

Bayesianismus und Falsifizierbarkeit

Michael Goerz (goerz@physik.fu-berlin.de)*

30. September 2006

Zusammenfassung

Dieses Essay untersucht, wie sich der Bayesianismus zur zweiwertigen Wahrheitstheorie verhält, was daraus für absolute Wahrheit und absolute Falschheit folgt, und was dies für den primitiven Induktivismus und den Falsifikationismus bedeutet.

Es wird sich zeigen, dass die Formel von Bayes Wahrheitswerte von 1 grundsätzlich verbietet, Wahrheitswerte von 0 allerdings prinzipiell zulässt, auch wenn es gegen die praktische Möglichkeit einer solchen Falsifizierung in Form der Quine-Duhem-These einen weitergehenden Einwand gibt.

Gerade im Lichte dieses Einwands werde ich darauf hinweisen, dass die Ablösung von den zweiwertigen Kategorien durchaus als sinnvoll und vielversprechend angesehen werden kann.

1 Einleitung

In einem vorherigen Essay¹ habe ich dafür argumentiert, dass Falsifizierbarkeit eine notwendige Bedingung für wissenschaftliche Theorien darstellt. Dabei habe ich allerdings nur die klassische zweiwertige Logik in Betracht gezogen, in der „Wissen“ ganz streng als absolute Wahrheit, als 0 oder 1 definiert ist. Wie der Falsifikationismus mit seiner Zurückweisung des primitiven Induktivismus zeigt, ist absolute Bestätigung („Verifizierung“) einer Theorie grundsätzlich nicht möglich. Umgekehrt kann aber eine Theorie falsifiziert werden; man kann, wenn ihre Vorhersagen nicht mit der Realität übereinstimmen, mit absoluter Gewissheit sagen, dass die Theorie falsch ist.

Allerdings gibt es an dieser Sicht eine nicht von der Hand zu weisende Kritik, auf die ich schon in meinem vorherigen Essay eingegangen bin: nach der Quine-Duhem-These kann eine Theorie in der Praxis gar nicht verifiziert werden, da sie unauflösbar mit weiteren Annahmen, etwa dem „Hintergrundwissen“ verknüpft ist. Die Falsifikation kann sich nur auf das Gesamtpaket, die zu testende Theorie *einschließlich* des Hintergrundwissens beziehen. Die Eleganz des Falsifikationismus erhält durch diese Kritik einen erheblichen Beigeschmack. Folgt man der Kritik streng, muss man zu dem Ergebnis kommen, dass es weder absolute Gewissheit über die Wahrheit, noch über die Falschheit einer Theorie geben kann.

*Dieses Essay wurde als 2. Hausarbeit für das Seminar „Einführung in Kontroversen der Wissenschaftstheorie“ bei Gregor Betz (Dozent), Freie Universität Berlin, SoSe 2006, erstellt.

¹(Goerz, 2006)

In diesem Lichte erscheint es fraglich, ob der strenge, absolut zweiwertige Wissensbegriff noch irgendeinen Gewinn verspricht. Es ist allerdings eine interessante Alternative in Betracht zu ziehen: der subjektive Bayesianismus². Dieser beruht auf dem Wahrscheinlichkeitstheorem von Bayes, welches angibt, wie sich die Wahrscheinlichkeit für die Wahrheit einer Hypothese ausgehend von einer willkürlich festgelegten Priorwahrscheinlichkeit auf der Basis von experimentellen Beobachtungen verändert. Das Theorem lautet:

$$\begin{aligned}
 P(h|e) &= P(h) \cdot \frac{P(e|h)}{P(e)} \\
 &= P(h) \cdot \frac{P(e|h)}{P(h)P(e|h) + P(\neg h)P(e|\neg h)} \\
 &= P(h) \cdot \frac{P(e|h)}{P(h)P(e|h) + [1 - P(h)]P(e|\neg h)}
 \end{aligned}$$

