Das schwache Gesetz der großen Zahlen

Auch weil immer wieder behauptet wird, das schwache Gesetz der großen Zahlen stelle das empirische Gesetz der großen Zahlen sicher oder es stelle sicher, dass die relative Häufigkeit gegen die Wahrscheinlichkeit konvergiere, sei hier dargestellt, welche Aussagen dieses Gesetz tatsächlich macht.

Das schwache Gesetz der großen Zahlen lautet:

$$\lim_{n\to\infty} P\left(|h_n(E) - p| > \varepsilon\right) = 0$$

dabei ist $h_n(E)$ die relative Häufigkeit eines Ereignisses E nach n Versuchsdurchführungen, P die Wahrscheinlichkeit des Ereignisses, das in der Klammer beschrieben ist, p die Wahrscheinlichkeit des Ereignisses E und ε ist eine beliebige reelle Zahl größer als 0.

In Worten bedeutet das: Die Wahrscheinlichkeit P, dass der Betrag der Differenz der relativen Häufigkeit $h_n(E)$ eines Ereignisses und der Wahrscheinlichkeit p dieses Ereignisses größer ist als eine Zahl ε, konvergiert gegen 0, falls die Anzahl der Versuchsdurchführungen n gegen unendlich geht.

Etwas einfacher – wenn auch ungenauer – ist folgende Formulierung: Die Wahrscheinlichkeit P, dass die relative Häufigkeit $h_n(E)$ eines Ereignisses nicht in der Nähe der Wahrscheinlichkeit des Ereignisses liegt, geht gegen 0, falls der Versuch immer weiter durchgeführt wird.

Noch einfacher gesagt sehen wir hier: Es wird immer wahrscheinlicher, dass die relative Häufigkeit eines Ereignisses in der Nähe der Wahrscheinlichkeit dieses Ereignisses liegt, wenn wir den Zufallsversuch immer öfter wiederholen.

Wenden wir dieses Gesetz nun auf den Münzwurf an:
Das Ereignis E soll als einziges Element das Ergebnis H enthalten. Dann ist die Wahrscheinlichkeit p dieses Ereignisses gleich $0,5$. Um uns die Sache konkret vorstellen zu können, wählen wir für n eine bestimmte Zahl, sagen wir 20. Wir betrachten also den 20-fachen Münzwurf. Jede Folge aus H und T der Länge 20 ist ein mögliches Ergebnis e des 20-fachen Münzwurfs. Jeder dieser Folgen ordnen wir nun die relative Häufigkeit von H zu. Die Folge

$(H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H)$
hat die relative Häufigkeit
$h_{20}(H) = \frac{20}{20} = 1,$
die Folge
$(T,T;H;H;H;T,H;T;T;T;H;H;T;H;H;H;H;T;T;H)$
hat die relative Häufigkeit
$h_{20}(H) = \frac{11}{20} = 0,55$
und die Folge
$(T;T;T;T;T;T;T;T;T;T;T;T;T;T;T;T;T;T;T;T)$
hat die relative Häufigkeit
$h_{20}(H) = \frac{0}{20} = 0.$

Wir wählen nun ein ε aus. Laut schwachem Gesetz der großen Zahlen soll dieses ε irgendeine Zahl sein, die größer als 0 ist. Wir entscheiden uns für $0,1$.
Mit Hilfe von $\varepsilon = 0,1$ und der relativen Häufigkeit $h_{20}(H)$ jedes Ergebnisses können wir nun Ereignisse definieren, z. B. das Ereignis, das die Ergebnisse enthält, deren relative Häufigkeit von H um mehr als $0,1$ von der Wahrscheinlichkeit $p(H) = 0,5$ abweicht. Dann sind z.B. die Ergebnisse

