4. ZUFALLSGRÖSSEN

4.1. DISKRETE ZUFALLSGRÖSSEN

(4.1.1) Überblick

Eine *Zufallsgrösse* X ist eine Abbildung

$$X : \Omega \to \mathbb{R} ,$$

die jedem Element des Ergebnisraumes Ω eine Zahl zuordnet. Zufallsgrössen treten also immer dann auf, wenn ein Ergebnis eines Zufallsexperiments mit einer Zahl verbunden ist, in der Praxis also z.B. bei Messungen oder Zählungen an zufällig ausgewählten Objekten. *(4.1.3)*

Eine Zufallsgrösse heisst *diskret*, wenn sie nur endlich oder abzählbar unendlich viele Werte $x_1, x_2, \ldots$ annimmt. In (4.2) werden die *stetigen* Zufallsgrössen besprochen. *(4.1.4)*

Jeder Wert x_i der diskreten Zufallsgrösse wird mit einer bestimmten Wahrscheinlichkeit

$$p_i = P(X = x_i)$$

angenommen. Unter der *Verteilung* von X versteht man die Menge der Paare (x_i, p_i). Sie kann auch durch das *Stabdiagramm*, durch die *Wahrscheinlichkeitsfunktion* *(4.1.5)* *(4.1.6)* *(4.1.6)*

$$f(x) = P(X = x)$$

oder durch die *Verteilungsfunktion* *(4.1.7)*

$$F(x) = P(X \leq x)$$

beschrieben werden.

(4.1.2) Einleitende Beispiele

Beispiel 4.1.2.A

Auf einer Wiese soll das Vorkommen einer bestimmten Pflanzenart untersucht werden. Zu diesem Zweck teilt der Forscher die Wiese in viele gleich grosse Parzellen ein und bestimmt in jeder dieser Parzellen die Anzahl der Pflanzen dieser Art. Diese Anzahl hängt erfahrungsgemäss vom Zufall ab: Es wird Parzellen geben, in denen keine solche Pflanze gedeiht, in andern wachsen vielleicht fünf oder zehn Exemplare. Wie ist nun diese Abhängigkeit vom Zufall zu verstehen? Das Zählen selbst ist bestimmt nicht ein zufälliger Prozess. Wenn unser Forscher nämlich eine Parzelle ausgewählt hat, dann kommt er auf ein eindeutig festgelegtes Resultat. Der Zufall steckt also vielmehr in der Auswahl der Parzelle. (Natürlich spielt auch die Grösse dieser Parzelle eine Rolle, doch diese soll für den Verlauf der Untersuchung ein für allemal fest gewählt sein.)

Sobald man von zufälligen Vorgängen spricht, stellt sich sofort die Frage nach der Wahrscheinlichkeit der Ergebnisse. Bei einer häufig vorkommenden Pflanzenart wird es viel wahrscheinlicher sein, dass auf einer Parzelle fünf Exemplare wachsen, als dass keines vorkommt.

Wie wir bereits zur Genüge gesehen haben, gehört zur Beschreibung eines Zufallsexperiments und zur Bestimmung von Wahrscheinlichkeiten stets ein Wahrscheinlichkeitsraum Ω. In unserm Beispiel ist dies gerade die Menge aller Parzellen. Neu an der Sache ist nun, dass zu jedem Ergebnis $\omega \in \Omega$, also auf Deutsch zu jeder Parzelle, eine Zahl gehört (nämlich die Anzahl Pflanzen), der unser eigentliches Interesse gilt. ⊠

<u>Beispiel 4.1.2.B</u>

Eine andere Gelehrte will die Körperlänge der Einwohner von Zürich erforschen. Dazu wählt sie zufällig eine Anzahl von Zürcher(inne)n aus und misst ihre Grösse. Ähnlich wie im Beispiel 4.1.2.A steckt hier der Zufall in der Auswahl der Person, die Körperlänge selbst ermittelt sich anschliessend auf klar bestimmte Weise. Der zum "Experiment" gehörende Ergebnisraum ist also die Menge aller Zürcher(innen), von Bedeutung aber ist vor allem die zum Ergebnis (also zur Person) gehörende Körperlänge, also, wie schon oben, eine Zahl. ⊠

<u>Beispiel 4.1.2.C</u>

Vergleichbare Situationen treten bei Glücksspielen auf, die ja Prototypen von Zufallsereignissen sind. Ein einfaches Beispiel:

Ich werfe dreimal eine Münze. Die Anzahl der Köpfe soll mein Gewinn sein. Der Ergebnisraum Ω besteht aus acht Ergebnissen, die man zusammen mit dem "Gewinn" wie folgt tabellieren kann:

Ergebnis	KKK	KKZ	KZK	KZZ	ZKK	ZKZ	ZZK	ZZZ
Gewinn	3	2	2	1	2	1	1	0

Von Interesse ist hier eigentlich nicht das Ergebnis, also etwa KKK oder KZK aus Ω, sondern vielmehr der Gewinn, denn es ist mir nämlich egal, ob ich meine zwei Franken mit der Kombination KKZ oder mit KZK gewonnen habe. In der Tabelle sieht man jetzt besonders deutlich, wie zu einem Zufallsergebnis die uns interessierende Zahl gehört. ⊠

<u>Beispiel 4.1.2.D</u>

Das folgende Würfelspiel hat etwas eigenartige Regeln, nämlich:

- Bei gerader Augenzahl gewinne ich, und zwar das Doppelte der gewürfelten Augenzahl.

- Bei ungerader Augenzahl verliere ich, und zwar das Dreifache der gewürfelten Zahl.

Der Ergebnisraum Ω besteht hier aus den Augenzahlen 1, 2, 3, 4, 5 und 6, seine Elemente sind also — im Gegensatz zu den Beispielen 4.1.2.A, 4.1.2.B und 4.1.2.C — selbst Zahlen. Wenn man einen Verlust als negativen Gewinn notiert, dann lässt sich der Sachverhalt mit der folgenden Tabelle beschreiben:

Ergebnis	1	2	3	4	5	6
Gewinn	-3	4	-9	8	-15	12

Wiederum haben wir einerseits die Ergebnisse aus Ω und anderseits die zugehörigen Zahlen, d.h., die Gewinne. ⊠

Beispiel 4.1.2.E

Eine Sendung von Massenartikeln, z.B. von Schrauben, wird auf ihre Brauchbarkeit geprüft. Wir ordnen jedem intakten Artikel (willkürlich) die Zahl 1, einem Ausschussartikel dagegen die Zahl 0 zu. Der Ergebnisraum Ω besteht aus allen Artikeln aus der Sendung, das Zufallsexperiment besteht im Auswählen eines Artikels. Zu jedem Ergebnis gehört eine Zahl, nämlich 1 oder 0, die hier aber bloss eine Art Code für "brauchbar" bzw. "Ausschuss" ist.

(4.1.3) Zufallsgrössen

Den obigen fünf Beispielen ist etwas gemeinsam: Jedes Mal haben wir einen Ergebnisraum Ω; unser primäres Interesse liegt aber nicht bei den Ergebnissen ω aus Ω, sondern bei gewissen Zahlen, die zu diesen Ergebnissen gehören, im Sinne der folgenden Tabelle:

Ergebnis $\omega \in \Omega$	*Zugehörige Zahl*
Parzelle	Anzahl Pflanzen
Person	Körperlänge
Münzenwurf	Anzahl Köpfe
Augenzahl	Gewinn
Artikel	0 oder 1

Allgemein lässt sich dieser Sachverhalt dadurch ausdrücken, dass man sagt, es sei eine *Abbildung* $\Omega \to \mathbb{R}$ gegeben. (Zum Begriff der Abbildung im allgemeinen Sinn finden Sie im ersten Band in (26.9) eine Bemerkung.) Traditionellerweise nennt man eine solche Abbildung eine *Zufallsgrösse* und bezeichnet sie mit X oder auch einem andern Buchstaben, meist vom Ende des Alphabets, wie etwa T, Y oder Z.

Die formale Definition lautet:

Es sei Ω ein Ergebnisraum. Unter einer *Zufallsgrösse* (auf Ω) versteht man eine Abbildung

$$X : \Omega \to \mathbb{R} \ .$$

Bemerkungen

1) Über den Ergebnisraum wird nichts vorausgesetzt. Er kann — wie in den Beispielen von (4.1.2) — endlich, aber auch abzählbar oder sogar überabzählbar unendlich sein. Beispiele hierzu folgen.

2) Wir wollen an zwei Beispielen illustrieren, wie die Abbildung $X : \Omega \to \mathbb{R}$ in konkreten Fällen aussehen kann.

   In Beispiel 4.1.2.C ist X die Anzahl Köpfe bei einem dreifachen Münzenwurf. So ist etwa $X(\text{KKK}) = 3,\ X(\text{KKZ}) = X(\text{KZK}) = 2$ usw.

   In Beispiel 4.1.2.D ist $X(k)$ der Gewinn, wenn beim Würfeln die Augenzahl k fällt, es ist also $X(1) = -3,\ X(2) = 4$ usw. Die im Beispiel angegebene Tabelle ist also einfach eine Wertetabelle für die Abbildung X, vgl. auch Bemerkung 5) weiter unten.

3) Sehr oft wird die Bezeichnung *"Zufallsvariable"* statt "Zufallsgrösse" verwendet. Mir scheint dies nicht sehr glücklich, denn X ist eine Abbildung und keine "Variable" (die Rolle der Letztern wird von $\omega \in \Omega$ gespielt). Sie ist auch nur insofern vom "Zufall" abhängig, als das Argument $\omega \in \Omega$ als zufälliges Ereignis aufgefasst wird. Der Wert $X(\omega)$ selbst ist vollständig bestimmt und nicht mehr zufällig. Ist in Beispiel 4.1.2.A die Parzelle einmal gewählt, so ist die Anzahl der darauf wachsenden Pflanzen eindeutig festgelegt.
   Aus diesen Gründen ziehe ich den Namen "Zufallsgrösse" vor. Allerdings müssen Sie auch den andern Begriff kennen, weshalb in diesem Skript gelegentlich auch von Zufallsvariablen die Rede sein wird.

4) Die folgende Vorstellung kann manchmal hilfreich sein: Eine Zufallsgrösse X lässt sich stets als Gewinn (oder Verlust) bei einem "Glücksspiel" auffassen. Das Spiel besteht darin, ein Element $\omega \in \Omega$ zufällig zu wählen. Dabei ist dann $X(\omega)$ der Gewinn oder der Verlust.

5) Zufallsgrössen kann man auch ganz willkürlich festlegen, ohne dass man an einen bestimmten konkreten Anlass denkt. Besteht Ω aus den Elementen $\omega_1,\ \omega_2, \ldots, \omega_n$, so kann man n beliebige Zahlen $x_1,\ x_2, \ldots, x_n$ wählen und durch

$$X(\omega_i) = x_i, \qquad i = 1, \ldots, n$$

   eine Zufallsgrösse auf Ω festsetzen. Dies entspricht der tabellarischen Darstellung

Ergebnis ω	ω_1	ω_2	$\ldots$	ω_n
$X(\omega)$	x_1	x_2	$\ldots$	x_n .

6) Aus gegebenen Zufallsgrössen kann man neue bilden. Ist X wie in Beispiel 4.1.2.D durch die Tabelle

ω	1	2	3	4	5	6
$X(\omega)$	-3	4	-9	8	-15	12

gegeben, so kann man eine neue Zufallsgrösse Y dadurch bilden, dass der durch X beschriebene Gewinn verdoppelt und um 1 erhöht wird. Würfelt man also eine "1", so ist der Gewinn gleich $2 \cdot (-3) + 1 = -5$ (also ein Verlust). Allgemein erhält man $Y(\omega)$ mit der Formel

$$Y(\omega) = 2X(\omega) + 1 ,$$

was folgende Tabelle ergibt:

ω	1	2	3	4	5	6
$Y(\omega)$	-5	9	-17	17	-29	25 .

Man schreibt dann kurz $Y = 2X + 1$. Auf ähnliche Weise kann man weitere Zufallsgrössen wie $Z = X^2$ usw. angeben.

7) Man kann auch mehrere Zufallsgrössen, die auf demselben Ergebnisraum Ω definiert sind, kombinieren: Es sei X die Erfahrungs-, Y die Prüfungsnote in Mathematik. (Unter Ω ist hier die Menge aller Maturand(inn)en zu verstehen; das Zufallsexperiment besteht in der Auswahl einer Person.) Dann sind auch $Z = X + Y$ und $M = \frac{1}{2}(X + Y)$ Zufallsgrössen (M ist, abgesehen von Rundungseffekten, gerade die Note im Maturzeugnis). Sie werden mit den Formeln

$$Z(\omega) = X(\omega) + Y(\omega),\ M(\omega) = \frac{1}{2}(X(\omega) + Y(\omega)), \quad \omega \in \Omega = \{\text{Maturand(inn)en}\}$$

berechnet. Hier steht ω also für eine Person.

8) Obwohl Zufallsgrössen definitionsgemäss Abbildungen von einem Ergebnisraum Ω nach $\mathbb{R}$ sind, sieht man in der Praxis sehr oft davon ab, diesen Raum Ω ausdrücklich anzugeben. Man sagt vielmehr einfach (etwa im Zusammenhang mit den Beispielen 4.1.2.A und 4.1.3.B):

- "Die Zufallsgrösse X bezeichnet die Anzahl der Pflanzen der betreffenden Art pro Parzelle."

oder

- "Die Zufallsvariable X sei die Körperlänge von Zürchern."

Es ist dann der Leserin oder dem Leser überlassen, das zugehörige Ω selbst zu konstruieren und X als Abbildung zu identifizieren, falls dies überhaupt nötig ist, was aber meistens gar nicht der Fall sein wird.

Hierzu noch ein weiteres Beispiel. Die Feststellung

- "X bezeichnet die Anzahl der Mädchen in Familien mit drei Kindern"

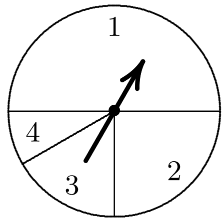
ist so aufzufassen: Wir betrachten alle Familien mit drei Kindern (z.B. im Kanton Zürich oder in der Schweiz); diese bilden den Ergebnisraum Ω. Für jede solche Familie (abstrakt: für jedes $\omega \in \Omega$) ist dann $X(\omega)$ die Anzahl der Mädchen, wobei $X(\omega)$ natürlich die Werte 0, 1, 2 oder 3 annehmen kann.

(4.1.4) Diskrete und stetige Zufallsgrössen

Für uns sind zwei verschiedene Arten von Zufallsgrössen besonders wichtig. Wir besprechen sie anhand von zwei Beispielen.

