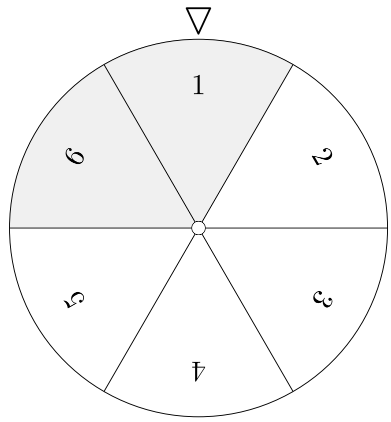
Wahrscheinlichkeitsrechnung, kurze Einführung

Die erste naheliegende Idee hinsichtlich Wahrscheinlichkeit ist,
dass sie etwas mit der Anzahl der bestehenden Möglichkeiten zu tun hat.
In zwei Stapeln mit 3 und 5 Schachteln befindet sich jeweils eine Schachtel mit einem Geschenk.
Du darfst eine der 8 Schachteln wählen. Welchem Stapel entnimmst du die Schachtel?



Das Würfeln entspricht dem Drehen eines Glücksrades.
Die Wahrscheinlichkeiten der sechs möglichen Ergebnisse betrachten wir als gleich.
Hierin steckt die Annahme, dass aufgrund der Symmetrie die Ergebnisse die gleichen Chancen haben und damit langfristig ungefähr gleichhäufig auftreten. Dies ähnelt einer Unschuldsvermutung, die solange aufrecht erhalten werden kann, bis etwas anderes naheliegt.

Die Betrachtung eines Ereignisses „Das Ergebnis ist entweder 1 oder 6."
führt zur naheliegenden Festlegung (Laplace 1749 - 1827):

Unter der Annahme, dass alle Ergebnisse die gleiche Chance haben einzutreffen, gilt:

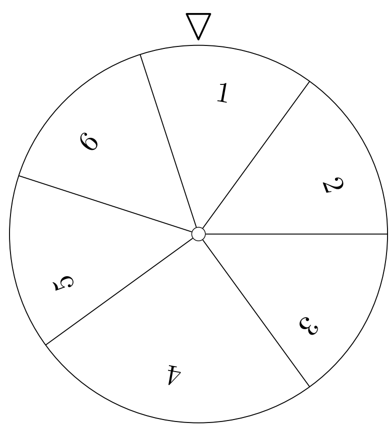
$$P(\textit{Ereignis}) = \frac{\textit{Anzahl der Ergebnisse, aus denen das Ereignis besteht}}{\textit{Anzahl aller möglichen Ergebnisse}}$$

$$P(\textit{Ereignis}) = \textit{Anteil des Ereignisses an allen Ergebnissen}$$

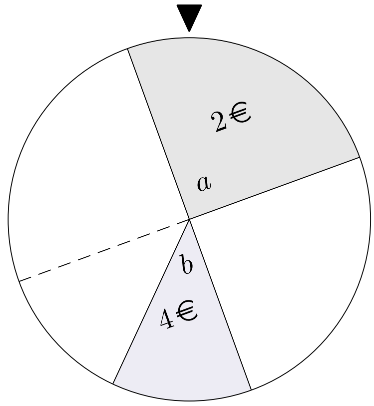
Die Wahrscheinlichkeit *(engl. probability)* wird mit einem P abgekürzt.

$P(\text{„}\mathit{1}\text{“}) = \frac{1}{6}, \quad P(\text{„}\mathit{1\ oder\ 6}\text{“}) = \frac{1}{3}$

Auch für unsymmetrische Wurfobjekte ist es stets denkbar, dass das Zufallsgeschehen durch ein Glücksrad gesteuert wird, bei dem die Anteile für das Auftreten der elementaren Ereignisse nicht mehr gleich sind.



Erwartungswert



Der Sektor a des Glücksrads bringt einen Gewinn von 2 € ,
der Sektor b das Doppelte. Um den fairen Einsatz zu ermitteln,
ist der durchschnittlich zu erwartende Gewinn pro Spiel zu berechnen.