$(H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H;H)$ und
$(H;H;H;H;H;H;H;H;T;H;H;H;H;H;H;H;T;H;H)$
Elemente dieses Ereignisses und
$(T,T;H;H;H;T,H;T;T;T;H;H;T;H;H;H;H;T;T;H)$

ist nicht Element dieses Ereignisses, weil die relative Häufigkeit $h_{20}(H) = \frac{11}{20} = 0,55$ nicht um mehr als $0,1$ von $0,5$ abweicht.
Dieses Ereignis wollen wir folgendermaßen bezeichnen:

$$\left|h_{20}(H) - 0,5\right| > 0,1$$

Jedes Ergebnis e des 20-fachen Münzwurfs - also jede 20er-Folge aus H und T - hat die Wahrscheinlichkeit

$$p(e) = \frac{1}{2^{20}}.$$

Da alle Ergebnisse die gleiche Wahrscheinlichkeit haben, können wir einem Ereignis dessen Wahrscheinlichkeit zuordnen, indem wir die Anzahl der Ergebnisse in diesem Ereignis durch 2^{20} teilen.
So hat auch unser Ereignis

$\left|h_{20}(H) - 0,5\right| > 0,1$ eine Wahrscheinlichkeit, die wir mit

$P\left(\left|h_{20}(H) - 0,5\right| > 0,1\right)$ bezeichnen wollen.

Wir können nun eine Folge von Wahrscheinlichkeiten bilden, indem wir für n die Zahlen $21;22;23;\ldots$ einsetzen. Diese Folge hat dann die Folgenglieder

$$P(|h_{21}(H) - 0,5| > 0,1),$$
$$P(|h_{22}(H) - 0,5| > 0,1),$$
$$P\left(\left|h_{23}(H) - 0,5\right| > 0,1\right) \text{ usw.}$$

Würden wir diese Wahrscheinlichkeiten ausrechnen, würden wir feststellen, dass sie immer kleiner werden. Wir würden sogar feststellen, dass diese Wahrscheinlichkeiten der 0 beliebig nahe kommen, wenn wir nur n groß genug wählen. Das heißt dann nichts anderes, als dass der Grenzwert dieser Folge von Wahrscheinlichkeiten gleich 0 ist. Diese Tatsache bezeichnen wir mit

$$\lim_{n\to\infty} P(|h_n(H) - 0,5| > 0,1) = 0.$$

Damit haben wir das, was das schwache Gesetz der großen Zahlen aussagt, wenn E das Ereignis ist, dass beim Münzwurf H eintritt, $p = 0,5$ ist und $\varepsilon = 0,1$ gewählt wird.

Umgangssprachlich formuliert bedeutet das, dass die Wahrscheinlichkeit, mit der die relative Häufigkeit von H um mehr als $0,1$ von der Wahrscheinlichkeit von H – nämlich $0,5$ – abweicht, immer geringer wird und sogar gegen 0 geht, wenn wir den Münzwurf nur häufig genug wiederholen.

In den folgenden Abbildungen sehen wir zu einem willkürlich gewählten ε ein paar Verteilungen der möglichen Ergebnisse in Abhängigkeit von der Anzahl der Versuchsdurchführungen. Je häufiger der Versuch durchgeführt wird, desto größer wird der Anteil der Ergebnisse, die innerhalb der ε-Umgebung liegen. Entsprechend steigt mit zunehmender Anzahl der Versuchsdurchführungen die Wahrscheinlichkeit, ein Ergebnis zu werfen, das innerhalb der ε-Umgebung liegt.

