

1 Der Wahrscheinlichkeitsbegriff

1.1 Die empirische Wahrscheinlichkeit

Im Zusammenhang mit statistischen Größen untersucht man etwa die Häufigkeitsverteilungen der Ergebnisse bei Münz- und Würfelwurf.

Exemplarische Ergebnislisten sind etwa:¹

Wurf mit einer Münze			Wurf mit einem Würfel						
Kopf	Zahl	Wurfzahl	1	2	3	4	5	6	Wurfzahl
4	6	10	3	1	1	2	1	2	10
7	3	10	5	5	2	3	3	2	10
49	51	100	16	18	24	19	14	9	100
43	57	100	15	15	21	19	18	12	100
492	508	1000	162	156	174	184	160	154	1000
518	482	1000	160	162	166	159	174	179	1000
5001	4999	10000	1689	1701	1656	1664	1594	1696	10000
4966	5034	10000	1653	1653	1710	1675	1646	1663	10000
49862	50138	100000	16413	16723	16688	16674	16771	16731	100000
50002	49998	100000	16412	16824	16675	16749	16592	16748	100000

An diesen ersten Beispielen hat man gesehen, dass die relative Häufigkeit beim Münzwurf für beide Ergebnisse jeweils ca. 0,5 beträgt, wenn man nur lange genug wirft. Ähnlich verhält es sich beim Würfelwurf: hier beträgt die relative Häufigkeit für alle drei Ergebnisse jeweils ca. 0,167.

Das sind die Ergebnisse, die man umgangssprachlich als Wahrscheinlichkeit der Würfe bezeichnet: $\frac{1}{2}$ beim Münzwurf und $\frac{1}{6}$ beim Würfelwurf.



Adresse: Eduard-Spranger-Berufskolleg, 59067 Hamm

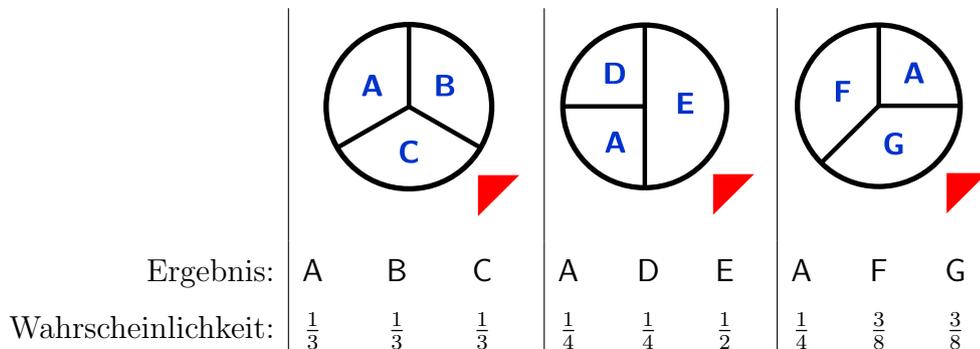
E-Mail: mail@frank-klinker.de

Version: 20. August 2025

¹Erzeugt mit dem [Würfelsimulator der TU Clausthal](#)

Die gleichen Wahrscheinlichkeiten leitet man üblicherweise aus der Geometrie der geworfenen Objekte her: in beiden Experimenten sind alle Ergebnisse "gleichwertig". Ohne auf die relativen Häufigkeiten zurückzugreifen, wendet man diese Idee auch in anderen Situationen an.

Beispiel 1. Wahrscheinlichkeiten von Glücksrädern anhand des Anteils des Kreisrandes:



Bemerkung 2. Diese empirische Art die Wahrscheinlichkeit von Ergebnissen eines Zufallsexperiments zu ermitteln wird oft verwendet und erweist sich als nützlich.

Insbesondere achtet man darauf, dass alle einzelnen Wahrscheinlichkeiten zusammenaddiert 1 ergeben.

Das entspricht der Tatsache, dass die relativen Häufigkeiten sich ebenfalls zu 1 addieren.

Dass das tatsächlich korrekt ist, formuliert das **Schwaches Gesetz für relative Häufigkeiten**.²

Die relativen Häufigkeiten eines Zufallsexperiments nähern sich mit steigender Versuchszahl der erwarteten Wahrscheinlichkeit an.