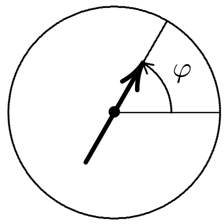
Beispiel 4.1.4.A

Wir betrachten ein Glücksrad. Der Gewinn bei diesem Spiel sei gerade die angeschriebene Zahl. Wir können ihn als Wert einer Zufallsgrösse X auffassen, wobei Ω aus allen möglichen Endlagen des Zeigers besteht. Je nach Zeigerstellung nimmt dann X einen der vier Werte 1, 2, 3, 4 an. ⊠



Beispiel 4.1.4.B

Beim selben Glücksrad können wir uns aber auch für den Winkel φ interessieren, der sich aus der Endlage des Zeigers ergibt (siehe Figur), wobei wir festlegen, dass $0 \leq \varphi < 2\pi$ sein soll, damit Eindeutigkeit herrscht. Diesen Winkel können wir als Zufallsgrösse Y auffassen. ⊠



Im Beispiel 4.1.4.A nimmt die Zufallsgrösse X nur endlich viele Werte an (nämlich 1, 2, 3, 4); man spricht von einer diskreten Zufallsgrösse. Die Zufallsgrösse Y dagegen kann jeden Wert aus dem Intervall $[0, 2\pi)$ annehmen; sie ist stetig. Etwas genauer definiert man:

Eine Zufallsgrösse heisst *diskret*, wenn sie nur endlich oder abzählbar unendlich viele Werte annimmt.

Die Beispiele 4.1.2.A, 4.1.2.C, 4.1.2.D und 4.1.2.E liefern diskrete Zufallsgrössen. Auch 4.1.2.B (Körperlänge von Zürchern) gehört zu diesem Typ; wenigstens, wenn die Körperlänge wie üblich durch Auf- oder Abrunden "diskretisiert" worden ist.

In Kapitel 7 werden wir im Zusammenhang mit der so genannten Poisson-Verteilung eine Zufallsgrösse kennen lernen, welche nicht bloss endlich (wie in den bisherigen Beispielen), sondern abzählbar unendlich viele Werte annimmt. (Es sei in Erinnerung gerufen, dass eine Menge "abzählbar unendlich" heisst, wenn ihre Elemente in eine Folge angeordnet werden können, vgl. (2.2.1.5) oder Beispiel 3.2.3.B.)

Der Begriff der diskreten Zufallsgrösse, der zur Wahrscheinlichkeitsrechnung gehört, ist also das genaue Analogon des Begriffs des diskreten Merkmals, wie es in der beschreibenden Statistik definiert wurde (vgl. (2.2.1.5)).

Dagegen ist die präzise Definition der stetigen Zufallsgrösse etwas komplizierter und wird erst im Teilkapitel 4.3 gegeben. Sie dürfen sich zwar für den Moment durchaus eine Analogie zum "stetigen Merkmal" von (2.2.1.5) vorstellen und sich merken, dass eine stetige Zufallsgrösse alle reellen Zahlen aus einem bestimmten Intervall als Werte annehmen kann, dass dies aber noch nicht die ganze Wahrheit ist.

Um Missverständnissen vorzubeugen, sei noch bemerkt, dass es auch Zufallsgrössen gibt, welche weder diskret noch stetig sind. Derartige Fälle werden wir aber nicht behandeln.

(4.1.5) Die Verteilung einer diskreten Zufallsgrösse

Bis jetzt haben wir unsere Zufallsgrössen nicht mit dem Begriff der Wahrscheinlichkeit in Beziehung gebracht. Nun ist es an der Zeit, dies zu tun. Wir betrachten zur Illustration noch einmal Beispiel 4.1.2.C, wobei wir in der Tabelle nun zusätzlich noch die Wahrscheinlichkeiten $P(\omega)$ der Ergebnisse anführen.

ω	KKK	KKZ	KZK	KZZ	ZKK	ZKZ	ZZK	ZZZ
$X(\omega)$	3	2	2	1	2	1	1	0
$P(\omega)$	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

Im Grunde interessiert uns aber, wie schon im Beispiel 4.1.2.C erwähnt, nicht das Ergebnis (also z.B. KKK), sondern der Wert der Zufallsgrösse, also der Gewinn (z.B. 3). Hier gibt es nur noch die Möglichkeiten 0, 1, 2, 3, entsprechend den vier Werten, welche unsere Zufallsgrösse annimmt.

Ein Gewinn von 1 ergibt sich genau dann, wenn KZZ oder ZKZ oder ZZK herauskommt, also — etwas gelehrter formuliert — wenn das Ereignis

$$E = \{\text{KZZ, ZKZ, ZZK}\}$$

eintritt. Dieses hat die Wahrscheinlichkeit $P(E) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$. Entsprechend bestimmt man die übrigen Wahrscheinlichkeiten in der folgenden Tabelle:

(∗)

Wert x von X	0	1	2	3
Wahrscheinlichkeit	1/8	3/8	3/8	1/8

Nun führen wir eine wichtige neue Bezeichnung ein, die man immer dann verwendet, wenn man Wahrscheinlichkeiten im Zusammenhang mit Zufallsgrössen angeben will:

Die Wahrscheinlichkeit dafür, dass die Zufallsgrösse X den Wert x annimmt, bezeichnet man mit

$$\boxed{P(X = x)\ .}$$

Im Beispiel der Tabelle (∗) hat man also konkret die folgenden Beziehungen:

$$P(X = 0) = \frac{1}{8},\ P(X = 1) = \frac{3}{8},\ P(X = 2) = \frac{3}{8},\ P(X = 3) = \frac{1}{8}\ .$$

Ganz entsprechend schreibt man

$$\boxed{P(X \leq x)\ .}$$

für die Wahrscheinlichkeit dafür, dass X einen Wert $\leq x$ annimmt. Im Falle der Tabelle $(*)$ ist dann beispielsweise

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}\ ,$$

denn $X \leq 1$ tritt genau dann ein, wenn $X = 0$ oder $X = 1$ ist. Der Grund dafür, dass wir die beiden Wahrscheinlichkeiten addieren dürfen, liegt darin, dass die Ereignisse "die Zufallsgrösse nimmt den Wert 0 an" und "die Zufallsgrösse nimmt den Wert 1 an" unvereinbar sind und dass daher Regel 4 angewendet werden kann.

Es gibt noch weitere Varianten, die wohl kaum einer detaillierten Erklärung bedürfen, wie etwa

$$P(1 \leq X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{7}{8}\ ,$$
$$P(X > 1) = P(X = 2) + P(X = 3) = \frac{1}{2} \quad \text{usw.}$$

Diese neuen Bezeichnungsweisen werden wir von nun an ständig verwenden.

Schon bisher hatten wir unsere Zufallsgrössen in Tabellenform dargestellt. Dies wollen wir nun von einem allgemeineren Standpunkt aus erneut tun. Wir betrachten dazu eine diskrete Zufallsgrösse X und gehen zunächst davon aus, dass sie nur endlich viele Werte annimmt. Die gesamte Information über X ist dann in der folgenden Tabelle enthalten (ein konkretes Beispiel dafür ist die Tabelle $(*)$):

Wert x von X	x_1	x_2	$\dots$	x_n
$P(X = x)$	$P(X = x_1)$	$P(X = x_2)$	$\dots$	$P(X = x_n)$

Statt $P(X = x_i)$ schreiben wir meist kürzer p_i. Die obige Tabelle nimmt dann folgende Form an:

Wert x von X	x_1	x_2	$\dots$	x_n
$P(X = x)$	p_1	p_2	$\dots$	p_n

Die x_i und die p_i können dabei beliebige Zahlen sein, abgesehen davon, dass die p_i als Wahrscheinlichkeiten die Bedingungen $0 \leq p_i \leq 1$ und $\sum_{i=1}^{n} p_i = 1$ erfüllen müssen. Die zweite Forderung ergibt sich übrigens daraus, dass die Ereignisse $X = x_1, \dots, X = x_n$ paarweise unvereinbar sind und dass ihre Vereinigung der ganze Ergebnisraum ist.

Eine entsprechende Aussage gilt selbstverständlich auch für den Fall, wo X abzählbar unendlich viele Werte annimmt. Die Tabelle hat dann folgendes Aussehen,

Wert x von X	x_1	x_2	x_3	$\ldots$
$P(X = x)$	p_1	p_2	p_3	$\ldots$

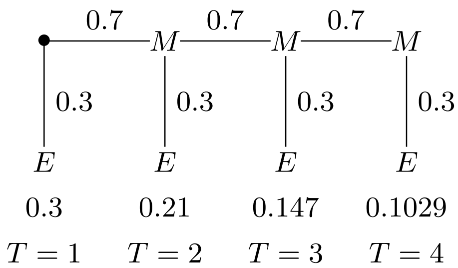
wobei diesmal $\sum_{i=1}^{\infty} p_i = 1$ sein muss (konvergente unendliche Reihe, vgl. Axiom 3).

Die in diesen Tabellen enthaltene Information über die Zufallsgrösse X nennt man die *Verteilung* von X. Etwas präziser, aber auch etwas abstrakter definiert man die Verteilung der diskreten Zufallsgrösse X als die Menge $\mathcal{V}$ aller Paare $\big(x_i,\, P(X = x_i)\big)$. Ein Element von $\mathcal{V}$ ist also ein Paar, bestehend aus x_i, d.h., einem der möglichen Werte für X und der Wahrscheinlichkeit dafür, dass X gerade diesen Wert annimmt. Diese Angaben können statt in Tabellenform auch graphisch dargestellt werden, siehe (4.1.6) und (4.1.7). Wir führen nun zwei Beispiele vor.

Beispiel 4.1.5.A

Wenn ich einen Auskunftsdienst anrufe, komme ich erfahrungsgemäss mit einer Wahrscheinlichkeit von 30% durch. Ich rufe mehrmals hintereinander an, bis ich Erfolg habe. Spätestens nach fünf Versuchen gebe ich aber auf. Wie ist die Zufallsgrösse $T =$ “Anzahl Anrufe” verteilt?

Es ist klar, dass T die Werte 1 bis 5 annehmen kann. Die zugehörigen Wahrscheinlichkeiten für 1 bis 4 Anrufe entnimmt man am besten einem Baumdiagramm (E bedeutet Erfolg, M Misserfolg), die Wahrscheinlichkeit für $T = 5$ ist dann die Ergänzung auf die Summe 1.



Die Verteilung von T ist also durch die folgende Tabelle gegeben:

t	1	2	3	4	5
$P(T = t)$	0.3	0.21	0.147	0.1029	0.2401

Daraus liest man beispielsweise sofort ab:

$$P(2 \leq T \leq 4) = 0.21 + 0.147 + 0.1029 = 0.4599 \; . \qquad \boxtimes$$

Beispiel 4.1.5.B

In einer Tüte hat es 3 weisse und 7 rote Zucker-Eili. Ich nehme mit einem Griff vier Eili heraus. Die Zufallsgrösse W sei die Anzahl weisser Eili, die ich dabei erwischt habe. Gesucht ist die Verteilung von W.

Offenbar kann W die Werte 0, 1, 2, 3 annehmen. Wie gross sind die zugehörigen Wahrscheinlichkeiten? Nach den Regeln der Kombinatorik gibt es $\binom{10}{4}$ Möglichkeiten, aus 10 "Objekten" (hier Zuckereiern) deren 4 herauszugreifen (auf die Reihenfolge kommt es hier nicht an), vgl. (1.5). Damit unter den vier ausgewählten Eiern kein weisses ist, muss ich meine vier Eier unter den sieben roten ausgewählt haben, wofür es $\binom{7}{4}$ Möglichkeiten gibt. Es folgt

$$P(W=0) = \frac{\binom{7}{4}}{\binom{10}{4}} = \frac{1}{6}\,.$$

Um genau ein weisses Ei zu haben, muss ich unter den drei weissen eines, unter den sieben roten drei auswählen. Man findet

$$P(W=1) = \frac{\binom{3}{1}\cdot\binom{7}{3}}{\binom{10}{4}} = \frac{1}{2}\,.$$

Die Situation ist übrigens genau dieselbe wie beim Zahlenlotto (Beispiel 3.4.4.A, vgl. auch (4.1.9.c)); den drei weissen Eiern entsprechen die 6 Gewinnzahlen, den 7 roten die 39 Nicht-Gewinnzahlen. Entsprechend folgt weiter

$$P(W=2) = \frac{\binom{3}{2}\cdot\binom{7}{2}}{\binom{10}{4}} = \frac{3}{10}, \qquad P(W=3) = \frac{\binom{3}{3}\cdot\binom{7}{1}}{\binom{10}{4}} = \frac{1}{30}\,.$$

In Tabellenform also

k	0	1	2	3
$P(W=k)$	1/6	1/2	3/10	1/30

⊠

Zum Schluss noch eine Klarstellung von Begriffen. Weiter oben (im Anschluss an die eingerahmte Formel $P(X \le x)$) war die Rede vom Ereignis "die Zufallsgrösse nimmt den Wert 0 (bzw. 1) an". Diese Sprechweise ist insofern nicht ganz präzis, als sich der Begriff "Ereignis" auf eine Teilmenge des Ergebnisraums Ω bezieht. Mit "die Zufallsgrösse nimmt den Wert 0 an" war natürlich das Ereignis {ZZZ} (kein Kopf), mit "sie nimmt den Wert 1 an" das Ereignis {KZZ, ZKZ, ZZK} (ein Kopf) gemeint. Die abgekürzte Sprechweise ist aber sehr zweckmässig, da im Zusammenhang mit Zufallsgrössen der Ergebnisraum Ω sehr oft überhaupt nicht erwähnt wird. Missverständnisse sind kaum zu befürchten.

Für jene, die es genau formuliert haben wollen: Wenn $X : \Omega \to \mathbb{R}$ eine Zufallsgrösse im abstrakten Sinn ist, dann betrachten wir für irgendeine reelle Zahl x die Teilmengen (die Ereignisse) $E = \{\omega \in \Omega \mid X(\omega) = x\}$ bzw. $F = \{\omega \in \Omega \mid X(\omega) \leq x\}$. Offensichtlich ist dann $P(X = x) = P(E)$ bzw. $P(X \leq x) = P(F)$, und es liegen nun wirklich Wahrscheinlichkeiten von Ereignissen E, $F \subset \Omega$ vor.

