Aus:
Andreas Bartels /
Manfred Stöckler (Hg.)
Wissenschaftstheorie
Paderborn 2007

KAPITEL 5

Induktion und Bestätigung

Jacob Rosenthal

1. Einleitung

Wissenschaftliche Theorien, Modelle oder Gesetzesaussagen müssen sich empirisch bewähren, sie werden durch Beobachtungen bestätigt, widerlegt oder in Frage gestellt. Dabei verhält es sich so, dass sie über einen bestimmten Bereich der Wirklichkeit Behauptungen mit Allgemeinheitsanspruch aufstellen und damit weit über die tatsächlich angestellten Beobachtungen hinausgehen. Es fragt sich, wie sich ein solcher Schritt rechtfertigen lässt. Es geht in diesem Kapitel um den erkenntnistheoretischen Status wissenschaftlicher Allaussagen, unter die sich Theorien, Modelle und Gesetzesaussagen als wichtige Spezialfälle subsumieren lassen. In welchem Verhältnis stehen diese Allaussagen zu den einzelnen Beobachtungen, an denen sie sich bewähren müssen? Können sie jemals beanspruchen, gesichertes Wissen darzustellen, oder werden sie durch die vorliegenden Beobachtungsdaten nur wahrscheinlich gemacht, oder handelt es sich gar nur um spekulative Hypothesen, bei denen man sich grundsätzlich darauf einzustellen hat, dass bereits die nächste Beobachtung sie wieder umstößt? Es sind diese und verwandte Fragen, die im folgenden einführend behandelt werden.

deduktive und induktive Schlüsse

Es handelt sich hierbei um einen Spezialfall des sogenannten *Induktionsproblems*. Bei einem gültigen *deduktiven* Schluss wird eine Konklusion k aus Prämissen $p_1, p_2, \ldots, p_n$ logisch abgeleitet (deduziert). Die Konklusion ist dabei logisch in den Prämissen enthalten; wenn diese wahr sind, ist auch die Konklusion wahr. Es ist nicht möglich, dass die Prämissen wahr, die Konklusion aber falsch ist. Bei *induktiven* Schlüssen geht demgegenüber der Gehalt der Konklusion über den Gehalt der Prämissen hinaus. Es ist möglich, dass der Schluss gültig und die Prämissen alle wahr, die Konklusion aber dennoch falsch ist. Induktive Schlüsse heißen deshalb auch „wahrheitserweiternde" Schlüsse. Man holt hier etwas als Konklusion heraus, das man nicht schon mit den Prämissen implizit hineingesteckt hatte. Das Induktionsproblem besteht in seiner allgemeinsten Form in der Frage, wie sich derartige Schlüsse rechtfertigen lassen, und in welcher Beziehung (die ja nicht das logische Enthaltensein ist) bei ihnen die Konklusion zu den Prämissen steht. In wissenschaftstheoretischen Kontexten werden die Prämissen oft als „Evidenzen" bezeichnet und entsprechend mit $e_1, e_2, \ldots, e_n$ abgekürzt. Man stellt sich dabei vor, dass es sich um Sätze oder Aussagen handelt, die Beob-

achtungen referieren. Die Konklusion wird als „Hypothese" bezeichnet und mit *h* abgekürzt. An diese Konventionen werde ich mich im folgenden halten. Es sind zunächst einmal nicht mehr als Bezeichnungen, die in unserem Kontext recht verbreitet sind.

2. Typen induktiver Schlüsse

Zur Diskussion steht also das Verhältnis einer Hypothese *h* zu gegebenen Evidenzen $e_1, e_2, \ldots, e_n$, wobei der Gehalt der Hypothese typischerweise nicht in dem Gehalt der Evidenzen auf-, sondern über ihn hinausgeht. Sehen wir uns nun, um das sehr allgemeine Schema etwas zu konkretisieren, verschiedene Typen induktiver Schlüsse an.

Zunächst einmal muss die Hypothese nicht unbedingt eine Allaussage sein; das Schema ist insofern allgemeiner, als die Eingangsproblematik andeutet. Ein anderer wichtiger Fall ist, dass die Hypothese eine (singuläre) *Vorhersage* ist, man also von bereits gemachten direkt auf eine künftige Beobachtung oder auf einen anderen singulären Tatbestand schließt. In der Regel wird aber eine Vorhersage aus einer allgemeinen (Gesetzes-)Aussage plus singulären Randbedingungen *deduktiv* abgeleitet, während sich die besagte Gesetzesaussage induktiv auf vorangegangene Beobachtungen stützt. Der Fall der Allaussage (und hier besonders die Unterfälle von Gesetzmäßigkeiten, Theorien oder Modellen) ist daher wissenschaftstheoretisch besonders bedeutsam.

numerative Induktion

Was die Struktur der Hypothese im Verhältnis zur Struktur der Evidenzen angeht, so ist der einfachste Fall, dass die Hypothese eine *Extrapolation* oder *Verallgemeinerung* der Evidenzen darstellt. Alle bisher beobachteten Raben waren schwarz, alle bisher gefundenen Smaragde grün, daraus schließen wir nun, dass auch der nächste Rabe bzw. Smaragd, dem wir begegnen, schwarz bzw. grün sein wird. Oder wir schließen sogar, dass alle Raben schwarz bzw. alle Smaragde grün sind. Man spricht hier von *enumerativer* („aufzählender") Induktion, vom typischen „Lernen aus Erfahrung". Dieser Typ ist sowohl systematisch, wegen seiner übersichtlichen Struktur, als auch historisch bedeutsam, es ist der klassische Fall eines induktiven Schlusses. Auch Schlüsse der folgenden Form zählt man dazu: „*n*% aller bisher beobachteten Fälle von *R* waren auch Fälle von *S*, also haben ungefähr *n*% aller *R* das Merkmal *S*." Die meisten Behandlungen des Induktionsproblems stellen die aufzählende Induktion in den Vordergrund, und nicht selten wird „Induktion" überhaupt nur darauf bezogen. Wir müssen also eine weite und eine engere Verwendung des Begriffs unterscheiden: Induktion im Gegensatz zu Deduktion, wie oben eingeführt, und Induktion als Verallgemeinerung. Die besondere Bedeutung dieses Typs sollte aber nicht darüber hinwegtäuschen, dass wissenschaftliche Gesetzesaussagen, Theorien oder Modelle kaum jemals in diesem Verhältnis zu den bestätigenden Evidenzen stehen. Denn es handelt sich dabei zwar um Allaussagen, aber um solche, die die besagten Evidenzen, und ihre Verallgemeinerung gleich mit, *erklären* sollen. So werden in der Newtonschen

Mechanik Geschwindigkeitsänderungen (Beschleunigungen) von Körpern durch die Einwirkung von Kräften erklärt. Aber gerade durch die Einführung dieses für die Theorie charakteristischen Begriffs und der für ihn gültigen Prinzipien geht die Theorie entscheidend über die Beobachtungen und ihre bloße Verallgemeinerung hinaus.

Bei den Beispielen für induktives Schließen wird natürlich durchweg vorausgesetzt, dass es sich bei den Hypothesen nicht um begriffliche Wahrheiten handelt. Würde die Schwärze zu unserem Begriff eines Raben gehören, dann wäre sofort klar, dass und warum alle Raben schwarz sind. „Alle Raben sind schwarz" wäre dann eine Aussage wie „Alle Rappen sind schwarz", und um *das* zu wissen, braucht man natürlich keinerlei Beobachtungen anzustellen, geschweige denn aus Beobachtungen zu verallgemeinern. Es würde sich dann um eine begrifflich oder *analytisch* wahre Aussage handeln, um eine Aussage, die wahr wäre allein aufgrund der Bedeutungen der in ihr auftretenden Begriffe. Wir gehen hier immer implizit davon aus, dass die Hypothesen wie auch die Evidenzen keine solchen, sondern *synthetische* Aussagen sind. So offensichtlich dieser Punkt ist, er verweist doch auf problematisches Terrain, nämlich auf die Frage, wie sich denn erkennen lässt, oder sogar, aufgrund wovon überhaupt feststeht, ob eine bestimmte Aussage analytischen oder synthetischen Charakter hat. Woran entscheidet sich denn, dass die Schwärze im Begriff des Rappen enthalten ist, nicht aber im Begriff des Raben? Was sind Bedeutungen, wie kommt ein Begriff zu seiner Bedeutung, und ist diese etwas scharf Umrissenes? Seit W.V.O. Quines einflussreichem Angriff auf die analytisch-synthetisch-Unterscheidung und seiner damit verbundenen These von der prinzipiellen Unbestimmtheit der Bedeutung ist dieser Fragekomplex ein zentrales Thema der Sprachphilosophie.

Abduktion

Das Verhältnis von bestätigenden Evidenzen zu Theorien, Modellen oder Gesetzesaussagen ist nicht eines aufzählender Induktion, sondern der sogenannten ↑*Abduktion*: des Schlusses auf die beste Erklärung (vgl. den Beitrag von *Bartels: Wissenschaftlicher Realismus* in diesem Band). Ein solcher liegt z.B. auch vor, wenn bei einem Verbrechen von vorhandenen Indizien auf eine bestimmte Person als Täter geschlossen wird. Schlüsse auf die beste Erklärung gehen in einem anderen Sinn über die vorliegenden Evidenzen hinaus als Verallgemeinerungen, oft implizieren sie aber derartige Verallgemeinerungen. Die Abduktion eignet sich auch besonders, um zwei weitere wichtige Aspekte unseres Themas explizit zu machen. Erstens eine Abgrenzung. Es geht hier nicht um die Frage, wie man, von Beobachtungen ausgehend, auf aussichtsreiche Hypothesen kommen kann, es geht nicht um *Entdeckung*. Sondern es geht um *Rechtfertigung*, darum, in welchem Sinne und wie stark vorliegende Evidenzen eine *gegebene* Hypothese stützen. Die Unterscheidung zwischen Entdeckungs- und Rechtfertigungskontext ist eine Standardunterscheidung in der Wissenschaftstheorie, die an verschiedenen Punkten wichtig wird. Beim Schluss auf die beste Erklärung ist klar, dass einem

Entdeckungs- versus Rechtfertigungskontext

Erklärungen und gar gute Erklärungen für ein bestimmtes Phänomen oft nicht leicht in den Sinn kommen, dass man sie aber haben muss, um sich Gedanken über ihre Stützung durch Evidenzen machen zu können. Zweitens ist es auch umgekehrt möglich, dass einem mehrere gute, aber inkompatible Erklärungen für bestimmte Evidenzen einfallen, und dieser Punkt ist für das Thema „Bestätigung" nun von höchster Relevanz. Es können konkurrierende Erklärungen für einen Phänomentyp vorliegen, Erklärungen, die man jeweils als korrekt zu akzeptieren geneigt wäre, wenn da nicht noch die Alternative wäre. Die *prima facie* bestehende Möglichkeit, dass ein und dieselben Evidenzen mehrere inkompatible Hypothesen jeweils in hohem Maße stützen, stellt eine erhebliche Verlegenheit für die Bestätigungstheorie dar. Das Problem ergibt sich, anders als man denken könnte, auch bei der aufzählenden Induktion. Es gibt prinzipiell immer mehrere Möglichkeiten, bestimmte Beobachtungen zu verallgemeinern, und eine Lösung des klassischen Problems der enumerativen Induktion setzt voraus, dass man alle diese Möglichkeiten bis auf eine mit guten Gründen ausschließen kann. In besonders krasser Form wird uns dieses Problem bei Goodmans „neuem Rätsel der Induktion" entgegentreten.

