In: Markus Schrenk (Hrsg.), *Handbuch Metaphysik*, Verlag J.B. Metzler, Stuttgart 2017, 312–318.

Wahrscheinlichkeit

Prof. Dr. Jacob Rosenthal, Universität Konstanz

Mit gewissen Ereignissen oder Sachverhalten ist nicht definitiv, sondern mehr oder minder stark zu rechnen. Man sagt dann, es sei mehr oder weniger wahrscheinlich, dass dies-und-jenes geschehen werde oder der Fall sei. Wir wissen vieles nicht sicher, haben aber bestimmte Indizien und modifizieren unsere Aussagen entsprechend durch Ausdrücke wie „wahrscheinlich", „sehr wahrscheinlich", „extrem unwahrscheinlich", etc. Derartige Qualifikationen sind ursprünglich nichts anderes als ein Reflex unseres unvollständigen Wissens. Wir wollen uns in Bezug auf einen bestimmten Sachverhalt weder festlegen noch des Urteils ganz enthalten: Da machen wir eine Wahrscheinlichkeitsaussage.

Die Wahrscheinlichkeitsrechnung

Der entscheidende Schritt über diese alltäglichen Verwendungen hinaus, der wesentlich dafür verantwortlich ist, dass es in einem Band über Metaphysik den Eintrag „Wahrscheinlichkeit" geben kann, ist die Entdeckung, dass sich Wahrscheinlichkeiten in bestimmten Kontexten *quantifizieren* und auf äußerst nützliche Weise *mathematisch kombinieren* lassen. Die Wahrscheinlichkeitsrechnung wurde im 17. Jahrhundert im Zusammenhang mit Glücksspielen entwickelt (siehe Hacking 1975, Schneider 1988). Wenn man einen Geldbetrag darauf zu setzen hat, dass bei einem Wurf mit zwei normalen Würfeln die Augensumme 7 resultiert, und sich überlegt, wie viel man dabei gegen zehn Taler höchstens einsetzen sollte, ist es wichtig zu wissen, dass das besagte Ereignis die Wahrscheinlichkeit 1/6 hat. Von solchen Anfängen aus sind die Wahrscheinlichkeitstheorie und die mathematische Statistik große Teilgebiete der Mathematik mit weitgehenden Resultaten und zahlreichen Anwendungen vor allem in den empirischen Wissenschaften geworden.

Im Rahmen der mathematischen Modellierung wird ein Ereignis als eine Menge von *Elementarereignissen* aufgefasst. Im Beispiel würde das Ereignis „Augensumme 7" durch die folgende Menge repräsentiert: {(6 1), (1 6), (5 2), (2 5), (4 3), (3 4)}. Aus Ereignissen in diesem Sinne kann man mittels Durchschnitts-, Vereinigungs- und Komplementbildung neue Ereignisse erzeugen. Mathematisch betrachtet man daher immer ein ganzes Wahrscheinlichkeits*maß* auf einer *Algebra* von Ereignissen. Ein Wahrscheinlichkeitsmaß **P** auf einer Algebra **A** ist eine nicht negative, normierte und additive Abbildung der Algebra in die reellen Zahlen:

(1) *Nichtnegativität*: **P**(E) ≥ 0 für jedes Ereignis E aus **A**.

(2) *Normiertheit*: **P**(S) = 1 falls S die Menge aller Elementarereignisse (das „sichere Ereignis") ist.

(3) *Additivität*: **P**(E ∪ F) = **P**(E) + **P**(F), falls E und F beliebige disjunkte Ereignisse aus **A**, also Ereignisse mit einer leeren Schnittmenge sind.

Diese Axiomatisierung geht auf Kolmogorov (1933) zurück und ist heute Standard. Die Begriffe der Algebra und des Wahrscheinlichkeitsmaßes darauf haben technische Bedeutung: Sie stellen sicher, dass man Ereignisse und ihre Wahrscheinlichkeiten stets in der erforderlichen Weise kombinieren kann. In unserem Beispiel gibt es 36 Elementarereignisse, nämlich alle möglichen Kombinationen von Augenzahlen der beiden Würfel. Die Wahrscheinlichkeit des Ereignisses „Augensumme 7" ergibt sich als die Summe der Wahrscheinlichkeiten der dazu gehörigen Elementarereignisse.

Eine wichtige Begriffsbildung ist die der *bedingten* Wahrscheinlichkeit. Die bedingte Wahrscheinlichkeit von E, gegeben F, oder die Wahrscheinlichkeit von E unter der Bedingung F, wird bezeichnet mit $\mathbf{P}(E|F)$ und ist wie folgt definiert:

(4) $\mathbf{P}(E|F) = \frac{\mathbf{P}(E \cap F)}{\mathbf{P}(F)}$ für beliebige Ereignisse E und F aus $\mathbf{A}$ mit $\mathbf{P}(F) > 0$.