Bei dieser Gelegenheit sei noch ein weiterer theoretischer Punkt angesprochen. In (3.3.9) ist erwähnt worden, dass in der allgemeinen Theorie nicht mehr jede Teilmenge von Ω als Ereignis zugelassen werden kann. Deshalb ist auch nicht mehr jede beliebige Abbildung $X : \Omega \to \mathbb{R}$ eine Zufallsgrösse. Man muss vielmehr verlangen, dass für jedes Intervall I aus $\mathbb{R}$ (wobei I hier auch bloss aus einer einzigen Zahl bestehen darf) die Menge $\{\omega \mid X(\omega) \in I\}$ ein Ereignis aus $\mathcal{E}$ ist, wobei $\mathcal{E}$ die zugelassene Menge von Ereignissen bezeichnet (vgl. (3.3.9)). Im Rahmen dieses Skripts spielt diese Einschränkung aber keine Rolle.

(4.1.6) Die Wahrscheinlichkeitsfunktion

Wir betrachten weiterhin eine diskrete Zufallsvariable X, welche die endlich oder abzählbar unendlich vielen Werte x_i mit der Wahrscheinlichkeit p_i annimmt. Ihre Verteilung ist wie in (4.1.5) durch die folgende Tabelle gegeben

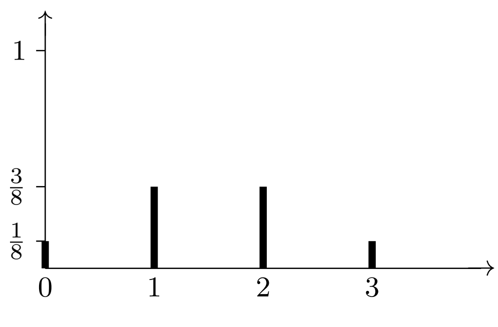
Wert x von X	x_1	x_2	x_3	$\dots$
$P(X = x)$	p_1	p_2	p_3	$\dots$

(mit $\sum p_i = 1$) oder als konkretes Beispiel:

(∗)

x_i	0	1	2	3
$P(X = x_i)$	1/8	3/8	3/8	1/8

Oft ist es zweckmässig, diese Angaben graphisch darzustellen. Dazu trägt man auf der x-Achse den Wert x_i und darüber eine vertikale Strecke der Länge p_i ab. Das folgende konkrete Beispiel enthält die Angaben der Tabelle (∗).



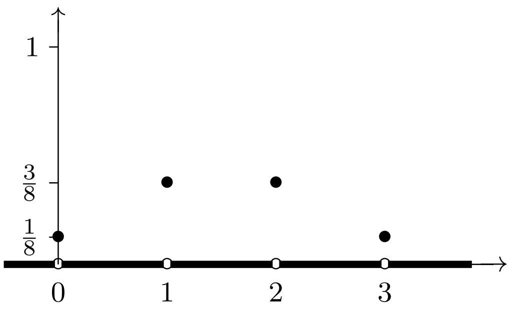
In Übereinstimmung mit (2.2.2.2) nennt man diese Zeichnung ein *Stabdiagramm* (oder auch Histogramm), wobei man statt einfachen Strichen auch gerne schmale Balken einzeichnet, vgl. (2.2.2.2)). Beachten Sie, dass bei den Stabdiagrammen der beschreibenden Statistik absolute oder relative Häufigkeiten, hier aber Wahrscheinlichkeiten abgetragen werden.

Die Verteilung kann auch mittels der so genannten Wahrscheinlichkeitsfunktion beschrieben werden.

Unter der *Wahrscheinlichkeitsfunktion* der diskreten Zufallsgrösse X versteht man die Funktion

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} p_i & x = x_i, \quad i = 1, 2, \ldots \\ 0 & \text{sonst} \end{cases} .$$

Der Graph dieser Funktion entspricht im wesentlichen dem Stabdiagramm, abgesehen davon, dass bei einem Funktionsgraphen nur die Endpunkte der vertikalen Strecken eingetragen werden.



Wir wollen noch eine zweite Formel für $f(x)$ angeben. Dazu beachten wir, dass einerseits für jeden von der Zufallsgrösse angenommenen Wert x_i die Beziehung $p_i = P(X = x_i)$ gilt und dass anderseits für die übrigen x (also für die $x \neq x_i$) die Wahrscheinlichkeit $P(X = x) = 0$ ist. Daraus ergibt sich die kurze Formel

$$f(x) = P(X = x) .$$

Beispiel 4.1.6.A

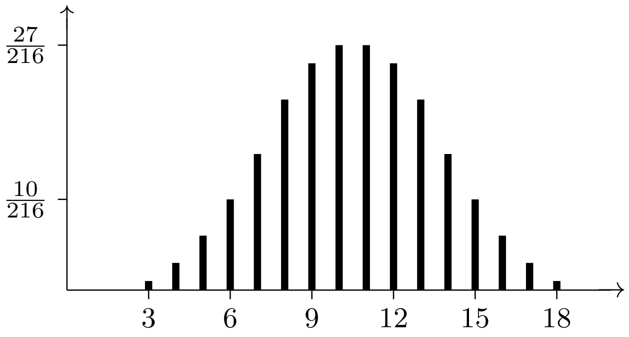
Wir würfeln mit drei Würfeln. Die Zufallsgrösse X bezeichne die Summe der drei Augenzahlen. Sie ist diskret, denn sie kann nur die 16 Werte 3, 4, …, 18 annehmen. Die zugehörigen Wahrscheinlichkeiten bestimmt man durch (etwas langweiliges) Auszählen:

Im Ganzen gibt es $6^3 = 216$ Möglichkeiten. Die Augenzahl 3 kann nur auf eine Art erzielt werden: $3 = 1+1+1$, daher ist $P(X = 3) = 1/216$. Für die Augenzahl 4 gibt es bereits 3 Möglichkeiten: $4 = 2+1+1 = 1+2+1 = 1+1+2$, also $P(X = 4) = 3/216$ usw.

Man kommt so schliesslich auf die nachstehende Tabelle:

$$\begin{array}{lllll} P(X=3) = P(X=18) = 1/216 & \qquad & P(X=7) = P(X=14) = 15/216 \\ P(X=4) = P(X=17) = 3/216 & & P(X=8) = P(X=13) = 21/216 \\ P(X=5) = P(X=16) = 6/216 & & P(X=9) = P(X=12) = 25/216 \\ P(X=6) = P(X=15) = 10/216 & & P(X=10) = P(X=11) = 27/216 \end{array}$$

Das Stabdiagramm sieht wie folgt aus (die Symmetrie ergibt sich direkt aus der Problemstellung):



⊠

(4.1.7) Die Verteilungsfunktion

Wir haben vorhin gesehen, dass die Wahrscheinlichkeitsfunktion einer Zufallsgrösse dem Stabdiagramm der beschreibenden Statistik entspricht. Wir haben in diesem Zusammenhang in (2.2.2.6) noch eine weitere Art zur Darstellung einer Verteilung angetroffen, nämlich die so genannte Summenhäufigkeitsverteilung. Auch diese hat hier ein Gegenstück, nämlich die *Verteilungsfunktion*.

Wie Sie sich erinnern werden, erhält man die Summenhäufigkeitsverteilung, indem man abzählt, wieviele Messwerte höchstens gleich einer gegebenen Zahl sind, und die entsprechende relative oder absolute Häufigkeit aufträgt.

Genau dasselbe tun wir nun auch im Fall einer Zufallsgrösse X, nur steht anstelle der Häufigkeit die Wahrscheinlichkeit. Wir interessieren uns also für die folgende Frage:

Wie gross ist die Wahrscheinlichkeit dafür, dass die Zufallsgrösse X einen Wert $\leq x$ annimmt, wobei x eine gegebene Zahl ist? Diese Wahrscheinlichkeit bezeichnen wir gemäss (4.1.5) mit

$$P(X \leq x) \; .$$

Eine ganz einfache Illustration ergibt sich am Beispiel der Tabelle (∗) von (4.1.6):

(∗)

x_i	0	1	2	3
$P(X = x_i)$	1/8	3/8	3/8	1/8

Wir überlegen uns nun, wie gross $P(X \leq x)$ für verschiedene x sein wird.

1) $x = 2$. Das Ereignis $X \leq 2$ tritt dann ein, wenn $X = 0$, 1 oder 2 ist. Die Wahrscheinlichkeit dafür ist $1/8 + 3/8 + 3/8 = 7/8$, d.h., es ist
$$P(X \leq 2) = 7/8 \; .$$
2) $x = 3$. Das Ereignis $X \leq 3$ tritt dann ein, wenn $X = 0$, 1, 2 oder 3 ist. Man findet so
$$P(X \leq 3) = 1/8 + 3/8 + 3/8 + 1/8 = 1 \; .$$
Dies ist nicht verwunderlich, denn das Ereignis $X \leq 3$ tritt *sicher* ein.
3) $x = -1$. Hier ist
$$P(X \leq -1) = 0 \; ,$$
denn dieses Ereignis ist *unmöglich*.
4) Beachten Sie, dass $P(X \leq x)$ für *alle* Werte von x berechnet werden kann und nicht nur für jene x, die in der Tabelle selbst vorkommen. So ist z.B.
$$\begin{aligned} P(X \leq 2.5) &= P(X = 0) + P(X = 1) + P(X = 2) = 7/8, \\ P(X \leq \sqrt{2}) &= P(X = 0) + P(X = 1) = 1/2, \quad \text{usw.} \end{aligned}$$

Die Wahrscheinlichkeit $P(X \leq x)$ ist also für jedes $x \in \mathbb{R}$ definiert und kann somit als Funktion
$$F : \mathbb{R} \to \mathbb{R}, \quad x \mapsto F(x) = P(X \leq x)$$
aufgefasst werden. Es ist klar, dass diese Funktion von der Zufallsgrösse X abhängig sein wird. Man nennt sie die Verteilungsfunktion von X.

Es ist wichtig zu beachten, dass die Formel $F(x) = P(X \leq x)$ für ganz beliebige Zufallsgrössen sinnvoll ist, obwohl unser obiges Beispiel einfachheitshalber eine diskrete Zufallsgrösse betraf.

Wir wiederholen die formale *Definition*:

> Wenn X eine beliebige Zufallsgrösse ist, dann heisst die durch
> $$F(x) = P(X \leq x)$$
> gegebene Funktion F die *Verteilungsfunktion* von X.

Obwohl, wie eben erwähnt, diese Definition für alle Zufallsgrössen gilt, bleiben wir für den Moment bei den diskreten Zufallsgrössen. Wir wollen für diesen Fall noch eine konkrete Formel für $F(x)$ angeben. Um $F(x)$ zu bestimmen, nimmt man genau wie im obigen Zahlenbeispiel alle von der Zufallsgrösse angenommenen Werte x_i, welche kleiner oder gleich x sind und addiert die zugehörigen Wahrscheinlichkeiten. In leicht verständlicher Summenschreibweise lässt sich dies so ausdrücken:

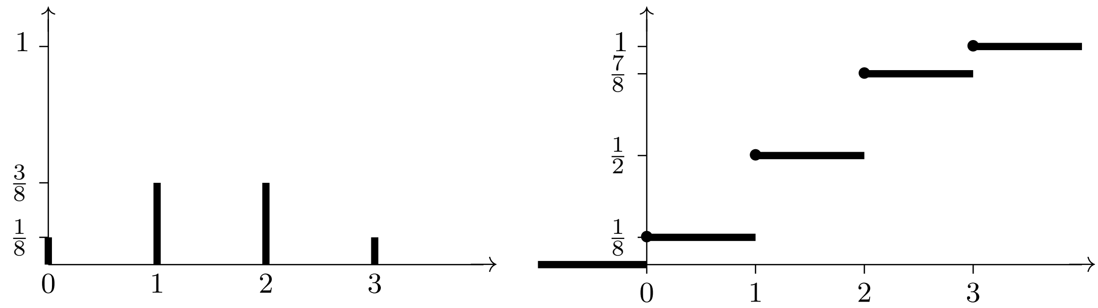
$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p_i \ .$$

Im Fall, wo die Zufallsgrösse abzählbar unendlich viele Werte annimmt, kann diese Summe auch unendlich, also eine Reihe, sein.

Eine weitere Variante erhält man, wenn man noch die Wahrscheinlichkeitsfunktion $f(x)$ von (4.1.5) mit ins Spiel bringt. Wegen $p_i = f(x_i)$ ist dann

$$F(x) = \sum_{x_i \leq x} f(x_i) \ .$$

Zum Schluss zeichnen wir noch den Graphen der Verteilungsfunktion $F(x)$ für den Fall einer diskreten Zufallsgrösse. Als konkretes Beispiel muss noch einmal die Tabelle $(*)$ von (4.1.6) herhalten.



Erläuterung

Für $x < 0$ ist $F(x) = P(X \leq x) = 0$, denn unsere Zufallsgrösse nimmt ja gar keine negativen Werte an. Erst an der Stelle $x = 0$ "passiert etwas", denn es ist $P(X = 0) = 1/8$. Somit springt $F(x)$ an der Stelle 0 auf den Wert 1/8 und bleibt dann konstant bis zur Stelle $x = 1$. Dort folgt erneut ein Sprung, denn $F(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = 1/8 + 3/8 = 1/2$. Der Wert 1/2 ist also sozusagen die bis zur Stelle 1 "angesammelte" oder "kumulierte" Wahrscheinlichkeit. Der Wert 7/8 im Intervall [2,3) erklärt sich genau gleich. Von $x = 3$ an haben wir dann die Wahrscheinlichkeit 1 für das sichere Ereignis angesammelt. (In (2.2.2.6) und (2.2.2.7) haben wir dasselbe für absolute und relative Häufigkeiten durchgeführt.)

Wie Sie sehen, bestimmen sich Wahrscheinlichkeitsfunktion und Verteilungsfunktion für diskrete Zufallsgrössen gegenseitig: Wenn man die eine kennt, kann man die andere berechnen. Die Verteilung einer diskreten Zufallsgrösse X kann also wahlweise durch

- die Tabelle,
- das Stabdiagramm,
- die Wahrscheinlichkeitsfunktion,
- die Verteilungsfunktion

gegeben werden.

(4.1.8) Allgemeines zur Verteilungsfunktion

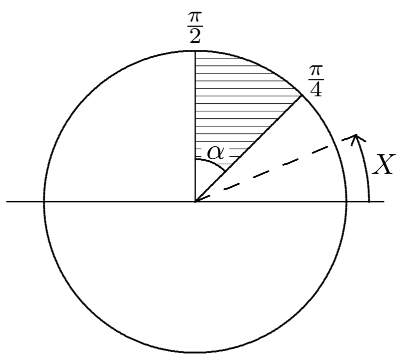
In diesem Abschnitt soll — auch im Hinblick auf Teilkapitel 4.3 — dargelegt werden, dass die Verteilungsfunktion auch für nicht diskrete Zufallsgrössen sinnvoll und nützlich ist. Dagegen werden wir gleich sehen, dass wir die Wahrscheinlichkeitsfunktion nur bei diskreten Zufallsgrössen verwenden können.