Als Abkürzung für die Wahrscheinlichkeit wird das Symbol³ P verwendet:

Münze: $P(K) = \frac{1}{2}, P(Z) = \frac{1}{2}.$

Würfel: $P(1) = \frac{1}{6}, P(2) = \frac{1}{6}, P(3) = \frac{1}{6}, P(4) = \frac{1}{6}, P(5) = \frac{1}{6}, P(6) = \frac{1}{6}.$

Glücksrad 1: $P(A) = \frac{1}{3}, P(B) = \frac{1}{3}, P(C) = \frac{1}{3}.$

Glücksrad 2: $P(A) = \frac{1}{4}, P(D) = \frac{1}{4}, P(E) = \frac{1}{2}.$

Glücksrad 3: $P(A) = \frac{1}{4}, P(F) = \frac{3}{8}, P(G) = \frac{3}{8}.$

1.2 Wahrscheinlichkeitsraum

Man löst sich im nächsten Schritt von den relativen Häufigkeiten und konzentriert sich nur auf die Wahrscheinlichkeiten.

²Später werden wir das etwas "mathematischer" formulieren.

³ P als Abkürzung für probability.

Das verbindet man dann mit den bisher erarbeiteten Begriffen wie Ergebnis, Ergebnismenge, Ereignis und Ereignisraum.

Da wir uns zunächst auf abzählbare (meistens sogar endliche) Ergebnismengen Ω beschränken, reicht die folgende Variante des Wahrscheinlichkeitsbegriffs:⁴

Wahrscheinlichkeitsverteilung eines Zufallsexperiments

$\Omega = \{e_1, e_2, e_3, \dots\}$ ist die abzählbare Ergebnismenge eines Zufallsexperimentes mit Ereignisraum $\mathcal{E} \subset \mathcal{P}(\Omega)$. Dann heißt P eine **Wahrscheinlichkeitsverteilung** des Zufallsexperiments, wenn folgende drei Bedingungen erfüllt sind:

- I. Für alle Ergebnisse e_i gibt es eine Wahrscheinlichkeit $p_i = P(e_i)$ mit $0 \leq p_i \leq 1$
- II. $p_1 + p_2 + p_3 + \dots = 1$
- III. Ist $A \in \mathcal{E}$ ein Ereignis, also insbesondere eine Teilmenge von Ω , dann ist

$$P(A) = \sum_{e_i \in A} p_i$$

(Ω, \mathcal{E}, P) zusammen heißt dann **Wahrscheinlichkeitsraum**.

Oft ist $\mathcal{E} = \mathcal{P}(\Omega)$ und \mathcal{E} enthält alle Teilmengen von Ω . Dann schreibt man kürzer (Ω, P) für den Ereignisraum.

Beispiel 3. Die Beispiele aus dem vorigen Abschnitt erfüllen die Eigenschaften einer Wahrscheinlichkeitsverteilung!

Man kann somit die Wahrscheinlichkeit spezieller Ereignisse berechnen:

- Mit einem gewöhnlichen Würfel eine 1, 2 oder 3 zu würfeln entspricht dem Ereignis $A = \{1, 2, 3\}$. Die Wahrscheinlichkeit dafür ist $P(A) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$.
- Die Wahrscheinlichkeit mit dem dritten Glücksrad F oder G zu drehen entspricht der Wahrscheinlichkeit des Ereignisses $A = \{F, G\}$, also $P(A) = \frac{3}{8} + \frac{3}{8} = \frac{3}{4}$.

1.3 Erste Rechenregeln

Aus den Eigenschaften 1-3 folgen direkt einige naheliegende Rechenregeln:

1. Wegen Eigenschaft II und III ist $P(\Omega) = 1$

2. Für zwei disjunkte⁵ Ereignisse A und B gilt wegen Eigenschaft III:

$$P(A \cup B) = P(A) + P(B) \quad (\text{Spezieller Additionssatz})$$

⁴Die Eigenschaften 1-3 sind vereinfachte Varianten der **Kolmogorov-Axiome**. (Andrei Nikolajewitsch Kolmogorow, *1903/†1987, [wikipedia: Andrei Nikolajewitsch Kolmogorow](https://de.wikipedia.org/wiki/Andrei_Nikolajewitsch_Kolmogorow))

⁵**disjunkt** bedeutet, dass A und B keine gemeinsamen Ergebnisse als Element besitzen, also $A \cap B = \emptyset$.