Im Beispiel ist z.B. die Wahrscheinlichkeit dafür, dass mindestens einer der Würfel eine 5 zeigt, unter der Bedingung, dass die Augensumme 7 ist, gleich $\frac{2/36}{6/36}$ oder 1/3.

Der Begriff der bedingten Wahrscheinlichkeit kann verschieden aufgefasst und eingeführt werden. Das hat Konsequenzen für den Status der Gleichung (4): Was die Standard-Wahrscheinlichkeitsrechnung als Definition ansieht, lässt sich bei bestimmten Interpretationen der Wahrscheinlichkeit als Theorem ableiten. Die Gleichung ist nur unter der Bedingung $\mathbf{P}(F) > 0$ sinnvoll, da nur dann der rechts stehende Quotient definiert ist. Diese Beschränkung ist manchmal störend und lässt sich überwinden, indem man den Begriff der bedingten Wahrscheinlichkeit als fundamental nimmt und „unbedingte" Wahrscheinlichkeiten als abgeleitet auffasst (siehe aktuell Hájek 2003).

Zwei Ereignisse E und F aus der Algebra $\mathbf{A}$ heißen *probabilistisch unabhängig*, falls die Wahrscheinlichkeit dafür, dass sie beide eintreten, gleich dem Produkt ihrer Wahrscheinlichkeiten ist:

(5) Unabhängigkeit: $\mathbf{P}(E \cap F) = \mathbf{P}(E) \cdot \mathbf{P}(F)$

Bei $\mathbf{P}(F) > 0$ ist diese Bedingung äquivalent dazu, dass die bedingte Wahrscheinlichkeit von E, gegeben F, gleich der „unbedingten" Wahrscheinlichkeit von E ist: $\mathbf{P}(E|F) = \mathbf{P}(E)$. Ebenso ist sie bei $\mathbf{P}(E) > 0$ äquivalent zu $\mathbf{P}(F|E) = \mathbf{P}(F)$. Salopp ausgedrückt sind zwei Ereignisse genau dann im Wahrscheinlichkeitssinne unabhängig, wenn das Eintreten des einen die Wahrscheinlichkeit des anderen nicht beeinflusst. Weiter ergibt sich im Falle $\mathbf{P}(E), \mathbf{P}(F) \neq 0$: $\mathbf{P}(E|F) > \mathbf{P}(E)$ ist äquivalent zu $\mathbf{P}(F|E) > \mathbf{P}(F)$. In diesem Fall heißen die Ereignisse E und F *positiv korreliert*. Grob gesagt erhöht dann das Eintreten des einen die Wahrscheinlichkeit des anderen. Ebenso ist $\mathbf{P}(E|F) < \mathbf{P}(E)$ äquivalent zu $\mathbf{P}(F|E) < \mathbf{P}(F)$. In diesem Falle heißen E und F *negativ korreliert*: Das Eintreten des einen reduziert die Wahrscheinlichkeit des anderen.

Zwei beliebige Ereignisse E und F aus einer Algebra $\mathbf{A}$ mit $\mathbf{P}(E), \mathbf{P}(F) \neq 0$ sind also entweder positiv korreliert oder negativ korreliert oder unkorreliert (probabilistisch unabhängig). Im Beispiel sind die Ereignisse „Augensumme 7" und „mindestens ein Würfel zeigt eine 5" leicht positiv korreliert, denn die Wahrscheinlichkeit des letzteren ist 11/36, gegeben das erstere beträgt sie aber 1/3 = 12/36.

Interpretationen der Wahrscheinlichkeit

Der Übergang zur Philosophie der Wahrscheinlichkeit geschieht mit der Frage, was quantifizierte Wahrscheinlichkeitsaussagen bedeuten, aufgrund wovon sie wahr oder falsch sind, oder auch, welchen ontologischen Status Wahrscheinlichkeiten haben. Bei den Interpretationen der Wahrscheinlichkeit kann man klassische, logische, subjektive, Häufigkeits-, Beste-System-, Propensity- und Spielraumtheorie unterscheiden. Im philosophischen Rahmen werden oft nicht Ereignisse, sondern *Propositionen* (Aussageinhalte, Sachverhalte; siehe Kap. VI.B.2) als dasjenige genommen, dem eine bestimmte Wahrscheinlichkeit zugeschrieben wird. An die Stelle der Vereinigung von Mengen tritt dann die logische Disjunktion, an die Stelle des Durchschnitts die Konjunktion, an die Stelle des Komplements die Negation. Es ist nicht nötig, sich zwischen diesen Auffassungen zu entscheiden.

