Variationen zu je $k=2$ Buchstaben bilden, wenn Buchstabenwiederholungen (MM, OO, SS und TT) zulässig sind.

Durch die Wiederholung von Elementen kann $k > n$ sein. Für $n=2$ Elemente und k Klassen ist

$$_{2}V_{k} = 2^{k}\ . \qquad (1.7)$$

Beim Werfen mit einer Münze z. B. haben wir $n=2$ Elemente (Zahl und Adler). Diese lassen sich auf $2^3=8$fache Weise in Dreiervariationen anordnen. Dies sind die Zahl-Adler-Abfolgen, die sich bei $k=3$ Münzwürfen ergeben können. Bei 5 Würfen wären also $2^5=32$ Abfolgen möglich. Um welche Abfolgen es sich hier jeweils handelt, lässt sich leicht dem Pascalschen Dreieck (▪ Tab. 1.1) entnehmen.

Kombinationen

Wenn wir aus n Elementen alle Gruppen von k Elementen bilden, erhalten wir alle Kombinationen von n Elementen zur k-ten Klasse. Zwei Kombinationen sind verschieden, wenn sie sich mindestens in einem Element unterscheiden. 123, 124, 234 etc. sind damit unterschiedliche Dreierkombinationen der Elemente 1234, aber nicht 123, 213, 312 etc. Dies wären Permutationen der Kombination 123.

Die in Gl. (1.5) errechnete Zahl der Permutationen von n Elementen zur k-ten Klasse umfasst sowohl alle Kombinationen als auch deren Permutationen. (Die Buchstabenabfolgen MORG, MOGR, MROG etc. wurden hier als verschiedene Permutationen gezählt.) Bei der Bestimmung der Anzahl der Kombinationen lassen wir die Permutationen von Buchstaben außer Acht, d. h. deren Reihenfolge ist beliebig. Wir wissen aus Gl. (1.4), dass jede Kombination zu k Elementen k!-fach permutiert werden kann. Die Anzahl der Kombinationen mal der Anzahl der Permutationen aus jeder Kombination muss also die Gesamtzahl der Permutationen von n Elementen zur k-ten Klasse gemäß Gl. (1.5) ergeben. Bezeichnen wir mit ${}_{n}C_{k}$ die Anzahl der Kombinationen von n Elementen zur k-ten Klasse, so können wir schreiben:

$$k! \cdot {}_{n}C_{k} = {}_{n}P_{k}\,. \tag{1.8}$$

Setzen wir den Wert für ${}_{n}P_{k}$ aus Gl. (1.5) ein und lösen die Gleichung nach ${}_{n}C_{k}$ auf, so erhalten wir den Ausdruck für die Berechnung der Kombinationszahl

$$ {}_{n}C_{k} = \frac{n!}{k!(n-k)!}\,. \tag{1.9}$$

Statt des verbleibenden Bruches auf der rechten Seite der Gleichung schreibt man meist das Symbol $\binom{n}{k}$, das von Euler eingeführt wurde und deshalb auch *Eulersches Symbol* genannt und als „n über k" gelesen wird:

$$ {}_{n}C_{k} = \binom{n}{k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}\,. \tag{1.10}$$

Aus dem Wort MORGEN lassen sich also 15 Kombinationen mit jeweils 4 verschiedenen Buchstaben bilden:

$$ {}_{6}C_{4} = \frac{6\cdot 5\cdot 4\cdot 3}{4\cdot 3\cdot 2\cdot 1} = 15\,.$$

Aus Gl. (1.9) ergibt sich, dass $\binom{n}{k} = \binom{n}{n-k}$. So würden wir unsere Aufgabe, ${}_{6}C_{4}$ zu berechnen, auch so bewältigen: ${}_{6}C_{4} = {}_{6}C_{6-4} = {}_{6}C_{2} = \binom{6}{2} = 6\cdot 5/2\cdot 1 = 15$.

Setzen wir in Gl. (1.9) $n=k$, so ist $\binom{n}{n} = 1$, andererseits muss dann aber $\binom{n}{n} = \binom{n}{n-n} = \binom{n}{0}$ ebenfalls 1 sein.

Ein weiteres Beispiel: Das Blatt eines Skatspielers repräsentiert eine Zehnerkombination aus den 32 Karten des Spiels. Danach kann ein Spieler im Verlauf seines Lebens höchstens

$$ {}_{32}C_{10} = \frac{32\cdot 31\cdot 30\cdot 29\cdot 28\cdot 27\cdot 26\cdot 25\cdot 24\cdot 23}{10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1} = 64\,512\,240$$

verschiedene Blätter erhalten, eine Möglichkeit, die er in der Tat wohl kaum ausschöpfen kann und bei der die Spielregel, über den sog. Skat nochmals Karten austauschen zu können, noch gar nicht berücksichtigt ist.