3. Aus $A \cup \bar{A} = \Omega$ und $A \cap \bar{A} = \emptyset$ folgt mit 2 und 1

$$\boxed{P(\bar{A}) = 1 - P(A)} \quad (\text{Wahrscheinlichkeit des Gegenereignisses})$$

4. Aus 3 folgt insbesondere $\boxed{P(\emptyset) = 0}$

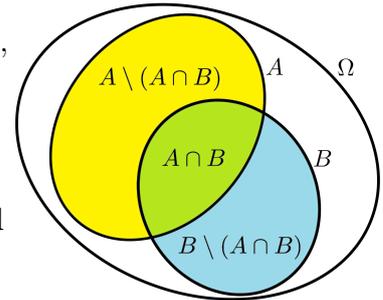
5. Gilt für zwei Ereignisse $B \subset A$, dann ist $A = B \cup (A \setminus B)$ eine disjunkte Zerlegung von A . Insbesondere liefert 2 in diesem Fall

$$\boxed{P(A \setminus B) = P(A) - P(B)} \quad (\text{wenn } B \subset A)$$

6. Wenn die Ereignisse A und B nicht disjunkt sind, dann ist

$$A \cup B = (A \setminus (A \cap B)) \cup (A \cap B) \cup (B \setminus (A \cap B))$$

eine disjunkte Zerlegung mit $(A \cap B) \subset A$ und $(A \cap B) \subset B$, vgl. mit der Skizze.



Damit gilt

$$\begin{aligned} P(A \cup B) &\stackrel{2}{=} P(A \setminus (A \cap B)) + P(A \cap B) + P(B \setminus (A \cap B)) \\ &\stackrel{5}{=} (P(A) - P(A \cap B)) + P(A \cap B) + (P(B) - P(A \cap B)) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

Zusammengefasst gilt:

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)} \quad (\text{Additionssatz})$$

1.4 Laplace-Experimente

Die **Laplace-Experimente** sind wichtige Spezialformen von Zufallsexperimenten.⁶

Ein Zufallsexperiment mit endlicher Ergebnismenge $\Omega = \{e_1, e_2, \dots, e_n\}$ heißt Laplace-Experiment, wenn alle Ergebnisse die gleiche Wahrscheinlichkeit haben, nämlich

$$p_i = \frac{1}{n}$$

Beispiel 4. Der Wurf eines gewöhnlichen Würfels ($p_i = \frac{1}{6}$), der Münzwurf ($p_i = \frac{1}{2}$), das Glücksrad 1 ($p_i = \frac{1}{3}$) und das Ziehen von Kugeln aus einer Urne ($p_i = \frac{1}{\text{Anz. der Kugeln}}$) sind Laplace-Experimente.

⁶Pierre-Simon Laplace *1749/†1827, [wikipedia: Pierre-Simon Laplace](https://de.wikipedia.org/wiki/Pierre-Simon_Laplace)

Wegen der einfachen Ergebniswahrscheinlichkeiten erhält man die Wahrscheinlichkeiten von Ereignissen, indem man dieses "abzählt":

$$p(A) = \frac{\text{Anzahl der Ergebnisse, für die } A \text{ eintritt}}{\text{Anzahl aller möglicher Ergebnisse}} = \frac{\#A}{\#\Omega}$$

Diese Formel heißt auch **Laplace-Formel**.

2 Mehrstufiger Zufallsversuche und Wahrscheinlichkeitsbäume

Bisher wurden Wahrscheinlichkeiten einfacher Zufallsexperimente behandelt. Das wird nun auf mehrstufige Experimente erweitert. Dabei wird das Baumdiagramm eine wesentliche Rolle spielen.

Als Einstieg wird der Wurf mit einem quaderförmigen Würfel (zwei kleine und 4 größere Seiten) mit dem anschließenden Wurf einer verbogenen Münze kombiniert.

- Für das Würfelexperiment ist $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$ mit $p_1 = p_6 = 0,08$ und $p_2 = p_3 = p_4 = p_5 = 0,21$
- Für das Münzexperiment ist $\Omega_2 = \{Z, K\}$ mit $p_K = 0,4$ und $p_Z = 0,6$

Nutzt man einen empirischen Zugang mit insgesamt 1000 kombinierten Würfeln, dann bedeutet das z. B.:

- In 21% aller Würfe wird durchschnittlich eine 2 gewürfelt, das sind hier dann 210 2er-Würfe.
- Außerdem wird in 40% aller Würfe mit der Münze Kopf geworfen. Das sind von den 210 2er-Würfeln dann 84 Würfe, die mit einem Kopf-Wurf kombiniert sind.
- Bei insgesamt 1000 Würfeln hat man also 8,4% (2, K)-Würfe.

