Das starke Gesetz der großen Zahlen

Beispiel: Münzwurf

Der betrachtete Zufallsversuch sei das einmalige Werfen einer Münze. Die beiden möglichen Ergebnisse seien *Kopf* und *Zahl*. Diese beiden Ergebnisse werden wir auch mit H und T abkürzen (entsprechend der englischsprachigen Ausdrücke *heads* und *tails*). Nach jedem Münzwurf soll die relative Häufigkeit von H gebildet werden, indem die Anzahl der bisher aufgetretenen Ergebnisse H durch die Anzahl aller bisherigen Versuchsdurchführungen geteilt wird. (Wir hätten uns auch für die relativen Häufigkeiten von T entscheiden können, was zu analogen Überlegungen führen würde.) Werfen wir also eine Münze n-mal ($n \in \mathbb{N}$), so entsteht gleichzeitig eine Folge der Länge n von relativen Häufigkeiten.

Stellen wir uns nun vor, eine Münze werde unbegrenzt oft geworfen. Dann besteht die Ergebnismenge aus allen unendlichen Folgen aus H und T. Zu jedem Ergebnis gibt es eine zugehörige unendliche Folge $(h_1(H), h_2(H), h_3(H), \ldots)$ von relativen Häufigkeiten von H (die wir im Folgenden kurz mit $(h_i(H))_{i \in \mathbb{N}}$ bezeichnen werden). Unter diesen Folgen relativer Häufigkeiten gibt es welche, die einen Grenzwert haben, die also analytisch konvergieren, und unter diesen konvergierenden Folgen gibt es welche, deren Grenzwert gleich $0,5$ ist.

Sei nun $\mathscr{M}$ die Menge, die aus allen Ergebnissen besteht, deren zugehörige Folgen relativer Häufigkeiten gegen $0,5$ konvergieren und sei $\overline{\mathscr{M}}$ die Menge, die aus allen anderen Ergebnissen besteht.

Stellen wir uns weiter vor, alle Elemente der beiden Mengen befänden sich in einer großen Box und wir zögen zufällig ein Element. Das starke Gesetz der großen Zahlen besagt nun: Die Wahrscheinlichkeit, ein Element aus $\mathscr{M}$ zu ziehen, ist gleich 1 und die Wahrscheinlichkeit, ein Element aus $\overline{\mathscr{M}}$ zu ziehen, ist gleich 0, also

$$\mathbf{P}(\mathscr{M}) = 1 \ \wedge \ \mathbf{P}\left(\overline{\mathscr{M}}\right) = 0$$

Das starke Gesetz der großen Zahlen

Schauen wir uns das starke Gesetz der großen Zahlen etwas genauer an, um zu verstehen, dass die obige Behauptung tatsächlich in diesem Gesetz steht.

Die Formulierung des starken Gesetzes der großen Zahlen lautet bei Wikipedia:

> Ist eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsvariablen gegeben, so sagt man, dass diese Folge dem starken Gesetz der großen Zahlen genügt, wenn der Mittelwert der zentrierten Zufallsvariablen
>
> $$\overline{X}_n := \frac{1}{n} \sum_{i=1}^{n} (X_i - E(X_i))$$
>
> fast sicher gegen 0 konvergiert. Das bedeutet, dass
>
> $$P\left(\lim_{n \to \infty} \overline{X}_n = 0\right) = 1 \text{ ist.}$$

(...) Ist eine unabhängige Folge von Zufallsvariablen $(X_n)_{n\in\mathbb{N}}$ mit endlicher Varianz gegeben und gilt

$$\sum_{i=1}^{\infty} \frac{\operatorname{Var}(x_n)}{n^2} < \infty$$

so genügt $(X_n)_{n\in\mathbb{N}}$ dem starken Gesetz der großen Zahlen. Dabei wird die obige Bedingung auch *Kolmogorow-Bedingung* genannt.

Überlegen wir uns nun, was dieses Gesetz über den Münzwurf aussagt. Ohne weiteren Beweis werden wir dabei davon ausgehen, dass die Kolmogorov-Bedingung für den Münzwurf erfüllt ist.

Die Ergebnismenge beim einfachen Münzwurf ist $\Omega = \{H; T\}$.

Es gilt: $p(H) = 0,5$ und $p(T) = 0,5$.

