Möglichkeiten und Grenzen der Formel von Bayes

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

15. April 2007

Zusammenfassung

Dieser Text erläutert die verschiedenen Varianten der Bayesischen Formel und ihre jeweiligen Probleme. Dabei lege ich mich auf eine schwache Formulierung des Theorems fest, die die Berechnung relativer Bestätigungsgrade auf der Basis empirischer Daten erlaubt.

Desweiteren erläutere ich, wie der Bayesianismus von verschiedenen fortgeschritteneren mathematischen Konzepten, d.h. dem Übergang zum Kontinuum, profitieren kann und führe einen Formalismus ein, der der tatsächlichen Anwendung gerecht wird. Auf dieser Basis führe ich eine Fallstudie am Planckschen Strahlungsgesetz durch.

Schließlich versuche ich eine gravierende Unzulänglichkeit des Bayesianismus aufzudecken; zentrales Postulat ist die These, dass „Verständnis" als nichtempirisches Element wesentliches Ziel der Wissenschaft ist. Der Bayesianismus wird diesem Ziel nicht gerecht.

Inhalt

*Dieser Text wurde als Hausarbeit für das Hauptseminar „Empirische Unterbestimmtheit und Theoriendynamik" bei Prof. Holm Tetens, Freie Universität Berlin, WiSe 2006/07, erstellt.

1 Einleitung

Das Bayesische Theorem ist ein recht einfach zu beweisender Satz[1] der Wahrscheinlichkeitslehre, der eine Aussage über bedingte Wahrscheinlichkeiten $\mathrm{P}(A|B)$, bzw. $\mathrm{P}(x_1 = A|\ x_2 = B)$, macht. Berechnet wird die Wahrscheinlichkeit, dass die Zufallsvariable x_1 den Wert A annimmt, wenn schon bekannt ist, dass x_2 den Wert B annimmt.

Es ist sehr populär geworden, das Theorem auf wissenschaftliche Hypothesen und von ihnen vorhergesagte Beobachtungsereignisse anzuwenden. Die „Wahrscheinlichkeit" einer Hypothese ist dann der subjektive Überzeugungsgrad von ihrer Wahrheit. Der Satz von Bayes gibt dann an, wie jemand seinen Überzeugungsgrad im Lichte eines Beobachtungsereignisses verändern muss. Die Identifikation von Überzeugungsgraden und mathematischen Wahrscheinlichkeiten fällt nicht vom Himmel. Eine mögliche Grundlage ist die Annahme, dass sich Überzeugungsgrade durch einen Wettquotienten quantifizieren lassen. Wie Glymour erläutert:[2]

> Wir können für eine Person den Glaubensgrad an einen Satz P messen, indem wir für einen festen Betrag v den höchsten Betrag u finden, sodass die Person u bezahlen würde um $u + v$ zu erhalten falls P wahr ist und nichts zu erhalten falls P falsch ist. [...] Wir erhalten

$$\mathrm{Prob}(P) = u/(u + v) \tag{1}$$

Es lässt sich nun zeigen, dass eine Person, die ihre Glaubensgrade anders als dem Bayesischen Theorem folgend anpasst, einem „Dutch Book" unterliegen kann: Es ist möglich, gegen sie eine Reihe von Wetten zu konstruieren, sodass sie unabhängig vom Ausgang Geld verliert. In diesem Sinne kann das Theorem, angewendet auf Überzeugungsgrade, den Anspruch auf die Beschreibung von *Rationalität* erheben.

Es ist zu bemerken, dass die Interpretation und Anwendung von Wahrscheinlichkeiten auf philosophische Begriffe keineswegs unstrittig ist, so sind viele Philosophen der Meinung, man solle den Begriff der Wahrscheinlichkeit in seinem wohldefinierten Rahmen – den relativen Häufigkeiten – lassen, und dass es nicht sinnvoll sei, ihn beispielsweise mit subjektiven Glaubensgraden zu identifizieren. Eine ausführliche Darstellung der philosophischen Debatte um die Interpretation von Wahrscheinlichkeiten ist bei (Gillies, 2000) zu finden.

Das Bayesische Theorem, in der Form in der ich es verwenden möchte, lautet

$$\mathrm{P}(h_i|e) \quad = \quad \mathrm{P}(h_i) \cdot \frac{\mathrm{P}(e|h_i)}{\sum\limits_{j=1}^{n} \mathrm{P}(h_j)\mathrm{P}(e|h_j)} \tag{2}$$

Es drückt aus, wie sich die Wahrscheinlichkeit der i-ten aus einer Liste von n alternativen Hypothesen im Lichte des Beobachtungsereignisses e verändert. Dabei ist $\mathrm{P}(h_i|e)$ die Wahrscheinlichkeit der Hypothese unter Berücksichtigung des Ereignis e (Posteriorwahrscheinlichkeit), $\mathrm{P}(h_i)$ ist die Priorwahrscheinlichkeit,

[1] siehe (Papoulis, 1984, S. 38 f.)