1.2 Wahrscheinlichkeitsverteilungen

1.2.1 Verteilungsformen von Zufallsvariablen

Das n-fache Werfen einer Münze stellt einen beliebig oft wiederholbaren Vorgang dar, der nach einer ganz bestimmten Vorschrift ausgeführt wird und dessen Ergebnis vom Zufall bestimmt ist. Einen Vorgang dieser Art bezeichnen wir als **Zufallsexperiment.** Die Zahl x zur Kennzeichnung des Ergebnisses eines Zufallsexperimentes (z. B. x=3-mal Adler) stellt dabei eine **Realisierung der Zufallsvariablen** X dar. Kann die Zufallsvariable nur bestimmte Zahlenwerte annehmen, wie 0, 1, 2, 3, 4 als Anzahl der „Adler" beim Wurf von 4 Münzen, dann handelt es sich um eine **diskrete** Zufallsvariable; kann sie (u. U. auch nur innerhalb gewisser Grenzen) alle möglichen Werte annehmen, wie der Fußpunkt eines einmal gerollten Zylinders alle Werte zwischen 0 und $2r\pi$, dem Umfang des Zylinders, dann spricht man von einer **stetigen** Zufallsvariablen. Zufallsvariablen werden im Allgemeinen mit lateinischen Großbuchstaben (X, Y, A, B) bezeichnet, wenn die Gesamtheit aller möglichen Werte gemeint ist, z. B. X = alle natürlichen Zahlen zwischen 0 und 4 oder Y = alle reellen Zahlen zwischen 0 und $2r\pi$; sie werden mit lateinischen Kleinbuchstaben (x, y, a, b) symbolisiert, wenn bestimmte, durch Zufallsexperimente gewonnene Werte (Realisationen) gemeint sind, z. B. x = (3, 0, 2) oder y = (6,2r; 1,76r; 0,39r; 3,14r) im Falle der obigen beiden Experimente.

Wahrscheinlichkeitsfunktion

Bei einer diskreten Zufallsvariablen ordnet die Wahrscheinlichkeitsfunktion f(X) jeder Realisation x_i eine Wahrscheinlichkeit p_i zu:

$$f(X) = \begin{cases} p_i & \text{für } X = x_i \\ 0 & \text{für alle übrigen } x\ . \end{cases}$$

Für x = 3-mal Adler beim Werfen von n = 4 Münzen beträgt die Wahrscheinlichkeit nach ► Abschn. 1.1.5 $f(x=3)=4/16$. Durch die Wahrscheinlichkeitsfunktion ist die Wahrscheinlichkeitsverteilung oder kurz die *Verteilung einer Zufallsvariablen* vollständig bestimmt. Die Summe der Wahrscheinlichkeiten aller möglichen Realisationen einer diskreten Zufallsvariablen ist 1: $\sum f(x_i)=1$.

Wird eine Wahrscheinlichkeitsverteilung von einer stetigen Variablen X gebildet, dann resultiert analog eine stetige Wahrscheinlichkeitsverteilung, die nicht durch Einzelwahrscheinlichkeiten, sondern durch eine sog. **Dichtefunktion** f(X) mathematisch beschrieben wird, deren Integral – wie oben die Summe – gleich 1 ist: $\int f(X)dX=1$. Hier kann die Wahrscheinlichkeit, dass ein mögliches Ergebnis realisiert wird, nur auf ein bestimmtes Intervall J der Dichtefunktion bezogen werden: Man kann also – um dies am Zufallsexperiment des Zylinderrollens zu veranschaulichen – fragen, wie groß die Wahrscheinlichkeit ist, dass der Zylinder in einem Intervall zwischen den Marken 3,14r und 6,28r des Zylinderumfanges aufliegen werde. Diese Wahrscheinlichkeit ist im vorliegenden Fall (einer stetigen Gleichverteilung) mit p = 0,5 ebenso groß wie die Wahrscheinlichkeit, dass der Fußpunkt des Zylinders nach dem Rollen zwischen 0,00r und 3,14r liegen werde.

Verteilungsfunktion

Wahrscheinlichkeitsverteilungen lassen sich auch so darstellen, dass sie angeben, wie groß die Wahrscheinlichkeit P ist, dass in einem Zufallsexperiment die Variable einen Wert kleiner oder gleich x annimmt. Aus derartigen Verteilungen lassen sich damit einfach die in ► Abschn. 1.1.6 behandelten Überschreitungswahrscheinlichkeiten P ablesen. Diese Darstellungsform der Wahrscheinlichkeiten einer Zufallsvariablen bezeichnet man als *Verteilungsfunktion* F(X). Bei diskreten Zufallsvariablen erhält man sie – wie das folgende Beispiel zeigt – durch fortlaufende

Summation (Kumulation) der Punktwahrscheinlichkeiten der Wahrscheinlichkeitsfunktion. Für das Werfen von $n=4$ Münzen erhält man:

Anzahl der „Adler" $= X_i$	0	1	2	3	4
Wahrscheinlichkeitsfunktion $f(x_i)$	1/16	4/16	6/16	4/16	1/16
Verteilungsfunktion $F(x_i)$	1/16	5/16	11/16	15/16	16/16.

Formalisiert man das Vorgehen der fortlaufenden Summierung bis jeweils zum Variablenwert x_k, so ergibt sich für diskrete Wahrscheinlichkeitsverteilungen

$$F(x_k) = \sum_{i=0}^{k} f(x_i)\,. \tag{1.11 a}$$

Die Verteilungsfunktion stetiger Zufallsvariablen F(X) erhält man in entsprechender Weise, wenn man statt von 0 bis x_k zu summieren von $-\infty$ bis x_k integriert:

$$F(x_k) = \int_{-\infty}^{x_k} f(X)dX\,. \tag{1.11 b}$$

Die stetige Wahrscheinlichkeitsverteilung unseres Zylinderbeispiels beginnt zwar bei $x=0$ (und nicht bei $x=-\infty$), doch können wir in gleicher Weise argumentieren: Die Wahrscheinlichkeit, einen Variablenwert von 3,14r oder einen niedrigeren Wert zu „errollen" ($x \leq 3{,}14$ r), beträgt $P=F(3{,}14\text{ r})=0{,}5$, die Wahrscheinlichkeit eines Wertes $x \leq 4{,}71$ r ist 0,75 und die Wahrscheinlichkeit eines Wertes $x \leq 6{,}28$ r ist 1,00.