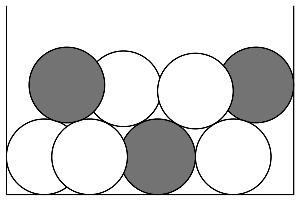
Angenommen, das Rad würde n-mal gedreht,
a erschiene (idealerweise) $\frac{1}{4} \cdot n$-mal, b $\frac{1}{8} \cdot n$-mal.
Der Gesamtgewinn wäre dann $\frac{1}{4} \cdot n \cdot 2 + \frac{1}{8} \cdot n \cdot 4$,
und der durchschnittliche Gewinn pro Spiel

$$\frac{\frac{1}{4} \cdot n \cdot 2 + \frac{1}{8} \cdot n \cdot 4}{n} = \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 4 = 1\,.$$

Der faire Einsatz beträgt also 1 € .

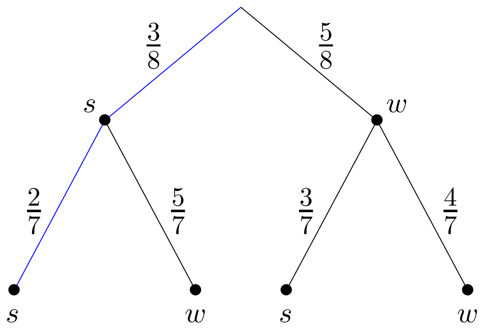
Um den durchschnittlich zu erwartenden Gewinn pro Spiel,
den sogenannten Erwartungswert, zu berechnen, werden die möglichen Gewinne
mit ihren Wahrscheinlichkeiten multipliziert und addiert.
Verluste können als negative Gewinne berücksichtigt werden.

Pfadwahrscheinlichkeiten

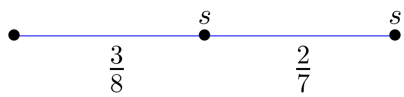


Zwei Kugeln werden nacheinander ohne Zurücklegen gezogen.
Mit welcher Wahrscheinlichkeit ist die zweite Kugel schwarz?

Jedem Elementarereignis (Ergebnis) entspricht ein Pfad
im Baum: $(s, s),\ (s, w),\ (w, s),\ (w, w)$



Wie berechnet sich die Wahrscheinlichkeit eines Pfades?
Betrachten wir hierzu zum Beispiel den Pfad



Bei der Berechnung lassen wir uns von
der Häufigkeitsinterpretation der Wahrscheinlichkeit leiten. Nehmen wir dazu an, dass der zweistufige Versuch n-mal wiederholt wird. Welcher Anteil der Wiederholungen wird diesen Pfad einschlagen?
$\frac{3}{8}$ der Wiederholungen werden den Zweig s einschlagen und $\frac{2}{7}$ der $\frac{3}{8} \cdot n$ Fälle werden von s nach s gehen.
Deshalb werden $\frac{2}{7} \cdot \frac{3}{8} \cdot n$ der n Wiederholungen entlang des Pfades (s, s) ablaufen, und der Anteil beträgt $\frac{2}{7} \cdot \frac{3}{8}$.

Die Wahrscheinlichkeit eines Pfades ist das Produkt der Wahrscheinlichkeiten längs des Pfades.

$P(\text{“zweite Kugel ist schwarz“}) = \frac{3}{8} \cdot \frac{2}{7} + \frac{5}{8} \cdot \frac{3}{7} = \frac{3}{8}$

Vierfeldertafel

In einer Studie wurde ein Medikament getestet. Die Ergebnisse sind in der Tabelle dargestellt.

Dabei bedeuten:

M: Medikament erhalten
$\overline{M}$: Placebo erhalten
G: gesund geworden
$\overline{G}$: nicht gesund geworden

	G	$\overline{G}$	*Summe*
M	2000	500	2500
$\overline{M}$	400	1200	1600
Summe	2400	1700	4100