Im Fall einer diskreten Zufallsvariablen X haben wir in (4.1.7) die Wahrscheinlichkeitsfunktion mit der Formel

$$f(x) = P(X = x)$$

beschrieben.

Es gibt nun Fälle von (nicht diskreten) Zufallsgrössen, wo diese Definition nichts einbringt. Betrachten wir etwa ein Glücksrad. Die Zufallsgrösse X (mit $0 \leq X < 2\pi$) sei der Winkel (in der angegebenen Weise gemessen), bei welchem der Zeiger stehen bleibt. Der hier verwendete Wahrscheinlichkeitsbegriff ist geometrisch. Die Wahrscheinlichkeit dafür, dass der Zeiger in einem gegebenen Sektor vom Öffnungswinkel α stehen bleibt, ist proportional zu α, und zwar ist sie (für $0 \leq \alpha < 2\pi$) gleich $\alpha/2\pi$. Konkret etwa



$$P(\frac{\pi}{4} \leq X \leq \frac{\pi}{2}) = \frac{\pi/4}{2\pi} = \frac{1}{8} \ .$$

Entsprechend ist die Wahrscheinlichkeit dafür, dass der Zeiger *genau* beim Winkel x anhält, gleich Null:

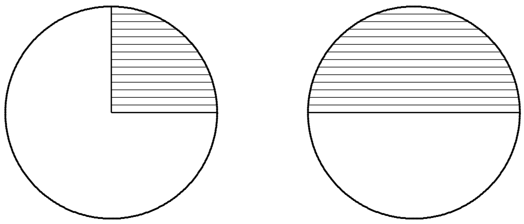
$$P(X = x) = 0 \quad \text{für alle } x,\ 0 \leq x < 2\pi \ ,$$

und dies, obwohl das Elementarereignis $X = x$ nicht unmöglich ist. In (3.3.3), vgl. den dortigen Hinweis, haben wir übrigens eine ähnliche Erscheinung angetroffen.

Die "Moral" ist die, dass im nicht diskreten Fall die Wahrscheinlichkeitsfunktion $f(x) = P(X = x)$ nicht brauchbar ist, da sie stets den Wert 0 annimmt. (An ihre Stelle wird in (4.3.3) die "Dichtefunktion" treten.)

Dagegen ist die Verteilungsfunktion $F(x) = P(X \leq x)$ durchaus sinnvoll. So ist beispielsweise

$$P(X \leq \tfrac{\pi}{2}) = \tfrac{1}{4}, \qquad P(X \leq \pi) = \tfrac{1}{2}.$$



Dieses Beispiel soll belegen, dass die Verteilungsfunktion für ganz beliebige (nicht nur für diskrete) Zufallsvariable Sinn macht. So betrachtet ist sie also bedeutsamer als die Wahrscheinlichkeitsfunktion. Aus diesem Grund wurde in (4.1.7) die Definition der Verteilungsfunktion für eine beliebige Zufallsgrösse angeschrieben. In Teilkapitel 4.3 und Kapitel 5 wird dann die Verteilungsfunktion eine wichtige Rolle spielen.

(4.1.9) Anhang: Einige diskrete Verteilungen

In der Mathematik kommen gewisse Funktionen besonders häufig vor oder sind speziell interessant und haben deshalb eigene Namen, wie etwa die Exponentialfunktion, der Sinus usw. Ganz ähnlich ist es mit den diskreten Verteilungen, wo wir bis jetzt einige allgemeine Beispiele gesehen haben. Manche Verteilungen sind aber besonders wichtig und erfreuen sich deshalb ebenfalls eines eigenen Namens. Dazu gehören die *Binomialverteilung* und die *Poisson-Verteilung*, denen eigene Teilkapitel (4.2) und (7.3) gewidmet sind.

Hier wollen wir als Ergänzung drei andere Verteilungen, denen wir schon mehrmals begegnet sind, bei ihrem in der Literatur gebräuchlichen Namen nennen. Zudem geben wir gleich auch Erwartungswert und Varianz an, obschon diese Konzepte erst in Kapitel 5 zur Verfügung stehen - in (5.4) repetieren wir diese Punkte dann nochmals. Die Anschauung von Kapitel 2 mag dazu erstmal genügen. Vorab doch noch die Definitionen:

$$E(X) := \sum_i x_i P(X = x_i)$$

und

$$V(X) := \sum_i (x_i - E(X))^2 P(X = x_i) \ .$$

a) <u>Die diskrete Gleichverteilung</u>

Man sagt, dass die diskrete Zufallsgrösse X einer *diskreten Gleichverteilung* folgt, wenn sie alle vorkommenden Werte mit derselben Wahrscheinlichkeit annimmt:

x_i	x_1	x_2	$\ldots$	x_n
$P(X = x_i)$	$\frac{1}{n}$	$\frac{1}{n}$	$\ldots$	$\frac{1}{n}$

Beispiele dazu sind der Münzenwurf ($n = 2$), das Würfelspiel ($n = 6$) oder das Roulette ($n = 37$); Ehrlichkeit der betreffendenen Spielgeräte vorausgesetzt.

Der Erwartungswert bzw. die Varianz der Gleichverteilung sind gegeben durch

$$E(X) = \frac{1}{n}(x_1 + x_2 + \ldots + x_n), \qquad V(X) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \Big(\frac{1}{n}\sum_{i=1}^{n} x_i\Big)^2 .$$

Die erste Formel ist offensichtlich, die zweite folgt au (2.2.3.7 b).

b) <u>Die geometrische Verteilung</u>

Ein Beispiel zu dieser Verteilung haben wir in 3.5.6.D kennen gelernt. Allgemein heisst eine Zufallsgrösse X *geometrisch verteilt* (mit dem Parameter p ($0 < p < 1$)), wenn sie die (abzählbar vielen) Werte 1, 2, 3, ... annimmt und wenn gilt

$$P(X = k) = p \cdot (1-p)^{k-1}, \quad k = 1,\ 2,\ 3, \ldots$$

In 3.5.6.D (Würfeln, bis eine Fünf erscheint) liegt eine solche Verteilung mit dem Parameter $p = 1/6$ vor. Dort wird auch ein Zusammenhang zur geometrischen Reihe hergestellt, welcher den Namen erklärt.

Man kann zeigen, dass für die geometrische Verteilung mit dem Parameter p die Beziehungen

$$E(X) = \frac{1}{p}, \quad V(X) = \frac{1-p}{p^2}$$

gelten.

c) <u>Die hypergeometrische Verteilung</u>

Diese Verteilung mit ihrem imposanten Namen* haben wir bereits mehrfach angetroffen, nämlich in den folgenden Fällen:

- Beispiel 3.4.4.A (Zahlenlotto),
- Beispiel 4.1.5.B (Zucker-Eili),
- Beispiel 4.1.9.D (Anzahl Under beim Jassen).

In diesen Beispielen geht es jedes Mal um folgendes: Man hat N Objekte (45 Lottozahlen, 10 Zucker-Eili, 36 Jasskarten), wovon M “gut” oder “interessant” sind (6 Gewinnzahlen, 3 weisse Eili, 4 Under). Aus den gesamten N Objekten werden nun n ausgewählt (6 angekreuzte Zahlen, 4 Eili, 9 Karten eines “Blattes”). Wie gross ist jetzt die Wahrscheinlichkeit $P(X = k)$ dafür, dass unter den n ausgewählten Objekten genau k “gute” sind, nämlich k Gewinnzahlen unter den 6 angekreuzten ($k = 0, \ldots, 6$), k weisse Eili unter den 4 ausgewählten ($k = 0, \ldots, 3$), oder k Under unter den 9 Karten

---

* Dieser Name erklärt sich dadurch, dass es in der Mathematik den Begriff der “hypergeometrischen Reihe” gibt, der etwas mit dieser Verteilung zu tun hat.

($k = 0, \ldots, 4$)? Allgemein gilt $\max(0, n + M - N) \leq k \leq \min(M, n)$ (überlegen Sie sich selber mit obigen drei Beispielen, dass die Anzahl der “Guten” nicht beliebig sein kann).

Die Überlegung ist genau dieselbe wie beim Lotto (Beispiel 3.4.4.A, beachten Sie aber auch die beiden andern Beispiele): Wir haben dort den Fall $N = 45,\ M = 6,\ n = 6,\ k = 4$ ausführlich behandelt und gefunden, dass die Wahrscheinlichkeit für einen Vierer

$$\frac{\binom{6}{4}\binom{39}{2}}{\binom{45}{6}}$$

beträgt. Ganz analog erhält man mit den allgemeinen Grössen die Formel

$$P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \ .$$

Die so definierte Verteilung der Zufallsgrösse X nennt man die *hypergeometrische Verteilung*. Die beiden andern Beispiele könnten ebenfalls mit dieser Formel behandelt werden.

Wir erwähnen noch ohne Beweis die Formeln für Erwartungswert und Varianz der hypergeometrischen Verteilung. Mit der Abkürzung

$$p = \frac{M}{N} \quad \text{ist} \quad E(X) = np, \quad V(X) = np(1-p)\frac{N-n}{N-1} \ .$$

Im Beispiel der Jasskarten 4.1.9.D haben wir $E(X)$ berechnet. Dort ist $M = 4,\ N = 36,\ n = 9$, und unsere Formel liefert $E(X) = 9 \cdot \frac{4}{36} = 1$, wie früher schon angegeben.

In (4.2.4) werden wir die hypergeometrische Verteilung und die Binomialverteilung vergleichen.

4.2. DIE BINOMIALVERTEILUNG

(4.2.1) Überblick

Ein wichtiges konkretes Beispiel einer diskreten Verteilung ist die *Binomialverteilung*. Sie tritt in der folgenden Situation auf: *(4.2.2)*

Ein Experiment mit nur zwei möglichen Ausgängen (Erfolg oder Misserfolg) wird n-mal durchgeführt, wobei die einzelnen Versuche voneinander unabhängig sind. Die Wahrscheinlichkeit eines Erfolgs bei einem einzelnen Versuch sei gleich p, ferner setzt man $q = 1 - p$.

Die Zufallsgrösse X = "Anzahl der Erfolge" ist dann binomial verteilt (mit den Parametern n und p), d.h., X nimmt die Werte

$$0,\ 1,\ 2, \ldots, n$$

mit der Wahrscheinlichkeit

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \quad (k = 0,\ 1, \ldots, n)$$

an.

(4.2.2) Wie kommt man auf die Binomialverteilung?

Bis jetzt haben wir allgemein über diskrete Verteilungen gesprochen. Zu jedem Wert x_i der Zufallsgrösse X gehörte eine Wahrscheinlichkeit $P(X = x_i)$, über die wir aber — abgesehen von ein paar Beispielen — wenig Konkretes sagten. Es gibt nun aber Verteilungen, die in der Praxis immer wieder auftreten und die besondere Namen haben. Dabei sind dann die möglichen Werte $x_1, x_2, \ldots$ von X von vornherein festgelegt, und es gibt Formeln für die Wahrscheinlichkeiten $P(X = x_i)$.

Eines der wichtigsten Beispiele hierzu ist die *Binomialverteilung*, die wir nun ausführlich behandeln wollen.

Sie können sich das Vorgehen wie etwa bei der Besprechung von Funktionen vorstellen. Zuerst geht es um allgemeine Dinge, wie etwa graphische Darstellung oder Differenzierbarkeit, anschliessend wendet man sich speziellen Funktionen zu, z.B. der Quadrat- oder der Exponentialfunktion.

Um die Binomialverteilung einzuführen, betrachten wir ein "Experiment" (im allgemeinen Sinn von (3.1.2)), das nur zwei Ergebnisse zulässt. Ein solches Experiment heisst manchmal ein "Bernoulli-Experiment". Konkrete Beispiele von solchen Paaren von Ergebnissen sind etwa

1	0
Kopf	Zahl
gesund	krank
weiblich	männlich
eine “Sechs”	keine “Sechs”
E	$\bar{E}$

wobei E im letzten Fall irgendein Ereignis ist.

Wir werden meist von *Erfolg* bzw. *Misserfolg* sprechen und zur Abkürzung die Zeichen 1 bzw. 0 gebrauchen.

Die Wahrscheinlichkeit für “Erfolg” bzw. “1” bezeichnen wir stets mit p, jene für “Misserfolg” bzw. “0” mit q. Natürlich ist $q = 1 - p$. Es wird aber nicht etwa vorausgesetzt, dass $p = q$ sei.

Nun wiederholen wir unsern Versuch (das “Einzelexperiment”) n-mal hintereinander, wobei die einzelnen Wiederholungen *unabhängig* voneinander sein sollen, d.h., das Ergebnis eines Einzelversuchs hängt nicht davon ab, wie die vorangegangenen Versuche ausgegangen sind. Diese Voraussetzung über die Unabhängigkeit ist wichtig. Ob man sie in einem konkreten Beispiel als gegeben annehmen darf, muss die Person entscheiden, die den Versuch durchführt und die Wahrscheinlichkeitsrechnung anwenden will.

Wir interessieren uns nun für die Anzahl Erfolge bei diesen n Wiederholungen. Diese Anzahl kann als Wert einer Zufallsgrösse X aufgefasst werden. Zwei einfache Beispiele hierzu:

- Ich würfle zehnmal hintereinander. Das Einzelexperiment ist hier ein einzelner Wurf. Eine “Sechs” rechne ich als Erfolg, alles andere als Misserfolg. Die Zufallsgrösse X ist dann gegeben durch
  $$X = \text{Anzahl “Sechsen” in zehn Würfen}$$
  und kann die ganzzahligen Werte 0 bis 10 annehmen.
- Ich untersuche eine Klasse von 20 Schülern auf Linkshänder. Das Einzelexperiment ist hier die Untersuchung eines einzelnen Schülers. Wenn ich “Linkshändigkeit” (willkürlich) als Erfolg rechne, dann ist
  $$X = \text{Anzahl Linkshänder}$$
  und X kann ganzzahlige Werte zwischen 0 und 20 annehmen.

Es dürfte klar sein, dass uns jetzt vor allem die Wahrscheinlichkeiten interessieren, mit denen die Werte unserer Zufallsgrösse angenommen werden. Um auf eine entsprechende Formel zu kommen, abstrahieren wir insofern etwas, als wir “Erfolg” mit “1”, “Misserfolg” mit “0” abkürzen und von der speziellen Bedeutung vorläufig absehen wollen.