3. Das klassische Induktionsproblem

Das klassische Induktionsproblem betrifft die aufzählende Induktion und wurde zuerst von David Hume deutlich formuliert.[1] Auch wenn alle bisher beobachteten Raben ohne Ausnahme schwarz gewesen sind, so ist es doch ohne weiteres möglich – nämlich ohne Widerspruch denkbar, klar und deutlich vorstellbar – dass der nächste Rabe oder sogar alle Raben, denen wir zukünftig begegnen, eine andere Farbe haben. Auch wenn bislang jeden Tag aufs neue die Sonne aufgegangen ist, so ist es doch auch hier ohne weiteres vorstellbar, dass sie es künftig nicht mehr tun wird. Dies gilt für alles, was wir aus der Erfahrung gelernt haben oder gelernt zu haben glauben, sogar und insbesondere für die von uns angenommenen Naturgesetze. Was ist es, das uns glauben lässt, es werde die Zukunft der Vergangenheit gleichen, oder genauer: die von uns (noch) nicht beobachteten den beobachteten Fällen? Hume zufolge lassen sich für diesen Glauben keine Gründe angeben. Es besteht lediglich eine durch die wiederholte Erfahrung hervorgerufene und immer weiter verstärkte psychologische Verknüpfung zwischen Vorstellungen, z.B. zwischen der Vorstellung eines Raben und der Vorstellung seiner Schwärze, die uns eine subjektive Gewissheit verleiht, dass jeder und insbesondere jeder künftig von uns beobachtete Rabe schwarz sei. Diese Gewissheit hat aber keine vernünftige Grundlage, denn – und jetzt kommt das Argument – wenn man die aufzählende Induktion dadurch rechtfertigen wollte, dass die Erfahrung ja zeige, dass man mit dergleichen Schlüssen Erfolg habe, dann würde man

duktiver Rechtfertigungszirkel oder -regress

[1] Am besten in der 1748 erschienenen *An Enquiry Concerning Human Understanding* (dt.: *Eine Untersuchung über den menschlichen Verstand*), Kap. 4 und 5.

damit wiederum einen Induktionsschluss vollziehen. Man würde aus der Tatsache, dass sich die Konklusionen früherer induktiver Schlüsse als wahr erwiesen haben, ableiten, dass dies auch bei zukünftigen Schlüssen so sein werde, und dadurch die aufzählend-induktive Schlussmethode zu rechtfertigen versuchen. Aber dieser Schluss von dem vergangenen auf den zukünftigen Erfolg induktiven Schließens wäre gerade selbst wieder eine aufzählende Induktion, deren Rationalität gerade in Frage steht. Auf diesem Wege gerät man also in einen Begründungszirkel.

Wir müssen hier übrigens Acht geben, wenn wir von dem „Erfolg" früherer Induktionsschlüsse sprechen. Falls deren Konklusion eine Allaussage war und nicht nur die Extrapolation von Beobachtungen auf endlich viele weitere Fälle, kann die Wahrheit dieser Konklusion ja auch später nicht festgestellt worden sein, und *insofern* sind wir auch nicht berechtigt zu sagen, der frühere Induktionsschluss sei erfolgreich gewesen. Was hier mit „Erfolg" gemeint ist, ist natürlich, dass sich die weiteren Fälle, die später beobachtet worden sind, der Hypothese gemäß verhalten haben. Das bedeutet, dass die Idee, Induktion induktiv, durch Verweis auf frühere Erfolge, zu rechtfertigen, unangesehen aller weiteren Kritikpunkte nur solche Induktionsschlüsse betreffen kann, deren Konklusion sich nur auf endlich viele Fälle bezieht. Das ist von vornherein eine gravierende Beschränkung. Aber auch wenn sie nicht bestünde, gilt jedenfalls, dass, da die Rationalität des aufzählend-induktiven Schlussprinzips eben in Frage steht, wir in einen Zirkel geraten, wenn wir die Rechtfertigung dadurch erbringen wollen, dass wir von seinem bisherigen Erfolg auf den zukünftigen schließen. Statt eines Zirkels kann man bei diesem Begründungsversuch auch eine Verschiebung des Problems diagnostizieren: Induktionen erster Stufe, die Raben, Sonnenaufgänge oder sonstige normale Gegenstände der Erfahrung betreffen, werden unter Rekurs auf eine Induktion zweiter Stufe gerechtfertigt, die Induktionen erster Stufe betrifft und aus ihrem vergangenen ihren zukünftigen Erfolg ableitet. Nun steht aber natürlich die Berechtigung dieses neuen, zweitstufigen Schlussprinzips in Frage usw. Wie es häufig vorkommt, kann sich das Problem entweder als *Zirkel* oder als *Regress* darstellen.

Die Situation verbessert sich auch nicht, wenn man für die Konklusion induktiver Schlüsse statt Sicherheit bloß Wahrscheinlichkeit in Anspruch nimmt. Dies ist wohl von vornherein plausibler, aber es fragt sich, aufgrund wovon es z.B. wahrscheinlich sein soll, dass die Sonne auch künftig aufgeht oder auch künftig beobachtete Raben schwarz sein werden. Eine bloße subjektive (Einschätzung der) Wahrscheinlichkeit steht hier nicht zur Debatte, denn dass wir uns der Konklusionen unserer induktiven Schlüsse subjektiv sicher oder ziemlich sicher sind, sie also wenigstens für wahrscheinlich wahr *halten*, bestreitet Hume ja nicht. Wie aber könnte man diesen Wahrscheinlichkeitsurteilen ein objektives Fundament verleihen? Wieder kommt einem der Verweis auf den vergangenen Erfolg induktiven Schließens in den Sinn: Weil man in der Vergangenheit mit solchen Schlüssen häufig Erfolg hatte, schließt man, dass es sich generell und insbesondere auch in der Zukunft so verhalten werde, dass die meisten derartigen Schlüsse eine wahre Konklusion aufweisen, und deshalb der einzelne zur Diskussion stehende Schluss

Wahrscheinlichke induktiver Konklusionen

wahrscheinlich eine wahre Konklusion besitzt. Aber dies ist natürlich wiederum induktiv geschlossen. Aus der vergangenen soll auf die zukünftige Erfolgsbilanz geschlossen werden, sei diese Erfolgsbilanz nun eine 100%ige, was unseren Verallgemeinerungen Sicherheit, oder bloß eine hohe, was ihnen Wahrscheinlichkeit verleihen soll.

Es ist wichtig, sich diesen zusätzlichen Punkt klar zu machen, um die Reichweite des von Hume entdeckten skeptischen Problems richtig einzuschätzen. Würde die Humesche Argumentation nur zeigen, dass wir uns der Konklusion aufzählend-induktiver Schlüsse nicht begründetermaßen *sicher* sein können, dann würde sich das Problem in Grenzen halten. Sicherheit erwarten wir ja meistens gar nicht. Wir sind uns von vornherein nicht vollständig sicher, dass man z.B. die Farben von Raben oder Smaragden verallgemeinern darf, auch wenn wir sehr viele davon beobachtet haben. Bei Naturgesetzen ist es anders: Wir sind uns schon sicher, dass auch morgen die Schwerkraft wie bisher wirksam sein wird, und es würde uns deshalb irritieren, hier rationalerweise nur eine hohe Wahrscheinlichkeit annehmen zu dürfen. Aber diese Revision unserer Überzeugungen wäre relativ harmlos, verglichen mit der, die uns Hume tatsächlich zumutet. Denn der Verweis auf Erfolg in der Vergangenheit rechtfertigt eben noch nicht einmal, die Konklusionen aufzählend-induktiver Schlüsse als *wahrscheinlich* wahr anzusehen, egal, wie viele Fälle beobachtet worden sind. Der diagnostizierte Zirkel bzw. Regress stellt sich hier ganz genauso ein, weil dieses „wahrscheinlich" eine objektive Bedeutung haben soll. Dass wir die Konklusionen typischer induktiver Schlüsse als wahrscheinlich oder sogar sicher wahr *behandeln*, wird, wie gesagt, von Hume weder bestritten noch kritisiert. Er beharrt nur darauf, dass diese *subjektive* Gewissheit oder Wahrscheinlichkeit das Produkt der Gewöhnung und nicht rational begründbar ist.

avid Hume (1711–1776) ist einer der Hauptvertreter des britischen *Empirismus*, einer enkrichtung, die die Bedeutung der (Sinnes-)Erfahrung für unser Wissen hervorhebt und eses in möglichst großem Umfang auf jene zu gründen unternimmt, anstatt auf sogenann- „reines", nämlich erfahrungsunabhängiges Denken. Es gibt auf den britischen Inseln eine ıge Tradition empiristischen Philosophierens, die bis ins Mittelalter zurückreicht und bis ute anhält. Als „britischen Empirismus" im engeren Sinne bezeichnet man aber nur die ochphase im späten 17. und 18. Jahrhundert, die sich besonders mit den Namen John ocke, George Berkeley und David Hume verbindet. Wichtige Vorläufer im Mittelalter sind oger Bacon, Johannes Duns Scotus und Wilhelm von Ockham, in der frühen Neuzeit ancis Bacon und Thomas Hobbes. Mit Hume, der ihr eine spezifisch skeptische Wenıng gibt, erreicht diese Denktradition ihren ersten Kulminationspunkt. Anders als seine orläufer erkennt Hume sehr deutlich, dass sich auf der Basis sinnlicher Erfahrung allein esentliche Komponenten unseres Weltbildes *nicht* rechtfertigen lassen. Dies betrifft neben duktiven Schlüssen vor allem den Kausalitätsbegriff. Hume hält einerseits am Empirismus

fest, ist aber andererseits über dessen Reichweite desillusioniert und wird so zu skeptischen Konklusionen gedrängt. Für diesen Skeptizismus ist wiederum charakteristisch, dass Hume ihn nicht allzu schwer nimmt: Er betrifft nur die philosophische Theorie, während sich ansonsten jeder Mensch, auch der Philosoph, in seinen Denkweisen von fest eingewurzelten Gewohnheiten leiten lässt. Diese, z.B. eben das induktive Verallgemeinern oder das Ursache-Wirkungs-Denken, das notwendige Verknüpfungen in der Natur annimmt, lassen sich psychologisch erklären, in ihrem Zustandekommen nachvollziehen, aber nicht vernünftig begründen. Das macht jedoch nichts aus, die Natur leitet uns durch diese Gewohnheiten sicherer als unsere Verstandestätigkeit es vermöchte. Bedeutende spätere Vertreter der britisch-empiristischen Traditionslinie sind John Stuart Mill, Bertrand Russell und Alfred Ayer. Sie stehen stark unter dem Einfluss Humes, dessen Stellenwert in der englischen Philosophie des 19. und frühen 20. Jahrhunderts dem Kants in der deutschen Philosophie dieser Zeit durchaus vergleichbar ist.