Unsere alltäglichen Wahrscheinlichkeitsaussagen weisen einen *epistemischen* und *objektivistischen* Grundzug auf. Sie beziehen sich ebenso auf Vergangenes wie auf Zukünftiges. Wir können mit derselben Leichtigkeit sagen: „Wahrscheinlich hat es sich so-und-so abgespielt“ wie „Wahrscheinlich wird dies-und-das geschehen“. Es gibt dabei kein Indiz für eine Bedeutungsverschiebung und deshalb auch keines dafür, dass sich die Wahrscheinlichkeit auf eine in den Dingen selber gelegene, „ontische“ Offenheit bezieht. Nichtsdestoweniger treten solche Aussagen mit einem Objektivitätsanspruch auf: Angesichts der vorliegenden Informationen ist das-und-das so-und-so wahrscheinlich, und wer etwas anderes behauptet, irrt sich.

Die klassische Theorie. Die frühen Anwendungen der Wahrscheinlichkeitsrechnung, von denen sie ihren Ausgang nahm, teilen diese Züge. Sehr deutlich wird dies bei Laplace, mit dessen Werk die *klassische Theorie* kulminiert. Wahrscheinlichkeiten in Glücksspielen werden nach dem „Indifferenzprinzip“ oder „Prinzip vom unzureichenden Grunde“ auf die Elementarereignisse verteilt: Wir haben beim Werfen eines Würfels keinen Grund, mit einer Augenzahl eher zu rechnen als mit irgendeiner anderen und müssen deshalb die Wahrscheinlichkeit gleich auf die sechs möglichen Resultate verteilen. Beim Werfen mit zwei Würfeln bestehen 36 mögliche Kombinationen, die gleich zu behandeln sind. Die klassische Theorie führt alle numerischen Wahrscheinlichkeitszuschreibungen in ähnlicher Weise auf Abzählungen fundamentaler „gleichmöglicher“ Fälle zurück. Eine vollgültige Interpretation von Wahrscheinlichkeitsaussagen ergibt sich daraus nicht, nur eine Zurückführung auf den Terminus „gleichmöglich“, der nun zum Kristallisationskern der Deutungsprobleme wird. Die klassische Theorie fasst diesen Begriff epistemisch, im Sinne informationeller Symmetrie, auf.

Die logische Theorie. Wie die klassische Theorie, aber viel allgemeiner als diese, möchte auch die *logische Theorie* eindeutig bestimmte Wahrscheinlichkeiten angesichts gewisser Informationen behaupten (siehe Keynes 1920 und in der formalen Durchführung Carnap 1950). Ihrer Auffassung nach gibt es zu jeder konsistenten Menge **M** von Propositionen und jeder Proposition A einen eindeutig bestimmten Grad logischer Stützung oder Bestätigung von A durch **M**, und das ist die Wahrscheinlichkeit von A, gegeben **M**. Auch dabei wird die wesentliche Arbeit von Indifferenzprinzipien geleistet, nur so kommt man in diesem Rahmen zu bestimmten numerischen Wahrscheinlichkeiten. Derartige Prinzipien sind zwar in gewissen Anwendungsfällen plausibel, können aber die auf sie gesetzten weitgehenden Fundierungshoffnungen nicht erfüllen. Jede nichtlineare Transformation einer kontinuierlichen Größe X verwandelt eine Gleichverteilung der Wahrscheinlichkeit über die möglichen Werte von X in eine Ungleichverteilung über die Werte der transformierten Größe. Es gibt keinen hinreichend prinzipiellen logischen (und manchmal überhaupt keinen) Grund, die Wahrscheinlichkeiten eher bezüglich X als bezüglich einer der unendlich vielen transformierten Größen gleich zu verteilen (sog. Bertrandsche Paradoxien). Eine Gleichverteilung in einer Hinsicht bedingt also ganz von selber Ungleichverteilungen in anderen Hinsichten. Zu einer Zuweisung von Wahrscheinlichkeiten, die dem Allgemeinheitsanspruch der logischen Theorie entspricht, kommt man nur, wenn man eine strukturell einfache Spielzeugsprache als globalen Beschreibungsrahmen wählt. Die resultierenden Wahrscheinlichkeiten sind dann relativ zu dieser Sprache.