Wie groß ist die Wahrscheinlichkeit, dass

a) eine Person gesundet, von der man weiß, dass sie das Medikament erhalten hat, $P(G\,|\,M) = \frac{2000}{2500} = 80{,}0\,\%$

b) eine Person nicht gesundet, da sie das Placebo erhalten hat, $P(\overline{G}\,|\,\overline{M}) = \frac{1200}{1600} = 75{,}0\,\%$

c) eine gesund gewordene Person das Medikament nicht erhalten hat? $P(\overline{M}\,|\,G) = \frac{400}{2400} = 16{,}7\,\%$

d) eine Testperson gesundet? $P = \frac{2400}{4100} = 58{,}5\,\%$

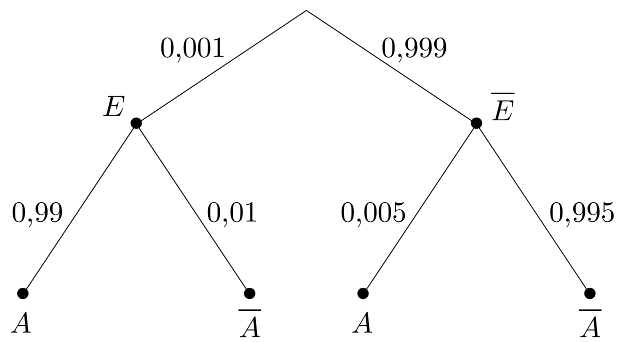
Bedingte Wahrscheinlichkeit

In einem Laden ist eine Alarmanlage eingebaut. Bei Einbruch gibt sie mit 99 %-iger Wahrscheinlichkeit Alarm. Wenn in einer bestimmten Nacht kein Einbruch stattfindet, gibt sie falschen Alarm mit Wahrscheinlichkeit 0,005 (Eine Maus berührt die Anlage oder ähnliches). Die Einbruchswahrscheinlichkeit für eine Nacht sei 0,001. Die Anlage hat gerade Alarm gegeben.
Wie groß ist die Wahrscheinlichkeit, dass gerade ein Einbruch stattfindet?

Wir stellen den zugehörigen Baum auf:

Es bedeuten:

E	Einbruch
$\overline{E}$	kein Einbruch
A	Alarm
$\overline{A}$	kein Alarm



Jedem Elementarereignis (Ergebnis) entspricht ein Pfad im Baum:
$(E, \overline{A})$, (E, A), $(\overline{E}, A)$, $(\overline{E}, \overline{A})$

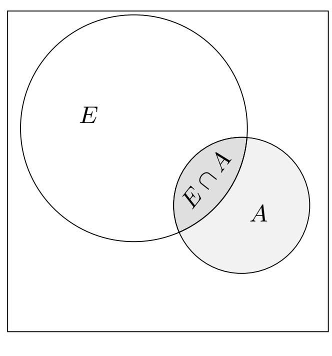
Da wir aber davon ausgehen, dass ein Alarm vorliegt, sind nur die Pfade (E, A) und $(\overline{E}, A)$ zu berücksichtigen. Um die Wahrscheinlichkeit für einen Einbruch zu bestimmen, falls Alarm vorliegt, müssen wir uns fragen, wie groß bei n-maliger Wiederholung des Versuchs die Anzahl des Pfads (E, A) im Verhältnis zur Summe der Anzahlen der Pfade (E, A) und $(\overline{E}, A)$ ist.

Die Pfade (E, A) und $(\overline{E}, A)$ werden mit der Häufigkeit $0{,}001 \cdot 0{,}99n + 0{,}999 \cdot 0{,}005n$ eingeschlagen, der Pfad (E, A) nur $0{,}001 \cdot 0{,}99n$ mal.

Für die gesuchte Wahrscheinlichkeit ergibt sich $\quad P = \dfrac{0{,}001 \cdot 0{,}99}{0{,}001 \cdot 0{,}99 + 0{,}999 \cdot 0{,}005} = 0{,}1654 \;$ (n gekürzt)

Diese Wahrscheinlichkeit heißt *bedingte Wahrscheinlichkeit von E unter der Bedingung A, wir schreiben hierfür:* $P(E \mid A)$

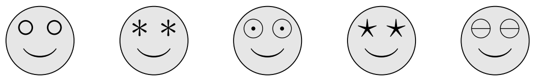
$$P(E|A) = \frac{P(E \cap A)}{P(A)}$$



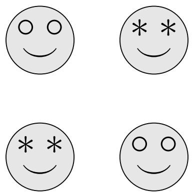
Die bedingte Wahrscheinlichkeit $P(E|A)$ kann als Neueinschätzung der Wahrscheinlichkeit von E interpretiert werden, wenn die Information vorliegt, dass das Ereignis A bereits eingetreten ist. Dann ist von E nur noch derjenige Teil zu berücksichtigen, der sich in A abspielt, also $E \cap A$, und dieser Teil ist in Bezug zu A zu bringen. In der Grafik ist $P(E|A)$ der Anteil der dunkelgrau gefärbten Fläche an der Fläche von A.