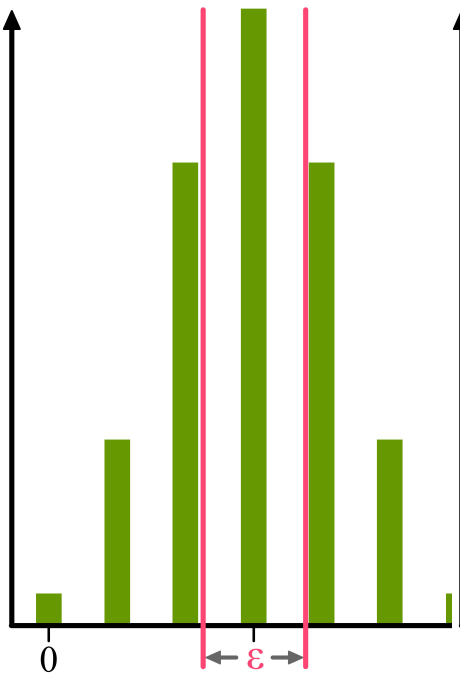

Abb. 1 6er Verteilung

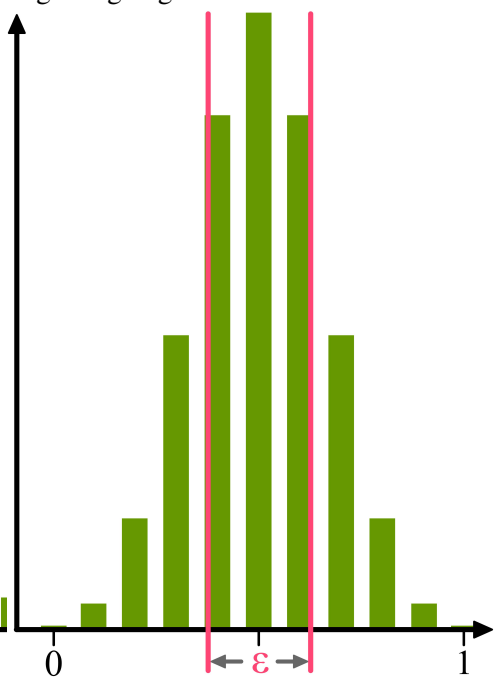

Abb. 2 10er Verteilung

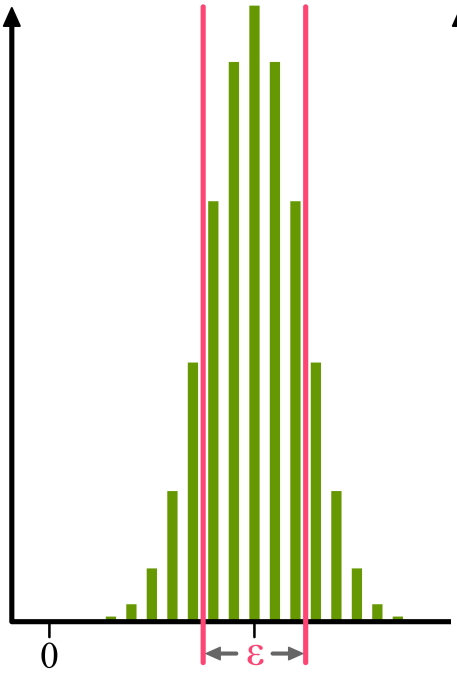

Abb. 3 20er Verteilung

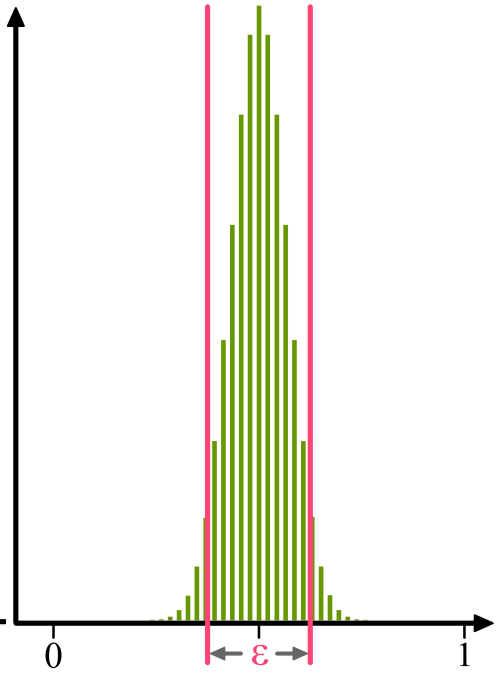

Abb. 4 40er Verteilung

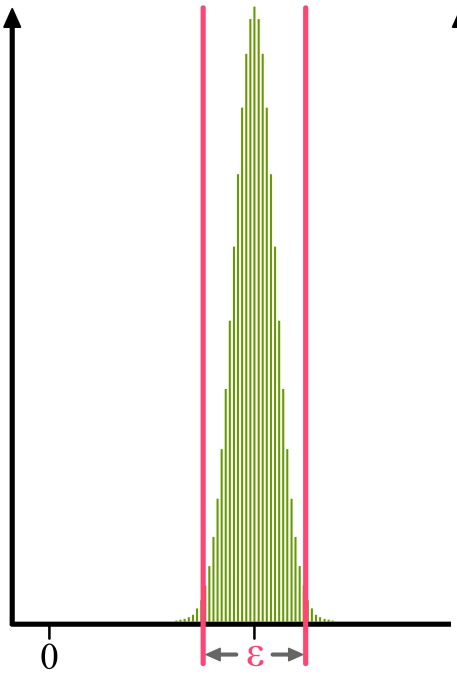

Abb. 5 100er Verteilung

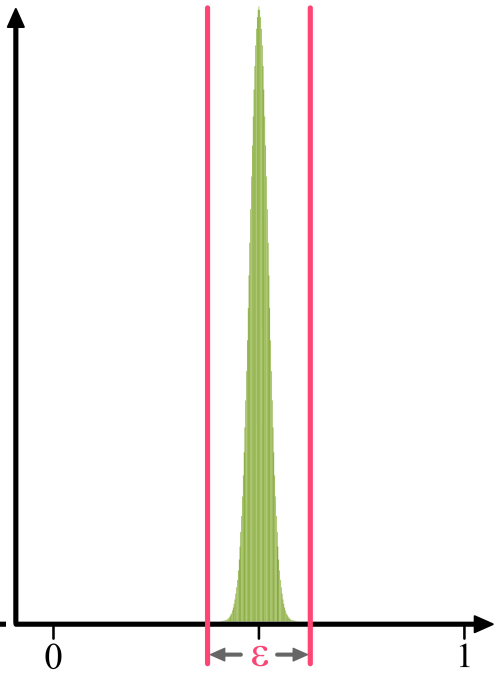

Abb. 6 500er Verteilung

Da wir ε frei wählen können, können wir nicht nur $0,1$, sondern auch z. B. $0,01$ oder $0,001$ usw. für ε einsetzen. Das bedeutet: Auch wenn wir die Abweichung der relativen Häufigkeit des Ergebnisses H von der Wahrscheinlichkeit von H sehr klein wählen, wird die Wahrscheinlichkeit für eine solche Abweichung immer geringer, falls wir den Zufallsversuch nur oft genug wiederholen. Im Gegenzug heißt das: Wenn die Anzahl der Durchführungen des Münzwurfs immer größer wird, wird die Wahrscheinlichkeit, mit der die relative Häufigkeit von H beliebig nahe bei $0,5$ liegt, auch immer größer und sie geht sogar gegen 1, wenn wir den Münzwurf unbegrenzt oft wiederholen.