Dabei ist $P(h|e)$ die Wahrscheinlichkeit der Hypothese unter Berücksichtigung des Ereignis e (Posteriorwahrscheinlichkeit), $P(h)$ ist die Priorwahrscheinlichkeit, mit der jemand im Vorhinein an die Wahrheit der Hypothese glaubt, $P(e|h)$ ist die Wahrscheinlichkeit, die die Hypothese für das Auftreten des Ereignis e voraussagt, und $P(e|\neg h)$ ist die Wahrscheinlichkeit, die Alternativtheorien für das Auftreten des Ereignis e voraussagen. Da der hier vorgestellte Bayesianismus rein subjektiv verstanden wird, müssen diese Alternativtheorien nicht alle möglichen Theorien umfassen, sondern nur diejenige, die dem jeweilige Forscher bekannt sind. Eine leichte Einschränkung dieser Aussage werde ich zu einem späteren Zeitpunkt anbringen. Der Einfachheit halber fasse ich alle Alternativhypothesen zu einer einzigen Alternativhypothese zusammen ($\neg h = h_1 \vee h_2 \vee h_3, \dots$), prinzipiell müsste über alle Theorien summiert werden.

Das Theorem kann als Formulierung eines nicht-absoluten Wissensbegriffs verstanden werden. Es drückt Grade der subjektiven Überzeugung aus. Dank dieses neuen Formalismus kann es nun gelingen, dem Induktivismus zu neuem Aufwind zu verhelfen: Wenn wir schon nicht *mit Sicherheit* vom Einzelnen auf das Kollektiv schließen können, so können wir nach Ansicht der Bayesianer doch zumindest vom diesem Schluss behaupten, dass er *wahrscheinlich* wahr ist, oder besser gesagt, dass es gute, streng logische Gründe gibt, an seine Wahrheit zu glauben. Tatsächlich erhöht sich die Wahrscheinlichkeit der These „Alle Raben sind schwarz“ mit jedem beobachteten Raben.

Man muss sich eine besondere Eigenschaft der beiden Randpunkte 0 und 1 des Definitionsbereichs vor Augen führen: Während die Werte der Bayesischen Formel grundsätzlich nur Grade der subjektiven Überzeugung angeben, nicht aber etwas über die tatsächliche Wahrheit der Theorie aussagen, verhält sich dies an den Extremstellen anders: eine hundertprozentige Gewissheit (im streng mathematischen Sinne) bedeutet auch absolute Gewissheit über die Wahrheit der Theorie, und wenn absolute Gewissheit über die Wahrheit einer Theorie besteht, dann folgt daraus, dass die Theorie auch tatsächlich wahr ist. Anders gesagt, es ist nicht möglich, etwas im strengen Sinne zu wissen, was nicht wahr ist. Für den entgegengesetzten Extrempunkt, die Falschheit, gilt dies entsprechend.

Was ist nun das Verhalten der Bayesischen Formel zwischen diesen ausgezeichneten Punkten? Wie verhält sie sich zum primitiven Induktivismus – ist

²siehe (Chalmers, 2001, Kapitel 12); zu dieser und im folgenden betrachteten probabilistischen Thesen auch besonders (Gillies, 2000)

es für den Bayesianer möglich, absolute Gewissheit über die Wahrheit einer Theorie zu gewinnen? Umgekehrt, kann eine Theorie falsifiziert werden?

Wie ich zeigen werde, ist die Verifizierung, also das Erreichen des Wertes 1, auch mit der Bayesischen Formel nicht möglich. Bei der Falsifizierbarkeit folgt die Theorie dem normalen Falsifikationismus, unter Ausklammerung der Quine-Duhem-These ist sie möglich: das Gesamtpaket von Hypothese und Hintergrundwissen kann im strengen Sinne falsifiziert werden.

2 Wahrheit der Priorwahrscheinlichkeit

Es ist nicht möglich, einer Hypothese die Priorwahrscheinlichkeit 1 zuzuweisen. Dies folgt mathematisch direkt aus dem Bayesischen Theorem: angenommen, $P(h) = 1$, dann gilt sofort:

$$P(h|e) = \frac{P(e|h)}{P(e|h)}$$

Dies bedeutet, dass bei einer Priorwahrscheinlichkeit von 1 die Posteriorwahrscheinlichkeit auch 1 ist, eine einmal gesicherte Behauptung nicht mehr widerlegt werden kann (was ein massives Problem bei der Aufrechterhaltung von Falsifizierbarkeit darstellen würde). Es gibt jedoch eine Ausnahme, nämlich genau $P(e|h) = 0$, wenn ein Ereignis eintritt, welches von der Hypothese ausgeschlossen wurde (genau der für die Falsifizierung interessante Fall). Dann erhält man als Ergebnis nicht 1, sondern den mathematisch unzulässigen unbestimmten Ausdruck $\frac{0}{0}$. Das Zulassen von Priorwahrscheinlichkeiten gleich 1 kann also zu einem mathematisch verbotenen Zustand führen und muss daher von vorneherein ausgeschlossen werden.

