# Stochastik

**Barth, Friedrich**

**München, [20]03**

7. 5. Axiomatische Definition der Wahrscheinlichkeit durch Kolmogorow

urn:nbn:de:hbz:466:1-83580

Auch *Abraham de Moivre* setzt die Gleichleichtigkeit aller Fälle ausdrücklich voraus, als er 1711 seine Abhandlung *De Mensura Sortis*, ganz im Stile seiner Zeit und aller seiner Vorgänger, mit der Definition des Verhältnisses der Wahrscheinlichkeiten für Eintreten und Nicht-Eintreten eines Ereignisses beginnt, da Spieler nur dieses Verhältnis interessierte*.

»Si $p$ sit numerus casuum quibus eventus aliquis contingere possit, et $q$ numerus casuum quibus possit non-contingere; tam contingentia quam non-contingentia eventus suum habent probabilitatis gradum: Quod si casus omnes quibus eventus contingere vel non-contingere potest, sint aeque faciles; probabilitas contingentiae, erit ad probabilitatem non-contingentiae ut $p$ ad $q$.«

1718 formuliert aber *de Moivre* in *The Doctrine of Chances* bereits

»The Probability of an Event is greater, or less, according to the number of Chances by which it may Happen, compar'd with the number of all the Chances, by which it may either Happen or Fail.«

1738 fügt er in der 2. Auflage eine explizite Definition der Wahrscheinlichkeit hinzu:

»Wherefore, if we constitute a Fraction whereof the Numerator be the number of Chances whereby an Event may happen, and the Denominator the number of all the Chances whereby it may either happen or fail, that Fraction will be a proper designation of the Probability of happening.«

Und eine Seite weiter lautet es noch präziser

»[...] that it is the comparative magnitude of the number of Chances to happen, in respect to the whole number of Chances either to happen or to fail, which is the true measure of Probability.«

Also wörtlich die von *Laplace* gegebene Definition! Zwar fehlt hier die Einschränkung, daß alle Fälle gleich möglich sein müssen – damals stillschweigend meist vorausgesetzt – aber in einem anschließenden Beispiel weist *de Moivre* wieder ausdrücklich darauf hin.

## 7.5. Axiomatische Definition der Wahrscheinlichkeit durch *Kolmogorow*

Als *Jakob Bernoulli* (1655–1705) im 1. Teil seiner *Ars conjectandi Huygens*' Abhandlung kommentiert, spürt er, daß ein Maß für die Wahrscheinlichkeit fehlt**. Er ergänzt den oben zitierten Satz 3 durch Bildung des Quotienten $\frac{p}{p+q}$, wobei $p$ die Anzahl der Fälle angibt, in denen man etwas gewinnen kann, und $q$ die Anzahl der Fälle, in denen man nichts gewinnt, und verwendet diesen Quotienten als Maß für die Wahrscheinlichkeit, ohne ihn jedoch so zu benennen. Erst im 4. Kapitel des 4. Teils kommt er auf diese Quotientenbildung nochmals zurück und schreibt:

»Und hier scheint uns gerade die Schwierigkeit zu liegen, da nur für die wenigsten Erschei-

* Die erste Aufgabe, bei der nach der Wahrscheinlichkeit in unserem Sinne gefragt wird, fanden wir in der 2. Auflage des *Essay* von *Montmort*. Dort ist ein Brief von *Nikolaus Bernoulli* an *Montmort* vom 30. 12. 1712 abgedruckt. *Nikolaus* stellt das »Problème I: Plusieurs Joueurs dont le nombre est $n+1$ jouent une poulle, on demande quelle est la probabilité que chacun a de gagner la poulle.«

** Wahrscheinlichkeit als meßbarer Begriff erscheint zum ersten Mal in *La Logique ou l'art de penser* 1662, die Logik von Port Royal, die sicherlich von *Pascal* inspiriert wurde. Nach deren lateinischem Titel *Logica sive Ars cogitandi* ist vermutlich *Bernoullis Ars conjectandi* geprägt.

