# Kap. 2 Anhang 3: Das Gesetz der großen Zahl – Eine Kurzeinführung

Will man den quantitativen Wahrscheinlichkeitsbegriff anschaulich erläutern, kommt man nicht umhin, von *relativen Häufigkeiten* zu reden. Auch „subjektive“ Wahrscheinlichkeiten sind ohne einen solchen Bezug nicht zu denken: Eine Wahrscheinlichkeit von 50%, ob nun subjektiv empfunden oder durch genau protokollierte quantitative Erfahrungen begründet, *bedeutet*: Unter den gegebenen Umständen oder bei Berücksichtigung aller verfügbaren Informationen zu den wirksamen Randbedingungen ist in einem von zwei Fällen mit dem Eintreten des Ereignisses zu rechnen.

Es geht um ein sogenanntes *Zufallsexperiment*. Um einen Vorgang, dessen Anfangsbedingungen einerseits hinreichend genau wiederholbar sind, zugleich aber einen gewissen Spielraum für unterschiedliche Abläufe und Endergebnisse lassen. Und es gibt jeweils ein genau umrissenes *Ergebnis* des Experiments. Die Gesamtheit der möglichen Ergebnisse wird mathematisch durch eine Menge $\Omega$ repräsentiert. Oft hat man Grund, nicht das ganz konkrete Einzelergebnis zu registrieren, sondern seine Zugehörigkeit zu einer irgendwie logisch definierten Ergebnis-Klasse, einem „Ereignis“. Die Gesamtheit der sinnvollen Ereignisse bei einem Experiment ergibt, wenn man beliebige aussagenlogische Verknüpfungen bildet, eine *Algebra,* oder sogar, wenn noch Grenzübergänge hinzukommen, eine *Sigma-Algebra* $\mathfrak{A}$. Den Ereignissen werden *Wahrscheinlichkeiten* zugeordnet, und diese haben die anschaulich-inhaltliche Bedeutung relativer Häufigkeit der Ereignisse beim Wiederholen des modellierten Zufallsexperiments; siehe oben. Formal:

Gegeben ein Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$ und ein Ereignis $A$, sei das zugehörige Zufallsexperiment $n$-mal wiederholt und die *relative Häufigkeit* $h_n(A)$ des Ereignisses $A$ betrachtet. Dann hat $\mathrm{P}(A) = p$ die *anschauliche Bedeutung*: $h_n(A) \approx p$ für große $n$.

Wie passt man die Wahrscheinlichkeitsfunktion $\mathrm{P}$ des Modells an das modellierte Zufallsexperiment an? Bei grundlegenden Wahrscheinlichkeiten, die *nicht* aus Symmetriegründen rein kombinatorisch durch Abzählen der „günstigen Fälle“ unter einer größeren Zahl gleichwahrscheinlicher „möglicher Fälle“ (Laplace-Annahme) berechenbar sind, bleibt gar nichts anderes übrig, als sie *empirisch* durch *beobachtete* relative Häufigkeiten zu bestimmen.

Das *Gesetz der großen Zahl*, zuerst in Jakob Bernoullis Werk *Ars conjectandi* formuliert und bewiesen (etwa: „Mutmaßungskunst“, 1713 postum erschienen), macht klar, dass *alle* Wahrscheinlichkeiten, auch die errechneten der verwickeltesten indirekt formulierten Ereignisse, anschaulich stets die Bedeutung eines *Grenzwertes der relativen Häufigkeiten* bei unbegrenzter Zahl von Wiederholungen des Zufallsexperiments haben. Es zeigt damit, dass die Gesetze der Wahrscheinlichkeitsrechnung und das mathematische Modell eines Wahrscheinlichkeitsraumes dem ursprünglichen intuitiven Konzept *Wahrscheinlichkeit* wirklich gerecht werden.

Allerdings erfordert die Deutung des Gesetzes als Bestätigung der Übereinstimmung von Wahrscheinlichkeiten und relativen Häufigkeiten ein *intuitives Vorverständnis*, da es ja selbst nur eine *Wahrscheinlichkeits*-Aussage macht. Man entkommt der unmittelbar anschaulichen Deutung der Grundbegriffe nicht.

Die präzise Formulierung des Gesetzes ist subtil. Die schon von Jakob Bernoulli formulierte *schwache* Form ist heute leicht zu beweisen; die *starke* entstand erst zu Beginn des 20. Jahrhunderts insbesondere durch Ideen von Émile Borel. Verfeinerte Versionen stammen von Markov, Bernstein und Kolmogorov.

## Schwaches Gesetz:

Zu jedem $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass $\mathrm{P}(|h_n(A) - p| < \varepsilon) > 1 - \varepsilon$ für alle $n \geq n_0$.

Man sagt auch, $h_n(A)$ strebe *in Wahrscheinlichkeit* („*stochastisch*“) gegen $p$; in Zeichen: $h_n(A) \xrightarrow{\mathrm{P}} p \; (n \to \infty)$.

## Starkes Gesetz:

Zu jedem $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass $\mathrm{P}\big(|h_k(A) - p| < \varepsilon \; (n_0 \leq k \leq n)\big) > 1 - \varepsilon$ für alle $n \geq n_0$.

Man sagt auch, $h_n(A) \to p$ *fast sicher*, oder: $\mathrm{P}\big(h_n(A) \to p\big) = 1$.

Letztere Formulierungen sind aber nur dann restlos verständlich (im Gegensatz zur eingerahmten finiten Darstellung des Gesetzes), wenn geklärt ist, wie man die Wahrscheinlichkeiten bei Mengen von gedachten unendlich langen Ketten von Versuchswiederholungen definiert. Mit den Überlegungen, die zum starken Gesetz führen, hängt ein Satz von Borel zusammen, der besagt, dass bei einer rein zufällig ausgewählten reellen Zahl mit Wahrscheinlichkeit 1 alle Ziffern ihrer Dezimalentwicklung dieselbe relative Häufigkeit 1/10 besitzen. Im (rein fakultativen!) Anhang 5 folgt eine mathematisch genaue Darstellung des starken Gesetzes.

Vertiefende Literatur:

**A. Rényi**, *Probability Theory*, Dover 1998 (urspr. North-Holland/Akadémiai Kiadó 1970);
**W. Feller,** *Introduction to Probability Thoery and its Applications, Vol. I*, Wiley 1968