Erwartungswerte

Oft stellt sich die Frage, wieviele Realisationen einer bestimmten Art man bei einem Zufallsexperiment zu erwarten hat, beim Münzenwurf etwa, wie oft man bei N Würfen mit $n=4$ Münzen $x=i$ „Adler" zu erwarten hat. Kennt man die Wahrscheinlichkeitsfunktion der Zufallsvariablen, dann bildet man einfach

$$E(N,\ x=i) = Nf(x=i) = N \cdot p_i\ .$$

Mittels dieser Gleichung wären die theoretisch zu erwartenden Häufigkeiten $E(x_i)$ der Ergebnisse von N Zufallsexperimenten vorauszusagen: Werfen wir $n=4$ Münzen $N=128$-mal, so erwarten wir

$E(x=0) = 128 \cdot (1/16) = 8$-mal „0 Adler",

$E(x=1) = 128 \cdot (4/16) = 32$-mal „1 Adler",

$E(x=2) = 128 \cdot (6/16) = 48$-mal „2 Adler",

$E(x=3) = 128 \cdot (4/16) = 32$-mal „3 Adler"

und

$E(x=4) = 128 \cdot (1/16) = 8$-mal „4 Adler".

Mit dieser theoretisch zu erwartenden Häufigkeitsverteilung könnten wir die Ergebnisse eines tatsächlich durchgeführten Experimentes – 128-mal 4 Münzen werfen – vergleichen und feststellen, wie gut Beobachtung und Erwartung übereinstimmen, wie gut sich die beobachtete der erwarteten Häufigkeitsverteilung anpasst (▶ Abschn. 5.1.3).

Ebenso oft stellt sich die Frage, welchen durchschnittlichen Wert die Zufallsvariable X bei vielen Versuchen annimmt. Dieser Wert wird als *Erwartungswert einer Zufallsvariablen* X bezeichnet. Für diskrete Zufallsvariablen errechnet man den Erwartungswert E(X) nach folgender Gleichung:

$$E(X) = \sum_i f(x_i) \cdot x_i \, . \tag{1.12 a}$$

Der Erwartungswert E(X) der Zufallsvariablen „Anzahl der Adler" bei einem Wurf mit $n=4$ Münzen lautet damit

$$\begin{aligned} E(X) &= 0 \cdot 1/16 + 1 \cdot 4/16 + 2 \cdot 6/16 + 3 \cdot 4/16 + 4 \cdot 1/16 \\ &= 2 \, . \end{aligned}$$

Bei stetigen Zufallsvariablen errechnet man den Erwartungswert nach folgender Beziehung:

$$E(X) = \int_{-\infty}^{+\infty} X \cdot f(X) d(X) \, . \tag{1.12 b}$$

Für den Erwartungswert einer Zufallsvariablen verwendet man auch das Symbol μ. μ bzw. E(X) kennzeichnen damit den *Mittelwert* bzw. die *„zentrale Tendenz"* einer Verteilung. Ein weiteres wichtiges Maß zur Charakterisierung der Verteilung einer Zufallsvariablen ist die *Varianz* σ^2. Mit ihr wird die Unterschiedlichkeit, die die Werte einer Zufallsvariablen X aufweisen, beschrieben:

$$\sigma^2 = \sum_i (x_i - \mu)^2 \cdot f(x_i) \, . \tag{1.13 a}$$

Betrachten wir den Ausdruck $x_i - \mu$ als eine neue Zufallsvariable, erkennt man unter Bezug auf Gl. (1.12), dass die Varianz mit dem Erwartungswert der quadrierten Abweichung $(X-\mu)^2$ identisch ist:

$$\sigma^2 = E(X-\mu)^2 \, . \tag{1.14}$$

Im oben genannten Münzwurfbeispiel errechnen wir eine Varianz von

$$\begin{aligned} \sigma^2 &= (0-2)^2 \cdot 1/16 + (1-2)^2 \cdot 4/16 + (2-2)^2 \cdot 6/16 \\ &\quad + (3-2)^2 \cdot 4/16 + (4-2)^2 \cdot 1/16 \\ &= 1 \, . \end{aligned}$$

Ist die Zufallsvariable stetig, errechnet man die Varianz nach folgender Beziehung:

$$\sigma^2 = \int_{-\infty}^{+\infty} (X-\mu)^2 f(X) dX \, . \tag{1.13 b}$$

1.2.2 Die Binomialverteilung

Mit dem Münzbeispiel haben wir eine Wahrscheinlichkeitsverteilung verwendet, die für gleich mögliche Ereignisse („Z" und „A") gilt. Diese Verteilung heißt Binomialverteilung für gleich wahrscheinliche Alternativereignisse. Die Wahrscheinlichkeitsfunktion für die Zufallsvariable X (z. B. Häufigkeit für das Ereignis „Zahl") lautet:

$$p(X) = \binom{n}{X} \cdot \left(\tfrac{1}{2}\right)^n . \tag{1.15}$$

Diese Verteilung wurde bereits im Pascalschen Dreieck (■ Tab. 1.1) tabelliert. Die Zahlenwerte im Dreieck entsprechen dem 1. Faktor $\binom{n}{x}$. In der rechten Randspalte finden wir den Kehrwert 2^n des 2. Faktors der Gl. (1.15). Wir können danach die Punktwahrscheinlichkeiten, x-mal „Zahl" zu werfen, für Würfe mit beliebig vielen (n) Münzen berechnen.

Es ist nun der allgemeine Fall zu betrachten, dass die beiden Ereignisse nicht gleich wahrscheinlich sind.