nungen und fast nirgends anders als in Glücksspielen dies möglich ist; die Glücksspiele wurden aber [...] so eingerichtet, daß die Zahlen der Fälle, in welchen sich Gewinn oder Verlust ergeben muß, im voraus bestimmt und bekannt sind, und daß alle Fälle mit gleicher Leichtigkeit eintreten können. Bei den weitaus meisten anderen Erscheinungen aber, welche von dem Walten der Natur oder von der Willkür der Menschen abhängen, ist dies keineswegs der Fall.«

Der klassische Wahrscheinlichkeitsbegriff bewährt sich also sehr bei der Analyse von Glücksspielen, ist aber kaum tragfähig für Probleme aus Technik und Wirtschaft, bei denen es praktisch unmöglich ist, die Ergebnisse so festzulegen, daß sie uns als »gleich wahrscheinlich« erscheinen. Auch bei der bereits von *Jakob Bernoulli* vorgenommenen Anwendung der Wahrscheinlichkeitsrechnung auf Krankheiten und Todesfälle lassen sich »gleich mögliche« Fälle nicht auszählen.

Die Schwierigkeiten bei der statistischen und auch der klassischen Definition der Wahrscheinlichkeit rühren davon her, daß sie »Wahrscheinlichkeit« durch eine explizite Definition inhaltlich erfassen wollten. In der modernen Mathematik geht man solchen Schwierigkeiten dadurch aus dem Weg, daß man die Theorie axiomatisch begründet und die Begriffe darin implizit definiert. So treibt man Geometrie mit Punkten und Geraden, ohne explizit definiert zu haben, was Punkte und Gerade sind. Wichtig sind ihre Eigenschaften, die in den Axiomen der Geometrie festgelegt sind. Für die Wahrscheinlichkeitstheorie wurde ein solcher axiomatischer Aufbau von *Andrei Nikolajewitsch Kolmogorow** (1903–1987) in seiner 1933 in Berlin erschienenen Arbeit *Grundbegriffe der Wahrscheinlichkeitsrechnung* vorgeschlagen. Die mathematische Festlegung des Wahrscheinlichkeitsbegriffs orientiert sich dabei auch an der experimentell zugänglichen relativen Häufigkeit, aber sie ist allgemein genug, um auch eine Grundlage für den subjektiven Wahrscheinlichkeitsbegriff abzugeben. *Kolmogorow* hat gezeigt, daß 3 geeignet ausgewählte Eigenschaften der relativen Häufigkeit genügen, um »Wahrscheinlichkeit« so zu definieren, daß damit eine tragfähige Grundlage für eine in der Praxis brauchbare Theorie aufgebaut werden kann. Nach *Kolmogorow* wird auf der Ereignisalgebra die Wahrscheinlichkeit $P(A)$ eines Ereignisses $A$ als Funktionswert einer reellwertigen Funktion $P$ definiert. Ist der Ergebnisraum $\Omega$ endlich, so lassen sich die Forderungen von *Kolmogorow* wie folgt formulieren:

> Eine Funktion $P\colon A \mapsto P(A)$ mit $A \in \mathfrak{P}(\Omega)$ und $P(A) \in \mathbb{R}$ heißt *Wahrscheinlichkeitsverteilung*, wenn sie folgenden Bedingungen genügt:
>
> **Axiom I:** $P(A) \geqq 0$ (Nichtnegativität)
> **Axiom II:** $P(\Omega) = 1$ (Normierung)
> **Axiom III:** $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$ (Additivität)

Nichtnegativität und Normierung entsprechen den Eigenschaften (1) und (4) für relative Häufigkeiten aus **4.2.** Dem Additionsaxiom für unvereinbare Ereignisse liegt die entsprechende Eigenschaft (6) für relative Häufigkeiten zugrunde.
Man könnte auf die Idee kommen, an Stelle von Eigenschaft (6) die Eigenschaft (5)