n-Fakultät

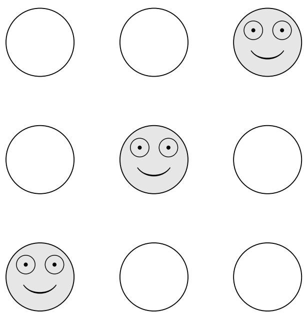
Wie viele Möglichkeiten gibt es, n Kinder in einer Reihe zu platzieren, z. B. für $n = 5$?



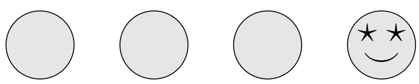
Für $n = 2$ gibt es 2 Möglichkeiten.



Für $n = 3$ hat das 3. Kind 3 Möglichkeiten, die beiden restlichen Plätze können jeweils auf 2 Weisen besetzt werden. Insgesamt gibt es $2 \cdot 3 = 6$ Möglichkeiten.



Für $n = 4$ hat das 4. Kind 4 Möglichkeiten, die restlichen 3 Plätze können jeweils auf $2 \cdot 3 = 6$ Weisen besetzt werden. Insgesamt gibt es $2 \cdot 3 \cdot 4 = 24$ Möglichkeiten.

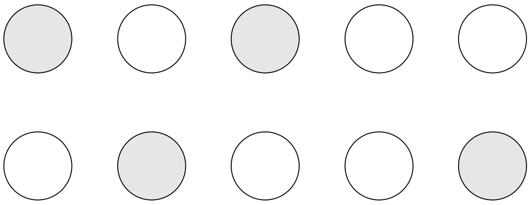


Für $n = 5$ ergeben sich $2 \cdot 3 \cdot 4 \cdot 5 = 120$ Möglichkeiten.

Für das Produkt $2 \cdot 3 \cdot 4 \cdot 5$ schreiben wir kürzer $5!$ und allgemein:

$1 \cdot 2 \cdot 3 \cdot \ ... \ \cdot (n-1) \cdot n = n!$

Binomialkoeffizient $\binom{n}{k}$

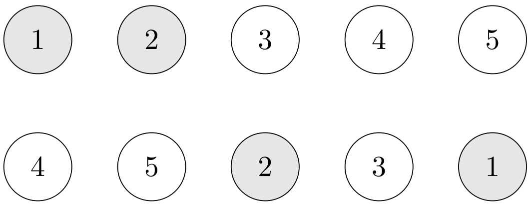


1. Muster

2. Muster

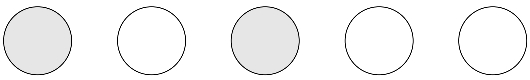
Wir wollen herausfinden, wie viele Muster der abgebildeten Art es gibt.
Von $n = 5$ Plätzen werden hier $k = 2$ ausgewählt.
Die Anzahl wird mit $\binom{n}{k}$ bezeichnet (sprich n über k). Wie groß ist diese Anzahl nun?

Wenn die 5 Elemente (Kugeln) nummeriert sind, können wir die Anzahl $n! = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n$ der Permutationen (Vertauschungen) bestimmen. Für die 1 gibt es n Plätze, für die 2 sind es $n-1$ Plätze, usw. Die Färbung der Elemente 1 und 2 verändert nicht die Anzahl:



...

Die Anzahl der Permutationen kann auf eine zweite Art bestimmt werden.
Wir gehen von der (noch unbekannten) Anzahl $\binom{n}{k}$ der Muster aus.



Die grau gefärbten Plätze lassen sich auf 2-fache Weisen mit den Zahlen 1 und 2 belegen,
die restlichen auf $3! = 6$ Arten mit 3, 4 und 5.
Insgesamt erhaltenen wir mit $\binom{5}{2} \cdot 2! \cdot 3!$ alle $5!$ Permutationen. Daraus folgt $\binom{5}{2} = \dfrac{5!}{2! \cdot 3!}$

und mit derselben Überlegung allgemein: $\binom{n}{k} = \dfrac{n!}{k! \cdot (n-k)!}$

$\binom{n}{k}$ ist auch die Anzahl der Möglichkeiten, mit der man aus n Plätzen k auswählen kann.