Übrigens liegt die Begründung der Konvergenz in Wahrscheinlichkeit der relativen Häufigkeit bei mehr als zwei Ausgängen des zugrunde liegenden Zufallsversuchs in den Multinomialverteilungen: Je gleichmäßiger die Zahlen im Nenner sind, desto kleiner wird das Produkt der Fakultäten. Dann wird die Anzahl der Permutationen am größten. Haben wir z. B. die Kugeln $1;2;\ldots;6$ in der Box (also wie beim Würfeln mit einem idealen Würfel), berechnen wir die Wahrscheinlichkeiten der verschiedenen relativen Häufigkeit bezüglich des Zählmaßes (Zähldichte) mit der Mulitnomialverteilung

$$p(n_1;\ldots;n_6) = \binom{n}{n_1;\ldots;n_6}\left(\frac{1}{6}\right)^{n_1}\cdot\ldots\cdot\left(\frac{1}{6}\right)^{n_6}$$

wobei n die Anzahl der Ziehungen ist. Wegen $n_1+\ldots+n_6 = n$ kommen die unterschiedlichen Wahrscheinlichkeiten nur aufgrund der Multinomialkoeffizienten

$$\binom{n}{n_1;\ldots;n_6} = \frac{n!}{n_1!\cdot\ldots\cdot n_6!}$$

zustande.
Ziehen wir z.B. 60-mal, ist der Nenner des Multinomialkoeffizienten dann am kleinsten, wenn $n_1 = \ldots = n_6 = 10$ gilt.
Aber dann ist auch die relative Häufigkeit z.B. der „2“ gleich $\frac{1}{6}$.