Die möglichen Ergebnisse eines "Gesamtversuchs", bestehend aus n unabhängigen Wiederholungen des Einzelversuchs, sind dann alle möglichen Folgen von total n Einsen und Nullen; für $n = 5$ erhält man etwa

$$00000,\ 00001,\ 00010,\ \ldots, 11111\ .$$

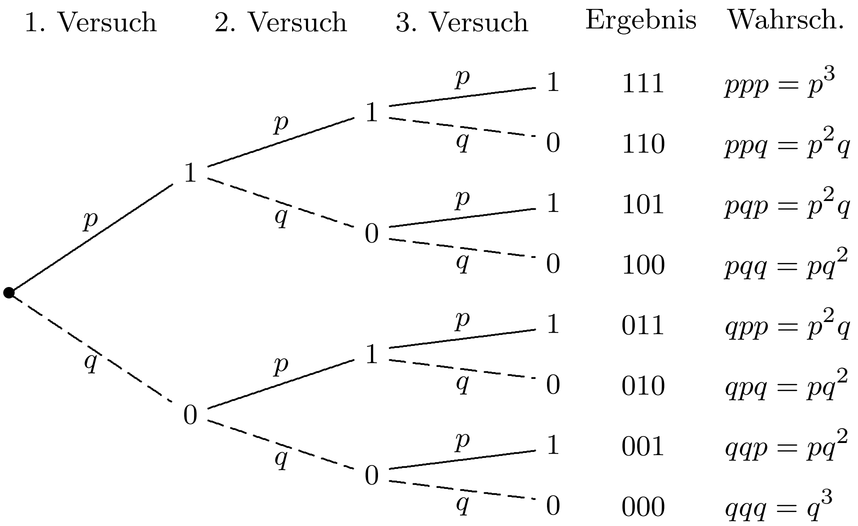
Die erste Zifferngruppe stellt also eine Reihe von fünf Misserfolgen hintereinander dar, bei der zweiten tritt nach vier Misserfolgen endlich ein Erfolg ein usw.

Da jede der n Stellen der Folge entweder 0 oder 1 sein kann, gibt es total 2^n mögliche Folgen, dies in Übereinstimmung mit den Regeln der Kombinatorik (1.6).

Beachten Sie noch, dass uns weniger die einzelnen Folgen interessieren als vielmehr die Anzahl der Erfolge. Im obigen Beispiel gibt es fünf Möglichkeiten dafür, dass genau ein Erfolg zustande kommt, nämlich:

$$00001,\ 00010,\ 00100,\ 01000,\ 10000\ .$$

Nun wollen wir aber endlich die uns interessierenden Wahrscheinlichkeiten bestimmen. Die Wahrscheinlichkeit des Auftretens einer bestimmten Folge lässt sich am besten mit einem Baumdiagramm (vgl. (3.5.6)) berechnen. Wir zeichnen ein Beispiel für den Fall $n = 3$; das Einzelexperiment wird hier also dreimal wiederholt.



Erläuterung

Wir haben ausdrücklich vorausgesetzt, dass die einzelnen Wiederholungen der Versuche unabhängig voneinander sind. Die Wahrscheinlichkeit eines Erfolgs ("1") ist also jedes Mal $= p$, jene eines Misserfolgs ("0") ist $= q$, wie das auch in der Figur angeschrieben wurde. Durch Multiplikation der Wahrscheinlichkeiten entlang jedes Zweigs des Baumes kommt man dann auf die rechts angeführten Wahrscheinlichkeiten der einzelnen Ergebnisse. Beispielsweise haben die Folgen 100, 010 und 001 je die Wahrscheinlichkeit

pq^2. Das uns mehr interessierende Ereignis E “ein Erfolg in drei Versuchen” setzt sich aus diesen drei Ergebnissen (Elementarereignissen, (3.2.5)) zusammen

$$E = \{100,\ 010,\ 001\}$$

und hat infolgedessen die Wahrscheinlichkeit $3pq^2$. Für die Zufallsgrösse

$$X = \text{Anzahl Erfolge in drei Wiederholungen}$$

gilt daher

$$P(X = 1) = 3pq^2 \ .$$

Für beliebiges n geht die Überlegung ganz analog: Eine Folge der Länge n mit k Einsen hat $n - k$ Nullen, und die Wahrscheinlichkeit ihres Auftretens beträgt

$$(*) \qquad p^k q^{n-k} \ .$$

Unsere Zufallsgrösse X bedeutet nun

“Anzahl Erfolge in n Versuchen”

und kann die Werte $k = 0,1,2,\ldots,n$ annehmen. Wir wollen jetzt $P(X = k)$ berechnen.

$k = 0$: Es gibt nur eine Folge mit 0 “Einsen”, nämlich $000\ldots0$, und die zugehörige Wahrscheinlichkeit ist $P(X = 0) = q^n$.

$k = 1$: Es gibt n verschiedene Folgen, in welchen genau einmal eine “Eins” auftritt: $1000\ldots0$, $0100\ldots0$, $0010\ldots0$, $\ldots$,$0000\ldots1$.
Jede davon hat die Wahrscheinlichkeit pq^{n-1}. Die Wahrscheinlichkeit dafür, dass $X = 1$ ist, ist daher n-mal grösser: $P(X = 1) = npq^{n-1}$.

$k = 2$: Nach den Regeln der Kombinatorik (siehe (1.5), wo auch die Binomialkoeffizienten $\binom{n}{k}$ erläutert sind) gibt es $\binom{n}{2}$ verschiedene Folgen mit 2 “Einsen” (und $n - 2$ “Nullen”), denn man kann die zwei Plätze für die “Eins” auf $\binom{n}{2}$ Arten auswählen. Jede davon hat die Wahrscheinlichkeit p^2q^{n-2}, und wir erhalten $P(X = 2) = \binom{n}{2}p^2q^{n-2}$.

k: Nun sollte der allgemeine Fall klar sein: Es gibt $\binom{n}{k}$ verschiedene Folgen mit genau k “Einsen”, und jede davon hat die Wahrscheinlichkeit p^kq^{n-k}. Wir erhalten somit wegen $(*)$

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \ .$$

Damit ist nun unser weiter oben formuliertes Problem gelöst: Wenn das Einzelexperiment die Erfolgswahrscheinlichkeit p hat, dann ist die Wahrscheinlichkeit dafür, dass in n unabhängigen Wiederholungen des Versuchs genau k Erfolge auftreten, gegeben durch

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, 2, \ldots, n \ .$$

Von unserer Zufallsgrösse X kennen wir also die Werte, die sie annimmt $(0, 1, \ldots, n)$ sowie die zugehörigen Wahrscheinlichkeiten (nämlich $\binom{n}{k}p^k q^{n-k}$). Mit andern Worten: Wir kennen die Verteilung von X (4.1.5).

Diese Verteilung nennt man nun die *Binomialverteilung* (nicht "Binomi**n**alverteilung"). Sie hängt von den Grössen n und p ab, die man auch die *Parameter* der Verteilung nennt. (Der Wert q ist durch $q = 1 - p$ festgelegt.) Für jedes $n \in \mathbb{N}$ und jedes p mit $0 \leq p \leq 1$ erhalten wir eine andere Binomialverteilung.

Zur Erleichterung der Rechenarbeit gibt es Tabellen, in denen die Werte von $P(X = k) = \binom{n}{k}p^k q^{n-k}$ in Abhängigkeit von n, p und k aufgeführt sind. Ein nicht sehr umfangreiches Exemplar finden Sie in (6.1.2).

Der Name "Binomialverteilung" rührt einfach davon her, dass die Wahrscheinlichkeiten $P(X = k)$ gerade die einzelnen Summanden sind, wenn man die binomische Formel (siehe (26.4) im ersten Band) auf $(p+q)^n$ anwendet:

$$(p+q)^n = \binom{n}{0}p^0 q^n + \binom{n}{1}p^1 q^{n-1} + \ldots + \binom{n}{k}p^k q^{n-k} + \ldots + \binom{n}{n}p^n q^0 \; .$$

Wegen $p + q = 1$ ergibt sich daraus sofort

$$\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = 1 \; ,$$

wie es nach (4.1.5) auch sein muss.

Zur weitern Illustration geben wir auf der nächsten Seite die Stabdiagramme der Binomialverteilung für die Fälle $n = 6, p = 0.1$, 0.2, 0.3, 0.4, 0.5 und 0.6 an:

Bemerkungen

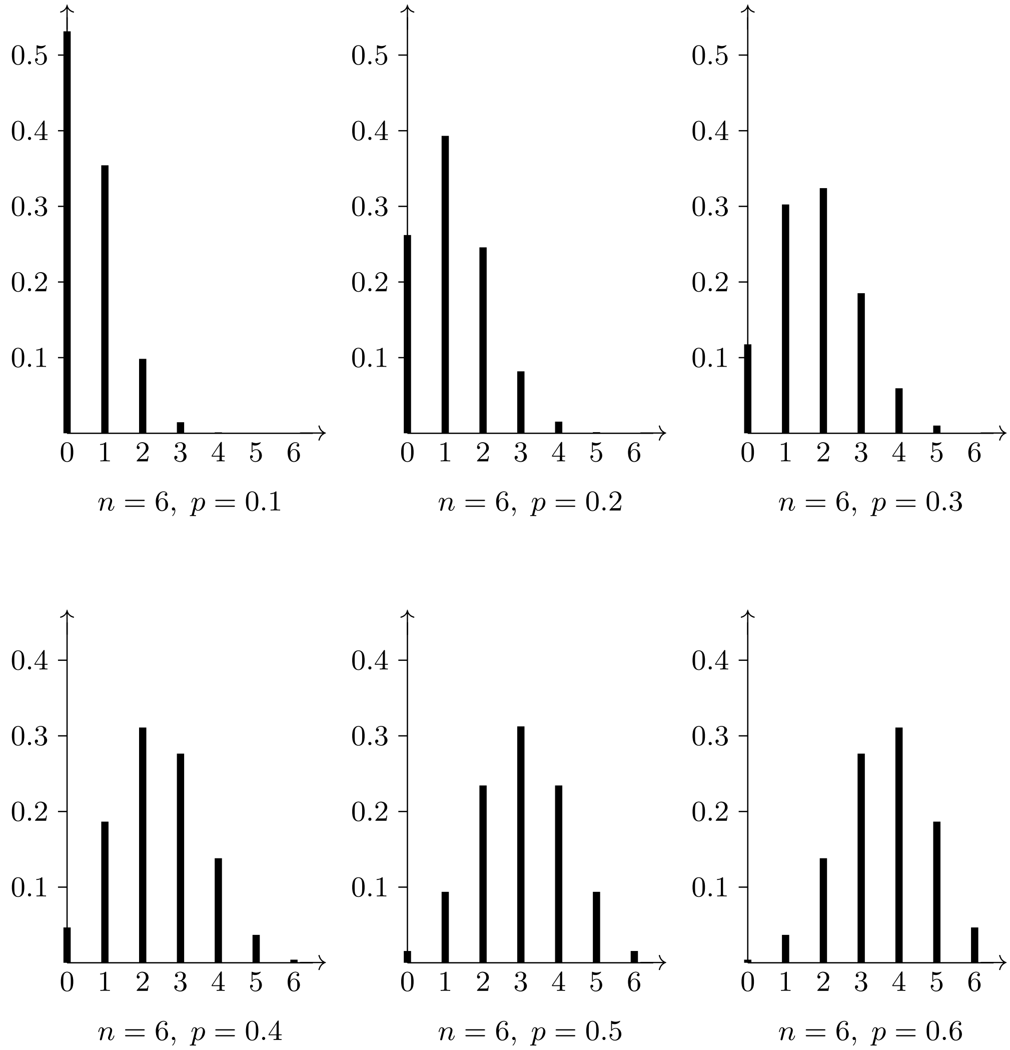
1) Im Fall $p = q = \frac{1}{2}$ wird die Formel etwas einfacher. Man erhält

$$P(X = k) = \binom{n}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k} = \binom{n}{k}\left(\frac{1}{2}\right)^n \; .$$

Wegen $\binom{n}{k} = \binom{n}{n-k}$ (siehe 1.5)) ist dann auch

$$P(X = n - k) = \binom{n}{k}\left(\frac{1}{2}\right)^n \; .$$

Dies bedeutet, dass für $p = q = \frac{1}{2}$ das Stabdiagramm symmetrisch ist. Dies überrascht nicht, denn in diesem Fall ist die Wahrscheinlichkeit von k Erfolgen genau dieselbe, wie die von k Misserfolgen (d.h., $n - k$ Erfolgen).



2) Die Histogramme für $p = 0.4$ und $p = 0.6$ gehen, wie man sehen kann, durch Spiegelung auseinander hervor. Es ist nämlich wegen $\binom{n}{k} = \binom{n}{n-k}$

$$\binom{n}{k} p^k q^{n-k} = \binom{n}{n-k} q^{n-k} p^k \ ,$$

und die erste Zahl ist die Wahrscheinlichkeit von k Erfolgen bei der Einzelerfolgswahrscheinlichkeit p, die zweite aber die Wahrscheinlichkeit von $n-k$ Erfolgen bei der Einzelerfolgswahrscheinlichkeit q.

Wir rekapitulieren die bisher behandelten Punkte.

<u>Zusammenfassung</u>

(1) Die diskrete Zufallsgrösse X folgt der Binomialverteilung mit den Parametern n und p ($n \in \mathbb{N}, 0 \leq p \leq 1$), wenn sie die Werte

$$0,\, 1,\, 2, \ldots, n-1,\, n$$

annimmt und wenn für die zugehörigen Wahrscheinlichkeiten gilt

$$P(X=k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0,\, 1,\, \ldots, n\ ,$$

wobei $q = 1 - p$ ist.

(2) In der Praxis tritt diese Binomialverteilung in der folgenden Situation auf: Ein Experiment mit nur zwei möglichen Ergebnissen (Erfolg/Misserfolg) wird n-mal wiederholt, wobei die einzelnen Versuche voneinander unabhängig sein sollen. Die Wahrscheinlichkeit eines Erfolgs bei einem einzelnen Versuch sei gleich p.

Dann folgt die Zufallsgrösse

$$X = \text{Anzahl der Erfolge}$$

der Binomialverteilung mit den Parametern n und p, d.h., es ist

$$P(X=k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0,\, 1,\, \ldots, n\ .$$

Es kommt nicht darauf an, wie die “Einzelwahrscheinlichkeit” p festgelegt wird. Es kann sich dabei z.B. um eine klassische Wahrscheinlichkeit oder um eine idealisierte relative Häufigkeit handeln (vgl. (3.1.4)).

(4.2.3) Beispiele zur Binomialverteilung

<u>Beispiel 4.2.3.A</u>

Ein (unverfälschter) Würfel wird zwölfmal geworfen. Wie gross ist die Wahrscheinlichkeit

a) genau dreimal eine Sechs,
b) höchstens drei Sechsen,
c) mindestens eine Sechs

zu werfen?