Wir definieren nun 4 Zufallsvariablen auf Ω: $X_1; X_2; X_3; X_4$.

Für alle $i \in \{1; 2; 3; 4\}$ gelte:

$X_i(H) = 1$ und $X_i(T) = 0$.

Dann ist $X_1 + X_2 + X_3 + X_4 := X_{1;...;4}$ eine Funktion von $\Omega \times \Omega \times \Omega \times \Omega \longrightarrow \mathbb{R}$
mit $X_{1;...;4}(\omega_1; \omega_2; \omega_3; \omega_4) = X_1(\omega_1) + X_2(\omega_2) + X_3(\omega_3) + X_4(\omega_4)$
mit $\omega_i \in \{H; T\}$ für alle $i \in \{1; \ldots; 4\}$.

Z.B. ist

$X_{1;...;4}((H; T; T; H)) = X_1(H) + X_2(T) + X_3(T) + X_4(H) = 1 + 0 + 0 + 1 = 2$

Die Wahrscheinlichkeit des Wertes einer (diskreten) Zufallsvariable X definieren wir wie üblich durch das durch die Werte der Zufallsvariable auf der Grundmenge Ω des zugrunde liegenden Zufallsversuchs induzierte Wahrscheinlichkeitsmaß, also:
$P(X = r) := p(\{\omega \in \Omega | X(\omega) = r\})$ für beliebiges $r \in \mathbb{R}$.

Die Zufallsvariablen $X_1; \ldots; X_4$ sollen unabhängig sein; es soll also gelten:

$$\begin{aligned} & P(X_1 = r_1 \wedge X_2 = r_2 \wedge X_3 = r_3 \wedge X_4 = r_4) \\ = \; & P(X_1 = r_1) \cdot P(X_2 = r_2) \cdot P(X_3 = r_3) \cdot P(X_4 = r_4) \end{aligned}$$

für beliebige Zahlen $r_1; \ldots r_4 \in \mathbb{R}$.
Z.B. gilt:

$$\begin{aligned} & P(X_1 = 1 \wedge X_2 = 0 \wedge X_3 = 0 \wedge X_4 = 1\ ) \\ = \; & P(X_1 = 1) \cdot P(X_2 = 0) \cdot P(X_3 = 0) \cdot P(X_4 = 1\ ) \\ = \; & 0,5 \cdot 0,5 \cdot 0,5 \cdot 0,5 \\ = \; & \tfrac{1}{16} \end{aligned}$$

Auf diese Weise hat das 4er-Tupel $(H; T; T; H)$ eine Wahrscheinlichkeit zugeordnet

bekommen.
Für unsere Zufallsvariablen X_i gilt:
$E(X_i) = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = 1 \cdot p(H) + 0 \cdot p(T) = 1 \cdot 0,5 + 0 \cdot 0,5 = 0,5.$
Die Subtraktion einer Zahl von einer Zufallsvariable ist punktweise definiert.
Sei also $r \in \mathbb{R}$, dann ist $X - r$ definiert durch

$$(X - r)(\omega) := X(\omega) - r.$$

Da $E(X_i)$ für alle $i \in \{1; \ldots; 4\}$ eine Zahl ist, gilt:

$$(X_i - E(X_i))(H) = X_i(H) - E(X_i) = 1 - 0,5 = 0,5$$

und

$$(X_i - E(X_i))(T) = X_i(T) - E(X_i) = 0 - 0,5 = -0,5.$$

Ebenso ist die Multiplikation einer Zufallsvariable mit einer Zahl punktweise definiert, also gilt für unsere Zufallsvariablen z.B.

$$\frac{1}{4} \cdot X_i(H) = \frac{1}{4} \cdot 1 = \frac{1}{4}$$

Dann ist z.B.

$$\frac{1}{4} \sum_{i=1}^{4} (X_i - E(X_i))((H;T;T;H)) = \frac{1}{4} \cdot (0,5 - 0,5 - 0,5 + 0,5) = 0.$$

Wir definieren: $\overline{X}_4 := \frac{1}{4} \sum_{i=1}^{4} (X_i - E(X_i)).$
Nun können wir uns auch vorstellen, was $P\left(\overline{X}_4 = 0\right)$ bedeuten soll. Es ist die Summe aller Wahrscheinlichkeiten derjenigen 4er-Tupel aus H und T, die genauso viele H wie T enthalten.
Diese Definition können wir auf beliebige natürliche Zahlen $n \in \mathbb{N}$ ausweiten und erhalten

$$\overline{X}_n := \frac{1}{n} \sum_{i=1}^{n} (X_i - E(X_i)).$$

$\overline{X}_n$ heißt der Mittelwert der zentrierten Zufallsvariable.