Herleitung der Wahrscheinlichkeitsfunktion

Ein Ereignis E habe die Realisationswahrscheinlichkeit $\pi(E) \neq 1/2$ und das alternative Ereignis $\bar{E}$ (lies: Non-E) die komplementäre Wahrscheinlichkeit $(\bar{E}) = 1 - (E)$. Nach dem Multiplikationssatz gelten dann für die Sukzession des Auftretens von E oder $\bar{E}$ in $n = 2$ Versuchen, wobei zur Veranschaulichung E das Würfeln einer „Sechs" und $\bar{E}$ das Würfeln einer anderen Augenzahl bedeuten möge, folgende Wahrscheinlichkeiten [für $\pi(E)$ schreiben wir vereinfachend π]:

$$\begin{aligned}
p(EE) &= \pi \cdot \pi \\
p(E\bar{E}) &= \pi \cdot (1 - \pi) \\
p(\bar{E}E) &= (1 - \pi) \cdot \pi \\
p(\overline{EE}) &= (1 - \pi) \cdot (1 - \pi) \ .
\end{aligned}$$

Lässt man die Reihenfolge der Ereignisse unberücksichtigt, ergeben sich die folgenden Wahrscheinlichkeiten:

$$\begin{aligned}
&p(EE) &&= \pi \cdot \pi &&= (1) \cdot \pi^2 &&= \tbinom{2}{2} \cdot \pi^2 \cdot (1 - \pi)^0 \\
&p(E\bar{E} \text{ oder } \bar{E}E) &&= \pi \cdot (1-\pi) + (1-\pi) \cdot \pi &&= (2) \cdot \pi \cdot (1-\pi) &&= \tbinom{2}{1} \cdot \pi^1 \cdot (1-\pi)^1 \\
&p(\overline{EE}) &&= (1-\pi) \cdot (1-\pi) &&= (1) \cdot (1-\pi)^2 &&= \tbinom{2}{0} \cdot \pi^0 \cdot (1-\pi)^2 \ .
\end{aligned}$$

In $n = 3$ Versuchen wären die entsprechenden Wahrscheinlichkeiten

$$\begin{aligned}
&p(EEE) &&= \pi \cdot \pi \cdot \pi &&= (1) \cdot \pi^3 &&= \tbinom{3}{3} \cdot \pi^3 \cdot (1-\pi)^0 \\
&p(EE\bar{E} \text{ oder } E\bar{E}E \text{ oder } \bar{E}EE) && &&= (3) \cdot \pi^2 \cdot (1-\pi) &&= \tbinom{3}{2} \cdot \pi^2 \cdot (1-\pi)^1 \\
&p(E\overline{EE} \text{ oder } \bar{E}E\bar{E} \text{ oder } \overline{EE}E) && &&= (3) \cdot (1-\pi)^2 \cdot \pi &&= \tbinom{3}{1} \cdot \pi^1 \cdot (1-\pi)^2 \\
&p(\overline{EEE}) &&= (1-\pi) \cdot (1-\pi) \cdot (1-\pi) &&= (1) \cdot (1-\pi)^3 &&= \tbinom{3}{0} \cdot \pi^0 \cdot (1-\pi)^3 \ .
\end{aligned}$$

Wir sehen, dass die eingeklammerten Faktoren den Zahlen der 2. und 3. Zeile des Pascalschen Dreiecks entsprechen, die sich als $\binom{n}{x}$ mit $x = n, \ldots, 0$ ergeben.

Verallgemeinern wir von $n = 3$ auf n Versuche, so erhalten wir folgende Wahrscheinlichkeiten für das x-malige Auftreten des Ereignisses E:

$$\begin{array}{lll}
x = n: & p\ (\text{n-mal E und 0-mal } \bar{E}) & = \binom{n}{n} \cdot \pi^n \cdot (1-\pi)^0 \\
x = n-1: & p\ (\text{n}-1\text{-mal E und 1-mal } \bar{E}) & = \binom{n}{n-1} \cdot \pi^{n-1} \cdot (1-\pi)^1 \\
x = n-2: & p\ (\text{n}-2\text{-mal E und 2-mal } \bar{E}) & = \binom{n}{n-2} \cdot \pi^{n-2} \cdot (1-\pi)^2 \\
\vdots & & \\
x = 0: & p\ (\text{0-mal E und n-mal } \bar{E}) & = \binom{n}{0} \cdot \pi^0 \cdot (1-\pi)^n \ .
\end{array}$$

Da mit $x = n, n-1, \ldots, 0$ alle möglichen Realisierungen der Zufallsvariablen X erschöpft sind, muss die Summe der Wahrscheinlichkeiten dieser Realisierungen 1 ergeben. Setzt man $p = \pi$ und $q = 1 - \pi$, muss wegen $p + q = 1$ folgende Gleichung gelten:

$$(p+q)^n = \binom{n}{n} \cdot p^n \cdot q^0 + \binom{n}{n-1} \cdot p^{n-1} \cdot q^1 \binom{n}{n-2} \cdot p^{n-2} \cdot q^2 + \ldots + \binom{n}{0} \cdot p^0 \cdot q^n \, .$$

Die rechte Seite dieser Verteilung stellt die Entwicklung des Binoms $p + q$ für die n-te Potenz dar und heißt deshalb *binomische Entwicklung*. Die Koeffizienten $\binom{n}{x}$ heißen *Binomialkoeffizienten*, die nach dem Pascalschen Dreieck einfach zu berechnen sind. Setzt man weiterhin $\binom{n}{n-x} = \binom{n}{x}$, wobei x die Zahlen $0, 1, \ldots, n$ durchläuft, so erhält man

$$p(x) = \binom{n}{x} \cdot \pi^x \cdot (1-\pi)^{n-x} = \frac{n!}{x!(n-x)!} \cdot \pi^x \cdot (1-\pi)^{n-x} \, . \tag{1.16}$$

Nach dieser Gleichung lässt sich die Wahrscheinlichkeit berechnen, genau x-mal E zu beobachten. Die Wahrscheinlichkeitsverteilung für alle Realisierungen der Zufallsvariablen X heißt *Binomialverteilung*. Ist $= 1 - \pi = 1/2$, geht Gl. (1.16) in Gl. (1.15) über.