[2] (Glymour, 1998, S. 589), eigene Übersetzung

mit der jemand im Vorhinein an die Wahrheit der Hypothese glaubt, $P(e|h_i)$ ist die Wahrscheinlichkeit, die die Hypothese h_i für das Auftreten des Ereignis e voraussagt, die „Likelyhood".

Es existieren verschiedene Formulierungen des Theorems, die jeweils verschiedene Interpretationen und auch Probleme nach sich ziehen. Ich wähle die in Gl. (2) angegeben Form, weil ich denke, dass sie gravierende Probleme dieser anderen Interpretationen umgeht, wie ich im Folgenden erläutern werde. Dabei reduziert sich die philosophische Bedeutung des Theorems allerdings. Es wird denn auch meine These sein, dass die Bedeutung des Bayesischen Theorems eher die eines statistischen Analysewerkzeugs ist und dass sich die Interpretation der Wahrscheinlichkeitsterme von „Glaubengrad" zu „Bestätigungsgrad" verschiebt.

2 Formulierungen des Bayesischen Theorems

2.1 Erwartbarkeit und das Old-Evidence-Problem

Die einfachste Form des Theorems ist wohl die folgende:

$$P(h|e) \quad = \quad P(h) \cdot \frac{P(e|h)}{P(e)} \tag{3}$$

Darin soll $P(e)$ die „Erwartbarkeit" der Beobachtung ausdrücken. Es ergibt sich sofort in Übereinstimmung mit unserer Intuition, dass die Vorhersage ansonsten unerwarteter Beobachtungen eine Theorie besonders stark bestätigen kann. Das Problem ist allerdings, dass es äußerst problematisch ist, den Begriff der Erwartbarkeit zu quantifizieren. Setzt man ihn intuitiv als subjektive Größe ein, bestimmt auf dieselbe Weise wie die Priorwahrscheinlichkeiten, so ist die logische Konsistenz gefährdet: Posteriorwahrscheinlichkeiten größer Eins sind möglich. $P(e)$ ist nicht unabhängig!

Ein weiteres Problem stellt sich: In der Regel werden zur Bestätigung von Theorien schon bekannte Daten herangezogen. Solche sind allerdings mit Sicherheit nicht unerwartet (der Ausgang der Messung steht ja schon fest).

> Alte Daten können in der Realität eine neue Theorie bestätigen, aber nach der Bayesianischen Dynamik können sie es nicht. Denn: angenommen e ist bekannt bevor die Theorie T zur Zeit t eingebracht wird. Da e zu t bekannt ist, gilt $\mathrm{Prob}_t(e) = 1$. Weiterhin: da $\mathrm{Prob}_t(e) = 1$, ist auch die Likelyhood von e gegeben T, $\mathrm{Prob}_t(e|T)$, eins. [...] Die abhängige Wahrscheinlichkeit von T gegeben e ist daher dieselbe wie die Priorwahrscheinlichkeit von T.[3]

Der Ausdruck $P(e)$ in seinem primitiven Verständnis ist also höchst problematisch. Die Versionen des Bayesischen Theorems unterscheiden sich denn auch nur in der (notwendigen) Explikation der Größe $P(e)$.

2.2 Vollständigkeit der Hypothesen

Nach dem Gesetz der vollständigen Wahrscheinlichkeiten müssen sich die Wahrscheinlichkeiten einer vollständigen Liste sich ausschließender Alternativen zu

[3] (Glymour, 1998, S. 600), eigene Übersetzung

Eins aufaddieren. Es lässt sich mit Sicherheit sagen, dass die Theorie h entweder zutrifft, oder nicht, und damit können wir sofort schreiben

$$\begin{aligned} \mathrm{P}(e) &= \mathrm{P}((h \cup \neg h) \cap e) \\ &= \mathrm{P}(h \cap e) + \mathrm{P}(\neg h \cap e) \\ &= \mathrm{P}(h)\mathrm{P}(e|h) + \mathrm{P}(\neg h)\mathrm{P}(e|\neg h) \end{aligned} \tag{4}$$