3 Absolute Wahrheit im Bayesianismus

Weiterhin lässt sich zeigen, dass auch die Posteriorwahrscheinlichkeit nicht 1 werden kann. Die Behauptung ist also, dass $P(h|e) < 1$.

Als Beweis am ausschlaggebendsten ist die Tatsache, dass die Bayesische Formel rekursiv ist, eine berechnete Posteriorwahrscheinlichkeit wird im nächsten Schritt zur Priorwahrscheinlichkeit. Wie soeben gezeigt, ist eine Priorwahrscheinlichkeit von 1 nicht zulässig. Daher darf auch eine Posteriorwahrscheinlichkeit von 1 nicht möglich sein.

Darüber hinaus lässt sich, immer noch unter der Voraussetzung $P(h) < 1$, eine plausible Annahme einführen, die dies untermauert, nämlich dass $P(e|\neg h) > 0$. Ohne diese Forderung erhält man für $P(e|\neg h) = 0$ unmittelbar

$$\begin{aligned} P(h|e) &= P(h) \cdot \frac{P(e|h)}{P(h)P(e|h)} \\ &= 1 \end{aligned}$$

Dieses Ergebnis würde also der Behauptung widersprechen. In Worten ausgedrückt bedeutet es: Wenn eine Theorie die einzig mögliche ist, die das Ereignis vorhersagt, und das Ereignis tritt ein, dann ist die Theorie wahr. Dies ist auch unmittelbar einsichtig, ebenso wie das erkenntnistheoretische Argument, mit

dem $P(e|\neg h) > 0$ gefordert werden kann: Es ist offensichtlich unsinnig anzunehmen, dass irgendeine Theorie die einzig mögliche ist, die ein Ereignis vorhersagen kann. Ich habe zuvor behauptet, dass eine Wissenschaftler bei der Berechnung von $P(e|\neg h)$ nur die ihm bekannten Theorien berücksichtigen muss. An dieser Stelle muss diese Behauptung nun eingeschränkt werden: Ich schlage vor, dass grundsätzlich eine der Alternativtheorien als „Die unbekannte Theorie“ postuliert wird. Damit ist eine noch nicht bekannte Theorie gemeint, die in irgendeinem beliebigen Maße (je nachdem, wie stark der Wissenschaftler an noch zu entdeckende relevante Theorien glaubt) das Ereignis e ebenfalls voraussagt. Damit ist dann automatisch stets $P(e|\neg h) > 0$. Wäre man nicht bereit, dieses Zugeständnis zu machen, käme man zu dem Ergebnis, dass eine Posteriorwahrscheinlichkeit von 1 sehr wohl erreicht werden könnte.

Angesichts der Tatsache, dass wie oben gezeigt aus $P(h) < 1$ schon $P(h|e) < 1$ folgt, wären wir auf ein echtes Paradoxon gestoßen. In diesem Sinne lässt sich sogar folgern, dass zur Vermeidung eines solchen Paradoxons die gemachte Annahme geradezu notwendig ist.

Es sei nun also der Fall $P(e|\neg h) > 0$ betrachtet. Dann gilt:

$$\begin{aligned}
 P(h|e) & \stackrel{P(e|h)=1}{\leq} P(h) \cdot \frac{1}{P(h) + [1 - P(h)] P(e|\neg h)} \\
 & < \frac{P(h)}{P(h)} \\
 & = 1
 \end{aligned}$$

Da nun gezeigt ist, dass absolute Wahrheiten im Bayesianismus nicht möglich sind, ist auch ein naiver Induktivismus widerlegt, der behauptet, dass sich aus Einzelbeobachtungen ein absolut wahrer Allsatz ableiten lässt.