Wir übersetzen zunächst die Aufgabenstellung in die Sprache der Wahrscheinlichkeitsrechnung. Jeder einzelne Wurf ist ein Zufallsexperiment, das aber in unserm Beispiel nicht etwa sechs mögliche Ergebnisse hat, sondern nur zwei. Wir interessieren uns nämlich nur dafür, ob eine Sechs herauskommt (Erfolg) oder keine Sechs (Misserfolg).

Die Wahrscheinlichkeiten dieser beiden Ereignisse sind verschieden; die Wahrscheinlichkeit für eine Sechs (also die Erfolgswahrscheinlichkeit) ist $p = 1/6$, die Wahrscheinlichkeit für einen Misserfolg, d.h., für das Gegenereignis "keine Sechs" ist $q = 1 - p = 5/6$. Diese Wahrscheinlichkeiten erhält man aufgrund unserer Annahme, dass der Würfel unverfälscht sei, als klassische Wahrscheinlichkeit.

Es handelt sich hier also um ein Bernoulli-Experiment im Sinne von (4.2.2), d.h., ein Experiment mit nur zwei möglichen Ergebnissen. Dieses Experiment wird nun zwölfmal wiederholt, wobei die einzelnen Wiederholungen sicher voneinander unabhängig sind, denn der Ausgang eines Wurfs hängt nicht von den vorangegangenen Würfen ab.

Die Anzahl der "Erfolge" (also der Sechsen) ist eine Zufallsvariable, welche hier die Werte 0, 1, ..., 12 annehmen kann, da zwölfmal gewürfelt wird. Wie eben erläutert wurde, sind alle Voraussetzungen von Punkt (2) der oben stehenden Zusammenfassung erfüllt. Wir wissen also:

Die Zufallsgrösse X = "Anzahl der Sechsen" folgt einer Binomialverteilung mit den Parametern $n = 12$ und $p = 1/6$. Es gilt deshalb

$$P(X = k) = \binom{12}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{12-k} .$$

Nun lassen sich die drei Teilaufgaben sofort lösen:

a) Hier ist gefragt, mit welcher Wahrscheinlichkeit X den Wert 3 annimmt. Wir setzen daher in der obigen Formel $k = 3$ und finden

$$P(X = 3) = \binom{12}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^9 = 0.1974 .$$

b) Höchstens drei Sechsen treten dann ein, wenn $X = 0$, 1, 2 oder 3 ist. Die gesuchte Wahrscheinlichkeit ist somit gleich

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) .$$

Wendet man die obige Formel für $k = 0$, 1, 2, 3 an und addiert, so erhält man

$$P(X \leq 3) = 0.8748 .$$

Beachten Sie übrigens, dass $P(X \leq 3)$ gerade der Wert der Verteilungsfunktion an der Stelle 3 ist.

c) "Mindestens eine Sechs" bedeutet

$$X = 1,\ X = 2,\ \ldots \text{ oder } X = 12 .$$

Die gesuchte Wahrscheinlichkeit ist also die Summe aller $P(X = k)$, $k = 1, \ldots, 12$. Die Berechnung ist mühsam; viel einfacher ist es, die Gegenwahrscheinlichkeit zu bestimmen:

Das Gegenereignis zu "mindestens eine Sechs" ist "keine Sechs" und hat die Wahrscheinlichkeit

$$P(X=0)=\binom{12}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{12}=\left(\frac{5}{6}\right)^{12}=0.1122\;.$$

Die ursprünglich gesuchte Wahrscheinlichkeit ist dann

$$P(X\geq 1)=1-0.1122=0.8878\;.$$

Der hier gebrauchte "Trick" mit der Gegenwahrscheinlichkeit ist natürlich auch sonst nützlich. ⊠

Die nächsten Beispiele wollen wir etwas knapper behandeln.

<u>Beispiel 4.2.3.B</u>

Wir denken uns zweihundert Familien mit je fünf Kindern. In wievielen von diesen Familien hat es (theoretisch) drei Mädchen und zwei Knaben? Dabei sei die Wahrscheinlichkeit für Knaben 0.513, für Mädchen 0.487 (vgl. Beispiel 3.1.3.C), also eine idealisierte relative Häufigkeit.

Das Zufallsexperiment besteht hier darin, dass man in einer Familie ein Kind herausgreift und feststellt, ob es ein Mädchen ("Erfolg") oder ein Knabe ("Misserfolg") ist (keine Diskriminierung beabsichtigt). Nach Voraussetzung ist $p=0.487$, $q=0.513$. Da es sich um Familien mit fünf Kindern handelt, ist $n=5$ (und nicht etwa $n=200$). Die Zufallsgrösse

$$X=\text{Anzahl der Mädchen}$$

ist also binomial verteilt mit den Parametern $n=5$ und $p=0.487$; dies allerdings nur unter der stillschweigend getroffenen Voraussetzung, dass die einzelnen Geburten voneinander unabhängig sind.

Die gesuchte und übrigens schon in (3.3.8) erwähnte Wahrscheinlichkeit $P(X=3)$ berechnet sich nach der üblichen Formel zu

$$P(X=3)=\binom{5}{3}0.487^3 0.513^2=0.304\;.$$

Bis jetzt hatte die Anzahl (200) der untersuchten Familien überhaupt keine Rolle gespielt. Interpretiert man nun aber die Wahrscheinlichkeit als idealisierte relative Häufigkeit, so werden theoretisch unter den 200 Familien

$$200\cdot 0.304=60.8\;,$$

also rund 61 Familien drei Mädchen und zwei Knaben haben.

Man erwartet natürlich nicht, dass bei einer effektiven Untersuchung von 200 Familien *genau* 61 Familien drei Mädchen haben werden, sondern man wird mit gewissen Abweichungen rechnen. Es ist Aufgabe der beurteilenden Statistik, nachzuprüfen, ob diese Abweichungen "zufällig" oder aber "signifikant" sind. Vgl. hierzu das Kapitel 9 über den χ^2-Test.

Man hätte in diesem Beispiel auch einfachheitshalber mit den Wahrscheinlichkeiten $p = q = 0.5$ arbeiten und die Tabelle (6.1.2) benützen können, wobei dann $P(X = 3) = 0.3125$ herausgekommen wäre. ⊠

Beispiel 4.2.3.C

Tulpenzwiebeln keimen nur zu einem gewissen Prozentsatz $P\%$. Die Firma "Frühlingsfreude" verkauft Packungen zu 12 Stück und möchte garantieren, dass höchstens eine Zwiebel nicht keimt. Wie gross muss P mindestens sein, damit die erwähnte Garantiebedingung mit einer Wahrscheinlichkeit von (mindestens) 95% erfüllt ist? Wir gehen wie folgt vor:

Die Wahrscheinlichkeit dafür, dass eine einzelne Zwiebel keimt, ist hier gleich $p = P/100$. Mit X bezeichnen wir die Anzahl der keimenden Zwiebeln. Diese Zufallsgrösse ist binomial verteilt mit $n = 12$ und unbekanntem p. Die gestellte Bedingung lautet $P(X \geq 11) \geq 0.95$, d.h. (wegen $q = 1 - p$)

$$\binom{12}{11} p^{11}(1-p) + \binom{12}{12} p^{12} = 12p^{11}(1-p) + p^{12} \geq 0.95$$

oder schliesslich, umgeformt, $12p^{11} - 11p^{12} \geq 0.95$. Die Gleichung $12p^{11} - 11p^{12} = 0.95$ kann nicht direkt mit einer Formel gelöst werden. Vielmehr braucht man ein Näherungsverfahren, z.B. das Newton-Verfahren ((21.2) im ersten Band). Man erhält $p \approx 0.9696$ und schliesst, dass P grösser als ungefähr 97% sein muss. ⊠

(4.2.4) Stichproben mit und ohne Zurücklegen

Beim Herausgreifen einer Stichprobe aus einer grösseren Menge von Objekten (z.B. zur Prüfung auf allfällige Ausschussteile) macht es einen Unterschied, ob man die Stichprobe "mit Zurücklegen" oder "ohne Zurücklegen" auswählt. Wir illustrieren das Problem am Beispiel 4.1.5.B. Dort waren in einer Tüte 3 weisse und 7 rote Zuckereier; wir griffen vier Stück heraus und interessierten uns für die Wahrscheinlichkeit dafür, dass man dabei 0, 1, 2 oder 3 weisse Eier erwischt. Dasselbe Resultat könnte man auch erzielen, indem man die vier Eili hintereinander entnimmt, aber *ohne Zurücklegen*, d.h., das gerade herausgenommene Ei wird nicht wieder in die Tüte getan. Dies führte in 4.1.5.B auf die folgenden Wahrscheinlichkeiten:

k	0	1	2	3
$P(W = k)$	1/6	1/2	3/10	1/30

In (4.1.9) haben wir erwähnt, dass hier eine so genannte hypergeometrische Verteilung vorliegt.

Ganz anders ist die Situation, wenn wir die Stichprobe *mit Zurücklegen* auswählen, d.h., wenn wir jedes Mal das gezogene Eili wieder in die Tüte zurücktun. Jetzt ist nämlich die Wahrscheinlichkeit dafür, bei einer einzelnen Auswahl ein weisses Ei zu erwischen, stets gleich $p = 0.3$, und wenn wir nach der Wahrscheinlichkeit fragen, in 4 Versuchen z.B. 2 weisse Eier zu ziehen, dann haben wir eine Binomialverteilung mit $n = 4,\ p = 0.3$ vor uns, daher ist

$$P(X = 2) = \binom{4}{2} 0.3^2 0.7^2 = 0.2646 \ ,$$

während die Wahrscheinlichkeit für zwei weisse Eier im Fall "ohne Zurücklegen" gemäss obiger Tabelle $= 0.3$ ist.

Für grosses n fällt der Unterschied zwischen Stichproben mit und ohne Zurücklegen allerdings nicht mehr stark in Betracht.

4.3. STETIGE ZUFALLSGRÖSSEN

(4.3.1) Überblick

Eine Zufallsgrösse X heisst *stetig*, wenn sich ihre Verteilungsfunktion (4.3.3)

$$F(x) = P(X \leq x)$$

als Integral einer Funktion $f(x)$ schreiben lässt:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt\ .$$

Dabei muss $f : \mathbb{R} \to \mathbb{R}$ folgende Bedingungen erfüllen:

1) f ist (stückweise) stetig,
2) $f(x) \geq 0 \quad$ für alle x,
3) $\int_{-\infty}^{\infty} f(t)\, dt = 1$.

Diese Funktion f heisst die *Dichtefunktion* von X. Dann lässt sich die Wahrscheinlichkeit dafür, dass X einen Wert in einem Intervall $[a, b]$ annimmt, als Integral der Funktion $f(x)$ schreiben: (4.3.3) (4.3.4)

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)\, dx\ .$$

Es ist aber zu beachten, dass $f(x)$ nicht etwa die Wahrscheinlichkeit dafür ist, dass X den Wert x annimmt. Vielmehr ist stets $P(X = x) = 0$. (4.3.4)

Beispiele für stetige Verteilungen sind:

- Stetige Gleichverteilung,
- Normalverteilung,
- t-Verteilung,
- χ^2-Verteilung.

(4.3.2) Einleitung

Wir haben bereits in (4.1.4) den Begriff der stetigen Zufallsgrösse erwähnt, ohne allerdings eine präzise Definition zu geben. Anschaulich gesprochen, beschreibt eine diskrete Zufallsgrösse ein diskretes, eine stetige Zufallsgrösse ein stetiges Merkmal im Sinne von (2.2.1.5).

Entsprechend nannten wir in (4.1.4) eine Zufallsgrösse diskret, wenn sie nur endlich (oder abzählbar unendlich) viele Werte annimmt. Eine stetige Zufallsgrösse dagegen nimmt (ev. auch bloss innerhalb eines bestimmten Intervalls) jeden Wert an. Für die theoretische Behandlung müssen wir allerdings diesen Grundgedanken in anderer Form ausdrücken, was nun erläutert werden soll.

Wir betrachten dazu ein stetiges Merkmal, etwa das Gewicht von Küken. Beim praktischen Wägevorgang werden die Gewichte gerundet (und somit diskretisiert (2.2.1.5)), d.h., es werden Klassen einer gewissen Breite (z.B. 1 Gramm) gebildet. Die relativen Häufigkeiten (absolute Häufigkeiten wären hier unpraktisch) der einzelnen Klassen stellen wir wie gewohnt im Histogramm (Figur 1) dar, wobei der Flächeninhalt der einzelnen Balken ein Mass für relative Häufigkeit ist (vgl. (2.2.2.4)):

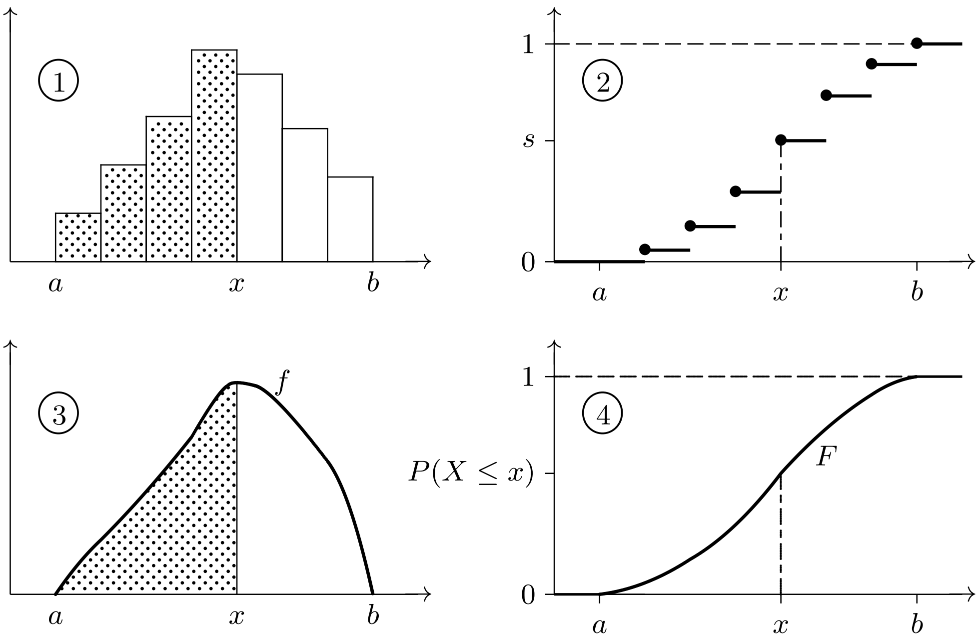
In Figur 1 entspricht der Inhalt des mit Punkten markierten Flächenstücks der relativen Häufigkeit s aller Küken mit Gewicht $\leq x$. Dies ist eine (relative) Summenhäufigkeit, die man auch im Diagramm der Summenhäufigkeitsverteilung (Figur 2) an der Stelle x abtragen kann.