So erhält man:

$$\begin{aligned} \mathrm{P}(h|e) &= \mathrm{P}(h) \cdot \frac{\mathrm{P}(e|h)}{\mathrm{P}(h)\mathrm{P}(e|h) + \mathrm{P}(\neg h)\mathrm{P}(e|\neg h)} \\ &= \mathrm{P}(h) \cdot \frac{\mathrm{P}(e|h)}{\mathrm{P}(h)\mathrm{P}(e|h) + [1 - \mathrm{P}(h)]\,\mathrm{P}(e|\neg h)} \end{aligned} \tag{5}$$

Das Problem der Inkonsistenz tritt hier nun nicht mehr auf. Der problematische Term ist hier aber $\mathrm{P}(e|\neg h)$. Die Theorie $\neg h$ ist als solches nicht berechenbar, und es gibt keine Möglichkeit festzustellen, mit welcher Wahrscheinlichkeit sie e voraussagt.

2.3 Die Catch-All-Hypothese

Zum Glück kann die Theorie $\neg h$ noch weiter ausgeführt werden. Sie besteht explizit aus allen Alternativtheorien zu h. Das einzige Problem daran ist, dass die Auflistung *aller* Alternativtheorien auch solche einschließen müsste, die noch gar nicht gefunden sind. Oft werden all jene Theorie zu einer „Catch-All-Hypothese“ h_k zusammengefasst. Bei einer Liste von n bekannten Alternativtheorien und der Catch-All-Hypothese erhält man so:

$$\mathrm{P}(h_i|e) = \mathrm{P}(h_i) \cdot \frac{\mathrm{P}(e|h_i)}{\sum\limits_{j=1}^{n} \mathrm{P}(h_j)\mathrm{P}(e|h_j) + \mathrm{P}(h_k)\mathrm{P}(e|h_k)} \tag{6}$$

Die Einführung der Catch-All-Hypothese löst das Problem aber in keiner Weise. Es ist nicht ersichtlich, wie die von h_k für e vorausgesagte Likelyhood berechnet werden könnte.

2.4 Relative Bestätigung

Da alle dieser Varianten keine Möglichkeit bieten, einen numerischen Wert für $\mathrm{P}(e)$ anzugeben, habe ich in meiner Schreibweise die Catch-All-Hypothese vernachlässigt. Damit sind alle Größen berechenbar. Auch das Old-Evidence-Problem tritt so nicht mehr auf. Allerdings bringt dieser Schritt einen erheblichen Preis mit sich.

Zum einen wird es nötig, bei der Aufstellung einer neuen Theorie die Berechnung der Kette von Posteriorwahrscheinlichkeiten von Grund auf neu durchzuführen. Global gesehen sind die Überzeugungen des Forschers dann in diesem Sinne nicht mehr konsistent: In der Wettanalogie könnte ein Dutch Book gegen ihn aufgesetzt werden, da alte Wetten nicht zurückgenommen werden können, wobei aber das Hinzufügen einer neuen These dies eigentlich erfordern würde.

Zum anderen ist man gezwungen, (vorübergehend) für die betrachteten Alternativtheorien Vollständigkeit zu postulieren:

$$\sum_i \mathrm{P}(h_i) \quad = \quad 1 \tag{7}$$

Dies ist aber nicht mehr streng konsistent zur Evaluierung der Überzeugungsgrade mittels der Wettanalogie: Es ist durchaus möglich, dass jemand von keiner der zur Verfügung stehenden Hypothesen sonderlich überzeugt ist, ohne allerdings eine explizite Alternativtheorie aufstellen zu können. Im Normalfall würde dies die Catch-All-Hypothese auffangen, in unserem Fall jedoch ist der Forscher gezwungen, Priorwahrscheinlichkeiten anzunehmen, die gar nicht seinem echten Überzeugungsgrad entsprechen. Selbst wenn man die Forderung der Vollständigkeit zurückweist, und die echten Überzeugungsgrade als Priorwahrscheinlichkeiten einsetzt, lässt sich sehr einfach zeigen, dass sich dann dennoch die Posteriorwahrscheinlichkeiten zu Eins aufaddieren. Da diese im nächsten Durchgang zu Priorwahrscheinlichkeiten werden, ist nichts gewonnen.