Der Subjektivismus. Eine mögliche Konsequenz daraus ist, den teilweise subjektiven Charakter von Wahrscheinlichkeiten einzuräumen. Diese hängen jedenfalls nicht *nur* von den vorliegenden Informationen ab. Daraus entsteht die *subjektive Theorie*, der zufolge Wahrscheinlichkeiten persönliche Glaubens- oder Überzeugungsgrade sind. Ihre Quantifizierung erfolgt prinzipiell über hypothetische Wetten: Welche Wetten einem als vorteilhaft, welche als fair und welche als nachteilig erscheinen, hängt damit zusammen, wie stark man mit dem Ereignis, auf das gewettet wird, rechnet. Die Axiome (1) bis (3) erscheinen dabei als Kohärenzanforderungen: Wer sie in seinen Überzeugungsgraden verletzt, setzt sich einem „dutch book" aus, einem System von Wetten, von denen ihm jede einzelne vorteilhaft zu sein scheint, aber insgesamt mit subjektiver Sicherheit ein Verlust resultiert. Daran zeigt sich, dass die Überzeugungsgrade nicht zusammenpassen. Wenn sie es aber tun, also kein „dutch book" gegen das Subjekt möglich ist, darf man tatsächlich von subjektiven *Wahrscheinlichkeiten* sprechen.

(1) bis (3) lassen sich als *synchrone* Kohärenzbedingungen für Überzeugungsgrade verstehen. Mithilfe der bedingten Wahrscheinlichkeit lässt sich auch eine *diachrone* Kohärenzforderung formulieren: Das Subjekt sollte, wenn ihm ein neuer Sachverhalt – nennen wir ihn „Evidenz E" – bekannt wird, seine persönlichen Überzeugungsgrade wie folgt modifizieren:

(6) $\mathbf{P}_{\mathrm{neu}}(\mathrm{A}) = \mathbf{P}_{\mathrm{alt}}(\mathrm{A}|\mathrm{E}) = \frac{\mathbf{P}_{\mathrm{alt}}(\mathrm{A} \wedge \mathrm{E})}{\mathbf{P}_{\mathrm{alt}}(\mathrm{E})}$ für alle Propositionen A in der Algebra **A**.

Auch für die Aktualisierung der Glaubensgrade durch Konditionalisierung lassen sich dutch-book-Argumente angeben, die aber in ihrem Status strittiger sind.

Der Bayesianismus. Bereits aus diesen einfachen Prinzipien ergibt sich eine Theorie subjektiver Wahrscheinlichkeiten als minimal rationaler (nämlich kohärenter) persönlicher Überzeugungsgrade und ihrer Veränderung im Lichte neuer Informationen: der *Bayesianismus*. Die Leistungsfähigkeit dieser Auffassung wird demonstriert in Howson und Urbach (1989) und Jeffrey (2004); kritischer ist Earman (1992). Das subjektivistische Element, das vortheoretisch als Zumutung erscheint, wird zum einen durch Konvergenzsätze gemildert: Unter gewissen Bedingungen konvergieren die persönlichen Überzeugungsgrade verschiedener Subjekte, falls diese fortgesetzt die gleichen Evidenzen erhalten. Zum anderen kann man weitere Rationalitätsbeschränkungen für Überzeugungsgrade ansetzen. Diese betreffen aufgrund der Konditionalisierungsregel (6) die a-priori-Wahrscheinlichkeiten, mit denen das Subjekt „vor aller Erfahrung“ beginnt. Hier werden wieder Indifferenzprinzipien relevant, insbesondere das sog. Prinzip der maximalen Entropie. Der *objektive Bayesianismus* (siehe Williamson 2010) ist der logischen Theorie der Wahrscheinlichkeit eng verwandt.

<u>*Subjektive versus objektive Wahrscheinlichkeit.*</u> In den meisten (nicht allen) Fällen, in denen wir einen bestimmten Wettquotienten für den richtigen halten, haben wir nicht primär entsprechend graduierte Überzeugungen (degrees of belief, die als psychologische Größen eine extreme Idealisierung darstellen), sondern einfache Überzeugungen über Grade (beliefs about degrees) – nämlich über Grade *objektiver* oder *ontischer* Wahrscheinlichkeit. Wir wollen mit quantifizierten Wahrscheinlichkeitsaussagen typischerweise nicht lediglich einen Überzeugungsgrad ausdrücken, sondern eine Aussage über die Wirklichkeit machen. Wir meinen etwa, dass die objektive Wahrscheinlichkeit, bei einem Wurf mit zwei Würfeln die Augensumme 7 zu erzielen, 1/6 beträgt. Deshalb, und nur deshalb, halten wir einen Einsatz von zwei Talern gegen zehn Taler darauf, dass dieses Ereignis eintritt, für fair. Uns ist bewusst, dass wir dabei irren können: Vielleicht ist die Wahrscheinlichkeit für das Resultat gar nicht 1/6, weil mit den Würfeln oder der Wurfprozedur etwas nicht in Ordnung ist.