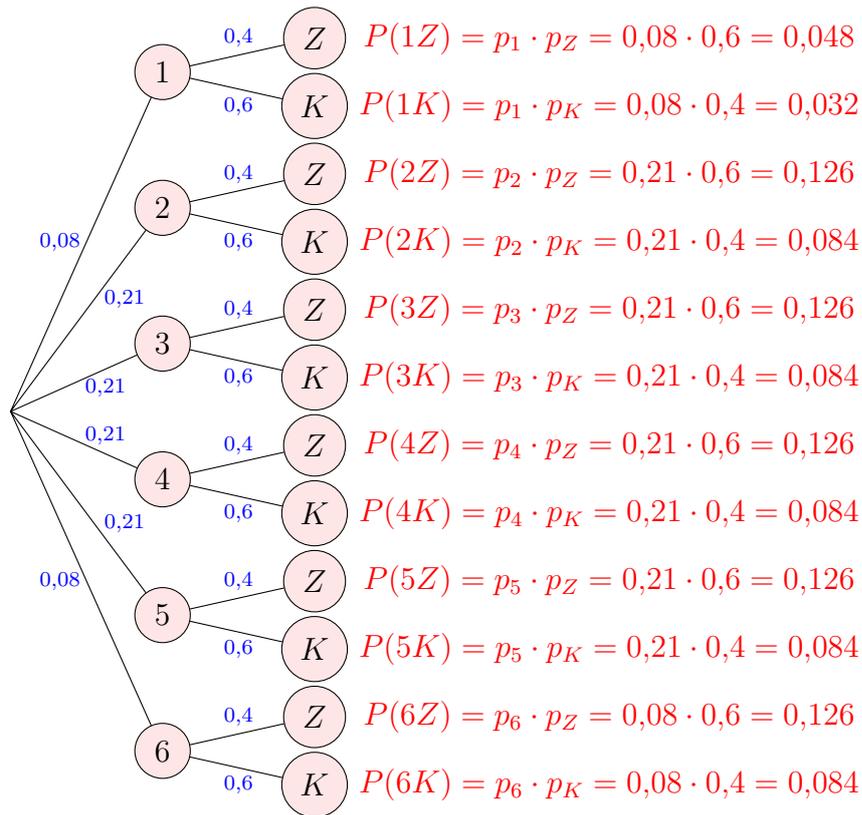
Das gleiche Ergebnis erhält man, wenn man $p_2 \cdot p_K$ berechnet: $p_2 \cdot p_K = 0,21 \cdot 0,4 = 0,084$.

Um das zu systematisieren ergänzt man das bekannte Baumdiagramm durch die Wahrscheinlichkeiten.

In dem Baumdiagramm beschreibt ein vollständiger Pfad von der Wurzel zum Ende ein Ergebnis des zweistufigen Zufallsexperiments.

1. Pfadmultiplikationsregel

Die Wahrscheinlichkeit eines Ergebnisses eines mehrstufigen Zufallsexperiments erhält man, indem man alle Wahrscheinlichkeiten entlang des vollständigen Pfades, der dieses Ergebnis beschreibt, multipliziert.



Aus den Rechenregeln für Wahrscheinlichkeiten ergeben sich folgende Regeln im Umgang mit Baumdiagrammen:

2. Pfadaddition:

Die Wahrscheinlichkeiten vollständiger Pfade werden addiert.

3. Erste Vollständigkeitsregel:

Die Wahrscheinlichkeiten der Äste, die von einem Knoten ausgehen addieren sich zu 1