Wie man zeigen kann (vgl. etwa Kreyszig, 1973, Abschn. 40) beträgt der Erwartungswert E(X) der Binomialverteilung $\mu = n \cdot \pi$ und die Varianz $\sigma^2 = n \cdot \pi \cdot (1 - \pi)$.

Verteilungsfunktion

Will man nicht Punktwahrscheinlichkeiten, sondern Überschreitungswahrscheinlichkeiten dafür ermitteln, dass $X \leq k$, bedient man sich zweckmäßiger der Verteilungsfunktion bzw. der Summenfunktion der Binomialverteilung. Für den speziellen Fall $\pi = 1 - \pi = 1/2$ lautet sie

$$P(X \leq k) = \sum_{x=0}^{k} \binom{n}{x} \cdot \left(\tfrac{1}{2}\right)^n \, . \tag{1.17}$$

Für beliebiges π lautet die Summenfunktion der Binomialverteilung entsprechend

$$P(X \leq k) = \sum_{x=0}^{k} \binom{n}{x} \cdot \pi^x \cdot (1-\pi)^{n-x} . \tag{1.18}$$

Diese Verteilung ist für ausgewählte Werte π tabelliert (▶ Tafel 1 des Anhangs).

Die Benutzung dieser Tafel sei anhand von Beispielen demonstriert. Bei einer Jahrmarktslotterie möge die Chance für ein Gewinnlos 10% ($\pi = 0{,}1$) betragen. Wie groß ist die Wahrscheinlichkeit, beim Kauf von $n = 15$ Losen mindestens 4-mal zu gewinnen? Wir entnehmen der Tafel für $n = 15$, $\pi = 0{,}1$ und $x = 4; 5 \ldots 15$:

$$P = 0{,}0428 + 0{,}0105 + 0{,}0019 + 0{,}0003 + 8 \cdot (0) = 0{,}0555 \ .$$

Oder als ein Beispiel für eine zweiseitige Fragestellung: Wie groß ist die Wahrscheinlichkeit, dass sich in einer Familie mit $n = 10$ Kindern höchstens 2 oder mindestens 8 Jungen befinden, wenn wir davon ausgehen, dass die Wahrscheinlichkeit für die Geburt eines Jungen bei $\pi = 0{,}5$ liegt? Für $n = 10$, $\pi = 0{,}5$ und $x = 0$; 1; 2 bzw. $x = 8$; 9; 10 entnehmen wir Tafel 1:

$$\begin{aligned} P' &= 0{,}0010 + 0{,}0098 + 0{,}0439 + 0{,}0439 + 0{,}0098 + 0{,}0010 \\ &= 2 \cdot 0{,}0547 = 0{,}1094 \ . \end{aligned}$$

Liegt die Wahrscheinlichkeit für die untersuchte Alternative im Bereich $\pi > 0{,}50$, benutzt man die andere Alternative und deren Häufigkeiten $n - x$ für die Ermittlung der Überschreitungswahrscheinlichkeit. Bezogen auf das 1. Beispiel ist die Wahrscheinlichkeit für mindestens 4 Gewinnlose ($\pi = 0{,}1$) mit der Wahrscheinlichkeit für höchstens $15 - 4 = 11$ Nieten ($\pi = 0{,}9$) identisch.

1.2.3 Die Normalverteilungsapproximation der Binomialverteilung

Wird die Anzahl der Versuche groß ($n > 50$), dann ermittelt man die Überschreitungswahrscheinlichkeiten bei nicht zu kleiner oder nicht zu großer Wahrscheinlichkeit der betrachteten Alternative ($0{,}1 < \pi < 0{,}9$) ökonomischer über die sog. Normalverteilung, der sich die Binomialverteilung mit wachsender Anzahl der Versuche schnell nähert. (Zur Bedeutung der Normalverteilung für die Statistik vgl. z. B. Bortz, 2005, Abschn. 2.5.1.) Die Gleichung für die Dichtefunktion der Normalverteilung lautet:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp[-(x-\mu)^2/2\sigma^2] \tag{1.19}$$

mit $\pi = 3{,}1416$. Ersetzt man die Parameter μ und σ durch die Parameter der Binomialverteilung, $\mu = np$ und, $\sigma = \sqrt{npq}$, so lautet die Gleichung für die Normalapproximation der Binomialverteilung

$$f(x) = \frac{1}{\sqrt{2\pi npq}} \cdot \exp[-(x-np)^2/2npq] \, . \tag{1.20}$$

Die Normalverteilung liegt als sog. *Standard-* oder *Einheitsnormalverteilung* mit $\mu = 0$ und $\sigma = 1$ tabelliert vor (▶ Tafel 2 des Anhangs):

$$f(u) = \frac{1}{\sqrt{2\pi}} \cdot \exp(-u^2/2) \, . \tag{1.21}$$

Hier kann zu jedem Wert x bzw. dem ihm entsprechenden Wert

$$u = \frac{x - np}{\sqrt{npq}} \tag{1.22}$$

die zugehörige Überschreitungswahrscheinlichkeit P abgelesen werden (► S. 34). Die Transformation überführt eine Verteilung mit beliebigem und σ in eine Verteilung mit $\mu = 0$ und $\sigma = 1$ (vgl. dazu etwa Bortz, 2005, unter dem Stichwort z-Transformation).

Die Wahrscheinlichkeit, beim Wurf von $n = 10$ Münzen mindestens $x = 8$ „Zahlen" zu erhalten, errechnen wir nach der exakten Binomialverteilung [Gl. (1.18)] zu $P = 0{,}0547$. Über die Normalverteilung erhalten wir für $p = q = 1/2$

$$u = \frac{8 - 10 \cdot (1/2)}{\sqrt{10 \cdot (1/2) \cdot (1/2)}} = 1{,}90$$

Diesem Abszissenwert u der Standardnormalverteilung entspricht nach Tafel 2 des Anhangs ein P-Wert von 0,0287, der im Verhältnis zum exakt ermittelten $P = 0{,}0547$ zu niedrig ausgefallen ist. Offenbar ist unsere Stichprobe mit $n = 10$ zu klein für die Normalverteilungsapproximation. Die Unterschätzung lässt sich allerdings – wie wir in ► Abschn. 5.1.1 sehen werden – mit Hilfe der sog. *Kontinuitätskorrektur* reduzieren.