Insgesamt sehen wir also, dass das Bayesische Theorem in der von mir gewählten Form nur einen vergleichenden Charakter hat. Es ist also ähnlich anzusehen wie die von Salmon aus dem Theorem hergeleitete Formel für den Vergleich zweier Theorien, die komplett auf die problematische Größe $\mathrm{P}(e)$ verzichtet:[4]

$$\frac{\mathrm{P}(T_1|E.B)}{\mathrm{P}(T_2|E.B)} \quad = \quad \frac{\mathrm{P}(T_1|B) \times \mathrm{P}(E|T_1.B)}{\mathrm{P}(T_2|B) \times \mathrm{P}(E|T_2.B)} \tag{8}$$

Auch hier wird nur berechnet, welche Theorie von beiden besser bestätigt wird, nicht aber, was der absolute Bestätigungsgrad der Theorien ist. Meine Formulierung der Bayesischen Formel in Gl. (2) hat demgegenüber den Vorteil, dass sie mehr als nur zwei Theorien gleichzeitig miteinander vergleichen kann.

3 Berechnung der Likelyhoods

3.1 Likelyhoods jenseits von Implikation und Widerspruch?

Die in der Bayesischen Formel auftretenden Likelyhoods $\mathrm{P}(e|h)$ spezifizierter Theorien erscheinen auf den ersten Blick unproblematisch im Vergleich zu den bisher diskutierten Kritiken. Dennoch hält die Frage nach ihrer Berechenbarkeit einige Philosophen davon ab, die Formel als besonders anwendbar in der Wissenschaftstheorie anzusehen. Oft wird nahegelegt, möglicherweise aus der Tradition des Falsifikationismus heraus, dass die Theorien ein Beobachtungsergebnis vorhersagen, oder mit ihm inkompatibel sind, dass h also entweder das Ereignis e oder ein anderes Ereignis $\neg e$ impliziert. Die Likelyhoods sind in diesen Fällen Eins und Null, was die Formel schlagartig recht langweilig werden lässt. Ein anderer Trivialfall sind Vorhersagen über einfache statistische Ereignisse, etwa Münzwürfe.

Es ist zunächst nicht offensichtlich, wie die Likelyhood in solchen Fällen berechnet werden kann, in denen eine Hypothese das Ereignis e nicht unmittelbar impliziert.

[4] siehe (Salmon, 1998, S. 569)

Dieses Problem lässt sich allerdings lösen, wenn man sich vor Augen führt, dass alle Gesetze der Physik implizit stochastisch sind; diese Betrachtung benötigt jedoch einen mathematischen Apparat, der über simple Münzwürfe oder ähnliches hinausgeht.

Zum einen ist es so, dass fast alle nichtelementaren Formeln direkt gut abschätzbare mathematische Näherungen beinhalten, sodass statistische Aussagen möglich werden.

Selbst wo das nicht der Fall ist, ist es eine Grundregel jeder experimentellen Wissenschaft, dass alle Messgrößen fehlerbehaftet sind und daher statistisch behandelt werden müssen.

3.2 Formalismus der Wahrscheinlichkeitsverteilungen

3.2.1 Likelyhoods für normalverteilte Daten

In der Realität werden Aussagen über Mess- und theoretische Größen durch Wahrscheinlichkeitsverteilungen gemacht. Ist eine solche Verteilung ϕ gegeben, berechnet man die Wahrscheinlichkeit, dass die Größe x im Intervall $[a:b]$ liegt als

$$\mathrm{P}(x \in [a:b]) = \int_a^b \phi(x)\,\mathrm{d}x \tag{9}$$

Bei experimentellen Daten wird als Wahrscheinlichkeitsverteilung in der Regel eine Normalverteilung (Gaußverteilung) angenommen:

$$\begin{aligned} \phi(x) &= \Gamma(\mu, \sigma^2) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \end{aligned} \tag{10}$$

Dabei ist μ der Mittelwert und σ die Breite der Verteilung.

Setzt man einen normalverteilten Messwert in eine Formel ein, so ist die Verteilung des Ergebnisses wiederum eine Gaußkurve. Hat das Gesetz die Form

$$y = h(x_1, x_2, \ldots x_i), \tag{11}$$

so gibt h selbst den Mittelwert μ_y von y an. Die Breite σ_y wird durch die Gaußsche Fehlerfortpflanzungsformel angegeben.

$$\sigma_y = \sqrt{\sum_i \left(\frac{\mathrm{d}h}{\mathrm{d}x_i}\sigma_{x_i}\right)^2} \tag{12}$$