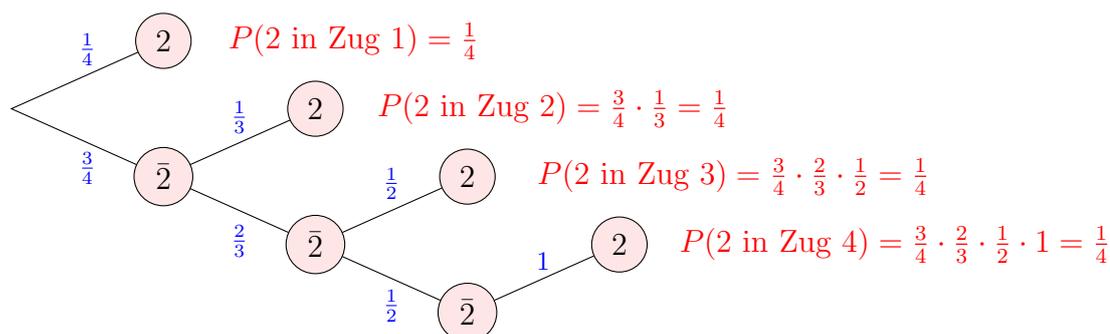
4. Zweite Vollständigkeitsregel:

Startet man von einem Knoten, dann addieren sich die Wahrscheinlichkeiten aller von dort startenden vollständigen Pfade zu 1

Beispiel 5. Aus einer Urne mit vier Kugeln (nummeriert mit 1, 2, 3 und 4) wird ohne Zurücklegen so lange eine Kugel gezogen, bis eine 2 erscheint.

Da die Anzahl der Kugeln sich nach jedem Zug ändert, ändern sich auf jeder Stufe auch die Wahrscheinlichkeiten. Wir wählen in jeder Stufe $\Omega = \{2, \bar{2}\}$. Das zugehörige

Baumdiagramm ist dann



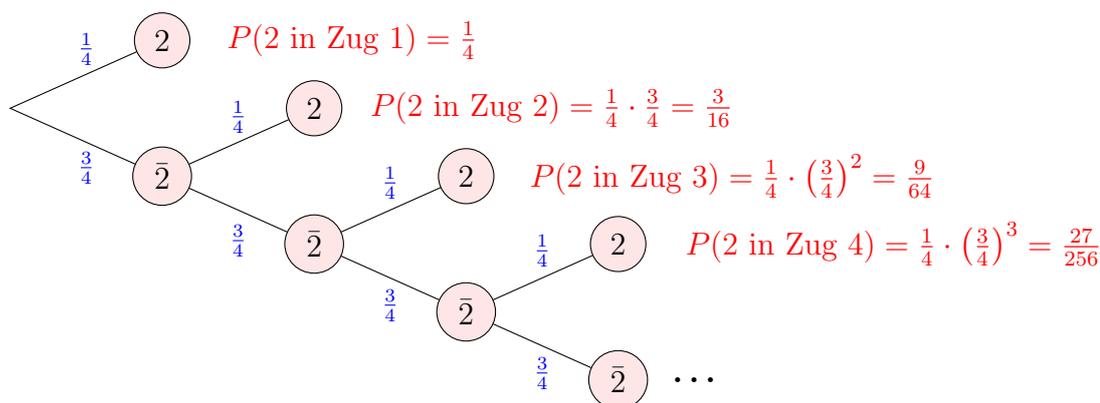
Man überzeugt sich davon, dass die Vollständigkeitsregeln erfüllt sind.

Man kann nun die Wahrscheinlichkeiten verschiedener Ereignisse berechnen.

Z. B. E : "Die 2 wird spätestens im dritten Zug gezogen":

$$P(E) = P(2 \text{ in Zug 1}) + P(2 \text{ in Zug 2}) + P(2 \text{ in Zug 3}) = \frac{3}{4}.$$

Beispiel 6. Wieder wird die Urne aus dem vorigen Beispiel genutzt und wieder soll die 2 gezogen werden. In diesem Zufallsexperiment wird die gezogene Kugel jedes Mal wieder zurückgelegt, sodass die Urne nie leer wird. Dadurch wird der Baum nicht enden, denn es kann ja passieren, dass in jedem Zug eine 1, 3 oder 4 gezogen wird. Außerdem ist die Wahrscheinlichkeiten auf jeder Stufe identisch:



Hier kann man allgemein die Wahrscheinlichkeit dafür angeben, dass die 2 im k -ten Zug gezogen wird:

$$P(2 \text{ in Zug } k) = \frac{1}{4} \cdot \left(\frac{3}{4}\right)^{k-1}.$$

Die Wahrscheinlichkeit für E_k : "Die 2 wird spätestens im k -ten Zug gezogen" berechnet man wie im vorigen Beispiel:⁷