Da die Normalverteilung symmetrisch ist, entspricht einem positiven u-Wert dieselbe Überschreitungswahrscheinlichkeit wie einem negativen u-Wert.

1.2.4 Die Polynomialverteilung

Lassen wir die Beschränkung auf 2 Ausprägungsarten fallen, so geht die Binomialverteilung in die *Polynomialverteilung* oder auch *Multinomialverteilung* über. Für m Ausprägungsarten mit den Wahrscheinlichkeiten π_1, π_2, … π_m ergibt sich die Punktwahrscheinlichkeit einer bestimmten Zusammensetzung einer Stichprobe des Umfanges n mit n_1 Elementen der ersten, n_2 Elementen der zweiten und n_m Elementen der m-ten Ausprägung zu

$$p(n_1, n_2, \ldots, n_m) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots n_m!} \cdot \pi_1^{n_1} \cdot \pi_2^{n_2} \cdot \ldots \pi_m^{n_m} \,. \tag{1.23}$$

Die Überschreitungswahrscheinlichkeit P, die beobachtete oder eine extremere Zusammensetzung der Stichprobe durch Zufall anzutreffen, ergibt sich zu

$$P = \sum p^* = \sum p(n_1^*, n_2^*, \ldots, n_m^*) \,, \tag{1.24}$$

wobei die p^* alle Punktwahrscheinlichkeiten für Anordnungen mit $n_1^*, n_2^*, \ldots, n_m^*$ Elementen bezeichnet, die kleiner oder gleich der Punktwahrscheinlichkeit der beobachteten Zusammensetzung sind.

Die Ermittlung von Punkt- und Überschreitungswahrscheinlichkeiten sei an einem Beispiel verdeutlicht. Angenommen, in einem akademischen Entscheidungsgremium befinden sich $n = 4$ Studenten, denen die folgenden Parteizugehörigkeiten nachgesagt werden:

Partei A: $n_1 = 0$,
Partei B: $n_2 = 1$
und
Partei C: $n_3 = 3$.

In der studentischen Population haben die 3 Parteien folgende Sympathisantenanteile: $\pi_1 = 0{,}5$, $\pi_2 = 0{,}3$ und $\pi_3 = 0{,}2$. Wir fragen nach der Wahrscheinlichkeit der Gremienzusammensetzung angesichts dieser Populationsverhältnisse. Nach Gl. (1.23) ergibt sich

$$p(n_1 = 0, n_2 = 1, n_3 = 3) = \frac{4!}{0! \cdot 1! \cdot 3!} \cdot 0{,}5^0 \cdot 0{,}3^1 \cdot 0{,}2^3 = 0{,}0096 \,.$$

Diese Wahrscheinlichkeit ist sehr gering und spricht nicht für eine „repräsentative“ Auswahl. Fragen wir – im Sinne der Überschreitungswahrscheinlichkeit –, wie wahrscheinlich diese und noch extremere Auswahlen sind (extremer im Sinne einer noch stärkeren Abweichung von der Populationsverteilung), benötigen wir die Punktwahrscheinlichkeiten der extremeren Auswahlen. In unserem Beispiel sind dies die Zusammensetzungen

$$n_1^* = 0,\; n_2^* = 4,\; n_3^* = 0 \text{ mit } p^* = 0{,}0081 \,,$$
$$n_1^* = 0,\; n_2^* = 0,\; n_3^* = 4 \text{ mit } p^* = 0{,}0016 \,.$$

Alle übrigen Zusammensetzungen haben eine größere Punktwahrscheinlichkeit als die angetroffene. Als Überschreitungswahrscheinlichkeit errechnen wir damit

$$P = 0{,}0096 + 0{,}0081 + 0{,}0016 = 0{,}0193 \,.$$

Die Polynomialverteilung spielt überall dort als Prüfverteilung eine Rolle, wo Elemente oder Ereignisse nach mehr als 2 Klassen aufgeteilt sind; sie wird, wie wir im nächsten Abschnitt sehen werden, durch eine andere, viel leichter zu handhabende Verteilung hinreichend gut angenähert, bei der die Bestimmung von Überschreitungswahrscheinlichkeiten keinerlei Mühe macht.

Ein Spezialfall der Polynomialverteilung ist die Gleichverteilung oder Rechteckverteilung, in der $\pi_1 = \pi_2 = \ldots \pi_m = 1/m$ für alle m Klassen ist. Die Punktwahrscheinlichkeit einer Stichprobe von n_1, n_2, …, n_m Elementen ist gegeben durch

$$\begin{aligned} p(n_1, n_2, \ldots, n_m) &= \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_m!} (1/m)^{n_1 + n_2 + \ldots + n_m} \\ &= \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_m!} \cdot \frac{1}{m^n} \,. \end{aligned} \tag{1.25}$$

Die Gleichverteilung für $m = 2$ Klassen ist die Binomialverteilung für $\pi = 0{,}5$. Nach der Terminologie von Gl. (1.15) entspricht $n_1 = x$ und $n_2 = n - x$, so dass

$$\begin{aligned} p(n_1, n_2) &= \frac{n!}{n_1 \cdot n_2!} \cdot \frac{1}{2^n} = \frac{n!}{x! \cdot (n-x)!} \cdot \frac{1}{2^n} \\ &= \binom{n}{x} \cdot \frac{1}{2^n} \,. \end{aligned}$$

1.2.5 Die χ^2-Approximation der Polynomialverteilung

Die Ermittlung der Überschreitungswahrscheinlichkeiten nach der Polynomialverteilung ist schon für kleine Stichproben sehr mühsam. Glücklicherweise geht sie bereits für relativ kleine Stichprobenumfänge in eine andere theoretische Verteilung, die χ^2-Verteilung, über, die von Pearson (1900) nach Überlegungen von Helmert (1876) erarbeitet wurde. Diese Verteilung liegt ebenfalls tabelliert vor (▶ Tafel 3 des Anhangs).