4. Lösungsvorschläge

Für das klassische Induktionsproblem sind sehr verschiedene Lösungsvorschläge unterbreitet worden. Ich will sie grob klassifizieren und für jeden Typ wenigstens ein Beispiel diskutieren. Damit wird nicht der Anspruch auf Vollständigkeit erhoben, denn Vorschläge desselben Typs können durchaus unterschiedlich vorgehen. Es geht lediglich darum, einen Eindruck davon zu vermitteln, auf welche grundsätzlichen Weisen man das Problem angehen kann und welche Schwierigkeiten dabei typischerweise auftreten. Zu denken ist im folgenden weiterhin immer an die *enumerative* Induktion.

Reaktionsmöglichkeiten auf das Problem

Zunächst einmal kann man versuchen, der Humeschen Herausforderung direkt zu begegnen und Gründe dafür zu liefern, dass aufzählend-induktive Schlüsse unter bestimmten, von uns kontrollierbaren Bedingungen sicher oder wahrscheinlich oder typischerweise wahre Konklusionen besitzen. Bei dieser Begründung kann man sich entweder a) auf bestimmte Erfahrungstatsachen berufen oder b) ein apriorisches (erfahrungsunabhängiges) Argument zu geben versuchen. Hume war der Auffassung, dass wahre Aussagen entweder *„matters of fact"* (Tatsachen) oder *„relations of ideas"* (Beziehungen zwischen Vorstellungen) konstatieren. Eine Rechtfertigung für Induktion dadurch zu geben, dass man ausschließlich *„relations of ideas"* betrachtet, also analytische Feststellungen trifft, hielt er für ausgeschlossen, da wir uns ohne weiteres denken können, dass unsere zukünftigen Beobachtungen ganz andere Resultate liefern als wir es gewohnt sind. Beruft man sich bei der Rechtfertigung von Induktion dagegen zusätzlich oder ausschließlich auf *„matters of fact"*, dann können das nur Tatsachen bezüglich des früheren Erfolgs induktiver Schlüsse sein, und die Argumentation wird zirkulär. Daher glaubte Hume, dass eine skeptische Schlussfolgerung

unausweichlich sei. Es ist *prima facie* plausibel, dass es nur die zwei genannten prinzipiellen Rechtfertigungsmöglichkeiten gibt, und dass sie beide aus den von Hume erkannten Gründen scheitern. Will man dem ausweichen, dann muss man entweder den Weg a) oder den Weg b) doch erfolgreich zu beschreiten versuchen, unter Widerlegung der Humeschen Bedenken, oder man muss nach einer dritten Begründungs- oder besser: überzeugenden Reaktionsmöglichkeit suchen. In der Tat ist das Induktionsproblem ein gutes Beispiel dafür, wie man sich gegenüber Fragestellungen, die nur bestimmte Antwortalternativen zuzulassen scheinen und eine dilemmatische Struktur haben, ganz unerwartet verhalten kann. Man ist, um ein bestimmtes Problem zu lösen, nicht gezwungen, sich auf den Rahmen oder Hintergrund einzulassen, der es hervorbringt, freilich bekommt dann „Lösung" einen uneigentlichen Sinn. So kann man im Falle des Induktionsproblems c) andere Formen von Rechtfertigung versuchen, die nicht auf die Wahrheit der Konklusion aufzählend-induktiver Schlüsse abzielen, d) zu zeigen versuchen, dass man induktive Schlüsse gar nicht braucht, oder e) das Verlangen nach einer Begründung als trivial oder sinnwidrig zurückweisen. Bei allen dreien wird Humes skeptischer Punkt im Effekt konzediert, da nicht gegen ihn argumentiert wird. Stattdessen wird auf sehr verschiedene Weise dafür argumentiert, dass er unsere Rechtfertigungspraxis nicht trifft, wir also rationalerweise weitermachen können wie bisher. Aus Platzgründen werde ich Positionen des Typs e) nicht diskutieren.[2]

empirische Rechtfertigung

a) Induktion durch Berufung auf bestimmte Erfahrungstatsachen zu rechtfertigen, heißt immer, zirkulär zu argumentieren: „Aus Erfahrung wissen wir, dass Induktion funktioniert." Man kann sich aber fragen, ob das die Begründung vollkommen entwertet. Immerhin *kann* man das Verfahren aufzählend-induktiven Schließens insofern mit positivem Ergebnis auf sich selbst anwenden, als man aus den vergangenen Erfolgen solcher Schlüsse die Vorhersage ableitet, dass auch zukünftige Induktionen eine wahre Konklusion aufweisen werden. Die Möglichkeit solcher Selbstanwendung ist nicht selbstverständlich, und erst recht nicht, dass sie zu einem positiven Resultat führt. Es wäre ja auch denkbar, dass eine Induktion auf der zweiten Stufe entweder gar kein Ergebnis hätte oder ein negatives, in welchem Falle sich Induktion selbst unterminieren würde. Ihre zweitstufige Anwendung würde uns dann zukünftige erststufige *Miss*erfolge erwarten lassen. Dass das Gegenteil der Fall ist, ist, wie gesagt, nicht selbstverständlich, und diese Tatsache kann von einem Vertreter der Rationalität von Induktion durchaus als eine Bestätigung seiner Position aufgefasst werden. Der Skeptiker wird sich davon freilich nicht beeindrucken lassen, denn gerade das, was er in Zweifel zieht, ist in dem Argument vorausgesetzt. Die Frage ist nun aber, ob deswegen auch *wir*, die wir keine Induktionsskeptiker sind, das besagte Argument für wertlos erklären müssen. Wir glauben eben von vornherein an die Rationalität oder Berechtigung induktiver Schlüsse, sie zählen zu unseren Begründungsressourcen, und wenn die

[2] Ein wichtiges Beispiel ist Strawson (1952), Kap. 9.

Frage nach ihrer Begründung aufgeworfen wird, dann zeigt sich, dass wir *mit eben diesen Ressourcen* tatsächlich eine solche liefern können.

Bei dieser Sichtweise der Argumentationssituation scheint ein Patt zu bestehen: Der Skeptiker ist durch das in Rede stehende Argument keineswegs gezwungen, die Rationalität von Induktion einzuräumen, aber wir, die Opponenten, müssen sie uns auch nicht ausreden lassen, denn wir können ja (unter Berufung auf Induktion) eine Rechtfertigung für Induktion angeben. Die Crux an dieser Strategie ist, dass sie den Rechtfertigungsbegriff relativiert. Es sieht danach aus, als würde die Rationalität oder das Gerechtfertigtsein einer Annahme von den Begründungsressourcen bestimmt, die jeweils zugrunde gelegt werden, und da diese sehr verschieden aussehen können, lässt sich im Prinzip allerhand rechtfertigen. Rechtfertigung wäre immer relativ zu einem Hintergrund, der im Extremfall von Subjekt zu Subjekt verschieden sein könnte.

apriorische Rechtfertigung

b) Ein apriorisches Argument für das Funktionieren induktiver Schlüsse sieht wie folgt aus. Nehmen wir wieder die Hypothese, dass alle Raben schwarz sind. Wenn sie falsch wäre, es also andersfarbige Raben gäbe, dann müsste doch wohl jemand schon einen gesehen haben. Genauer gesagt: Wenn es einen *nennenswerten* Prozentsatz an nicht-schwarzen Raben gäbe, dann wäre es *sehr wahrscheinlich*, dass schon einmal ein nicht-schwarzer Rabe beobachtet worden wäre. Wenn das nicht der Fall ist, können wir umgekehrt schließen, dass es sehr wahrscheinlich ist, dass zumindest ein hoher Prozentsatz (nahezu 100%) aller Raben schwarz ist. Soweit das Argument. Es beansprucht nicht, die Konklusion „alle Raben sind schwarz" zu etablieren, sondern etwas Schwächeres, nämlich „ein sehr hoher Anteil aller Raben ist schwarz", und es verleiht dieser Hypothese auch keine Sicherheit, sondern nur eine hohe Wahrscheinlichkeit. Aber das wäre ja schon eine ganze Menge.

Um diesen Anspruch zu beurteilen, müssen wir uns den Terminus „sehr wahrscheinlich" näher ansehen. Wenn man ihm eine subjektive Bedeutung verleiht, dann erhalten wir nichts über das hinaus, was Hume selbst sagt. „Sehr wahrscheinlich" hieße dann „wir rechnen sehr stark damit" (im Beispiel, dass alle oder fast alle Raben schwarz sind), und das wissen wir natürlich. Die Frage ist gerade, wie sich diese Einstellung begründen lässt. Es muss also „wahrscheinlich" hier heißen „objektiv wahrscheinlich". Die Idee ist in etwa, dass wir den Raben in ähnlicher Weise begegnen wie man nach gutem Durchmischen Kugeln aus einer Urne zieht, dass wir also unterstellen dürfen, dass alle Raben die objektiv gleiche Chance haben, von uns beobachtet zu werden. Stellen wir uns vor, wir würden allen Raben, die jemals gelebt haben, leben oder leben werden, eine Losnummer zuteilen, und dann nach jeweils gründlichem Mischen eine Nummer nach der anderen ziehen und die Farbe des zugehörigen Raben bestimmen. Wenn sich dann in vielen Zügen immer „schwarz" ergäbe, wäre es in der Tat objektiv sehr wahrscheinlich, dass ein sehr hoher Anteil aller Raben schwarz ist. Das Argument wäre gültig.

Man sieht nun aber, was die Schwierigkeit ist. Zukünftig lebende Raben können wir ja gar nicht beobachten. Es ist also nicht richtig, dass bei jedem Fall einer Raben-Beobachtung jeder Rabe dieselbe objektive Chance hat, gesichtet zu werden. Wenn wir dennoch mit dieser Modellannahme operieren, dann unterstellen wir, dass die von uns beobachteten Raben *repräsentativ* für alle Raben sind. Wir unterstellen insbesondere, dass unsere Chance, wie hoch oder niedrig sie auch immer sein mag, in der Zukunft einem nicht-schwarzen Raben zu begegnen, dieselbe bleibt wie bisher. So dass wir, was die Farbe von Raben angeht, uns bei der Beobachtung auf diejenigen Raben beschränken können, die während eines bestimmten Zeitabschnitts leben. Aber damit setzen wir gerade wieder voraus, was erst begründet werden sollte: dass uns die zukünftigen Raben kein anderes Bild vermitteln werden als die bisherigen es taten. Auch dieses apriorische Argument dafür, dass ein typischer Induktionsschluss mit hoher (objektiver!) Wahrscheinlichkeit eine wahre Konklusion besitzt, ist also zirkulär. Apriorisch ist der skizzierte Lösungsansatz deshalb, weil er ein zumindest *prima facie* erfahrungsunabhängiges Argument dafür anbietet, dass ein typischer induktiver Schluss, *gegeben* seine Prämissen, wahrscheinlich eine wahre Konklusion besitzt. *A priori* ist im Beispiel also nicht, dass alle (oder fast alle) Raben schwarz sind – dann bräuchte man ja gar nicht zu beobachten –, sondern dass, gegeben dass bisher viele Raben beobachtet wurden und alle schwarz waren, die Hypothese wahrscheinlich wahr ist.