Da wir mit relativen Häufigkeiten arbeiten, hat die Gesamtfläche des Histogramms in Figur 1 den Inhalt 1, entsprechend "steigt" die Summenhäufigkeit bis zum Wert 1.

Nun lassen wir in Gedanken die Klasseneinteilung immer feiner werden (z.B. durch Erhöhung der Messgenauigkeit, wobei wir aber die Anzahl der Messwerte erhöhen müssten*). Die Treppenkurve des Histogramms von Figur 1 geht dabei im Grenzfall in eine glatte Kurve über, die wir als Graph einer gewissen Funktion f auffassen können (Figur 3). Im Zuge dieser Idealisierung ersetzen wir gleichzeitig die relative Häufigkeit durch die Wahrscheinlichkeit. Der Wert des Merkmals (hier des Gewichts) wird dann zum Wert einer Zufallsgrösse X.

Entsprechend der Situation beim Histogramm hat die totale Fläche unter der Kurve $f(x)$ den Inhalt 1. Er ist hier als "Gesamt"wahrscheinlichkeit (besser: Die Wahrscheinlichkeit des sicheren Ereignisses) zu deuten (Figur 3).

* Eine feinere Klasseneinteilung bei gleichbleibender Zahl der Beobachtungsdaten würde ein weniger übersichtliches Histogramm mit "Lücken" liefern, vgl. die Histogramme in (2.2.2.3) und (2.2.2.4).



Aus der relativen Häufigkeit s aller Messwerte $\leq x$ wird bei diesem Übergang die Wahrscheinlichkeit dafür, dass $X \leq x$ ist; mit andern Worten:

Der Inhalt des in Figur 3 markierten Flächenstücks entspricht der Wahrscheinlichkeit

$$P(X \leq x) \; .$$

Nun stellen wir folgendes fest: Einerseits lässt sich dieser Flächeninhalt durch ein Integral beschreiben, nämlich durch

$$\text{(A)} \qquad P(X \leq x) = \int_a^x f(t)\,dt \; ,$$

anderseits ist $P(X \leq x)$ nach (4.1.8) gerade der Wert der Verteilungsfunktion F von X an der Stelle x :

$$\text{(B)} \qquad F(x) = P(X \leq x) \; .$$

Durch Zusammenfassung von (A) und (B) sehen wir, dass im Falle einer — vorläufig im anschaulichen Sinne — stetigen Zufallsgrösse X die Verteilungsfunktion $F(x)$ von X als Integral einer gewissen Funktion $f(x)$ berechnet werden kann:

$$\text{(C)} \qquad F(x) = \int_a^x f(t)\,dt \; .$$

Genau diese Eigenschaft werden wir in (4.3.3) zur allgemeinen Charakterisierung der stetigen Verteilungen gebrauchen.

Zunächst wollen wir uns aber noch überlegen, was bei der besprochenen Idealisierung aus der Summenhäufigkeitsverteilung wird: Die “Treppe” wird zu einer glatten Kurve, und an der Stelle x tragen wir nicht mehr die relative Summenhäufigkeit s ab, sondern die Wahrscheinlichkeit $P(X \leq x)$, die aber nach (B) gerade gleich $F(x)$ ist. Das heisst: aus der Summenhäufigkeits“kurve” wird der Graph der Verteilungsfunktion $F(x)$ von X, siehe Figur 4.

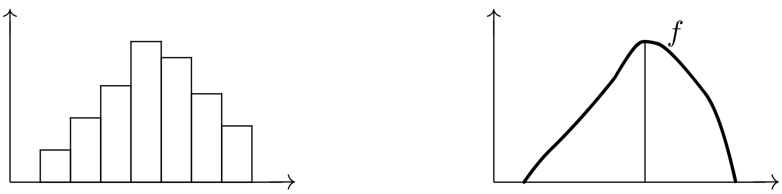
Bevor wir zur allgemeinen Theorie übergehen können, muss noch ein letzter Punkt geklärt werden: In der Formel (C) integrierten wir von a bis x. Die untere Integrationsgrenze a wird natürlich je nach betrachteter Kurve variieren, ja es gibt sogar Fälle, wo die Kurve die x-Achse gar nie berührt (z.B. bei der später zu besprechenden Normalverteilung). Aus diesem Grund muss man im allgemeinen Fall von $-\infty$ bis x integrieren, so dass man ein uneigentliches Integral im Sinne von (20.2), vergleiche erster Band, erhält. Die Formel lautet dann:

$$\text{(D)} \qquad F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt \,.$$

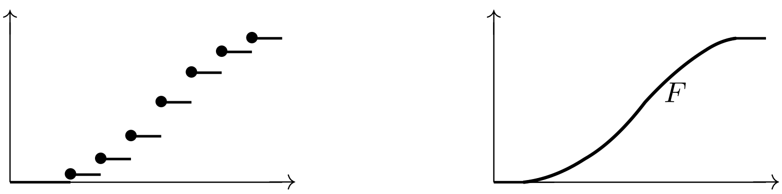
Nun können wir alles bisher Besprochene nochmals zusammenfassen:

Unsere Überlegungen haben uns gezeigt, dass die Verteilungsfunktion $F(x)$ durch ein bestimmtes Integral einer andern Funktion $f(x)$ beschrieben werden kann. Diese Funktion $f(x)$ heisst die *Dichtefunktion* (oder Wahrscheinlichkeitsdichte) der Zufallsgrösse X (oder der Verteilung). Durch die Angabe dieser Dichtefunktion $f(x)$ wird die Verteilung vollständig bestimmt.

Wir stellen noch die verwendeten Übergänge bei unserem Idealisierungsprozess zusammen:



Histogramm $\longrightarrow$ Dichtefunktion



Summenhäufigkeitsverteilung $\longrightarrow$ Verteilungsfunktion
relative Häufigkeit $\longrightarrow$ Wahrscheinlichkeit
Wert des quantitativen Merkmals $\longrightarrow$ Zufallsgrösse

(4.3.3) Definition der stetigen Zufallsgrösse

Wir sind nun in der Lage, genau zu sagen, was eine stetige Zufallsgrösse ist: Darunter versteht man eine Zufallsgrösse, deren Verteilungsfunktion in der oben besprochenen Weise als Integral einer "Dichtefunktion" f dargestellt werden kann. Diese Funktion f darf aber nicht ganz beliebig sein, sondern muss die in der folgenden zusammenfassenden *Definition* angegebenen Bedingungen erfüllen:

Eine Zufallsgrösse X heisst *stetig*, wenn es eine Funktion $f : \mathbb{R} \to \mathbb{R}$ gibt mit

1) f ist stetig (oder auch bloss stückweise stetig, vgl. (10.9) im ersten Band),
2) $f(x) \geq 0 \quad$ für alle x,
3) $\int_{-\infty}^{\infty} f(t)\, dt = 1$,

so dass für die Verteilungsfunktion $F(x) = P(X \leq x)$ von X gilt:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt \; .$$

Die Funktion f heisst die *Dichtefunktion* von X.

Bemerkungen

a) Die Bedingungen 1) und 2) der Definition postuliert man aufgrund der motivierenden Betrachtungen. Auch 3) ist vernünftig. Diese Bedingung besagt einfach, dass die Gesamtfläche unter dem Graphen von f den Wert 1 hat und wird verlangt, weil dieses Integral der Gesamtwahrscheinlichkeit entspricht. Die Wahl von ∞ als obere Integrationsgrenze begründet sich ähnlich wie jene von $-\infty$ als untere Grenze. Bei
$$\int_{-\infty}^{\infty} f(t)\, dt$$
handelt es sich um ein uneigentliches Integral im Sinne von (20.3) im ersten Band.

b) Das für uns wichtigste Beispiel ist die Dichtefunktion der so genannten "Normalverteilung", die wir in Kapitel 5 noch näher besprechen werden. Sie ist gegeben durch
$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (\sigma > 0) \; .$$
Die Bedingungen 1) und 2) der Definition sind sicher erfüllt, denn die Exponentialfunktion ist stetig und nimmt nur positive Werte an. Eine Diskussion dieser Funktion finden Sie in (13.6) wo auch gezeigt wird, dass die Bedingung 3) gilt.

c) Die Verteilungsfunktion $F(x)$ lässt sich aus der Dichtefunktion $f(x)$ per definitionem durch Integration berechnen. Umgekehrt kann man aber auch aus der Verteilungsfunktion die Dichtefunktion bestimmen. Aufgrund der "Tatsache I" von (11.2) im ersten Band ist nämlich (wenigstens dort, wo F differenzierbar ist)

$$F'(x) = f(x)$$

d.h., die Dichtefunktion ist die Ableitung der Verteilungsfunktion. (Das Faktum, dass im abzuleitenden Integral die untere Grenze $-\infty$ statt a ist, ändert an der Gültigkeit der benützten Tatsache nichts.) Somit bestimmen sich Dichtefunktion und Verteilungsfunktion gegenseitig. (Im Fall einer diskreten Zufallsgrösse stehen Wahrscheinlichkeitsfunktion und Verteilungsfunktion in derselben Beziehung. Vergleiche dazu (4.1.7).)

(4.3.4) Berechnung von Wahrscheinlichkeiten

Definitionsgemäss ist $P(X \leq x) = F(x)$. Ist die Verteilungsfunktion $F(x)$ bekannt, so lassen sich aber auch andere Wahrscheinlichkeiten berechnen. Wir geben die wichtigsten Formeln an.

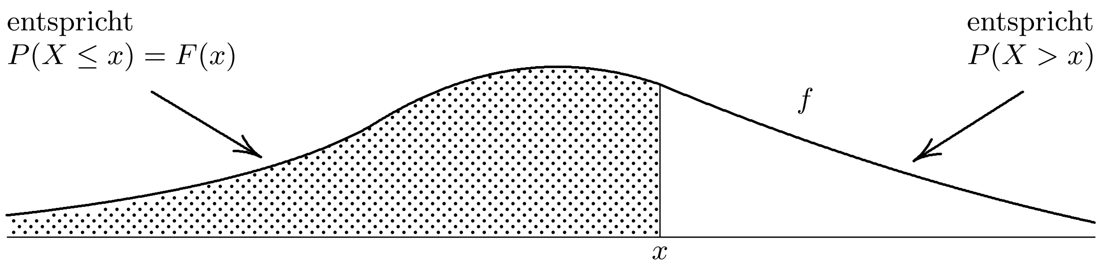
a) $\boxed{P(X > x) = 1 - F(x).}$

Da $X > x$ das Gegenereignis von $X \leq x$ ist, gilt nach Regel 6 von (3.3.6)

$$P(X > x) = 1 - P(X \leq x) \ ,$$

woraus durch Einsetzen von $P(X \leq x) = F(x)$ die Behauptung folgt.

Die Formel hat eine ganz anschauliche geometrische Interpretation:



Der totale Flächeninhalt unter der Kurve ist = 1.

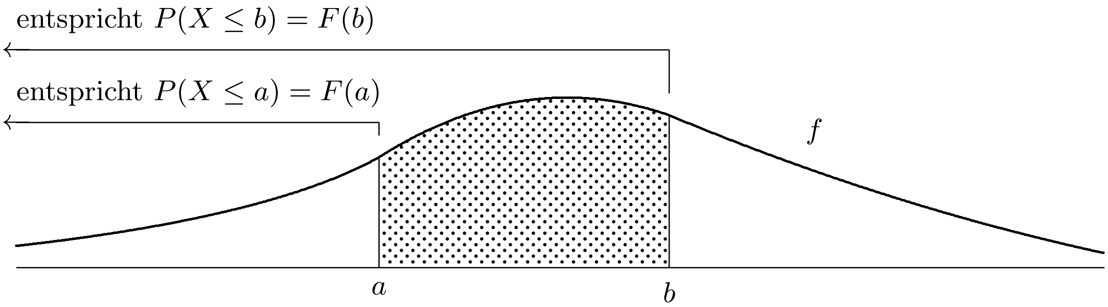
b) $\boxed{P(a < X \leq b) = F(b) - F(a).}$

Die Ereignisse $X \leq a$ und $a < X \leq b$ sind unvereinbar, und ihre Vereinigung ist das Ereignis $X \leq b$. Daraus ergibt sich nach der Regel 4 von (3.3.5):

$$P(X \leq a) + P(a < X \leq b) = P(X \leq b) \ .$$

Mit $P(X \leq a) = F(a)$, $P(X \leq b) = F(b)$ folgt die Behauptung durch Umformen.

Auch hier gibt es eine geometrische Deutung:



Wir halten noch eine Übereinstimmung mit der Integralrechnung fest. Nach dem Hauptsatz (11.4) aus dem ersten Band ist wegen $F'(x) = f(x)$

$$\int_a^b f(t)\,dt = F(b) - F(a) \ ,$$

dies entspricht dem Inhalt des markierten Flächenstücks und dieser wiederum der Wahrscheinlichkeit $P(a < X \leq b)$.

c) Sinngemäss ist zu setzen:

$$P(X = a) = \int_a^a f(t)\,dt = 0 \qquad \text{(vgl. (10.8.b))} \ .$$

Wir haben also wieder einmal den Fall vor uns, dass ein Ereignis die Wahrscheinlichkeit 0 hat, ohne dass es das unmögliche Ereignis ist, vgl. (3.3.3) oder (4.1.8).

Aus der Beziehung $P(X = a) = 0$ folgen weitere Formeln wie etwa

$$P(X < x) = P(X \leq x) = F(x)$$

$$P(a \leq X \leq b) = P(a < X \leq b) = F(b) - F(a)$$

die wir jeweils ohne speziellen Hinweis benützen werden. Merken Sie sich also:

Bei stetigen Zufallsvariablen kommt es nicht darauf an, ob ein "kleiner"- oder ein "kleiner oder gleich"-Zeichen vorliegt.

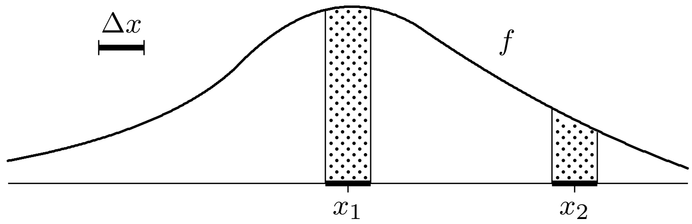
d) Die in a) bis c) angestellten Überlegungen zeigen uns, dass wir sämtliche uns interessierenden Wahrscheinlichkeiten im Zusammenhang mit der Zufallsgrösse X berechnen können, wenn wir nur die Verteilungsfunktion $F(x)$ kennen.