Die χ^2-Verteilung – genauer: die χ^2-Verteilung für k Freiheitsgrade – ist definiert als Verteilung der Summe der Quadrate von k unabhängigen Standardnormalvariablen $u_i = (x_i - \mu)/\sigma$ nach

$$\chi^2 = u_1^2 + u_2^2 + \ldots + u_k^2 . \tag{1.26}$$

Durch infinitesimale Ableitung lässt sich zeigen, dass die Ordinate f der χ^2-Verteilung im Punkt χ^2 der Abszisse gegeben ist durch

$$f(\chi^2) = K \cdot \chi^{k-1} \cdot e^{-\chi^2/2} \tag{1.27}$$

wobei die Konstante K den folgenden Wert annimmt:

$$K = \frac{1}{\left(\frac{k-2}{2}\right)! 2^{(k-2)/2}} .$$

Wie die Polynomialverteilung eine Verallgemeinerung der Binomialverteilung ist, so ist auch die χ2-Verteilung eine Verallgemeinerung der Normalverteilung: Entnimmt man jeweils nur $k = 1$ normalverteilte Zufallszahlen, dann geht der Ausdruck (1.27) in die Form

$$f(\chi^2) = K \cdot e^{-x^2/2} \tag{1.28}$$

über, die mit Gl. (1.21) identisch ist, wenn man χ^2 durch u^2 und K durch $1/\sqrt{2\pi}$ ersetzt.

Kritisch für die Bestimmung der zu einem bestimmten χ^2-Wert gehörenden Überschreitungswahrscheinlichkeit P ist die Zahl der *Freiheitsgrade* (Fg). In der Definitionsgleichung (1.26) ist $Fg = k$, also gleich der Zahl der unabhängigen u-Werte. Liegt aber $\sum u = \text{const.}$ fest, weil etwa $\mu_u = \sum u/k$ als Durchschnitt der u-Variablen gegeben ist, dann reduziert sich die Zahl der Freiheitsgrade um 1; dies ist auch bei m Klassen von Häufigkeiten f der Fall, wenn $\sum f = n = \text{const.}$

Wie Pearson gezeigt hat, ist auch der folgende Ausdruck approximativ χ^2-verteilt:

$$\chi^2 = \sum_{i=1}^{m} \frac{(b_i - e_i)^2}{e_i} . \tag{1.29}$$

Dabei sind b_i die in einer Kategorie i beobachteten und e_i die theoretisch erwarteten Häufigkeiten. Dieser Ausdruck ist χ^2-verteilt, wenn die erwarteten Häufigkeiten e_i genügend groß sind. Als Richtwerte für ein ausreichend großes e_i werden in der statistischen Literatur unter verschiedenen Bedingungen Werte $e_i = 5$, $e_i = 10$ oder $e_i = 30$ angegeben (▶ dazu Kap. 5).

Zur Verdeutlichung von Gl. (1.29) greifen wir erneut das in ▶ Abschn. 1.2.4 genannte Beispiel auf, nun allerdings mit einer größeren Stichprobe. Angenommen, von $n = 30$ Studenten sympathisieren $b_1 = 15$ mit Partei A, $b_2 = 11$ mit Partei B und

$b_3 = 4$ mit Partei C. Die theoretisch erwarteten Häufigkeiten erhalten wir, indem die auf S. 18 genannten π-Werte mit n multipliziert werden: $e_1 = 0{,}5 \cdot 30 = 15$, $e_2 = 0{,}3 \cdot 30 = 9$ und $e_3 = 0{,}2 \cdot 30 = 6$. Nach Gl. (1.29) resultiert damit ein χ^2 von

$$\chi^2 = \frac{(15-15)^2}{15} + \frac{(11-9)^2}{9} + \frac{(4-6)^2}{6} = 1{,}11\,.$$

Da die theoretischen Häufigkeiten in diesem Beispiel die gleiche Summe ergeben müssen wie die beobachteten, hat dieser χ^2-Wert $m-1=2$ (m = Anzahl der Kategorien) Freiheitsgrade. Tafel 3 des Anhangs ist zu entnehmen, dass für Fg = 2 ein $\chi^2 = 1{,}022$ eine Überschreitungswahrscheinlichkeit von $P = 0{,}60$ und ein $\chi^2 = 1{,}386$ eine Überschreitungswahrscheinlichkeit von $P = 0{,}50$ aufweisen. Demnach hat der empirisch ermittelte χ^2-Wert eine Überschreitungswahrscheinlichkeit, die zwischen 0,50 und 0,60 liegt. Daraus wäre zu folgern, dass die theoretische Verteilung nicht gravierend von der empirischen Verteilung abweicht (Näheres dazu ► Abschn. 5.1.3).