Solche Lösungsvorschläge vom Typ b) machen in der Regel Annahmen, die zweifelhafter sind als das, was durch sie erwiesen werden soll; im besten Fall machen sie, wie eben gesehen, eine Annahme, die das zu Zeigende vorwegnimmt. Mit weitergehenden begrifflichen Ressourcen operiert das folgende Argument: Die *beste Erklärung* dafür, dass alle von uns bisher beobachteten Raben schwarz waren, ist eine *notwendige Verknüpfung* zwischen dem Rabesein und der Schwärze. Wir dürfen aufgrund unserer Beobachtungen vermuten, dass ein Rabe mit *naturgesetzlicher Notwendigkeit* schwarz ist oder die Schwärze gar zu seinen *essentiellen Eigenschaften* (↑ *essentielle Eigenschaft*) zählt, die er nicht verlieren kann, ohne aufzuhören, ein Rabe zu sein. Daraus würde natürlich folgen, dass alle Raben schwarz sind und insbesondere jeder von uns in Zukunft zu beobachtende (vgl. z.B. Armstrong (1983), Kap. 4.5 und 6.7; Armstrong (1995); BonJour (1998), Kap. 7). Insofern in diesem Argument „beste Erklärung" nichts anderes bedeutet als „wahrscheinlichste Erklärung", erhalten wir eine enge Verwandtschaft mit dem vorhergehenden. Dieses sagte, es wäre sehr unwahrscheinlich, nur schwarze Raben zu beobachten, wenn nicht fast alle tatsächlich schwarz wären, das jetzige sagt, es wäre sehr unwahrscheinlich, nur schwarze Raben zu beobachten, wenn nicht Raben mit Notwendigkeit schwarz wären. Insofern dies die weitergehende Aussage ist, ist das Argument noch problematischer. Bei dem vorigen hatten wir immerhin eine – allerdings unrealistische – Modellannahme zur Verfügung, bei deren Zugrundelegung sich *zeigen* ließ, dass die Evidenzen (alle bisher beobachteten Raben waren schwarz) die Hypothese (alle oder nahezu alle Raben sind schwarz) objektiv wahrscheinlich machen. Ein solches Modell fehlt uns

hier; wir wissen nicht zu sagen, aufgrund wovon und in welchem Sinne es objektiv unwahrscheinlich wäre, lauter schwarze Raben zu beobachten, wenn die behauptete notwendige Verbindung *nicht* bestünde. Zumal wir ja mit dem Urnenmodell bereits eine Modellsituation an der Hand haben, in welcher die beste Erklärung für eine Evidenz der genannten Art ohne notwendige Verknüpfungen auskommt: Die beste Erklärung dafür, dass wir nur rote Kugeln aus einer Urne ziehen, ist, dass sie nur rote Kugeln enthält, und nicht, dass Kugeln in dieser Urne (oder Kugeln in Urnen überhaupt) mit Notwendigkeit rot sind.

Lassen wir also offen, ob „beste Erklärung" „wahrscheinlichste Erklärung" bedeutet oder etwas anderes. Grundsätzlich macht das jetzige Argument die Rechtfertigung enumerativer Induktion abhängig vom Funktionieren des Schlusses auf die beste Erklärung, der noch problematischer ist. Wieso dürfen wir annehmen, die beste Erklärung für etwas sei wahr oder (in einem objektiven Sinne) wahrscheinlich wahr? Darauf müssten wir eine Antwort finden, die nicht wiederum von aufzählender Induktion abhängt. Sicherlich könnte man zu argumentieren versuchen, die Erfahrung habe gezeigt, dass Schlüsse auf die beste Erklärung häufig die Wahrheit träfen, aber das verbietet sich hier, denn es soll ja gerade umgekehrt aufzählende Induktion mithilfe des Schlusses auf die beste Erklärung begründet werden. Auch wenn sich aber eine befriedigende, von enumerativer Induktion unabhängige, Begründung für Abduktion finden ließe, und wir außerdem konzedieren, dass eine notwendige Verbindung von Rabentum und Schwärze die beste Erklärung für unsere Beobachtungen darstellt, wäre immer noch nicht klar, wieso wir diese notwendige Verknüpfung in die Zukunft fortschreiben dürfen. Wieso dürfen wir mehr sagen als dass *bisher* Raben notwendig schwarz waren? Wir können freilich einen Notwendigkeitsbegriff zugrunde legen, der Änderungen in der Zeit ausschließt: Wenn zu irgendeinem Zeitpunkt Raben mit Notwendigkeit schwarz sind, dann sind sie es immer. Aber nun wird noch einmal deutlich zweifelhafter, wie eine derart aufgefasste Notwendigkeit als „beste Erklärung" aus den Evidenzen erschlossen werden kann. Wiederum droht hier eine bestimmte Gleichförmigkeit der Welt einfach vorausgesetzt zu werden, wobei die Voraussetzung diesmal in der Wahl der Begriffe verborgen ist.

pragmatische Rechtfertigung

Ein Vorschlag vom Typ c) stammt von Hans Reichenbach (vgl. Reichenbach (1938), §§ 38–40). Wir sind, so Reichenbach, durchaus darin gerechtfertigt, die Konklusionen typischer induktiver Schlüsse zu akzeptieren – aber nicht deshalb, weil wir mehr oder minder starke Hinweise auf ihre Wahrheit hätten, sondern weil wir nichts Besseres tun können. Es könnte sein, dass schon in der nächsten Minute all das, worauf wir uns bisher verlassen haben, sich als falsch erweist, der Eindruck von Stabilität und Gleichförmigkeit, den unsere Welt in bestimmten Hinsichten bisher auf uns gemacht hat, wie eine Seifenblase zerplatzt. Diese Möglichkeit können wir nicht nur nicht ausschließen, sondern nicht einmal mit Grund für objektiv unwahrscheinlich erklären. Wir haben keinerlei rationalen Hinweis darauf, dass diese Möglichkeit nicht realisiert wird, aber wenn sich die

Welt derart entwickeln sollte, dann können wir genauso gut auch weiter unsere Induktionen machen. Sie werden dann mehrheitlich falsche Konklusionen haben, aber wir haben auch keine bessere, erfolgversprechendere Methode, die Zukunft zu antizipieren. Ist die Welt (in einem bestimmten Bereich) völlig chaotisch, wird keine Vorhersagemethode erfolgreich sein, es sei denn aus purem Zufall. Alle Regelmäßigkeiten aber, die es gibt, können wir durch Induktion auch prinzipiell herausfinden. Insbesondere können wir Induktion auf den Erfolg von Vorhersagemethoden (einschließlich Induktion selber) anwenden und sagen: Wenn es überhaupt eine Methode gibt, die verlässlich funktioniert, dann wird sie auch durch Induktion herausgegriffen (wenn uns die Methode überhaupt in den Sinn kommt): per Induktion extrapolieren wir ihre Erfolge. In diesem Sinne ist also Induktion gerechtfertigt: als Strategie, die mindestens so erfolgreich ist wie jede Alternative. *Wie* erfolgreich, darüber können wir freilich keine begründete Vorhersage treffen.

Dieses Argument liefert eine pragmatische Rechtfertigung aufzählender Induktion. Die Situation erinnert an die sogenannte Pascalsche Wette, in der der Glaube an Gott auch nicht dadurch gerechtfertigt wird, dass Argumente für die Existenz Gottes beigebracht werden, sondern dadurch, dass man durch diesen Glauben nichts (oder nur wenig) verlieren, ggf. aber sehr viel gewinnen können soll. Das Reichenbachsche Argument, das sich gegenüber der hier gegebenen skizzenhaften Darstellung deutlich präzisieren lässt, ist gültig, aber in seiner Reichweite begrenzt. Unser Gefühl der Sicherheit oder weitgehenden Sicherheit, das wir bei vielen induktiv etablierten Hypothesen haben, wird auch nicht ansatzweise gerechtfertigt.

ypothetisch-deduktive Methode

d) In einem deutlichen Bruch mit allen Versuchen, Induktion auf die eine oder andere Weise zu rechtfertigen, versucht Karl Popper zu zeigen, dass wir weder in den Wissenschaften noch im Alltag überhaupt induktive Schlüsse benötigen (vgl. Popper (1934); Popper (1972), Kap. 1).[3] Popper ist der Auffassung, dass alles Schließen deduktiv erfolgt und setzt gegen induktivistische Auffassungen des Lernens aus Erfahrung die so genannte *hypothetisch-deduktive* Methode. Es ist laut Popper nicht so, dass Allaussagen über die empirische Wirklichkeit, und insbesondere wissenschaftliche Theorien, durch Beobachtungen mehr oder minder wahrscheinlich gemacht, gerechtfertigt oder sonstwie positiv gestützt werden können. Sie können nur negativ auffallen, nämlich falsifiziert werden, indem sie mit bestimmten Beobachtungen in Konflikt stehen. Empirische Allaussagen sind und bleiben kühne, man könnte auch sagen: spekulative Hypothesen, aus denen sich beobachtbare Konsequenzen logisch ableiten lassen, und das Beste, was sich von einer solchen Hypothese sagen lässt, ist, dass sie auf diesem Wege bisher nicht widerlegt worden ist. Daraus erklärt sich die Bezeichnung der Methodologie: „hypothetisch-deduktiv". Popper gibt damit die gesamte positive Seite der

[3] Und zwar soll das nicht nur für die aufzählende Induktion, sondern generell gelten.

Bestätigung auf. Ein Allsatz lässt sich auch aus noch so vielen einzelnen Beobachtungen nicht deduzieren, und mehr ist dazu nicht zu sagen. Seine Negation kann dagegen aus einer einzigen Beobachtung deduziert werden, und auf dieser Asymmetrie zwischen logischer Verifikation und Falsifikation von Allaussagen durch einzelne Beobachtungen möchte Popper das gesamte Lernen aus Erfahrung aufgebaut wissen. Daraus resultiert eine Beweislastumkehr: Eine Theorie muss nicht begründet, wahrscheinlich gemacht, positiv bestätigt werden, noch ist dies überhaupt möglich, sondern darf vertreten werden, solange sie nicht falsifiziert ist. Dieses Bild des Lernens aus Erfahrung und insbesondere des wissenschaftlichen Erkennens hat Ähnlichkeit mit der biologischen Evolution: Von einer jetzt lebenden Art lässt sich nur sagen, dass sie bisher der Selektion nicht zum Opfer gefallen ist, dass sie bisher allen ggf. wechselnden Anforderungen ihrer Umwelt gewachsen war. Schon morgen kann dies aber anders sein, und ebenso kann schon morgen eine Theorie, die bisher nie falsifiziert wurde, auf ein empirisches Datum treffen, das ihr den Garaus macht.