⁷An der Stelle (*) wird die **geometrische Summenformel** verwendet:

$$\sum_{\ell=0}^{k-1} q^\ell = 1 + q + q^2 + \dots + q^{k-2} + q^{k-1} = \frac{1 - q^k}{1 - q}$$

$$\begin{aligned}
P(E_k) &= P(2 \text{ in Zug } 1) + P(2 \text{ in Zug } 2) + P(2 \text{ in Zug } 3) + \dots + P(2 \text{ in Zug } k) \\
&= \frac{1}{4} + \frac{1}{4} \cdot \frac{3}{4} + \frac{1}{4} \cdot \left(\frac{3}{4}\right)^2 + \dots + \frac{1}{4} \cdot \left(\frac{3}{4}\right)^{k-1} \\
&= \frac{1}{4} \cdot \left(1 + \frac{3}{4} + \left(\frac{3}{4}\right)^2 + \dots + \left(\frac{3}{4}\right)^{k-1}\right) \\
&\stackrel{(*)}{=} \frac{1}{4} \cdot \frac{1 - \left(\frac{3}{4}\right)^k}{1 - \frac{3}{4}} \\
&= 1 - \left(\frac{3}{4}\right)^k
\end{aligned}$$

Die Frage "Wird die 2 irgendwann gezogen", ist eher theoretischer Natur.⁸ Dennoch kann man sich dieser Frage theoretisch nähern: Man kann das zugehörige Ereignis mit E_∞ identifizieren und dessen Wahrscheinlichkeit mit dem Wert $\lim_{k \rightarrow \infty} P(E_k)$:

$$P(E_\infty) = \lim_{k \rightarrow \infty} P(E_k) = \lim_{k \rightarrow \infty} \left(1 - \left(\frac{3}{4}\right)^k\right) = 1 - \lim_{k \rightarrow \infty} \left(\frac{3}{4}\right)^k = 1 - 0 = 1$$

Das bestätigt auch die zweite Vollständigkeitsregel für das unendlich große Baumdiagramm.

Der "Restbaum" sieht immer aus wie der gesamte Baum, egal von welchem Knoten man startet. Deshalb wird auch die erste Vollständigkeitsregel bestätigt.

3 Wahrscheinlichkeit einer Zufallsvariable

Gegeben ist eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ zum Zufallsexperiment mit der abzählbaren Ergebnismenge Ω . Weiterhin ist auch eine zugehörige Wahrscheinlichkeitsverteilung P gegeben.

Dann liefert das eine Wahrscheinlichkeitsverteilung $P_X : \mathbb{R} \rightarrow [0, 1]$ durch

$$P_X(k) = P(\{e \in \Omega \mid X(e) = k\}) = P(X = k).$$

Das heißt: die Wahrscheinlichkeit dafür, dass $X = k$ eintritt ist gleich der Wahrscheinlichkeit des Ereignisses $E_k \subset \Omega$, wobei $E_k = \{e \in \Omega \mid X(e) = k\}$ alle die Ergebnisse enthält, die durch X auf k abgebildet werden.

Insbesondere ist $P_X(k) = 0$, wenn k durch X gar nicht getroffen wird.

Durch die alternative Schreibweise $P_X(k) = P(X = k)$ lassen sich Ereignisse sehr elegant formulieren, siehe unten.

⁸Anschaulich ist klar, dass irgendwann 2 gewürfelt wird, man muss ja nur beliebig oft würfeln. Da man praktisch aber immer nur endlich viele Würfe durchführen kann, kann es wirklich passieren, dass nie die 2 fällt.

Ein Ergebnis $e \in \Omega$ kann durch X nicht auf zwei verschiedene Bilder abgebildet werden. Daher ist stets $E_{k_1} \cap E_{k_2} = \emptyset$. Mit dem Additionssatz ergeben sich deshalb folgende Rechenregeln:

$$\begin{aligned}
 P(k_1 \leq X \leq k_2) &= \sum_{k_1 \leq \ell \leq k_2} P(X = \ell) & P(k_1 < X \leq k_2) &= \sum_{k_1 < \ell \leq k_2} P(X = \ell) \\
 P(k_1 \leq X < k_2) &= \sum_{k_1 \leq \ell < k_2} P(X = \ell) & P(k_1 < X < k_2) &= \sum_{k_1 < \ell < k_2} P(X = \ell) \\
 P(X \leq k_2) &= \sum_{\ell \leq k_2} P(X = \ell) & P(X \geq k_1) &= \sum_{\ell \geq k_1} P(X = \ell) \\
 P(X < k_2) &= \sum_{\ell < k_2} P(X = \ell) & P(X > k_1) &= \sum_{\ell > k_1} P(X = \ell)
 \end{aligned}$$

Außerdem gilt

$$\sum_{k \in \mathbb{R}} P(X = k) = 1$$

sowie etwa

$$P(X \leq k) = 1 - P(X > k).$$

Hinweis: Ist Ω sogar endlich, dann ist auch die Menge der von X getroffenen Werte endlich. Damit sind dann auch die Summen alle endlich!