Für diesen Kontrast hat sich die Rede von subjektiven und objektiven Wahrscheinlichkeiten eingebürgert. Es handelt sich um den Gegensatz zwischen dem, was unsere Informationen (auf der Subjektseite) wie stark nahelegen, und dem, was in Wirklichkeit (auf der Objektseite) der Fall ist. Klarer wäre es, von epistemischen und ontischen Wahrscheinlichkeiten zu sprechen, um Verwechslungen mit dem objektiven Bayesianismus und der logischen Wahrscheinlichkeit zu vermeiden. Insbesondere braucht man eine ontische Deutung für die Wahrscheinlichkeitsaussagen der empirischen Wissenschaften, die als Behauptungen über die Wirklichkeit auftreten.

<u>*Die Spielraumtheorie*</u>. Was könnte aber eine Wahrscheinlichkeitsaussage aufgefasst als Aussage über die subjektunabhängige Wirklichkeit bedeuten? Bei den klassischen Glücksspielen wird man zuerst an physikalische Symmetrien als „Wahrmacher“ (siehe Kap. VI.A.3) denken. Wie informationelle Symmetrien epistemischen Wahrscheinlichkeiten zugrunde liegen, so physikalische Symmetrien ontischen. Die 36 möglichen Kombinationen von Augenzahlen beim Wurf zweier Würfel sind „gleichmöglich“ in einem *physikalischen* Sinne, wenn die Würfel symmetrisch aufgebaut sind und normal geworfen werden – egal, was wir darüber denken. Von dieser Grundidee ausgehend kommt man mit geeigneten Verallgemeinerungen zur *Spielraumtheorie* der Wahrscheinlichkeit. Wenn wir den Anfangsbedingungsraum (alle möglichen Konstellationen von Anfangsbedingungen) beim Würfeln überblicken könnten, würden wir zweierlei bemerken: erstens die labile Abhängigkeit des Ergebnisses von den genauen Anfangsbedingungen und zweitens einen über den gesamten Raum nahezu konstanten Anteil jedes bestimmten Ereignisses.

Wahrscheinlichkeiten erscheinen so als Anteile an Anfangsbedingungs- oder Zustandsräumen (in der Physik spricht man auch von „Phasenräumen“). Ein verfälschter Würfel lässt sich ebenso behandeln, bei ihm ist der Anteil, mit dem z.B. das Resultat „6“ im Anfangsbedingungsraum repräsentiert ist, nicht 1/6, sondern größer oder kleiner. Allgemein:

Es sei ***Z*** ein Typ von Zufallsexperiment mit dem assoziierten Anfangsbedingungsraum ***S***. Es sei E ein mögliches Ereignis bei der Durchführung von ***Z***. Wenn die Konstellationen von Anfangsbedingungen, die zu E führen, an jedem nicht zu kleinen Intervall in ***S*** ungefähr denselben Anteil *p* haben, dann ist *p* die objektive Wahrscheinlichkeit von E bei der Durchführung von ***Z***.

„Experiment“ ist hier in einem weiten Sinne zu verstehen, der auch „Experimente durch die Natur“ einschließt. Eine Konzeption dieser Art wurde zuerst von von Kries (1886) artikuliert, aktuelle Varianten sind Strevens (2011, 2013), Abrams (2012) und Vf. (2012). Diese Auffassung von Wahrscheinlichkeiten ist der sog. Methode der willkürlichen Funktionen eng verwandt (siehe Engel 1992), ihr paradigmatischer Anwendungsbereich sind physikalische Prozesse, die prinzipiell im Rahmen der klassischen Mechanik modellierbar sind. Die Spielraumtheorie ist grundsätzlich ebenfalls Paradoxien nach Art der Bertrandschen ausgesetzt. Diese sind hier jedoch viel schwerer zu konstruieren, da jede halbwegs normale Transformation der physikalischen Größen, die die Anfangsbedingungen charakterisieren, die Spielraum-Wahrscheinlichkeiten unverändert lässt. Wenn einem das an Objektivität nicht genügt und man zusätzlich Annahmen über die faktisch auftretenden Anfangsbedingungen macht, bewegt man sich in Richtung der *Häufigkeitstheorie*.

Die Häufigkeitstheorie ist trotz zahlreicher Probleme (siehe Hájek 1997, 2009) die Standardauffassung objektiver Wahrscheinlichkeit und wird es wegen ihres weiten Anwendungsbereiches bleiben. Während physikalische Symmetrien oder Anteile an möglichen Konstellationen von Anfangsbedingungen die Eintrittshäufigkeiten von Ereignissen in gewissen Fällen gut erklären können, sind es der Häufigkeitstheorie zufolge diese Häufigkeiten selber, an denen sich bemisst, ob eine Wahrscheinlichkeitsaussage wahr oder falsch ist. Auf die relativen Häufigkeiten von Ereignissen kommt es an; die Frage, was ihnen zugrunde liegt, ist sekundär. Eine klassische Darstellung dieser Auffassung ist von Mises (zuerst 1928). Sie besagt in etwa:

Wenn bei häufiger Wiederholung des Zufallsexperimentes ***Z*** das Ereignis E regellos, aber ungefähr mit der relativen Häufigkeit *p* auftritt, dann ist *p* die objektive Wahrscheinlichkeit von E bei der Durchführung von ***Z***.