Die Poppersche Sichtweise auf das Problem hat, ähnlich wie die Reichenbachsche, weitgehende skeptische Implikationen, die man nur dann vermeiden kann, wenn man gegen Poppers erklärte Absicht seine Methode um induktive Elemente ergänzt. So meint man selbstverständlich, dass eine in der Vergangenheit durch bestimmte Experimente widerlegte Theorie auch in der Zukunft in vergleichbaren Situationen scheitern werde, so dass man sie ein für allemal beiseite lassen kann. Aber dies wäre ein induktiver Schluss, wie ihn Popper gerade nicht zulassen möchte. Es ist für ihn nicht statthaft, von vergangenen Misserfolgen auf zukünftige zu schließen. Das heißt, dass die fragliche Theorie zwar falsifiziert worden ist, wir aber keinen Grund haben, sie für künftige Anwendungen, Prognosen usw. nicht doch zu verwenden nach dem Motto: „Bisher ist es immer schief gegangen, aber von jetzt ab wird es funktionieren." Die Falsifikation einer Theorie durch bestimmte Beobachtungen hat also, wenn man nicht bereit ist, die Misserfolge induktiv in die Zukunft fortzuschreiben, bei weitem nicht die intendierten Konsequenzen. Wir befinden uns mit Blick auf den nächsten Fall dann immer im Stadium völliger Unwissenheit: Hier sind verschiedene Theorien, darunter einige, die sich in der Vergangenheit nicht bewährt haben, aber das sagt nichts über ihre zukünftige Performanz. Und einige haben sich bewährt, aber das sagt auch nichts über ihre zukünftige Performanz. Ohne ein Urteil über die Reichweite und Leistungsfähigkeit der hypothetisch-deduktiven Methode zu fällen, kann man daher zumindest feststellen, dass sich die *vollständige* Verbannung induktiver Elemente aus dem Prozess der Hypothesenprüfung und -auswahl nicht durchhalten lässt.

Das „neue Rätsel d Induktion"

Soviel zu verschiedenen Lösungsvorschlägen für das Problem der aufzählenden Induktion. Dass man keine allgemeine Rechtfertigung für sie finden kann, ist weniger verwunderlich, wenn man Nelson Goodmans so genanntes „neues Rätsel der Induktion" betrachtet (vgl. Goodman (1955), Kap. 3.4). T_0 bezeichne den gegenwärtigen Zeitpunkt, und das Prädikat „grot" sei wie folgt definiert: Ein

Gegenstand ist grot, wenn er entweder vor oder zu T_0 beobachtet wird und sich dabei als grün erweist, oder nach T_0 beobachtet wird und sich dabei als rot erweist. Gewiss ein ungewöhnliches Prädikat, aber nichts Verbotenes, so könnte man meinen. Es bringt uns nun aber in folgende Verlegenheit: Wir können einerseits aus der Tatsache, dass alle bisher beobachteten Smaragde grün waren, induktiv schließen, dass wohl alle Smaragde grün sind und sich daher insbesondere die zukünftig beobachteten als grün erweisen werden. Oder wir können aus der Tatsache, dass sich alle bisher beobachteten Smaragde als grot erwiesen haben (so ist es nämlich), schließen, dass sie wohl alle grot sind und sich daher insbesondere die zukünftig beobachteten als grot erweisen werden – was bedeutet, dass wir prognostizieren, dass alle künftig beobachteten Smaragde rot sein werden! Diesen Trick kann man offenbar mit jeder Farbe machen. Wenn man bereit ist, auch solche Prädikate wie „grot" induktiv zu verallgemeinern, dann könnte man deshalb jede beliebige Farbe für die künftigen Smaragde vorhersagen. Jede derartige Prognose könnte sich darauf berufen, dass ja alle bisherigen Smaragde so-und-so – hier muss jetzt ein geeignetes Prädikat wie „grot" eingesetzt werden – gewesen sind. Diese Vorhersagen sind einerseits völlig beliebig und untereinander im Konflikt, scheinen aber andererseits alle dieselbe hohe induktive Stützung durch die vorhandenen Evidenzen beanspruchen zu können. Wenn das wahr wäre, wäre es natürlich von vornherein ein vergebliches Unterfangen, aufzählende Induktion rechtfertigen zu wollen. Wir müssen also Prädikate wie „grot" von der Verallgemeinerung ausschließen können.

Damit zeigt das Goodmansche „neue Rätsel der Induktion", dass von einer Rechtfertigung induktiver Schlüsse ohne weiteres gar nicht gesprochen werden kann. Es gibt nämlich stets sehr viele unvereinbare Möglichkeiten, vorhandene Evidenzen zu verallgemeinern, und es ist durchaus nicht so, dass sich diese Möglichkeiten immer erst durch spitzfindige Konstruktionen ergeben. Wenn wir uns ein Koordinatensystem denken, in dem eine empirische Größe in Abhängigkeit von einer anderen aufgetragen wird, dann ermitteln wir den funktionalen Zusammenhang zwischen diesen beiden durch einzelne Messdaten, die anschließend durch eine Kurve inter- und extrapoliert werden. Dies ist eine für die empirischen Wissenschaften typische und dort sehr häufig anzutreffende Art der Verallgemeinerung, aber natürlich gibt es für dergleichen immer unendlich viele Möglichkeiten. Durch endlich viele Messpunkte lassen sich stets unendlich viele Kurven zeichnen. Teil jeder erfolgreichen Rechtfertigung aufzählender Induktion muss daher eine Überlegung sein, die alle diese Möglichkeiten bis auf eine nach einem begründeten Kriterium ausschließt. Wenn wir uns fragen, wieso wir „grün" von den künftigen Smaragden prädizieren dürfen, dann ist klar, dass jede dafür vorgeschlagene Rechtfertigung solange wertlos ist, wie sie ebenso auf „grot" zutreffen würde. Analoges gilt für den Fall der unterschiedlichen Funktionen, die gegebene Messdaten verallgemeinern. Die zuvor besprochenen Standardrechtfertigungen von Induktion haben aber gar keine Mittel, zwischen diesen verschiedenen Prädikaten oder Funktionen zu unterscheiden. Sie

tun so, als stünde fest, auf welche Weise man Beobachtungen zu verallgemeinern hätte. Die von Goodman aufgeworfene Problematik ist tiefliegender als das klassische Induktionsproblem und deshalb auch schwerer zu entdecken gewesen. Ihre Auflösung ist eine notwendige Voraussetzung für eine Antwort auf die Humesche Herausforderung. Es scheint außerordentlich schwer zu sein, Unterscheidungskriterien der erforderlichen Art auch nur zu formulieren, geschweige denn zu begründen.

5. Subjektive Bestätigungstheorie: der Bayesianismus

Wir haben in der vorhergehenden Diskussion an verschiedenen Stellen gesehen, dass bei der Bestätigung von Hypothesen durch Evidenzen Gesichtspunkte eine Rolle spielen, die sich nur schwer objektivieren lassen. Anstatt zu versuchen, alle solchen subjektiven Faktoren auszuräumen, kann man eine Bestätigungstheorie aufbauen, die von vornherein mit ihrer Existenz rechnet und in deren Rahmen sich von selbst ergibt, in welchem Ausmaß man mit ihnen leben muss. Eine solche Theorie ist der Bayesianismus. Er liefert eine Quantifikation induktiver Stützung, die auf einem einzigen und im Grunde einfachen Prinzip basiert, aber trotzdem diverse, zunächst sehr heterogen scheinende Intuitionen einzufangen vermag, die man vortheoretisch in Bezug auf die Bestätigung einer Hypothese durch Evidenzen hat. Diese integrative Kraft zeichnet den Bayesianismus vor allen Konkurrenzpositionen aus. Der Preis dafür ist eine subjektive Grundtendenz, die gerade *nicht* dem entspricht, was man vortheoretisch zu meinen geneigt ist.

Bestätigungsgrade Wahrscheinlichkeit

Es geht ab jetzt wieder um Bestätigung und Induktion im allgemeinen Sinne des Wortes. Wenn man versucht, eine quantitative Bestätigungstheorie aufzustellen, in deren Rahmen präzise angegeben werden kann, wie stark eine gegebene Hypothese durch bestimmte Evidenzen gestützt oder bestätigt wird, dann liegt es sehr nahe, dafür auf die Wahrscheinlichkeitsrechnung zurückzugreifen. Die Idee ist einfach, dass die Hypothese durch die Evidenzen mehr oder weniger *wahrscheinlich* gemacht wird, und wenn man diesen Begriff jetzt spezifisch im Sinne der gleichnamigen mathematischen Theorie versteht, dann kann man deren ausgefeilten Apparat für den Zweck einer dann ebenso ausgefeilten quantitativen Bestätigungskonzeption einsetzen. Für das weitere müssen wir uns zunächst die Grundsätze der Wahrscheinlichkeitsrechnung vergegenwärtigen. Diese betreffen ausschließlich die *formalen* Eigenschaften von Wahrscheinlichkeiten. Die Mathematik kümmert sich nicht um die Frage, was es bedeutet, einer Aussage eine Wahrscheinlichkeit zuzuschreiben, und auch nicht darum, wie man zu dieser Wahrscheinlichkeit kommt. Sie hat es nur mit dem Kalkül selbst zu tun, nicht aber mit seiner Deutung oder Interpretation, die je nach Anwendungsgebiet anders aussehen kann. Drei dieser Interpretationsmöglichkeiten (Wahrscheinlichkeiten als subjektive Überzeugungsgrade, als logische Relationen zwischen Aussagen,

als relative Häufigkeiten auf lange Sicht) werden wir im weiteren Verlauf kennen lernen, und es gibt noch andere.[4] Einstweilen brauchen wir uns nicht auf eine bestimmte Deutung festzulegen.

Grundprinzipien der Wahrscheinlichkeitsrechnung

A) Axiome für Wahrscheinlichkeiten

Es seien a und b Aussagen (genauer: Aussageinhalte, Propositionen, Sachverhalte). $P(a)$ bezeichne die Wahrscheinlichkeit von a. Dann gilt:

(A1) $P(a) \geq 0$

(A2) $P(a) = 1$, falls a notwendig wahr ist.

(A3) $P(a \vee b) = P(a) + P(b)$, falls a und b sich ausschließen.

Zusammengefasst in Prosa: Die Wahrscheinlichkeit ist eine nicht negative, normierte und additive Funktion von einer geeigneten Menge von Aussagen in reelle Zahlen. Geeignet ist eine Aussagenmenge, falls sie unter Disjunktion, Konjunktion und Negation abgeschlossen ist.

Aus (A1) bis (A3) lässt sich leicht beweisen, dass gilt:

(F1) $P(\neg a) = 1 - P(a)$

(F2) $P(a_1 \vee a_2 \vee \ldots \vee a_n) = P(a_1) + P(a_2) + \ldots + P(a_n)$, falls die Aussagen $a_1, a_2, \ldots, a_n$ sich gegenseitig ausschließen.