Die Definition im obigen Kasten klingt zunächst etwas kompliziert. Sie ist jedoch sehr natürlich und entspricht dem, was man bisher gemacht hat. Zur Illustration kann das folgende Beispiel hilfreich ein.

Beispiel 7. 1. Betrachtet wird der Wurf eines einzelnen Würfels mit der Ergebnismenge $\Omega = \{1, 2, 3, 4, 5, 6\}$. Die Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ ist durch $X(i) = (i - 3)^2$ gegeben.

Die zugehörige Ereignistabelle kann man durch die Wahrscheinlichkeiten ergänzen:

k	$E_k = \{e \mid X(e) = k\}$	$\#E_k$	$P(X = k)$
0	{3}	1	$\frac{1}{6}$
1	{2, 4}	2	$\frac{1}{3}$
4	{1, 5}	2	$\frac{1}{3}$
9	{6}	1	$\frac{1}{6}$

Einige Ereignisse sind:

- $P(X < 9) = P(X = 0) + P(X = 1) + P(X = 4) = \frac{1}{6} + \frac{1}{3} + \frac{1}{3} = \frac{5}{6}$
(Die Wahrscheinlichkeit, dass X kleiner als 9 ist, beträgt ca. 83%)
- $P(|X - 3| \leq 2) = P(X = 1) + P(X = 4) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$
(Die Wahrscheinlichkeit, dass X von 3 höchstens den Abstand 2 hat, beträgt ca. 67%)

- $P(|X - 3| < 2) = P(X = 4) = \frac{1}{3}$

(Die Wahrscheinlichkeit, dass X von 3 einen Abstand kleiner als 2 hat, beträgt ca. 33%.)

2. Beim Wurf mit zwei Würfeln ist $\Omega = \{ij \mid 1 \leq i \leq 6, 1 \leq j \leq 6\}$. Hierbei handelt es sich um ein Laplace-Experiment, da jeder der 36 Würfe gleich wahrscheinlich ist mit $p_{ij} = \frac{1}{36}$.

Die Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ ordnet jedem Wurf die Summe der beiden Werte zu, also $X(ij) = i + j$. Weiter oben haben wir bereits die einem Bild $k \in \mathbb{R}$ zugeordneten Ereignisse gesammelt und können nun die Wahrscheinlichkeiten $P(X = k)$ ergänzen.

k	$E_k = \{e \mid X(e) = k\}$	$\#E_k$	$P(X = k)$
2	{11}	1	$\frac{1}{36}$
3	{12, 21}	2	$\frac{1}{18}$
4	{13, 22, 31}	3	$\frac{1}{12}$
5	{14, 23, 32, 41}	4	$\frac{1}{9}$
6	{15, 24, 33, 42, 51}	5	$\frac{5}{36}$
7	{16, 25, 34, 43, 53, 61}	6	$\frac{1}{6}$
8	{26, 35, 44, 53, 62}	5	$\frac{5}{36}$
9	{36, 45, 54, 63}	4	$\frac{1}{9}$
10	{46, 55, 64}	3	$\frac{1}{12}$
11	{56, 65}	2	$\frac{1}{18}$
12	{66}	1	$\frac{1}{36}$

Es gilt z. B.

- $P(3 < X < 7) = P(X = 4) + P(X = 5) + P(X = 6) = \frac{4}{9}$

Die Wahrscheinlichkeit, dass die Augensumme größer als 3 und kleiner als 7 ist, beträgt ca. 44%.

- $P(X \leq 10) = 1 - P(X > 10) = 1 - (P(X = 11) + P(X = 12)) = \frac{11}{12}$

Die Wahrscheinlichkeit, dass die Augensumme höchstens 10 ist, beträgt ca. 92%.