Der Begriff der Regellosigkeit lässt sich nach Kolmogorov und Chaitin komplexitätstheoretisch erfassen. Eine wichtige Binnendifferenzierung ist, ob *tatsächliche* oder *hypothetische* relative Häufigkeiten die Wahrmacher von Wahrscheinlichkeitsaussagen sein sollen. Erstere sind oft nicht in ausreichendem Maße vorhanden. Es liegt deshalb nahe, sich auf hypothetische Häufigkeiten zu beziehen, die sich ergeben *würden*, falls das Experiment häufig wiederholt *würde*. Dabei stützt man sich allerdings auf kontrafaktische Konditionalsätze über relative Häufigkeiten, für deren Wahrheit bei näherem Hinsehen nichts spricht. Beim Würfeln kann mehrfach hintereinander dieselbe Augenzahl kommen – das hat jeder schon erlebt. Wenn aber zwei, drei oder viel Mal hintereinander die „1" kommen kann, warum dann nicht noch öfter? Es gibt bei wiederholtem Werfen nicht *die* relative Häufigkeit, mit der sich die Augenzahl „1" einstellen *würde*, sondern sie *könnte* mit allen möglichen Häufigkeiten auftreten, auch in sehr langen Wurfserien. Dieser Sachverhalt wird durch die sog. Gesetze der großen Zahlen präzisiert, mathematischen Theoremen, die qualitativ folgendes besagen: Falls bei einem Zufallsexperiment ***Z*** das Ereignis E mit der Wahrscheinlichkeit *p* auftritt, dann liegt in einer *langen* Serie von *probabilistisch unabhängigen* Wiederholungen von ***Z*** die relative Häufigkeit von E *sehr wahrscheinlich sehr nahe* an *p*. Die Modellierung der Wiederholung von Zufallsexperimenten als probabilistisch unabhängig ist für die Anwendungen der Wahrscheinlichkeitsrechnung grundlegend. Es ist nicht zu sehen, wie man ohne spezielle Annahmen gerechtfertigt sein könnte, eine noch engere kontrafaktische Verbindung zwischen Wahrscheinlichkeiten und relativen Häufigkeiten von Ereignissen anzunehmen. Insofern dieser Zusammenhang seinerseits probabilistisch ist, kann die Häufigkeitstheorie Wahrscheinlichkeitsaussagen nur durch weitere derartige Aussagen erläutern und führt so in einen Regress oder begrifflichen Zirkel.

Die Beste-System-Theorie. Wenn man an der Häufigkeitstheorie dennoch festhalten möchte, dann besser in einer indirekten Variante wie der *Beste-System-Theorie* von Lewis (1994). Bei ihr spielen relative Häufigkeiten in regellosen Mustern von Ereignissen ebenfalls eine wesentliche (wenn auch nicht die alleinige) Rolle für die Wahrheitsbedingungen von Wahrscheinlichkeitsaussagen. Die Sache geht ungefähr so: Ontische Wahrscheinlichkeiten erwachsen aus probabilistischen Naturgesetzen. Naturgesetze beschreiben alles, was es empirisch manifest gibt – das gesamte raumzeitliche Universum – mit einer bestmöglichen Kombination an Einfachheit und Informationsstärke (siehe Kap. V.C.1). Wahrscheinlichkeiten können dadurch in Naturgesetze hineinkommen, dass durch sie ein erheblicher Gewinn an Einfachheit bei einem vergleichsweise geringen Opfer an Informationsstärke entsteht. So wäre es zwar extrem informativ, die Ergebnisse sämtlicher Würfelwürfe in der Geschichte des Universums aufzuzählen, aber viel einfacher und immer noch sehr informativ ist es zu sagen, dass beim normalen Werfen mit einem symmetrischen Würfel die Wahrscheinlichkeit für jede Augenzahl 1/6 ist.