B) Bedingte Wahrscheinlichkeit

(B1) Definition: Es sei $P(b) \neq 0$. Dann heißt der Quotient $\frac{P(a \wedge b)}{P(b)}$ die *bedingte* Wahrscheinlichkeit von a, gegeben b, und wird durch $P(a,b)$ abgekürzt.

Es handelt sich hierbei um die Wahrscheinlichkeit dafür, dass a der Fall ist, unter der Voraussetzung oder Bedingung, dass b der Fall ist. Die Definition ist intuitiv einleuchtend.

(B2) Definition: Vergleicht man $P(a,b)$ und $P(a)$, dann sind drei Fälle möglich: $P(a,b) > P(a)$, $P(a,b) < P(a)$ und $P(a,b) = P(a)$. Intuitiv gesprochen, bedeutet das, dass das der-Fall-sein von b die Wahrscheinlichkeit von a erhöht bzw. senkt bzw. unverändert lässt. b heißt dann *positiv relevant* bzw. *negativ relevant* bzw. *irrelevant* für a.

(B3) Es sei nun auch $P(a) \neq 0$. Dann ist $\frac{P(a \wedge b)}{P(b)} > P(a)$ äquivalent mit $\frac{P(a \wedge b)}{P(a)} > P(b)$, und d.h. es ist $P(a,b) > P(a)$ genau dann, wenn $P(b,a) > P(b)$ ist: b ist positiv relevant

[4] Das Thema der möglichen Bedeutungen von Wahrscheinlichkeitsaussagen kann hier nicht angemessen behandelt werden. Siehe dazu z.B. Gillies (2000) und zu objektiven Interpretationen auch Rosenthal (2004).

für a genau dann, wenn umgekehrt auch a positiv relevant für b ist. Die drei eben eingeführten Relevanzbeziehungen sind symmetrisch. Im Falle der Irrelevanz spricht man auch von der *Unabhängigkeit* von a und b. Sie ist genau dann gegeben, wenn $P(a \wedge b) = P(a) \cdot P(b)$ ist.

Der Begriff der bedingten Wahrscheinlichkeit liefert einen offensichtlichen Vorschlag für die *Veränderung* von Wahrscheinlichkeiten angesichts neuer Informationen: Wenn wir von einem bestimmten Wahrscheinlichkeitsmaß auf einer geeigneten Aussagenmenge ausgehen und dann erfahren, dass b der Fall ist, dann gilt für jede Aussage a unserer Menge: *Konditionalisier*

(K) $P_{neu}(a) = P_{alt}(a,b) = \frac{P_{alt}(a \wedge b)}{P_{alt}(b)}$

Wir erhalten die neuen Wahrscheinlichkeiten durch *Konditionalisierung* bzgl. der neu gewonnenen Information. Das neue Wahrscheinlichkeitsmaß auf der Aussagenmenge wird geliefert durch die alten bedingten Wahrscheinlichkeiten unter der Bedingung b. Wenn wir dieses Verfahren zur Modifikation von Wahrscheinlichkeiten akzeptieren – und es ist in der Tat plausibel – dann bleibt als zentrale Frage, woher wir unser ursprüngliches Wahrscheinlichkeitsmaß auf der Aussagenmenge nehmen und was dessen Status ist.

Um dieses Problem deutlicher zu fixieren, unterscheiden wir nun wieder zwischen Hypothesen h und (Aussagen über) Evidenzen e. Wenn eine Hypothese h hinreichend spezifisch formuliert ist, dann ist zu erwarten, dass die bedingten Wahrscheinlichkeiten $P(e,h)$ zumindest in gewissen Grenzen objektiv vorgegeben sind. *Gegeben*, dass eine bestimmte Hypothese richtig ist, lässt sich häufig objektiv sagen, wie wahrscheinlich es *unter dieser Voraussetzung* ist, ein bestimmtes Beobachtungsdatum e zu erhalten. Somit können wir Wahrscheinlichkeiten der Form $P(e,h)$ als vergleichsweise unproblematisch ansehen. Dies sind sie allerdings nur, wenn h überhaupt einschlägig für den Kontext ist, in welchem e auftritt, also über diesen Kontext eine substantielle Aussage macht. Ansonsten ist einfach $P(e,h) = P(e)$, die Wahrscheinlichkeit von e unabhängig von jeglicher Hypothese, und diese ist in hohem Maße willkürlich.

Nehmen wir nun weiter an, bei einer bestimmten (wissenschaftlichen) Problemstellung stünden n verschiedene Hypothesen $h_1, h_2, \ldots, h_n$ zur Auswahl, die sich wechselseitig ausschließend und zusammen alle Möglichkeiten erschöpfend seien. Die Lage ist also die, dass genau eine der Hypothesen wahr ist (nur wissen wir nicht, welche). In diesem Fall ist

$$\begin{aligned} P(e) &= P((e \wedge h_1) \vee (e \wedge h_2) \vee \ldots \vee (e \wedge h_n)) \\ &= P(e \wedge h_1) + P(e \wedge h_2) + \ldots + P(e \wedge h_n) \text{ wegen (F2)} \\ &= P(e,h_1) \cdot P(h_1) + P(e,h_2) \cdot P(h_2) + \ldots + P(e,h_n) \cdot P(h_n) \text{ wegen (B1).} \end{aligned}$$

essche Formel

Unter Berücksichtigung dieser sogenannten „Formel von der totalen Wahrscheinlichkeit" ergibt sich für die Veränderung der Hypothesenwahrscheinlichkeiten angesichts einer Evidenz e:

$$\text{(B)}\quad P_{neu}(h_i) = P_{alt}(h_i,e) = \frac{P_{alt}(h_i \wedge e)}{P_{alt}(e)} = \frac{P_{alt}(e,h_i)\cdot P_{alt}(h_i)}{P_{alt}(e)}$$

$$= \frac{P_{alt}(e,h_i)\cdot P_{alt}(h_i)}{P_{alt}(e,h_1)\cdot P_{alt}(h_1) + \ldots + P_{alt}(e,h_n)\cdot P_{alt}(h_n)} \quad \text{für jedes } i.$$

Dies ist (in einer speziellen Form) die sogenannte *Bayessche Formel*, so benannt nach Reverend Thomas Bayes (1702–1761). Sie ist ein Standardtheorem der Wahrscheinlichkeitsrechnung und namengebend für die hier diskutierte Methodologie. Man sieht an ihr folgendes: Bedingte Wahrscheinlichkeiten der Form $P(h,e)$, wie sie uns bei der Bestätigungsproblematik besonders interessieren, werden zurückgeführt auf bedingte Wahrscheinlichkeiten der Form $P(e,h)$ und auf unbedingte Wahrscheinlichkeiten der Form $P(h)$. Falls die Wahrscheinlichkeiten $P(e,h)$ nun in der Tat unproblematisch sind – insbesondere sich durch neue Evidenzen nicht ändern und die Indizes „alt" und „neu" bei ihnen daher wegfallen können – dann hängt die ganze Methode an den *a-priori-Wahrscheinlichkeiten* der Hypothesen $h_1, h_2, \ldots, h_n$. Die Frage ist, mit welchen Wahrscheinlichkeiten – nennen wir sie $P_{start}(h_1), P_{start}(h_2), \ldots, P_{start}(h_n)$ – für diese Hypothesen wir anfangen. Steht das einmal fest, dann folgt alles weitere durch wiederholte Anwendung der Bayesschen Formel. Es kommen sukzessive Evidenzen herein: $e_1, e_2, e_3, \ldots$, und jedes Mal werden die Hypothesenwahrscheinlichkeiten gemäß (B) aktualisiert. Dadurch erhalten wir zu jedem Zeitpunkt für jedes h_i eine bestimmte Wahrscheinlichkeit, die den Bestätigungsgrad von h_i im Lichte der bis dahin aufgelaufenen Evidenzen ausdrückt, aber eben auch von den *a-priori*-Wahrscheinlichkeiten $P_{start}(h_1), \ldots, P_{start}(h_n)$ abhängt. Wie sind diese also zu wählen?

An dieser Stelle sagt der Bayesianismus einfach, dass die Startwahrscheinlichkeiten für die verschiedenen Hypothesen rein subjektiver Natur sind. Sie spiegeln persönliche Überzeugungsgrade wider und können daher von Beobachter zu Beobachter verschieden sein. Wenn Hypothesen $h_1, h_2, \ldots, h_n$ vorgeschlagen werden, dann veranschlagt jeder nach seinem Gusto bestimmte Anfangsplausibilitäten für die verschiedenen Hypothesen und arbeitet mit diesen weiter, indem er sie im Lichte neuer Evidenzen immer wieder gemäß (B) modifiziert. Die Objektivität der Hypothesenwahrscheinlichkeiten kommt dabei dadurch zustande, dass für diesen Modifikationsprozess weitreichende *Konvergenzsätze* bewiesen werden können. Im Laufe der Zeit, d.h. je mehr Evidenzen hereinkommen, desto geringer wird der Einfluss der Startwahrscheinlichkeiten auf die jeweils aktuellen Wahrscheinlichkeiten $P_{neu}(h_i)$, und im Grenzfall sinkt er auf Null ab („*washing out of the priors*").

Konvergenz der ypothesenwahr- cheinlichkeiten

Genauer: Nehmen wir an, zwei epistemische Subjekte S und S' veranschlagen für die Hypothesen je eigene *a-priori*-Wahrscheinlichkeiten. Diese seien allerdings

alle von 0 und 1 verschieden, d.h. beide Subjekte sind sich wenigstens darin einig, dass alle Hypothesen prinzipiell in Frage kommen und keine von vornherein mit absoluter Sicherheit richtig ist (Dogmatismusverbot). Wenn nun S und S' dieselbe Reihe von Beobachtungen $e_1, e_2, e_3, \ldots$ machen, dann werden sich ihre jeweils aktuellen *a-posteriori-Wahrscheinlichkeiten* für die verschiedenen Hypothesen immer mehr einander angleichen und schließlich kaum noch unterschieden sein – ganz gleich, wie unterschiedlich die Startwahrscheinlichkeiten waren. Dies gilt allerdings nur, falls sich die Subjekte über alle Evidenzen einig sind (d.h. darüber, was beobachtet worden ist), und auch darüber, wie wahrscheinlich eine bestimmte Beobachtung, gegeben eine bestimmte Hypothese, ist (also keine Differenzen bei den Faktoren $P(e,h)$ bestehen). Wie viel Zeit die Konvergenz dann beansprucht, wie viele gemeinsame Informationen die Subjekte sammeln müssen, bis sie sich über die Hypothesenwahrscheinlichkeiten so gut wie einig sind, hängt davon ab, wie unterschiedlich ihre Ausgangspositionen waren. Der Einigungsprozess kann lange dauern, er verläuft aber, unter den angegebenen Bedingungen, immer erfolgreich. Wer freilich von vornherein seiner Sache ganz sicher ist, also für eine Hypothese die *a-priori*-Wahrscheinlichkeit 1 und für alle anderen 0 veranschlagt, der ist durch Erfahrung unbelehrbar: Die Wahrscheinlichkeiten 0 und 1 ändern sich durch Konditionalisierung nicht.