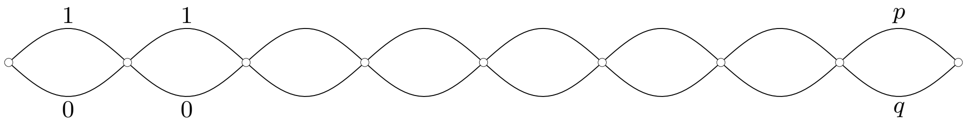
Bernoulli-Kette

Jacob Bernoulli (1654 - 1705)

Die folgende Aufgabe bereitet spätere Anwendungen der Wahrscheinlichkeitsrechnung vor, insbesondere das Erstellen von Tests zur Qualitätskontrolle von Erzeugnissen, zur Untersuchung der Wirksamkeit von Medikamenten oder zur Beantwortung biologischer Fragen wie:
Sind Ratten farbenblind?

In einer Urne befinden sich 4 schwarze und 6 weiße Kugeln. Wir mischen und entnehmen der Urne eine Kugel, notieren die Farbe und legen die Kugel wieder in die Urne zurück. Diesen Einzelversuch wiederholen wir 8mal. Wie groß ist die Wahrscheinlichkeit, dass unter den 8 Ziehungen keine (eine, zwei, ..., acht) schwarze Kugeln sind?

Ein Zufallsversuch mit zwei möglichen Ausfällen (Treffer 1, Fehlschlag 0) heißt Bernoulli-Versuch.
Die Wahrscheinlichkeiten werden mit p (Treffer) und q bezeichnet.
Wiederholt man einen Bernoulli-Versuch n-mal, so entsteht eine Bernoulli-Kette der Länge n.



Die Elementarereignisse einer Bernoulli-Kette der Länge n
bestehen aus allen 0-1-Folgen der Länge n.

Für die Aufgabe ist z. B. die Wahrscheinlichkeit des Elementarereignisses (1, 1, 0, 0, 1, 1, 0, 1)
$p^5 \cdot q^3$ mit $p = 0{,}4$ und $q = 1 - p = 0{,}6$.

Allgemein interessiert man sich bei einer Bernoulli-Kette für die Wahrscheinlichkeit, genau k Treffer zu erzielen. Sei X die Anzahl der Treffer für jedes Elementarereignis. Wir fragen nach der Wahrscheinlichkeit des Ereignisses $X = 2$.
Elementarereignisse für genau 2 Treffer sind z. B. (1, 1, 0, 0, 0, 0, 0, 0)
und (0, 0, 1, 0, 1, 0, 0, 0)

Hiervon gibt es $\binom{8}{2}$ Stück, die alle jeweils die Wahrscheinlichkeit $p^2 \cdot q^6$ haben, insgesamt erhalten wir:

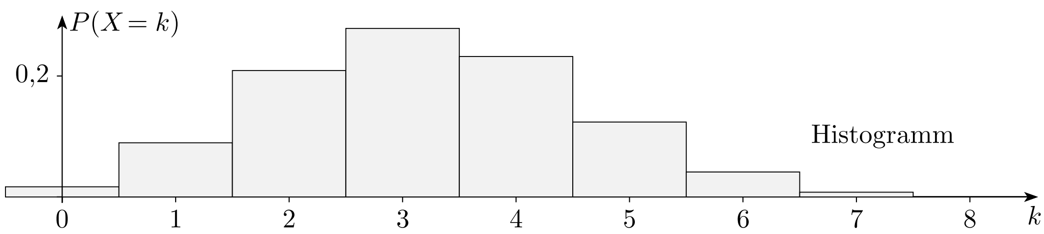
$$P(X = 2) = \binom{8}{2} \cdot p^2 \cdot q^6$$

Bei einer Bernoulli-Kette der Länge n gebe die Zufallsvariable X die Anzahl der Treffer an.
Die Trefferwahrscheinlichkeit sei p. Dann ist die Wahrscheinlichkeit für k Treffer