*Колмогоров (sprich: kɐlmɐgórɐf)

dem dritten Axiom zugrunde zu legen, da sie keine Voraussetzungen für die Ereignisse $A$ und $B$ fordert. Das ergäbe ein

$$\text{Axiom III}': \ P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Leider geht das aber schief, weil die Axiome I, II und III′ auch unerwünschte Wahrscheinlichkeitsverteilungen zulassen. So würde z.B. die Festsetzung $P(E) := 1$ für *alle* Ereignisse $E$ des Ereignisraums die Axiome I, II und III′ erfüllen:

I: $P(A) = 1 \geqq 0$

II: $P(\Omega) = 1$

III′: $1 = P(A \cup B) = P(A) + P(B) - P(A \cap B) = 1 + 1 - 1 = 1$.

Bei dieser Wahrscheinlichkeitsverteilung hätte die leere Menge und damit das unmögliche Ereignis die Wahrscheinlichkeit 1. Also kann man aus I, II und III′ bestimmt nicht mehr folgern, daß $P(\emptyset) = 0$, was aber für eine sinnvolle Anwendung wünschenswert ist, weil die Interpretationsregel $P(\emptyset) = 0$ nahelegt. Nach (3) gilt nämlich $h_n(\emptyset) = 0$.

*Wir* haben in Definition 42.1 die Wahrscheinlichkeit ebenfalls axiomatisch definiert. Die daraus gefolgerten Sätze 44.1, 44.2 und 44.3 sind gerade die drei Axiome von *Kolmogorow*. Umgekehrt läßt sich aus den drei *Kolmogorow*-Axiomen unsere Definition 42.1 herleiten. Es gilt nämlich

**Satz 81.1:** Ist $P$ eine Wahrscheinlichkeitsverteilung über einem endlichen Ergebnisraum $\Omega$, die den Axiomen von *Kolmogorow* genügt, dann gilt:

**1)** Für alle $\omega \in \Omega$ gilt $0 \leqq P(\{\omega\}) \leqq 1$.

**2)** Die Summe der Wahrscheinlichkeiten aller Elementarereignisse ist 1; kurz $\sum\limits_{\omega \in \Omega} P(\{\omega\}) = 1$.

**3)** Die Wahrscheinlichkeit des unmöglichen Ereignisses ist 0; kurz $P(\emptyset) = 0$.

**4)** Ist $A$ nicht das unmögliche Ereignis, so ist die Wahrscheinlichkeit $P(A)$ gleich der Summe der Wahrscheinlichkeiten derjenigen Elementarereignisse, deren Vereinigung das Ereignis $A$ ergibt; kurz
$A \neq \emptyset \Rightarrow P(A) = \sum\limits_{\omega \in A} P(\{\omega\})$.

**Beweis:**

**4)** Ist $A = \{a_1, a_2, \ldots, a_k\}$, so erhalten wir durch wiederholte Anwendung des 3. *Kolmogorow*-Axioms

$$\begin{aligned} P(A) &= P(\{a_1, a_2, \ldots, a_k\}) = \\ &= P(\{a_1\} \cup \{a_2, \ldots, a_k\}) = \\ &= P(\{a_1\}) + P(\{a_2, \ldots, a_k\}) = \\ &= P(\{a_1\}) + P(\{a_2\} \cup \{a_3, \ldots, a_k\}) = \\ &= P(\{a_1\}) + P(\{a_2\}) + P(\{a_3, \ldots, a_k\}) = \\ &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots = \\ &= P(\{a_1\}) + P(\{a_2\}) + \ldots + P(\{a_k\}). \end{aligned}$$

**3)** Ist $A = \emptyset$, so ist $A \cap B = \emptyset \cap B = \emptyset$. Somit ist für $A \cap B$ die Voraussetzung des 3. Axioms von *Kolmogorow* erfüllt, und wir erhalten einerseits $P(A \cup B) = = P(\emptyset \cup B) = P(B)$, andererseits $P(A \cup B) = P(A) + P(B) = P(\emptyset) + P(B)$. Der Vergleich der beiden rechten Seiten ergibt $P(\emptyset) = 0$.