So gelangt in eine zunächst einmal subjektive Bestätigungsmethodologie ein Element der Objektivität hinein, das den Bestätigungsgraden ihre Willkür wieder nimmt. Wir können dem Bayesianismus zufolge recht genau sagen, wie wahrscheinlich eine bestimmte Hypothese im Lichte bestimmter Evidenzen ist, und wer dem dann widersprechen wollte, der würde sich durch seinen Widerspruch auf extreme Startwahrscheinlichkeiten (extrem nahe an 1 oder 0) festlegen. In diesem Sinne also erfüllt der Bayesianismus das zunächst vielleicht unrealistisch scheinende Programm einer quantitativen Bestätigungskonzeption.

Die Bayesianische Methodologie ist ein florierendes Forschungsgebiet mit zahlreichen, z.T. tiefliegenden und überraschenden Anwendungsmöglichkeiten.[5] Zumindest zwei einfache Resultate möchte ich hier vorführen. Wir konzentrieren uns auf die Bestätigung einer Hypothese h durch ein Datum e und vergessen für den Moment wieder die kompliziertere Situation, die man für die Konvergenzresultate braucht, in der man n alle Möglichkeiten ausschöpfende Hypothesen simultan betrachtet.

Spezifität von Implikationen

Eine der Intuitionen, die man in Bezug auf Bestätigung hat, ist, dass eine erfolgreiche Vorhersage umso stärker für eine Hypothese spricht, je präziser oder spezifischer sie ausfällt. Wie spiegelt sich dies in der Bayesschen Formel? Es gilt:

$$\text{(B)}\quad P_{neu}(h) = P_{alt}(h,e) = \frac{P_{alt}(h \wedge e)}{P_{alt}(e)} = \frac{P(e,h)\cdot P_{alt}(h)}{P_{alt}(e)}.$$

[5] Eine gründliche, weitreichende Einführung geben Howson & Urbach (1989). Wesentlich konziser ist Jeffrey (2004). Bovens & Hartmann (2004) bringen anspruchsvolle Anwendungen.

Wenn nun e von h impliziert oder sehr wahrscheinlich gemacht wird, ist $P(e,h) \approx 1$ und damit

$$P_{neu}(h) \approx \frac{P_{alt}(h)}{P_{alt}(e)}.$$

Je spezifischer das von h implizierte Datum e ist (je kleiner z.B. das Intervall, das von h als Messergebnis prognostiziert wurde), umso weniger war vorher (unabhängig von jeglicher Hypothese) mit gerade dieser Evidenz zu rechnen, umso kleiner ist also $P_{alt}(e)$, und umso größer ist $P_{neu}(h)$ im Vergleich zu $P_{alt}(h)$, wenn e sich dann tatsächlich ergeben hat.

Diversität von Implikationen

Eine andere Intuition ist, dass verschiedenartige erfolgreiche Anwendungen eine Hypothese stärker bestätigen, als wenn sie sich immer nur in *einem* Typ von Situation bewährt. Es kommt für den Bestätigungsgrad nicht nur auf die Quantität (Vielzahl) und die Spezifität (dies sahen wir eben) der bestätigenden Instanzen an, sondern auch auf ihre Diversität (Vielfalt). Auch dies kommt aus der Bayesschen Formel heraus, und zwar wie folgt. Nehmen wir an, wir würden k Evidenzen $e_1, \ldots, e_k$ erhalten. Diese fassen wir formal zu dem komplexen Datum $e = e_1 \wedge \ldots \wedge e_k$ zusammen. Es sei von der Hypothese impliziert, also wieder $P(e,h) \approx 1$. Außerdem gilt aufgrund der Definition bedingter Wahrscheinlichkeiten:

$$P_{alt}(e) = P_{alt}(e_1 \wedge \ldots \wedge e_k) = P_{alt}(e_1) \cdot P_{alt}(e_2,e_1) \cdot P_{alt}(e_3,e_1 \wedge e_2) \cdot \ldots \cdot P_{alt}(e_k,e_1 \wedge \ldots \wedge e_{k-1})$$

Wenn nun die k Experimente deutlich verschiedenen Typs sind, dann kann man zuvor (unabhängig von irgendwelchen Hypothesen) nicht von dem Ausgang eines oder mehrerer von ihnen auf den Ausgang der anderen schließen. Die e_i sind dann voneinander unabhängig, und das Produkt wird zu $P_{alt}(e) = P_{alt}(e_1) \cdot P_{alt}(e_2) \cdot \ldots \cdot P_{alt}(e_k)$. Im entgegengesetzten Extremfall sind die e_i hoch korreliert, wenn etwa das gleiche Experiment k mal wiederholt wird. Während wir bei der ersten Durchführung noch eine echte Information erhalten – $P_{alt}(e_1)$ kann sehr klein sein – ist $P_{alt}(e_2,e_1)$ bereits recht groß, denn wir sind uns (unabhängig von irgendwelchen Hypothesen) ziemlich sicher, dass die erneute Durchführung des Experiments das beim ersten Mal erhaltene Resultat bestätigen wird. Die weiteren Faktoren des obigen Produkts rücken dann immer näher an Eins heran. Somit ist $P_{alt}(e)$ in dem Fall unabhängiger Evidenzen viel kleiner als im Fall abhängiger, und folglich $P_{neu}(h)$ entsprechend größer. Je weniger bestimmte Evidenzen voneinander abhängig sind (sich wechselseitig wahrscheinlich machen), umso stärker bestätigen sie eine Hypothese, von der sie gemeinsam impliziert werden.

In ähnlicher Weise können noch diverse weitere spezielle und allgemeine Intuitionen, die wir in Bezug auf Bestätigung haben, durch die Bayessche Formel eingefangen werden, womit sich erweist, dass *prima facie* ganz heterogene, intuitiv plausible Aussagen über Bestätigungstatsachen aus einem einheitlichen Prinzip gewonnen und als aus diesem entspringend gedacht werden können. Dies ist,

neben den erwähnten Konvergenzresultaten, die wesentliche Leistung des Bayesianismus.

„catch-all"-Hypothesen

Gerade an den Konvergenzresultaten kann man aber auch bestimmte Beschränkungen unmittelbar ablesen.[6] So musste dafür angenommen werden, dass die Hypothesen $h_1, h_2, \ldots, h_n$ sich wechselseitig ausschließen und erschöpfend sind. Das letztere ist eine sehr weitgehende Annahme, denn sie bedeutet, dass mit diesen endlich vielen Hypothesen alle Möglichkeiten (zwar nicht unbedingt explizit gemacht, aber doch) abgedeckt sind. Eine solche Hypothesenmenge wird man nur in speziellen Fällen aufstellen können – es sei denn, man bedient sich eines formalen Tricks und setzt $h_n = \neg(h_1 \vee h_2 \vee \ldots \vee h_{n-1})$ Eine solche so genannte *„catch-all"*-Hypothese besagt lediglich, dass alles, was ansonsten vorgeschlagen wurde, nicht richtig ist, sondern eben irgendetwas anderes. Die Bezeichnung „Hypothese" ist dafür im Grunde fehl am Platze, und der Preis für ihre Einführung ist, dass die Wahrscheinlichkeiten $P(e,h_n)$ völlig unbestimmt sind. Wie wahrscheinlich soll es denn sein, dass ein bestimmter Sachverhalt e unter der Bedingung besteht, dass „irgendetwas anderes" als $h_1, h_2, \ldots, h_{n-1}$ richtig ist? Wenn sich dazu nichts sagen lässt, dann spiegeln die Wahrscheinlichkeiten $P(e,h_n)$ auch nur bloße subjektive Überzeugungsgrade wider, und die Konvergenz der Wahrscheinlichkeiten der h_i kann nicht mehr gewährleistet werden. Damit das Konvergenz-Argument greift, muss vorausgesetzt werden, dass Wahrscheinlichkeiten der Form $P(e,h)$ nicht subjektiv sind, und dazu müssen zumindest die zur Diskussion stehenden Hypothesen sehr spezifisch formuliert sein. Das ist bei einer *„catch-all"*-Hypothese natürlich nicht der Fall. Aus diesem Grunde beschränkt sich die Leistungsfähigkeit der Bayesianischen Methodologie hinsichtlich der Gewinnung eindeutiger Hypothesenwahrscheinlichkeiten auf solche Kontexte, in denen sich die Menge aller möglicherweise wahren Hypothesen in einem *substantiellen* Sinne überblicken lässt.

subjektive Wahrscheinlichk

Man kann zum Bayesianismus auf zwei verschiedenen Wegen gelangen, die ich kurz andeuten möchte (dahinter verbergen sich jeweils wieder weit ausgearbeitete, facettenreiche Theorien). Entweder man interessiert sich für eine Theorie subjektiver Überzeugungsgrade. Diese beginnt mit der Beobachtung, dass verschiedene Menschen an verschiedene Sachverhalte mehr oder minder stark glauben – negativ formuliert, dass sie mehr oder minder überrascht wären, sollte sich das Gegenteil als wahr erweisen. Diese zunächst einmal grob abgestuften Überzeugungsgrade lassen sich über hypothetische Wetten exakt quantifizieren, und es lässt sich zeigen, dass diese (nun als reelle Zahlen daherkommenden) Grade in einem bestimmten Sinne genau dann kohärent sind, wenn sie den Axiomen (A1)–(A3) genügen. Damit sind *kohärente* persönliche Überzeugungsgrade als Wahrscheinlichkeiten im mathematischen Sinne erwiesen, und der Wahrscheinlichkeitskalkül erhält eine entsprechende subjektive Interpretation. Wie man es ausdrücken könnte, ist damit die Statik persönlicher Überzeugungsgrade charakterisiert. Man braucht

[6] Reichweite und Grenzen des Bayesianismus werden insbesondere von Earman (1992) ausgelotet.

nun aber auch eine Dynamik, d.h. eine Theorie über ihre Veränderung angesichts neuer Evidenzen, und dieses Desiderat führt über weitere Kohärenzargumente zum Verfahren der Konditionalisierung (K) und d.h. zum Bayesianismus.[7] Die erwähnten Konvergenzsätze zeigen nun, dass jeweils nur durch Kohärenzforderungen eingeschränkte Systeme persönlicher Überzeugungsgrade sich unter bestimmten Bedingungen im Laufe der Zeit einander angleichen werden. Die Konvergenz steht bei diesem Zugang aber nicht im Mittelpunkt des Interesses und ist sozusagen eine willkommene Zugabe.