(4.3.5) Dichtefunktion und Wahrscheinlichkeitsfunktion

Es sei X eine stetige Zufallsgrösse. Weil $P(X = x) = 0$ ist für alle x, ist es, im Gegensatz zu den diskreten Zufallsvariablen, nicht sinnvoll, eine "Wahrscheinlichkeitsfunktion" $P(X = x)$ im Sinne von (4.1.8) zu betrachten.

An deren Stelle tritt die Dichtefunktion $f(x)$. Allerdings darf man nicht glauben, dass $f(x) = P(X = x)$ sei, denn es ist ja stets $P(X = x) = 0$. Trotzdem lässt ein grosser Wert von $f(x)$ auf eine hohe Wahrscheinlichkeit schliessen, und zwar in folgendem Sinne:

Wir betrachten ein Intervall I der Länge Δx sowohl an der Stelle x_1 als auch an der Stelle x_2. Wenn nun $f(x_1) > f(x_2)$ ist, dann ist die Wahrscheinlichkeit , dass $X \in I$ ist, an der Stelle x_1 grösser als an der Stelle x_2:



(4.∞) Aufgaben

4−1 In Bemerkung 8) von (4.1.3) wird darauf hingewiesen, dass eine Zufallsgrösse oft durch einen umgangssprachlichen Satz beschrieben wird. Beispiele:

a) Die Zufallsgrösse X sei die Anzahl Reissnägel in einer Schachtel.
b) Die Zufallsvariable Y bezeichne die Brenndauer einer Wunderkerze am Christbaum.

Beschreiben Sie in den beiden Fällen ein passendes Zufallsexperiment, und geben Sie den Ergebnisraum Ω sowie die Abbildungen $X : \Omega \to \mathbb{R}$ bzw. $Y : \Omega \to \mathbb{R}$ in Worten an.

4−2 Eine diskrete Zufallsgrösse X ist durch die folgende Tabelle gegeben

x_i	-3	-1	0	2	4	6
p_i	0.2	0.1	0.1	0.15	0.25	a

a) Wie gross ist a? b) Zeichnen Sie den Graphen der Verteilungsfunktion von X. Berechnen Sie ferner c) $P(0 < X \leq 4)$, d) $P(\pi \leq X \leq 2\pi)$, e) $P(X^2 < 10)$.

4−3 Eine diskrete Zufallsgrösse X ist durch die folgende Tabelle gegeben:

x_i	-2	-1	0	1	4
p_i	0.1	0.2	0.3	0.1	0.3

a) Stellen Sie die entsprechende Tabelle für $Z = X^2 - 1$ auf. b) Berechnen Sie $P(3 \leq Z \leq 8)$.

4−4 Eine diskrete Zufallsgrösse X ist durch folgende Tabelle gegeben:

x_i	-4	-2	0	4	7
p_i	0.1	0.2	0.3	0.1	0.3

a) Zeichnen Sie den Graphen der Verteilungsfunktion $F(x)$ von X.
b) Berechnen Sie $P(X \geq 2)$ und $P(|X| \geq 3)$.

4−5 Ein Würfelspiel: Wir würfeln mit zwei Würfeln. Sind die beiden Augenzahlen gleich, dann ist der Gewinn X gleich dieser Augenzahl, also Fr. 1.– bis Fr. 6.–. Andernfalls ist der Gewinn gleich der Differenz zwischen der grösseren und der kleineren Augenzahl. Geben Sie die Verteilung von X an a) durch eine Tabelle, b) durch ein Stabdiagramm. c) Wie gross ist $P(X \leq 3)$?

4−6 In einem Karton mit 10 Eiern sind vier faul. Ich nehme die Eier einzeln heraus, bis ich ein brauchbares erwische. Die faulen werfe ich weg. Die Zufallsgrösse W gibt an, wieviele Eier ich bei diesem Vorgang wegwerfe. a) Geben Sie die Verteilung von W in Tabellenform an. b) Wie gross ist $P(W \leq 2)$? c) Zeichnen Sie den Graphen der Verteilungsfunktion von W.

4−7 Eine ausgewogene Münze wird dreimal geworfen. Die Zufallsgrössen X bzw. Y bezeichnen die Anzahl der Köpfe in den beiden ersten Würfen bzw. in den beiden letzten Würfen. a) Geben Sie die möglichen Werte von X und Y mit den zugehörigen Wahrscheinlichkeiten in Tabellen an. b) Dasselbe für die Zufallsgrösse $X + Y$. c) Berechnen Sie $P(X \leq 1)$ und $P(X + Y \leq 2)$.

4−8 Eine unverfälschte Münze wird so lange geworfen, bis eine der beiden Seiten zum zweiten Mal erschienen ist. Die Zufallsgrösse X bezeichne die Anzahl der dazu notwendigen Würfe. Bestimmen Sie a) den Ergebnisraum Ω, b) die Werte $X(\omega)$ für $\omega \in \Omega$, c) die Verteilung von X in Form einer (allerdings recht kleinen) Tabelle. d) Zusatzfrage: Wieviele Elemente hat Ω, wenn wir so lange werfen, bis eine der beiden Seiten zum dritten Mal erschienen ist?

4−9 Eine diskrete Zufallsgrösse X nimmt die abzählbar unendlich vielen Werte 0, 1, 2, ... mit den Wahrscheinlichkeiten $P(X = k) = \frac{c}{3^k}$ an. a) Wie gross muss c sein, damit durch diese Festsetzung tatsächlich eine Zufallsgrösse gegeben ist? b) Wie gross ist dann $P(X \leq 5)$? c) Mit welcher Wahrscheinlichkeit nimmt X einen ungeraden Wert an?

4−10 Zeichnen Sie das Stabdiagramm und den Graphen der Verteilungsfunktion der Binomialverteilung mit den Parametern $n = 4,\ p = 1/3$.

4−11 Die Zufallsgrösse X folge der Binomialverteilung mit $n = 8,\ p = 0.6$. Berechnen Sie a) $P(X \leq 3)$, b) $P(X \geq 6)$, c) $P(2 \leq X \leq 5)$.

4−12 Eine Maschine stellt Unterlagsscheiben mit einer Ausschussquote von 10% her. Was ist wahrscheinlicher: a) Kein Ausschuss unter 10 Scheiben, b) höchstens ein Ausschuss-Stück unter 20 Scheiben?

4−13 Wir würfeln mit einem fairen Würfel. Berechnen Sie die Wahrscheinlichkeit dafür, dass a) bei 12 Würfen genau zweimal, b) bei 30 Würfen genau fünfmal eine Sechs erscheint.

4−14 Bei Verkehrsflugzeugen mit vier Triebwerken müssen mindestens zwei funktionieren, damit sie landen können, bei solchen mit drei Triebwerken dagegen nur eines. Welcher Flugzeugtyp ist sicherer, wenn wir annehmen, dass alle Triebwerke dieselbe Ausfallwahrscheinlichkeit haben und die Ausfälle unabhängig voneinander sind?

4−15 Eine Konditorei produziert äusserlich nicht unterscheidbare Pralinés mit roter (60% der Gesamtproduktion) oder weisser (40%) Zucker-Fondant-Füllung, welche in hübsche Schachteln zu 12 Stück abgefüllt werden. a) Welches ist die wahrscheinlichste Anzahl rot gefüllter Pralinés pro Schachtel? Wie gross ist die betreffende Wahrscheinlichkeit? b) Ich kaufe gleich fünf Schachteln aufs Mal. Wie gross ist die Wahrscheinlichkeit dafür, dass jede dieser Schachteln höchstens drei weiss gefüllte Pralinés enthält?

4−16 Wie gross müsste die Wahrscheinlichkeit für eine Mädchengeburt sein, damit in einer Familie mit drei Kindern die Wahrscheinlichkeit für das Ereignis "zwei Mädchen, ein Knabe" genau 30% ist? Gesucht ist eine Lösung im Bereich zwischen 0.4 und 0.6.

4−17 Jemand behauptet, mittels Psychokinese den Fall von Münzen beeinflussen zu können. Als Test wird eine Münze zehnmal geworfen. Unser Freund gibt in 7 Fällen eine korrekte Vorhersage. Als kritische(r) Beobachter(in) fragen Sie sich: Wie gross ist denn die Wahrscheinlichkeit dafür, auch ohne spezielle Fähigkeiten so gut oder gar noch besser abzuschneiden?

4−18 Bei einem Multiple-Choice-Test werden vier Fragen mit je drei und eine Frage mit nur zwei möglichen Antworten gestellt. Wie gross ist die Wahrscheinlichkeit dafür, bei völlig zufälligem Ankreuzen drei oder mehr Fragen richtig zu beantworten?

4−19 Willy Würfel (vgl. Aufgabe 3−32) bietet Ihnen folgendes Spielchen an: Eine ehrliche Münze wird 20mal geworfen. Sie gewinnen, wenn 9- oder 10- oder 11-mal Kopf erscheint. Bestimmen Sie Ihre Gewinnwahrscheinlichkeit.

4−20 Zur Abwechslung eine Extremal-Aufgabe: Ein Glücksrad hat einen roten und einen blauen Sektor. Sie dürfen es zehnmal drehen und gewinnen einen schönen Preis, wenn dabei der Zeiger genau einmal im roten Sektor stehen bleibt. Wie gross muss der Öffnungswinkel des Sektors (in Grad) sein, damit Ihre Gewinnchance maximal ist? Wie gross ist dann diese Gewinnwahrscheinlichkeit?

4−21 Von allen Eiern, welche ein Produzent auf den Markt bringt, sind leider nur P% frisch, die andern sind bereits etwas angefault. Er verkauft dafür seine (zufällig abgefüllten) Zwölferpackungen zum Preis von zehn Eiern und garantiert, dass höchstens eines der Eier schlecht ist.
a) Es sei $P = 90\%$. Mit welcher Wahrscheinlichkeit trifft die garantierte Bedingung ein?
b) Unser Produzent möchte, dass bei (mindestens) 95% aller Packungen die Garantie zutrifft. Wie gross muss P mindestens sein?

4−22 Drei Produzenten A, B und C mit den Marktanteilen 50%, 30% und 20% stellen einen Artikel her, der in zufällig abgefüllten Packungen zu sechs Stück auf den Markt gelangt. Die einzelnen

Artikel sind mit den folgenden Wahrscheinlichkeiten defekt: Bei A 5%, bei B 10% und bei C 15%.
Mit welcher Wahrscheinlichkeit hat es in einer zufällig ausgewählten Sechserpackung genau ein defektes Stück, wenn das Erstellen der Packungen wie folgt geschieht?
a) "Zuerst mischen, dann abfüllen": Die einzelnen Artikel der drei Hersteller werden zuerst gemischt, und die Packungen werden anschliessend abgefüllt.
b) "Zuerst abfüllen, dann mischen": Jeder Hersteller packt seine eigene Produktion in Sechserpackungen; diese Packungen gelangen dann zufällig gemischt in den Verkauf.

4–23 a) Bestimmen Sie die Zahl c so, dass

$$f(x) = \begin{cases} 0 & x < 0 \\ c(1+x)^{-3} & x \geq 0 \end{cases}$$

die Dichtefunktion einer Zufallsvariablen X wird. b) Skizzieren Sie den Graphen von f, und stellen Sie die Wahrscheinlichkeit $p = P(0.5 \leq X \leq 1)$ graphisch dar. c) Berechnen Sie p.

4–24 a) Wie muss man die Zahl a wählen, damit

$$f(x) = \begin{cases} 0 & x < 1 \\ ax^{-2} & x \geq 1 \end{cases}$$

eine Dichtefunktion wird? Die zugehörige stetige Zufallsgrösse nennen wir X. b) Skizzieren Sie den Graphen von f, und stellen Sie die Wahrscheinlichkeit $p = P(1.5 \leq X \leq 2)$ graphisch dar. c) Bestimmen Sie die Verteilungsfunktion von X. d) Berechnen Sie p.

4–25 Kann man die Zahl c so wählen, dass

$$f(x) = \begin{cases} 1 + cx & \text{für } 0 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

eine Dichtefunktion ist? Wenn ja, wie gross ist dieses c?

4–26 a) Zeigen Sie, dass die nachstehende Funktion f eine Dichtefunktion ist:

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}x & 0 \leq x \leq 2 \\ 0 & x > 2 \end{cases}$$

b) Bestimmen Sie die Verteilungsfunktion $F(x)$ der zugehörigen Zufallsgrösse X, und zeichnen Sie ihren Graphen. c) Wie muss man y wählen, damit $P(X \geq y) = 0.05$ ist? d) Überlegen Sie sich, wie der Median von X zu definieren ist, und berechnen Sie ihn.

4–27 a) Zeigen Sie, dass

$$f(x) = \begin{cases} \frac{\pi}{2}\sin(\pi x) & 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

die Dichtefunktion einer Zufallsgrösse X ist. b) Geben Sie die Verteilungsfunktion $F(x)$ von X an. c) Für welche Zahl z ist $P(X < z) = 0.2$?

4–28 Die Dichtefunktion einer stetigen Zufallsgrösse X ist gegeben durch

$$f(x) = \begin{cases} a(2x - x^2) & 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

a) Wie gross muss a sein, damit f eine Dichtefunktion ist? b) Skizzieren Sie den Graphen von f.

4–29 Bestimmen Sie die Konstante c so, dass

$$f(x) = \begin{cases} 0 & x < 0 \\ cxe^{-x/2} & x \geq 0 \end{cases}$$

die Dichtefunktion einer Zufallsgrösse X ist. Berechnen Sie $P(X \geq 6)$.

4−30 Durch

$$f(x) = \begin{cases} 0 & x < 0 \\ 2x \cdot \exp(-x^2) & x \geq 0 \end{cases}$$

ist die Dichtefunktion einer Zufallsgrösse X gegeben. a) Stimmt's? b) Zeichnen Sie den Graphen von f (wo liegt das Maximum?) c) Berechnen Sie $P(X \geq 1)$. d) Für welche Zahl z ist $P(X \leq z) = P(X \geq z)$? Wie ist diese Zahl zu interpretieren?

4−31 Der nebenstehende Graph zeigt die Dichtefunktion einer so genannten Dreiecksverteilung. a) Bestimmen Sie die Höhe h des Dreiecks. b) Bestimmen Sie die Verteilungsfunktion F, und zeichnen Sie ihren Graphen. c) Wie gross ist die Wahrscheinlichkeit dafür, dass eine entsprechend verteilte Zufallsvariable einen Wert ≤ 1 annimmt?