1 Fehlinterpretationen

1) Die häufigste Fehlinterpretation ist die Behauptung, durch dieses Gesetz werde die Gültigkeit des empirischen Gesetzes der großen Zahlen sichergestellt. Das ist aber definitiv nicht der Fall, denn das schwache Gesetz der großen Zahlen behauptet nur, dass die *Wahrscheinlichkeit*, dass die relative Häufigkeit eines Ereignisses in der Nähe der Wahrscheinlichkeit dieses Ereignisses liegt, mit zunehmender Anzahl von Versuchsdurchführungen immer größere wird und gegen 1 konvergiert. Es behauptet *nicht*, dass die relative Häufigkeit eines Ereignisses sich der Wahrscheinlichkeit des Ereignisses annähern *muss*.

Auch um diesen Unterschied deutlich zu machen, gibt es sogar unterschiedliche Begriffe für die jeweiligen Konvergenzen. Da das schwache Gesetz der großen Zahlen die Konvergenz der Wahrscheinlichkeit (der Nähe der relativen Häufigkeit eines

Ereignisses zur Ereigniswahrscheinlichkeit) behauptet, nennt man sie „*Konvergenz in Wahrscheinlichkeit*“ oder auch „*Konvergenz nach Wahrscheinlichkeit*“. Demgegenüber gibt es die analytische Konvergenz, also die Konvergenz wie wir sie aus der Analysis gewohnt sind. So konvergiert z.B. die Folge $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ gegen 0 und zwar so, dass jedes Folgenglied der 0 näher liegt als das vorhergehende Folgeglied. Sagt man Schülern, das empirische Gesetz der großen Zahlen stelle sicher, dass sich die relative Häufigkeit der Wahrscheinlichkeit annähern müsse, verstehen sie diese Annäherung in der Regel als analytische Konvergenz, die ja von dem schwachen Gesetz der großen Zahlen eben *nicht* behauptet wird.

2) Eine andere Fehlinterpretation ist die Vorstellung: Da die Wahrscheinlichkeit für die Annäherung von relativer Häufigkeit und Ereigniswahrscheinlichkeit tatsächlich gegen 1 konvergiert, sei diese Annäherung irgendwann so wahrscheinlich, dass sie eintreten *muss*. Um zu verstehen, warum das falsch ist, können wir uns überlegen, was die Konvergenz der Wahrscheinlichkeit der Annäherung genau bedeutet: Zu jeder positiven Zahl $g \in \mathbb{R}$, die kleiner als 1 ist, gibt es eine Zahl $n_g \in \mathbb{N}$ für die gilt: Die Wahrscheinlichkeit, mit der ein Ergebnis in der ε-Umgebung liegt, ist für jede Anzahl von Versuchsdurchführungen, die größer als n_g ist, größer als g. Das bedeutet nicht, dass es irgendeine endliche Anzahl von Versuchsdurchführungen gibt, ab der die Wahrscheinlichkeit, mit der ein Ergebnis in der ε-Umgebung liegt, *gleich* 1 ist. Tatsächlich gibt es diese Anzahl für den Münzwurf auch nicht. Auf unsere Schaubilder angewendet heißt das: Für jede noch so große Anzahl n von Versuchsdurchführungen gibt es Ergebnisse, die außerhalb der ε-Umgebung liegen und zwar in der ganzen Bandbreite von 0 bis 1. Da der Zufallsversuch nicht weiß, welche Ergebnisse wir für extrem unwahrscheinlich halten, wird er das tun, was er immer macht: irgendein Ergebnis eintreten lassen - vielleicht auch eines außerhalb der ε-Umgebung.