logische rscheinlichkeit

Der andere Zugang ist, dass man nicht subjektive Überzeugungen und deren Änderung, sondern Bestätigungsrelationen modellieren möchte und dabei auf ein möglichst objektives Konzept von Stützung oder Bestätigung abzielt. Der Idealfall wäre es, wenn zu zwei beliebig vorgegebenen Aussageinhalten a und b ein objektiver Bestätigungsgrad $P(a,b)$, eine eindeutig bestimmte Wahrscheinlichkeit von a unter der Bedingung b, existieren würde. Setzte man insbesondere für b eine Tautologie ein, dann würde dieser Ausdruck die *a-priori*-Wahrscheinlichkeit von a liefern, also die Wahrscheinlichkeit, die man für a zu veranschlagen hätte, solange nichts weiter bekannt wäre. Dieses Programm wurde besonders von Rudolf Carnap unter dem Namen einer „induktiven Logik" verfolgt.[8] Aussagen über graduierte Bestätigungsrelationen sollten logisch wahr oder falsch sein, und daraus folgt natürlich insbesondere, dass die entsprechenden Grade $P(a,b)$ eindeutig und *a priori* bestimmbar sein müssen. Wie in der deduktiven Logik die Prämissen eines Schlusses und seine Konklusion Erfahrungssätze sein können, aber die Frage, ob die Prämissen die Konklusion implizieren, *a priori* zu beantworten ist, so sollte nun auch der Grad der Stützung, den a durch b erfährt, eine rein logische, erfahrungsunabhängige Angelegenheit sein. Dieses Programm einer induktiven Logik lässt sich, so kann man heute mit Bestimmtheit sagen, nicht erfüllen. Carnap war zwar imstande, eine globale Bestätigungsfunktion der gewünschten Art für bestimmte einfach strukturierte Sprachen bzw. Modellsituationen anzugeben, aber sogar in diesen haftet der Wahl des Wahrscheinlichkeitsmaßes eine erhebliche Willkür an. Der Anspruch auf die Objektivität oder gar den logischen Charakter der Bestätigungsgrade $P(a,b)$ ist nicht einlösbar. Statt sie mittels mehr oder weniger plausibler und jedenfalls nicht zwingender Prinzipien, die nur in bestimmten einfach strukturierten Kontexten überhaupt anwendbar sind, als eindeutig erweisen zu wollen, macht man besser reinen Tisch und räumt die grundsätzliche Subjektivität dieser Grade ein. Auch auf diese Weise landet man beim Bayesianismus. Umgekehrt führt beim Bayesianismus der Versuch, die *a-priori*-Wahrscheinlichkeiten der Hypothesen eindeutig zu fixieren, zu Überlegungen der Art, wie Carnap sie anstellte, und zu ähnlichen unüberwindlichen Schwierigkeiten.

[7] Die Bezeichnungen „Statik" und „Dynamik" übernehme ich von Spohn (1995), Kap. 4 und 5. Dort findet man auch ausgeführt, was ich hier über persönliche Überzeugungsgrade angedeutet habe.

[8] Das fulminante Hauptwerk dieses Programms ist Carnap (1950), nebst der Ergänzung Carnap (1952). Besser verdaulich ist die übersetzte Auswahl daraus: Carnap & Stegmüller (1959).

Frequentismus

Man muss sich also damit zufrieden geben, dass die Hypothesenwahrscheinlichkeiten nur durch die genannten Konvergenzresultate eine Art Objektivität erhalten, und wir haben gesehen, dass auch diese Konvergenz noch an verschiedene Voraussetzungen geknüpft ist. Wegen dieser Subjektivität ist die ganze Idee von Hypothesenwahrscheinlichkeiten von Seiten *frequentistischer* Methoden zur Bestätigung stark kritisiert worden. Die klassischen statistischen Verfahren, die sich mit den Namen Ronald A. Fisher, Jerzy Neyman und Egon Pearson verknüpfen, kommen ganz ohne die Zuordnung von Wahrscheinlichkeiten zu Hypothesen aus. Sie lehnen subjektive Wahrscheinlichkeiten ab und lassen nur solche zu, die zwanglos eine statistische oder Häufigkeits-Interpretation erlauben. Ihnen zufolge ist z.B. die „Wahrscheinlichkeit, mit diesem Würfel hier eine Sechs zu werfen" statthaft, weil sie sich interpretieren lässt als eine Aussage über eine relative Häufigkeit auf lange Sicht. Wenn nämlich der Würfel wiederholt geworfen würde, dann, so die Idee, würde sich auf die Dauer der Ausgang „sechs" mit einer bestimmten charakteristischen relativen Häufigkeit einstellen, und zwar tendenziell umso genauer, je länger die Wurfserie wäre (und im fiktiven Grenzfall unendlich wiederholten Werfens sogar ganz genau). Dieser charakteristische Prozentsatz oder Anteil von Würfen mit dem Ausgang „sechs" an allen Würfen ist die statistische Wahrscheinlichkeit der Sechs beim Werfen dieses Würfels. Dagegen darf von der „Wahrscheinlichkeit, dass dieser Würfel hier fair ist" oder gar der „Wahrscheinlichkeit, dass diese oder jene wissenschaftliche Theorie wahr ist" nicht gesprochen werden – bedeutungsvolle Wahrscheinlichkeitszuschreibungen setzen wiederholbare Zufallsexperimente voraus. Hypothesen können aufgrund statistischer Daten entweder akzeptiert oder verworfen werden, man kann ihnen aber nicht, oder nur in Ausnahmefällen, selbst eine Wahrscheinlichkeit zuordnen. Von diesem Standpunkt aus sind Wahrscheinlichkeiten der Form $P(h)$ oder $P(h,e)$, von Spezialfällen abgesehen, sinnlos, und auch von $P(e)$, der „Wahrscheinlichkeit schlechthin, dass e eintritt" kann keine Rede sein. Akzeptiert wird nur die Wahrscheinlichkeit, dass e der Fall ist, gegeben, dass die Hypothese h wahr ist, und diese wird auch nicht als eine bedingte Wahrscheinlichkeit aufgefasst – sonst wäre man wegen der Definition (B1) doch wieder bei Hypothesenwahrscheinlichkeiten – sondern als die statistische Wahrscheinlichkeit der Evidenz vor dem Hintergrund einer bestimmten Hypothese. Die Hypothese erzeugt einen Wahrscheinlichkeitsraum, innerhalb dessen die verschiedenen möglichen Evidenzen eine Wahrscheinlichkeit, die eine Häufigkeitsinterpretation gestattet, erhalten. Andere Wahrscheinlichkeiten gibt es nicht, und eine gegebene Hypothese muss anhand dieses Wahrscheinlichkeitsraumes und der faktisch eingetretenen Evidenz entweder angenommen oder abgelehnt werden.[9]

statistische Wahrscheinlichke

[9] Mayo (1996) bietet eine umfassende philosophische Verteidigung und Anwendung einer derartigen Bestätigungskonzeption auf wissenschaftstheoretische Problemstellungen. Zur Häufigkeitsinterpre-

Die verschiedenen frequentistischen Methodologien können hier nicht vorgestellt werden. Unter ihnen ist die so genannte Neyman-Pearson-Theorie der Standard in den Sozialwissenschaften und wird dementsprechend in jedem Lehrbuch der Statistik entwickelt, wie man es in den Sozialwissenschaften, besonders der Ökonomie, zugrunde legt. Frequentistische Ansätze stellen sowohl in der mathematischen Statistik im engeren Sinne als auch in der Bestätigungstheorie im weiteren Sinne eine echte Alternative zum Bayesianismus dar. Zwischen den beiden Schulen gab und gibt es heftige Dispute, beide bemühen sich durchaus erfolgreich, Beispiele zu konstruieren, in welchen die Methodologie der Gegenseite zu offensichtlich absurden Resultaten führt. Man sollte den Dissens aber nicht übertreiben. Auf der einen Seite bietet der Bayesianismus dadurch, dass er Hypothesenwahrscheinlichkeiten zulässt, das flexiblere und feinere Instrumentarium. Es ist eine starke und kaum durchzuhaltende Position, Ausdrucksweisen wie „die-und-die Hypothese ist aufgrund der-und-der ermittelten Daten wahrscheinlich wahr (falsch)" ganz aus der Bestätigungstheorie zu verbannen. Auf der anderen Seite ist es ein legitimes Anliegen, eine von subjektiven Wahrscheinlichkeiten, die der Preis für die Vorzüge des Bayesianismus sind, unabhängige Methodologie der Bestätigung zu entwickeln, und zu sehen, wie weit man allein mit Aussagen der folgenden Struktur kommt: „Das-und-das Datum ist unter Voraussetzung der-und-der Hypothese statistisch (im Sinne einer Häufigkeitsinterpretation) sehr wahrscheinlich bzw. unwahrscheinlich, und wir dürfen deshalb die Hypothese, falls dieses Datum auftritt, vorläufig akzeptieren bzw. zurückweisen."

Schließlich soll nicht unerwähnt bleiben, dass Bestätigungsmethoden nicht unbedingt an die mathematische Wahrscheinlichkeitsrechnung und Statistik anschließen müssen. Es gibt verschiedene Gründe, mit Bestätigungsgraden zu operieren, die andere formale Eigenschaften besitzen, d.h. keine Wahrscheinlichkeiten im Sinne der Axiome (A1)–(A3) sind.[10] Solche Alternativen mussten hier aus Platzgründen außen vor bleiben, haben aber gleichfalls ihre Berechtigung. Es gibt keine in jeder Hinsicht vorziehenswerte Methode.

Literaturhinweise

Zum Induktionsproblem sind die zitierten Klassiker besonders zu empfehlen: Hume (1748), Popper (1934), Popper (1972), Reichenbach (1938) und Goodman (1955). Als eines der zentralen epistemologischen Probleme wird es in zahlreichen einführenden Werken zur Erkenntnis- und Wissenschaftstheorie behandelt oder mitbehandelt. Nennen möchte ich beispielhaft Gillies (1993), das sich durch die Anführung von historischen Hintergründen und einschlägigen Fallbeispielen auszeichnet. Klärner (2003) ist eine gründliche neuere

tation von Wahrscheinlichkeiten und darauf aufbauender Methodologien zum Testen probabilistischer Hypothesen siehe immer noch grundlegend Hacking (1965).

10 Nennen möchte ich hier beispielhaft die Theorie der Rangfunktionen, siehe einführend Spohn (1995), Kap. 5.

Untersuchung der Abduktion (des Schlusses auf die beste Erklärung). Eine sehr gute qualitative Einführung in die Bestätigungsproblematik bietet Hempel (1966). Die quantitativen Bestätigungskonzeptionen erfordern immer formale Anstrengungen; eine sehr gute Einführung, die auch Alternativen zum Wahrscheinlichkeitskalkül behandelt, bietet Spohn (1995). Eine gründliche und umfassende Einführung in den Bayesianismus ist Howson & Urbach (1989). Das brilliante Buch von Jeffrey (2004), einem Pionier und Altmeister des Gebietes, ist wesentlich konziser, dadurch aber auch anspruchsvoller. Obwohl teilweise überholt und nicht als Einführungsbuch konzipiert, bietet Hacking (1965) eine sehr gute, sich durch zurückhaltende Formalisierung auszeichnende Darstellung und Diskussion verschiedener frequentistischer Bestätigungstheorien. Gillies (2000) ist eine empfehlenswerte Darstellung und Diskussion der verschiedenen Interpretationen des Wahrscheinlichkeitsbegriffs.