Wenn wir die Münze einmal werfen, können wir $\overline{X}_1$ bilden. Wenn wir die Münze zweimal werfen, können wir $\overline{X}_2$ bilden und mit jedem weiteren Wurf können wir $\overline{X}_3$, $\overline{X}_4$ usw. bilden. So entsteht dann eine Folge von Mittelwerten zentrierter Zufallsvariablen.
Lassen wir die Anzahl n der Münzwürfe über alle Grenzen wachsen, entstehen unendliche Folgen von $\overline{X}_n$, die einen Grenzwert haben können. Dabei sind für uns *die* Folgen der $\overline{X}_n$ interessant, für die gilt:

$$\lim_{n \to \infty} \overline{X}_n = 0.$$

Um uns ganz auf die Bedeutung des starken Gesetzes der großen Zahlen konzentrieren zu können, wollen wir nun so viele Formalitäten wie möglich weglassen. Im Fall des mehrmaligen Münzwurfs können wir uns eine Folge von $\overline{X}_1$, $\overline{X}_2$, $\overline{X}_3$, ... so vorstellen: Wir werfen die Münze einmal und bilden die relative Häufigkeit $h_1(H)$ von z.B. H. (Wir könnten auch T verwenden, was aber auf eine analoge Betrachtung hinaus laufen würde.) $h_1(H)$ ist dann entweder gleich 1 oder gleich 0. Dann werfen wir die Münze noch einmal und bilden $h_2(H)$, also die relative Häufigkeit von H in den ersten beiden Würfen. Dann werfen wir noch einmal und bilden $h_3(H)$ usw.
Wird die Münze (in unserer Vorstellung) endlos oft geworfen, entsteht eine unendliche Folge relativer Häufigkeiten, die einen Grenzwert haben kann. Dabei gilt:

$$\lim_{n\to\infty} h_n(H) = 0,5 \Leftrightarrow \lim_{n\to\infty} \overline{X}_n = 0.$$

Interessant sind also *die* Folgen von H und T, deren Grenzwert der relativen Häufigkeiten von H gleich $0,5$ ist. Dazu gehören Folgen wie

$(H;T;H;T;H;T;\ldots)$ oder auch
$(T;T;T;H;H;H;T;T;T;H;H;H;\ldots)$.

Es gibt aber auch Folgen, die nicht dazu gehören, weil deren relative Häufigkeit von H nicht gegen $0,5$, sondern gegen eine andere Zahl konvergiert, wie z.B.

$(H;H;H;H;H;H;\ldots)$ oder
$(T;H;H;T;H;H;T;H;H;T;\ldots)$ oder
$(T;H;H;T;H;H;H;T;H;H;H;H;T;\ldots)$.

Es gibt sogar unendlich viele Folgen, deren relative Häufigkeiten gegen eine andere Zahl als $0,5$ konvergieren.
Außerdem gibt es Folgen, deren relative Häufigkeiten gar nicht konvergieren. Solche Folgen können wir uns so vorstellen: Ab einem bestimmten $n \in \mathbb{N}$ tritt in den nächsten $2n$ Würfen nur noch H ein. Dann ist die relative Häufigkeit von H größer oder gleich $\frac{2}{3}$. Anschließend tritt in den nächsten $2 \cdot 3n$ Würfen nur noch T ein. Die relative Häufigkeit von H ist dann kleiner oder gleich $\frac{1}{3}$. Danach tritt in den nächsten $2 \cdot 3^2 n$ Würfen nur noch H ein usw.
Auch von diesen Folgen, deren relative Häufigkeiten gar nicht konvergieren, gibt es unendlich viele.