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}, \qquad q = 1 - p$$

Die Zufallsvariable X heißt binomialverteilt.

k	$P(X=k)$
0	0,017
1	0,090
2	0,209
3	0,279
4	0,232
5	0,124
6	0,041
7	0,008
8	0,000

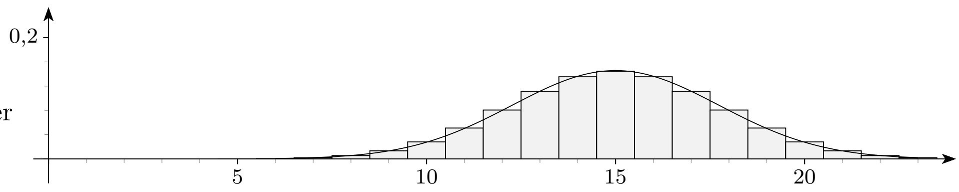


Sigma-Umgebung

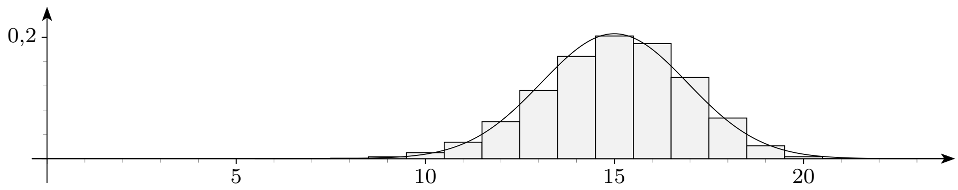
Vergleichen wir die beiden Binomialverteilungen:

$n = 30$
$p = 0{,}5$

(z.B. 30-maliges Werfen einer Münze, X Anzahl von "Zahl")



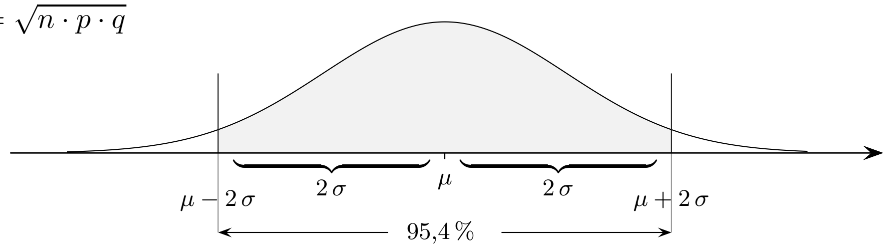
$n = 20$
$p = 0{,}75$



Der Erwartungswert beider Binomialverteilungen beträgt $\mu = n \cdot p = 15$. Jedoch sind die Wahrscheinlichkeiten, mit denen die Zufallsvariable X um den Erwartungswert streut, verschieden. Wie weit darf die Zufallsvariable vom Erwartungswert höchstens abweichen, damit kein Zweifel an der Unverfälschtheit der Münze (allgemeiner: an der zugrundegelegten Wahrscheinlichkeit) berechtigt ist? Um solche und ähnliche Fragen beantworten zu können, benötigen wir zunächst ein Maß für die Streuung um den Erwartungswert.
Wenn wir die Binomialverteilung durch eine symmetrische Kurve annähern, so charakterisiert die Lage der beiden Wendepunkte die Streuung um den Erwartungswert.

Der Abstand vom Erwartungswert zur x-Koordinate eines Wendepunkts heißt daher Standardabweichung und wird mit σ (lies: sigma) bezeichnet.

Mit Mitteln der Analysis kann $\sigma = \sqrt{n \cdot p \cdot q}$ bestimmt werden.



Für n-stufige Bernoulli-Versuche gilt, falls die Laplace-Bedingung $\sigma > 3$ erfüllt ist:
Die Wahrscheinlichkeit (genähert), dass die Anzahl der Treffer in der

1σ-Umgebung liegt, beträgt 68,3 % $\approx \frac{2}{3}$	*(1σ-Regel),*	normalcdf$(-1, 1)$
2σ-Umgebung liegt, beträgt 95,4 %	*(2σ-Regel),*	normalcdf$(-2, 2)$
3σ-Umgebung liegt, beträgt 99,7 %	*(3σ-Regel).*	normalcdf$(-3, 3)$