Diese Informativität besteht nur, wenn die *tatsächlichen* relativen Häufigkeiten beim Werfen solcher Würfel in der Nähe von 1/6 liegen, der Einfachheits-Vorzug nur, wenn die entsprechenden Ergebnisfolgen weitgehend regellos sind. Würden die Resultate beim Würfeln einer bestimmten Regel folgen, dann gäbe es eine einfache und noch deutlich informativere nicht-probabilistische Beschreibung dieses Aspekts unserer Welt. Wenn aber eine optimale Beschreibung des gesamten raumzeitlichen „Weltmosaiks" mit Wahrscheinlichkeiten operiert, dann gibt es ontische Wahrscheinlichkeiten im Sinne der Beste-System-Analyse, und die Wahrmacher der entsprechenden Aussagen sind auf eine indirekte Weise im Wesentlichen die faktischen relativen Häufigkeiten von Ereignissen. Dieser Zugang zu objektiven Wahrscheinlichkeiten ist gegenwärtig sehr populär. Albert

(2000) wendet ihn im Anschluss an Loewer auf die statistische Mechanik und damit potentiell auf makroskopische Wahrscheinlichkeiten insgesamt an. Ein Problem der Konzeption ist, dass sie mit ungedeckten Schecks operiert. Es ist ungeklärt, inwieweit man für Einfachheit und Informationsstärke, und erst recht für eine optimale Balance zwischen ihnen, Objektivität beanspruchen kann. Die Wahrscheinlichkeiten der Beste-System-Analyse sind, ohne dass dies intendiert wäre, womöglich doch epistemische, indem sie unter anderem von unserer kognitiven Ausstattung abhängen.

Die Propensity-Theorie. Alle bisher vorgestellten Konzeptionen objektiver Wahrscheinlichkeit sind reduktiv. Wie immer in der Metaphysik kann man aber das Gesuchte oder Gewünschte einfach in die Ontologie eintragen und behaupten, es sei da. Es ist dann eben „fundamental", „irreduzibel" oder „primitiv". Bei Wahrscheinlichkeiten führt dieses Manöver zur *Propensity-Theorie*, zuerst aufgestellt von Popper (1959, 1990). Ihr zufolge sind Wahrscheinlichkeiten „Propensitäten", nämlich Neigungen oder Tendenzen von experimentellen Arrangements zur Hervorbringung bestimmter Resultate:

Wenn bei einem Zufallsexperiment ***Z*** eine Tendenz der Stärke p zur Herbeiführung des Ereignisses E besteht, dann ist p die objektive Wahrscheinlichkeit von E bei der Durchführung von ***Z***.

Das ist ersichtlich ziemlich nichtssagend. Die Propensity-Theorie bietet zunächst einmal nur suggestive Ausdrücke an, und wenn man sich fragt, was es damit auf sich hat, wird man entweder auf den Wahrscheinlichkeitsbegriff oder auf andere Deutungen desselben zurückgeführt. So liegt es nahe, Propensitäten als Dispositionen (siehe Kap. III.A.5) bestimmter Art zu begreifen. Wenn man aber angeben soll, worin die Manifestationen dieser Dispositionen bestehen, dann kommen entweder relative Häufigkeiten oder erneut Wahrscheinlichkeiten ins Spiel: Nach Popper manifestieren sich Propensitäten in relativen Häufigkeiten auf lange Sicht, nach Mellor (1971) dagegen in Wahrscheinlichkeitsverteilungen. Auf diese Weise wird keine eigenständige Interpretation von Wahrscheinlichkeiten erreicht (siehe Eagle 2004, Vf. 2004).

Von der Propensity-Theorie bleibt im Kern folgendes: Ontische Wahrscheinlichkeiten werden als irreduzible Bestandteile der Wirklichkeit betrachtet. Die „Tendenzen" operieren in jedem Einzelfall eines echten Zufallsexperimentes, und tatsächliche relative Häufigkeiten ergeben sich daraus als ein (wiederum nur probabilistisch mit ihnen verbundenes) Symptom. Die Theorie erscheint so als eine Verallgemeinerung der Auffassung von Kausalbeziehungen als notwendigen Verknüpfungen, die konkret in der Welt vorliegen (siehe Kap. V.C.2). Auch über solche Verknüpfungen lässt sich nichts Informatives sagen, wie sie „funktionieren", bleibt verborgen. Vielleicht ist die Welt einfach so, dass sie Derartiges enthält und in ihrem Verlauf davon bestimmt wird.

Aus begrifflichen Gründen muss es aber jedenfalls einen Zusammenhang zwischen ontischen und epistemischen Wahrscheinlichkeiten geben. Wie er auszubuchstabieren und zu begründen ist, ist eine Frage, die sich für *jede* Konzeption objektiver Wahrscheinlichkeit stellt. Gemäß der Propensity-Theorie gibt es fundamentale Naturgrößen, denen eine nicht weiter erklärbare normative Kraft innewohnt: nämlich objektiv richtige Überzeugungsgrade zu liefern. Der Witz ist nun, dass man über Propensitäten kaum mehr sagen kann als eben dies und diesen Zusammenhang daher auch benutzen muss, um epistemisch auf sie zuzugreifen. Nach Mellor (1971) hat zuerst Lewis (1980) einen solchen Ansatz in voller Klarheit verfolgt. Die Schlüsselrolle spielt dabei sein viel diskutiertes „Hauptprinzip“ (Principal Principle).