Betrachten wir nun wieder die Menge der unendlichen Folgen, für die gilt:

$$\lim_{n\to\infty} \overline{X}_n = 0.$$

Im starken Gesetz der großen Zahlen ist von der Wahrscheinlichkeit P die Rede, die die Menge *der* Folgen der $\overline{X}_n$ hat, deren Grenzwert 0 ist. Wir brauchen also noch für unseren Münzwurf eine anschauliche Vorstellung, was mit dieser Wahrscheinlichkeit gemeint ist. Dazu stellen wir uns vor, dass wir alle unendlich langen Folgen von H und T in einem Behälter haben und wir nun eine dieser Folgen zufällig aus diesem herausgreifen. Um nicht in den Tiefen der Axiomatik und der Mengenlehre zu versinken, gehen wir einfach davon aus, dass wir eine Folge aus dem Behälter zufällig

auswählen können und wir befassen uns auch nicht damit, warum dem P aus dem starken Gesetz ein Zufallsversuch zugrunde liegt, den wir mit dem zufälligen Ziehen einer Folge aus diesem Behälter modellieren können, sondern wir stellen gleich die nächste - und entscheidende - Frage:
Wie groß ist die Wahrscheinlichkeit, mit der wir eine Folge ziehen, deren Grenzwert relativer Häufigkeiten $h_n(H)$ gleich $0,5$ ist?
Das starke Gesetz der großen Zahlen sagt: $P\left(\lim\limits_{n\to\infty} \overline{X}_n = 0\right) = 1$.
Auf unser Beispiel übertragen bedeutet das: Die Wahrscheinlichkeit, eine Folge mit dem Grenzwert $0,5$ zu ziehen, ist gleich 1.

Unterschied zwischen starkem und schwachem Gesetz

Das schwache Gesetz besagt, dass die Wahrscheinlichkeit außerhalb einer beliebigen – auch sehr kleinen – ε-Umgebung um $0,5$ kleiner wird als jede positive Zahl. Das starke Gesetz besagt, dass wirklich nur noch die Folgen aus H und T übrig bleiben, deren relative Häufigkeit gegen $0,5$ konvergiert. In der Abbildung liegen alle diese Ergebnisse an einer einzigen Stelle auf der x-Achse. Anschaulich können wir uns das so vorstellen: Das schwache Gesetz der großen Zahlen behauptet, dass die Wahrscheinlichkeit außerhalb jeder beliebigen ε-Umgebung um $0,5$ gegen 0 geht. Das starke Gesetz behauptet, dass das auch dann passiert, wenn ε gegen 0 geht.