1.2.6 Die Poisson-Verteilung

Wenn die Anzahl der Ereignisse n sehr groß und die Wahrscheinlichkeit des untersuchten Ereignisses π sehr klein sind, wird die Ermittlung binomialer Wahrscheinlichkeiten nach Gl. (1.16) sehr aufwendig. In diesem Falle empfiehlt es sich, die exakten binomialen Wahrscheinlichkeiten durch die Wahrscheinlichkeiten einer anderen Verteilung, der Poisson-Verteilung, zu approximieren. Die Wahrscheinlichkeitsfunktion der Poisson-Verteilung lautet:

$$p(x) = \frac{\mu^x}{x!} \cdot e^{-\mu} \tag{1.30}$$

mit $\mu = n \cdot \pi$ und $e = 2{,}7183$ (Basis der natürlichen Logarithmen).

Die Binomialverteilung geht in die Poisson-Verteilung über, wenn $n \to \infty$, $\pi \to 0$ und $n \cdot \pi = \text{const.}$ (vgl. dazu etwa Kreyszig, 1973, Abschn. 42). Varianz und Mittelwert sind bei der Poisson-Verteilung identisch: $\mu = \sigma^2 = n \cdot \pi$.

Die Poisson-Verteilung wird gelegentlich auch als *Verteilung seltener Ereignisse* bezeichnet. Ihre Berechnung sei im Folgenden an einem Beispiel verdeutlicht. (Weitere Anwendungen der Poisson-Verteilung findet man z. B. bei Hays, 1973, Abschn. 5.21).

An einem Roulettetisch werden an einem Abend n = 300 Spiele gemacht. Ein Spieler behauptet, dass an diesem Abend die Zahl 13 nicht häufiger als 2-mal fällt. Mit welcher Wahrscheinlichkeit hat der Spieler mit seiner Behauptung recht, wenn es sich um ein „faires" Roulette handelt, d. h. wenn $\pi = 1/37$?

Nach Gl. (1.30) errechnen wir $\mu = 300/37 = 8{,}11$ und

$$p(x=0) = \frac{8,11^0}{0!} \cdot e^{-8{,}11} = 0{,}0003$$

$$p(x=1) = \frac{8,11^1}{1!} \cdot e^{-8{,}11} = 0{,}0024$$

$$p(x=2) = \frac{8,11^2}{2!} \cdot e^{-8{,}11} = 0{,}0099\,.$$

Als Überschreitungswahrscheinlichkeit ergibt sich damit der Wert $P=0{,}0126$. Es empfiehlt sich also nicht, der Intuition des Spielers zu folgen.

Unter Verwendung von Gl. (1.16) lautet die exakte Überschreitungswahrscheinlichkeit nach der Binomialverteilung $P=0{,}0003+0{,}0023+0{,}0094=0{,}0120$. Die Poisson-Approximation kann damit bereits für n- und π-Werte in der Größenordnung des Beispiels als brauchbar angesehen werden.

1.2.7 Die hypergeometrische Verteilung

Wir haben nun abschließend noch eine Wahrscheinlichkeitsverteilung kennenzulernen, die sich dann ergibt, wenn Stichproben zweiklassiger Elemente aus einer endlich begrenzten Grundgesamtheit entnommen werden: die hypergeometrische Verteilung.

Die hypergeometrische Verteilung lässt sich anhand eines sog. Urnenmodells folgendermaßen herleiten: In einer Urne befinden sich K farbige und $N-K$ farblose Kugeln, insgesamt also N Kugeln. Die Wahrscheinlichkeit, eine farbige Kugel zu ziehen, ist damit $\pi=K/N$. Die Wahrscheinlichkeit, genau x farbige Kugeln in einer Stichprobe von n Kugeln zu finden, ergibt sich aus folgenden Überlegungen: Es bestehen $\binom{K}{x}$ Möglichkeiten, x farbige Kugeln aus den K insgesamt vorhandenen farbigen Kugeln herauszugreifen; es bestehen weiterhin $\binom{N-K}{n-x}$ Möglichkeiten, $n-x$ farblose Kugeln aus den insgesamt vorhandenen $N-K$ farblosen Kugeln herauszugreifen. Daher ergeben sich nach dem Multiplikationssatz $\binom{K}{x}\cdot\binom{N-K}{n-x}$ Möglichkeiten, aus den N Kugeln x farbige und $n-x$ farblose Kugeln zu ziehen. Da die Gesamtzahl aller Kombinationen für n Kugeln aus N Kugeln $\binom{N}{n}$ beträgt, ergibt sich die Wahrscheinlichkeit p(x) für x farbige Kugeln aus n Kugeln zu:

$$p(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \,. \tag{1.31}$$

Der Ausdruck p(x) entspricht einer Punktwahrscheinlichkeit. Die Überschreitungswahrscheinlichkeit, x oder weniger farbige Kugeln zu ziehen, bestimmt man als Summe der zugehörigen Punktwahrscheinlichkeiten: $P=p(x)+p(x-1)+\ldots+p(0)$.

Die hypergeometrische Verteilung hat einen Mittelwert von $n\cdot\pi$ und eine Standardabweichung von $\sqrt{n\cdot\pi\cdot(1-\pi)\cdot(N-n)/(M-1)}$; sie hat also das gleiche Mittel wie die Binomialverteilung, nur eine um den Faktor $\sqrt{(N-n)/(N-1)}$ kleinere Streuung. Sie geht in die Binomialverteilung über, wenn $N\rightarrow\infty$.

Auch diese Verteilung sei an einem Beispiel erläutert: Wenn wir berechnen wollen, wie hoch die Chance ist, im Zahlenlotto „6 aus 49“ den niedrigsten Gewinnrang ($x=3$ Richtige) zu haben, so wären einzusetzen: $K=6$ (Anzahl der möglichen Treffer), $x=3$ (Anzahl der tatsächlichen Treffer), $N=49$ (Anzahl der Kugeln im Ziehungsgerät), $n=6$ (Anzahl zu ziehender Kugeln), $N-K=43$ und $n-x=3$.

$$p(x=3) = \frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}} = \frac{20\cdot 12341}{13983816} = 0{,}0177 \,.$$