Die Projektionstheorie. Der Gedanke einer *Projektionstheorie* der objektiven Wahrscheinlichkeit liegt dann sehr nahe; Lewis selber spricht von „chance as objectified credence“. Einen solchen Ansatz verfolgen konsequent Logue (1995) und Spohn (2010), und bereits Skyrms (1980) konzipiert objektive Wahrscheinlichkeiten als „widerständige“ (resilient) Überzeugungsgrade. Zugänge dieser Art setzen freilich nicht nur voraus, dass die subjektive oder epistemische Seite des Wahrscheinlichkeitsbegriffs besser verstanden ist als die objektive oder ontische, sondern auch, dass sie unabhängig von dieser theoretisch erfasst werden kann.

Weiterführende Literatur

Childers, Timothy (2013): *Philosophy and Probability*. Oxford University Press.

Gillies, Donald (2000): *Philosophical Theories of Probability*. Routledge.

Hájek, Alan und Christopher Hitchcock (2016): *The Oxford Handbook of Probability and Philosophy*. Oxford University Press.

Schurz, Gerhard (2015): *Wahrscheinlichkeit*. De Gruyter.

Literatur

Abrams, Marshall (2012): „Mechanistic Probability“. In: *Synthese* 187, 343–375.

Albert, David (2000): *Time and Chance*. Harvard University Press.

Carnap, Rudolf (1950): *Logical Foundations of Probability*. University of Chicago Press.

Eagle, Antony (2004): „Twenty-One Arguments Against Propensity Analyses of Probability“. In: *Erkenntnis* 60, 371–416.

Earman, John (1992): *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. MIT Press.

Engel, Eduardo (1992): *A Road to Randomness in Physical Systems*. Springer Verlag.

Hacking, Ian (1975): *The Emergence of Probability*. Cambridge University Press.

Hájek, Alan (1997): „‘Mises Redux’ – Redux: Fifteen Arguments Against Finite Frequentism“. In: *Erkenntnis* 45, 209–227.

Hájek, Alan (2003): „What Conditional Probability could not be“. In: *Synthese* 137, 273–323.

Hájek, Alan (2009): „Fifteen arguments against hypothetical frequentism“. In: *Erkenntnis* 70, 211–235.

Howson, Colin und Peter Urbach (1989): *Scientific Reasoning. The Bayesian Approach*. Open Court, 3. Auflage 2006.

Jeffrey, Richard (2004): *Subjective Probability – The Real Thing*. Cambridge University Press.

Keynes, John Maynard (1920): *A Treatise on Probability*. Macmillan.

Kolmogorov, Andrej (1933): *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer Verlag.

von Kries, Johannes (1886): *Die Principien der Wahrscheinlichkeitsrechnung*. J.C.B. Mohr.

Lewis, David (1980): „A Subjectivist's Guide to Objective Chance“. Wieder abgedruckt in: David Lewis, *Philosophical Papers Vol. II*. Oxford University Press 1986, 83–132.

Lewis, David (1994): „Humean Supervenience Debugged“. In: *Mind* 103, 473–490.

Logue, James (1995): *Projective Probability*. Oxford University Press.

Mellor, Hugh (1971): *The Matter of Chance*. Cambridge University Press, 2. Auflage 2004.

von Mises, Richard (1928): *Wahrscheinlichkeit, Statistik und Wahrheit*. Springer Verlag, 4. Auflage 1971.

Popper, Karl (1959): „The Propensity Interpretation of Probability“. In: *The British Journal for the Philosophy of Science* 10, 25–42.

Popper, Karl (1990): *A World of Propensities*. Thoemmes.

Rosenthal Jacob (2004): *Wahrscheinlichkeiten als Tendenzen. Eine Untersuchung objektiver Wahrscheinlichkeitsbegriffe*. mentis Verlag.

Rosenthal, Jacob (2012): „Probabilities as Ratios of Ranges in Initial-State Spaces“. In: *Journal of Logic, Language and Information* 21, 217–236.

Schneider, Ivo (1988) (Hrsg.): *Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933*. Wissenschaftliche Buchgesellschaft.

Skyrms, Brian (1980): *Causal Necessity*. Yale University Press.

Spohn, Wolfgang (2010): „Chance and Necessity: From Humean Supervenience to Humean Projection“. In: Ellery Eells und James Fetzer (Hrsg.), *The Place of Probability in Science*, Springer Verlag, 101–131.

Strevens, Michael (2011): „Probability out of Determinism“. In: Claus Beisbart and Stephan Hartmann (eds.), *Probabilities in Physics*, Oxford University Press.

Strevens, Michael (2013): *Tychomancy. Inferring Probability from Causal Structure*. Harvard University Press.

Williamson, Jon (2010): *In Defence of Objective Bayesianism*. Oxford University Press.