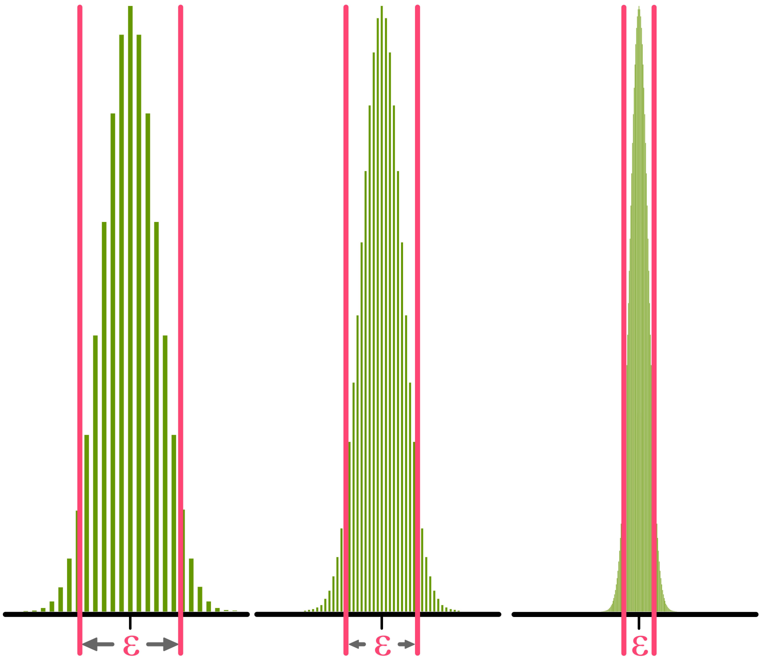

Abb. 1 Starkes Gesetz der großen Zahlen

In dieser Darstellung können wir uns den Unterschied zwischen schwachem und starkem Gesetz der großen Zahlen so vorstellen: Das schwache Gesetz der großen Zahlen besagt: Immer wenn wir uns ein bestimmtes $\varepsilon > 0$ mit $\varepsilon \in \mathbb{R}$ hernehmen, es um die Wahrscheinlichkeit eines Ereignisses legen und die Anzahl $n \in \mathbb{N}$ der Versuchsdurchführungen über alle Grenzen wächst, konvergiert das Verhältnis der Anzahl der Ergebnisse außerhalb der ε-Umgebung zu der Anzahl der Ergebnisse innerhalb der ε-Umgebung gegen 0.

Das starke Gesetz der großen Zahlen behauptet nun zusätzlich, dass das auch dann gilt, wenn ε gegen 0 geht.

Wir Menschen können ganz gut mit endlichen Ergebnismengen umgehen und können auf dieser Grundlage auch Wahrscheinlichkeiten berechnen. Bei unendlichen Mengen mag das aber ganz anders aussehen und möglicherweise kommen wir zu Resultaten, die wir nicht erwarten. Deshalb sei an dieser Stelle auf zwei Zusammenhänge verwiesen, die ganz anders sind, als wir es vom Rechnen im Endlichen gewohnt sind.

1) Wir können uns Fragen, wie wahrscheinlich es ist, dass beim mehrmaligen Münzwurf die relative Häufigkeit von H genau gleich z. B. $0{,}7$ ist. Wir werden feststellen, dass diese Wahrscheinlichkeit gleich 0 ist, wenn die Anzahl n der Versuchsdurchführungen nicht durch 10 teilbar ist. Betrachten wir nur Anzahlen von Versuchsdurchführungen der Form $10n$ (also nur Zahlen, die durch 10 teilbar sind), sehen wir zwar positive Wahrscheinlichkeiten, die aber gegen 0 konvergieren, falls n über alle Grenzen wächst. Wir können also behaupten: Die Wahrscheinlichkeit, mit

der die relative Häufigkeit genau gleich $0,7$ ist, konvergiert gegen 0.

Das gilt nicht nur für $0,7$, sondern entsprechend auch für alle anderen (rationalen) Zahlen, z.B. auch für $0,5$. Das heißt: Obwohl wir gerade eben mit dem starken Gesetz der großen Zahlen gezeigt haben, dass sich die gesamte Wahrscheinlichkeit (bildlich gesprochen) auf den Punkt $0,5$ auf der x-Achse konzentriert (wenn die Anzahl n der Versuche immer größer wird), geht trotzdem die Wahrscheinlichkeit, mit der die relative Häufigkeit genau gleich $0,5$ ist, gegen 0.

2) Wenn wir eine Münze einmal werfen, haben wir 2^1 mögliche Ergebnisse. Wenn wir eine Münze zweimal werfen, haben wir 2^2 mögliche Ergebnisse und wenn wir die Münze n-mal werfen, haben wir 2^n mögliche Ergebnisse. Da jeweils alle Ergebnisse die gleiche Wahrscheinlichkeit haben, wird mit zunehmender Anzahl von Versuchsdurchführungen die Wahrscheinlichkeit für jedes einzelne Ergebnis immer kleiner und konvergiert sogar gegen 0, falls n gegen unendlich geht. Betrachten wir also die Ergebnismenge, welche aus allen unendlichen Folgen aus H und T besteht, haben wir nur Ergebnisse vor uns, deren Wahrscheinlichkeit gleich 0 ist, was im endlichen Fall überhaupt keinen Sinn ergeben würde.

Dass alle Ergebnisse die Wahrscheinlichkeit 0 haben, ist sogar der übliche Fall, wenn wir Ergebnismengen betrachten, die überabzählbar viele Elemente enthalten. Dennoch muss die Wahrscheinlichkeit nicht „überall gleich“ sein: Wir können mit einer Dichtefunktion angeben, welche Bereiche eines Intervalls höhere und welche Bereiche niedrigere Wahrscheinlichkeitsdichten aufweisen sollen (siehe dazu das Beispiel weiter unten). Es gibt also Ergebnisse, über denen die Wahrscheinlichkeitsdichte größer ist als über anderen Ergebnissen, obwohl die Wahrscheinlichkeiten aller Eregebnisse gleich 0 sind.