4 Falsifikation im Bayesianismus

Sehr wohl möglich ist es allerdings, dass die Posteriorwahrscheinlichkeit 0 wird. Dies ist genau dann der Fall, wenn $P(e|h) = 0$ ist, wenn also ein Ereignis eintritt, welches von der Hypothese ausgeschlossen wurde. Offensichtlich gibt es dann auch keine Möglichkeit mehr, dass sich die Wahrscheinlichkeit der These nach einmal verändert. Die Priorwahrscheinlichkeit (welche im vorherigen Schritt die Posteriorwahrscheinlichkeit war) als Faktor gewährleistet, dass die Wahrscheinlichkeit konstant auf 0 bleibt.

An dieser Stelle könnte man in Rückblick auf den Beweis der Unmöglichkeit von Posteriorwahrscheinlichkeiten gleich 1 folgenden Einwand einbringen: Wenn es möglich ist, für eine Hypothese eine definitive Wahrscheinlichkeit 0 zu errechnen, dann müsste es doch auch möglich sein, eine Posteriorwahrscheinlichkeit von 1 zu erhalten, denn man müsste ja nur die Negation der widerlegten Hypothese betrachten, diese wäre dann automatisch mit Sicherheit bestätigt.

Dies ist jedoch ein Trugschluss, denn die logische Form wissenschaftlicher Sätze ist nicht beliebig: sie sind generell als bedingter Allsatz zu formulieren. Bei Carnap findet sich die folgende Formalisierung³:

$$\forall x : Fx \rightarrow Gx$$

³(Carnap, 1958, Abschnitt 9c. Universal conditionals)

Jeder Gegenstand der Art F hat die Eigenschaft G.

Es ist nun leicht zu sehen, dass die Negation eines Allsatzes

$$\exists x : \neg(Fx \rightarrow Gx)$$

selbst *kein* Allsatz mehr ist, und also auch nicht als wissenschaftliche Hypothese in Frage kommt.

5 Erwiderung auf die Kritik der reinen Subjektivität

Oft wird kritisiert, dass es notwendig sei, nachvollziehen zu können, warum ein bestimmter Wissenschaftler an eine bestimmte Hypothese geglaubt hat. So Chalmers:

Das [...] Problem ist die mangelnde Plausibilität des Gedankens, dass wir den Zugang zu privaten Überzeugungen benötigen, um zu erfassen, in welchem Sinne zum Beispiel die Wellentheorie des Lichts eine Verbesserung gegenüber ihrer Vorgängerin darstellt.

[...]

Wenn die Fortschrittlichkeit dieser Unternehmung als Konzentration auf den Umfang von Überzeugungen verstanden wird, dann stellt sich die Frage, auf wessen Überzeugung wir vertrauen und warum.⁴

Da der Bayesianismus eine streng subjektive Theorie ist, und es tendenziell nicht möglich ist, zu den Priorwahrscheinlichkeiten eines Wissenschaftlers Zugang zu erhalten, wird das Bemühen vereitelt, etwas über die Gültigkeit seiner Überzeugungen auszusagen.

Als Entgegnung sei hier auf Popper zurückgegriffen, der behauptet, dass diese Notwendigkeit in keinsten Weise besteht. Es interessiert eben *nicht*, warum ein bestimmter Wissenschaftler etwas geglaubt hat. Wie Popper immer wieder betont, gibt es keinerlei Anforderungen oder Notwendigkeiten einer Rechtfertigung an den kreativen Prozess eines Wissenschaftlers. Die falsifikationistische Haltung, dass sich eine Theorie an der Wirklichkeit messen muss, gewährleistet, dass nur das *Ergebnis* der Überlegung in Betracht gezogen werden muss:

1. Es gibt keine letzten Quellen der Erkenntnis. Jede Quelle, jede Anregung ist uns willkommen; aber jede Quelle, jede Anregung ist auch Gegenstand kritischer Überprüfung. Soweit es sich aber nicht um historische Fragen handelt, pflegen wir eher die behaupteten Tatsachen selbst zu prüfen, als den Quellen unserer Information nachzugehen.