**2)** Nach dem soeben bewiesenen Teil **4** dieses Satzes ist $P(\Omega) = \sum\limits_{\omega \in \Omega} P(\{\omega\})$. Axiom II besagt aber, daß $P(\Omega) = 1$ ist, woraus die Behauptung folgt.

**1)** Wegen Axiom I ist $P(\{\omega\}) \geqq 0$ für jedes $\omega \in \Omega$. Unter Verwendung des soeben bewiesenen Teils **2** erhalten wir noch $P(\{\omega\}) \leqq \sum\limits_{\omega \in \Omega} P(\{\omega\}) = 1$.

Für endliche Ergebnisräume sind die beiden Definitionen demnach äquivalent. Wir haben Definition 42.1 gewählt, weil das Belegen der Elementarereignisse mit Wahrscheinlichkeiten ein sehr anschaulicher Vorgang ist, ebenso wie das Zusammensetzen der Wahrscheinlichkeit eines Ereignisses aus den Wahrscheinlichkeiten seiner Elementarereignisse. Die Definition von *Kolmogorow* hat den Vorteil, daß sie sich so verallgemeinern läßt, daß sie auch für unendliche Ergebnisräume brauchbar wird. Das haben wir in diesem Buch aber nicht vor.
Die Festlegung der Funktionswerte $P(A)$, d.h. der Wahrscheinlichkeiten der Ereignisse, ist im Rahmen dieser Axiome völlig willkürlich. Man wird jedoch die Werte so festlegen, daß sie den jeweiligen Verhältnissen angepaßt sind. So wird man bei einem idealen Würfel auf Grund der Symmetrie für jede Augenzahl die Wahrscheinlichkeit $\frac{1}{6}$ a priori festlegen. Bei einem realen Würfel hingegen empfiehlt es sich, wie im Beispiel auf Seite 42 durchgeführt, die relativen Häufigkeiten der Augenzahlen in einer möglichst langen Versuchsserie zu bestimmen und diese relativen Häufigkeiten als Wahrscheinlichkeiten der Augenzahlen a posteriori zu verwenden. Dann ist nämlich die Interpretationsregel für Wahrscheinlichkeiten (Seite 44) anwendbar, wie *Jakob Bernoulli* im Hauptsatz seiner *Ars conjectandi*, dem »Gesetz der großen Zahlen«, gezeigt hat.

## Aufgaben

Die Behauptungen der Aufgaben **1.–6.** sollen rein formal aus den Axiomen von *Kolmogorow* hergeleitet werden.

**1.** Für alle Ereignisse $A$ gilt: $P(\bar{A}) = 1 - P(A)$.
**2.** $P(\emptyset) = 0$.
**3.** Für alle Ereignisse $A$ gilt: $P(A) \leqq 1$.
**4.** Für alle Ereignisse $A, B$ gilt:
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

**5.** Für paarweise unvereinbare Ereignisse $A_i$ $(i = 1, 2, \ldots, n)$ gilt die folgende Verallgemeinerung des Axioms III:
$$\boxed{P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).}$$

**6.** Es gilt das Monotoniegesetz für Wahrscheinlichkeiten: $A \subset B \Rightarrow P(A) \leqq P(B)$.

•**7. a)** Ein Axiomensystem heißt **widerspruchsfrei**, wenn es ein Modell gibt, das sämtliche Axiome erfüllt. Zeige, daß das Axiomensystem von *Kolmogorow* widerspruchsfrei ist anhand nebenstehenden Modells: $\Omega := \{\omega\}$ und $\begin{array}{c|cc} A & \emptyset & \Omega \\ \hline P(A) & 0 & 1 \end{array}$.

**b)** Ein Axiomensystem heißt **unvollständig**, wenn es mehrere, nicht isomorphe Modelle gibt. Begründe, daß das Axiomensystem von *Kolmogorow* unvollständig ist.