Fehlinterpretationen

1) Wenden wir uns wieder den beiden Eingangs erwähnten Mengen $\mathscr{M}$ und $\overline{\mathscr{M}}$ zu. $\mathscr{M}$ ist die Menge, die aus allen Folgen aus H und T besteht, deren zugehörige Folgen relativer Häufigkeiten gegen $0,5$ konvergieren und $\overline{\mathscr{M}}$ besteht aus allen anderen Folgen. Alle Elemente beider Mengen befinden sich in einem Behälter und der Zufallsversuch ist das einfache Ziehen einer Folge. Das starke Gesetz der großen Zahlen besagt, dass die Wahrscheinlichkeit, eine Folge zu ziehen, die in $\mathscr{M}$ ist, gleich 1 ist.

Eine beliebte Fehlinterpretation dieser Aussage ist, dass die Wahrscheinlichkeit von 1 bedeutet, dass im Behälter nur Folgen sind, deren relative Häufigkeiten von H gegen $0,5$ konvergieren, dass also $\overline{\mathscr{M}}$ leer ist. Wie wir aber oben gesehen haben, gibt es sogar unendlich viele Folgen im Behälter, deren relative Häufigkeiten nicht gegen $0,5$ konvergieren und auch unendlich viele, deren relative Häufigkeiten gar nicht konvergieren. Die Aussage $P\left(\lim\limits_{n\to\infty}\overline{X}_n = 0\right) = 1$ bedeutet insbesondere auch nicht, dass jede Folge von Münzwürfen „am Ende“ doch noch genau 50% H aufweisen muss.

Selbstverständlich ist damit auch jede Verallgemeinerung dieser Fehlinterpretation falsch wie z. B. die folgende: Jede Folge relativer Häufigkeiten eines Ereignisses konvergiert gegen die Wahrscheinlichkeit des Ereignisses.

2) Dass die Wahrscheinlichkeit, eine Folge aus $\mathscr{M}$ zu ziehen, gleich 1 ist, bedeutet nicht, dass wir keine andere Folge zufällig aus dem Behälter ziehen können, denn *jede* Folge, die sich im Behälter befindet, hat die Wahrscheinlichkeit von exakt 0. D.h.: *Immer*, wenn wir eine Folge ziehen, tritt ein Ergebnis mit der Wahrscheinlichkeit 0 ein.

In gewisser Weise läuft dieses starke Gesetz der großen Zahlen unserem gesunden Menschenverstand zuwider. Das könnte daran liegen, dass wir „im Endlichen“ denken. In diesem Fall gilt: Hat ein Ereignis die Wahrscheinlichkeit 1, dann enthält es alle Ergebnisse der Ergebnismenge, deren Wahrscheinlichkeit größer als 0 ist und hat ein Ereignis die Wahrscheinlichkeit 0, dann kann es nicht eintreten. Z. B. können wir für den Zufallsversuch „Einmaliges Werfen eines Würfels“ die Ergebnismenge $\{1;2;3;4;5;6;7;8\}$ vereinbaren. Das Ereignis $\{1;2;3;4;5;6\}$ hat dann die Wahrscheinlichkeit 1 und das Ereignis $\{7;8\}$ hat die Wahrscheinlichkeit 0. Es kann nicht eintreten, weil der Würfel keine Seiten hat, die mit 7 oder 8 beschriftet werden können. Für Ergebnismengen mit überabzählbar vielen Elementen, wie z.B. die Menge aller unendlichen Folgen aus H und T, gilt das aber nicht.

Um eine gute Vorstellung davon zu haben, warum das so ist, können wir uns ein typisches Beispiel ansehen, welches in der Schulmathematik vorkommt. Dabei geht es darum, zufällig eine reelle Zahl auszuwählen. Die Wahrscheinlichkeit soll nicht überall gleich sein, sondern soll durch eine Funktion angegeben werden, die Wahrscheinlichkeitsdichte genannt wird. Beispielhaft ist in Abb. 2 eine typische Wahrscheinlichkeitsdichtefunktion auf dem Intervall $[0;1]$ dargestellt. Die Wahrscheinlichkeit, mit der sich die zufällig gewählte Zahl in dem Intervall $[0,4;0,5]$ befindet, ist dann gleich der grün markierten Fläche, also dem Wert des bestimmten Integrals der Wahrscheinlichkeitsdichte in den Grenzen von $0,4$ bis $0,5$.