2. Die Fragen der Wissenschaftslehre haben mit Quellen eigentlich nichts zu tun. Was wir fragen ist vielmehr, ob eine Behauptung wahr ist – das heißt, ob sie mit den Tatsachen übereinstimmt.⁵

⁴(Chalmers, 2001, S. 151)

⁵(Popper, 1984, „Über die sogenannten Quellen der Erkenntnis“)

Jeder kann dann subjektiv bewerten, inwieweit es Gründe gibt, diese Ergebnisse anzunehmen oder insbesondere zu verwerfen. Die empirische Bestätigung oder Falsifikation steht dabei an erster Stelle.

Darüber hinaus gewährleistet die Eigenschaft der Konvergenz der Bayesischen Wahrscheinlichkeit⁶, dass intersubjektive (stets vorläufige) Erkenntnisse erreicht werden können: Selbst wenn Wissenschaftler mit unterschiedlichen Priorwahrscheinlichkeiten beginnen, so muss sich ihre Überzeugung doch demselben Wert annähern, wenn sie eine ausreichende Anzahl derselben Bestätigungen berücksichtigen. Gerade bei der Untersuchung, welchen Wert eine ältere Theorie hat, kann man dabei auf die Unzahl der historischen Experimente zurückgreifen, welche objektiv zu Verfügung steht und ohne weiteres gewährleisten sollten, dass heute zwei Wissenschaftler unabhängig von allen ihren Anfangsüberzeugungen beispielsweise an die Gravitationstheorie glauben können.

Unter welchen Bedingungen genau eine solche intersubjektive „scientific community“ denkbar ist, müsste an anderer Stelle untersucht werden.

6 Zusammenfassung und Schlussbemerkungen

Wir haben also gesehen, dass der Bayesianismus, ebenso wie der Falsifikationismus, die Erkenntnis absoluter Wahrheit ablehnt. Den Wahrheitswert 1 für eine Hypothese anzunehmen ist schlicht mathematisch nicht möglich.

Auch ansonsten ist der Bayesianismus nicht inkompatibel zum Falsifikationismus, die Art und Weise, in der ein Wahrheitswert 0 möglich ist, stimmt genau überein, sofern Falsifikation überhaupt möglich sein soll.

Darüber hinaus aber bietet das Bayesische Theorem die Chance, sich von den strengen Wahrheitskategorien wahr und falsch zu lösen. Dies erscheint insbesondere dann sinnvoll, geradezu zwingend, wenn man der Kritik der Quine-Duhem-These folgt und auch das Erkennen absoluter Falschheit verbietet. Dann kann Wissenschaft nur noch als Entwicklung (intersubjektiver) Glaubensgrade beschrieben werden. Ich halte diese Möglichkeit nicht für unbefriedigend. Eine Orientierung am kritischen empirischen Maßstab ist nach wie vor gültig, die Theorie muss sich an der Wirklichkeit messen. Der Bayesianismus lädt gerade auch dazu ein, eine Theorie immer wieder kritisch zu überprüfen und ernsthaft eine „Falsifizierung“ zu versuchen. Die „neue Berufsethik“⁷ Poppers, nach der der Fehler ein zentraler und guter Bestandteil von Wissenschaft ist, bietet auch dem Bayesianer einen guten Leitfaden.

Literatur

[Carnap 1958] CARNAP, Rudolf: *Introduction to Symbolic Logic and its Applications*. Dover, New York, 1958

[Chalmers 2001] CHALMERS, Alan F.: *Wege der Wissenschaft. Einführung in die Wissenschaftstheorie*. Springer, Berlin, 2001

[Gillies 2000] GILLIES, Donald: *Philosophical Theories of Probability*. Routledge, London, 2000

⁶Eine ausführliche Beschreibung siehe (Howson und Urbach, 1989)

⁷(Popper, 1984, „Duldsamkeit und intellektuelle Verantwortlichkeit“)

- [Goerz 2006] GOERZ, Michael: *Falsifikationismus als Notwendigkeit*. 2006
- [Howson und Urbach 1989] HOWSON, Colin ; URBACH, Peter: *Scientific Reasoning: The Bayesian Approach*. Open Court, 1989
- [Popper 1984] POPPER, Karl: *Auf der Suche nach einer besseren Welt*. Piper, München, 1984