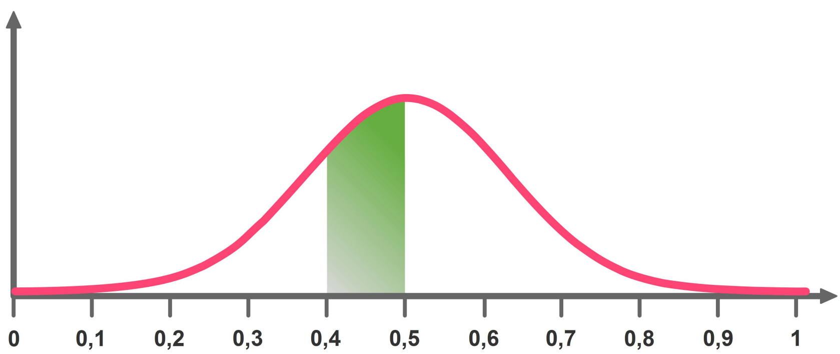

Abb. 2 Wahrscheinlichkeitsdichte

Weil jede einzelne Zahl – also jeder Punkt auf der Zahlengerade – keine Länge hat, ist das Integral über dieser Zahl gleich 0. Damit hat jede einzelne Zahl die Wahrscheinlichkeit von exakt 0 und deshalb gilt: *Immer*, wenn eine Zahl zufällig ausgewählt wird, tritt ein Ergebnis ein, welches die Wahrscheinlichkeit 0 hat. Wir können auch mehrere einzelne Zahlen oder sogar abzählbar unendlich viele einzelne Zahlen (wie z. B. alle rationalen Zahlen) zu einem Ereignis zusammenfassen, ohne dass die Wahrscheinlichkeit dieses Ereignisses größer als 0 werden würde. Erst wenn wir alle (überabzählbar vielen) Zahlen eines Intervalls mit einer Länge, die größer als 0 ist, zu einem Ereignis zusammenfassen, steigt die Wahrscheinlichkeit des Ereignisses über 0.

So entsteht auch die intuitiv schwer zugängliche Tatsache, dass die gesamte Ergebnismenge zwar die Wahrscheinlichkeit 1 hat, diese Menge aber nicht etwa aus allen Ergebnissen besteht, deren Wahrscheinlichkeit größer als 0 ist, sondern nur aus Ergebnissen besteht, deren Wahrscheinlichkeit exakt gleich 0 ist.

3) Eine weitere Fehlinterpretation vermengt quasi die beiden oben angesprochenen Fehler. Die Argumentation ist: Da die Wahrscheinlichkeit, eine Folge zu ziehen, deren relative Häufigkeiten gegen $0,5$ konvergieren, gleich 1 ist, kann man beim konkreten mehrmaligen Münzwurf sicher sein, ein Anfangsstück einer Folge erwischt zu haben, deren relative Häufigkeiten gegen $0,5$ konvergieren.

Das ist also wieder die Behauptung, die relative Häufigkeit müsse sich irgendwann doch noch bei $0,5$ einpendeln. Aber zum einen sagt das starke Gesetz der großen Zahlen nichts über endliche viele Versuchsdurchführungen aus und zum anderen liegt hier wohl der Denkfehler vor, man könne einem endlichen Anfangsstück einer Folge von Ergebnissen eines wiederholten Münzwurfs ansehen, ob es sich um eine konvergente Folge handele. Da aber jedes noch so lange, endliche Anfangsstück einer Folge auf deren Konvergenzverhalten überhaupt keinen Einfluss hat, kann das nicht sein.

Zudem hat das Werfen einer Münze nichts mit der Auswahl einer unendlichen Folge aus einer überabzählbaren Menge von Folgen zu tun. Beides hat auf so vielen Ebenen nichts miteinander zu tun, dass hier nur ein kurzer Hinweis genannt werden soll: Wenn man mit dem Werfen einer Münze eine Auswahl aus einer überabzählbaren Menge durchführen könnte, bräuchten wir vermutlich das Auswahlaxiom nicht mehr.