Im einfachsten Fall $y = h(x)$ reduziert sich dies auf

$$\sigma_y = \frac{\mathrm{d}h}{\mathrm{d}x}\sigma_{x_i} \tag{13}$$

Nimmt man zum Beispiel die Formel für den zurückgelegten Weg bei konstanter Beschleunigung

$$s(t) = \frac{1}{2}at^2 \tag{14}$$

und misst man t mit der Breite σ_t, so ist die Breite von s

$$\sigma_s = at \cdot \sigma_t \tag{15}$$

Hat man nun also aus einer Formel für gegebene Parameter einen erwarteten Messwert mit der Verteilung $\Gamma(y, \sigma_y^2)$ ermittelt, so ist die Likelyhood den tatsächlichen Wert mit der Verteilung $\Gamma(e, \sigma_e^2)$ zu messen wiederum durch die Gaußsche Fehlerfortpflanzung gegeben: Wir suchen die Wahrscheinlichkeit, dass die Differenz $y - e$ innerhalb einer Toleranz t Null wird, dass also vorhergesagter und gemessener Wert tatsächlich übereinstimmen. Man erhält:

$$\mathrm{P}(e|h) = \int_{-t}^{t} \Gamma(y - e, (\sigma_y + \sigma_e)^2) \,\mathrm{d}x \tag{16}$$

Es bietet sich an, für die Toleranz

$$t = \frac{1}{2}(\sigma_y + \sigma_e) \tag{17}$$

anzunehmen.

3.2.2 Parameterabhängige Posteriorwahrscheinlichkeiten

Da $h = h(x)$, spricht nichts dagegen, von $\mathrm{P}(e|h(x))$ zu $\mathrm{P}(e|h)(x)$ überzugehen. Es ist also möglich, einen Überzeugungsgrad für eine Theorie h in einem bestimmten Wertebereich von x zu berechnen. Zum Beispiel könnte man feststellen, dass die Newtonsche Mechanik im Bereich niedriger Geschwindigkeiten einen sehr hohen Bestätigungsgrad erreichen kann, der erst bei Geschwindigkeiten nahe der Lichtgeschwindigkeit abnimmt. Es ist nicht zwingend nötig, die Theorie als Ganzes zu bewerten. Ich werde genau eine solche Art von Analyse in Abschnitt 4 durchführen. Das Vorgehen ist auch noch für mehr als einen Parameter möglich, allerdings sind die Posteriorwahrscheinlichkeiten dann entsprechend multidimensional und damit nicht mehr besonders anschaulich.

3.2.3 Kontinuierliche Hypothesenräume

Wie wir gesehen haben, gewinnt das Bayesische Theorem an Klarheit, wenn man von diskreten zu kontinuierlichen Größen übergeht. Dieser Ansatz lässt sich weiterführen zu einer Formel für eine unendliche Menge von Theorien – ein Modell mit kontinuierlichen Parametern. Die Formel kann dann angeben, welche Parameterwerte einen hohen Bestätigungsgrad haben und welche nicht. Dabei geht die Summe im Nenner zum Integral über, nach (Christensen u. a., S. 17) muss man nun schreiben:

$$p(\theta|y) = \frac{L(\theta|y)p(\theta)}{\int L(\theta|y)p(\theta) \,\mathrm{d}\theta} \tag{18}$$

Dabei steht θ für die Parameterverteilung, L gibt die Likelyhoods an.

Aus diesen Ideen lässt sich ein umfangreiches und mathematisch höchst anspruchsvolles Werkzeug zur statistischen Analyse und zur Bewertung und Variation von Theorien aufbauen. Sehr umfangreich werden viele dieser Konzepte in (Gregory, 2005) diskutiert.

4 Eine Fallstudie: das Plancksche Strahlungsgesetz

Ich möchte die in Abschnitt 3.2.1 und 3.2.2 dargestellten Methoden auf ein historisches Beispiel anwenden: die Schwarzkörperstrahlung. Es geht dabei um das Energiespektrum der Strahlung eines heißen Körpers.[5] Es gab drei mögliche Theorien zur Beschreibung dieser Strahlung. Das aus heutiger Sicht korrekte Gesetz stammt von Planck und ist ein Meilenstein quantenmechanischer Überlegungen, die beiden Alternativen stammen von Rayleigh-Jeans (aus der klassischen Elektrodynamik) und von Wien (aus der Thermodynamik).

Wiensches Strahlungsgesetz, Plancksches Strahlungsgesetz und Rayleigh-Jeans-Gesetz lauten:[6]

$$\rho_{\mathrm{W}}(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \exp(-\hbar\omega / k_B T) \tag{19}$$

$$\rho_{\mathrm{P}}(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar\omega / k_B T) - 1} \tag{20}$$

$$\rho_{\mathrm{RJ}}(\omega, T) = \frac{\omega^2}{\pi^2 c^3} k_B T \tag{21}$$

Der Verlauf der Kurven ist in Abb. 1 gezeigt. Wie man sieht, liefert das Rayleigh-

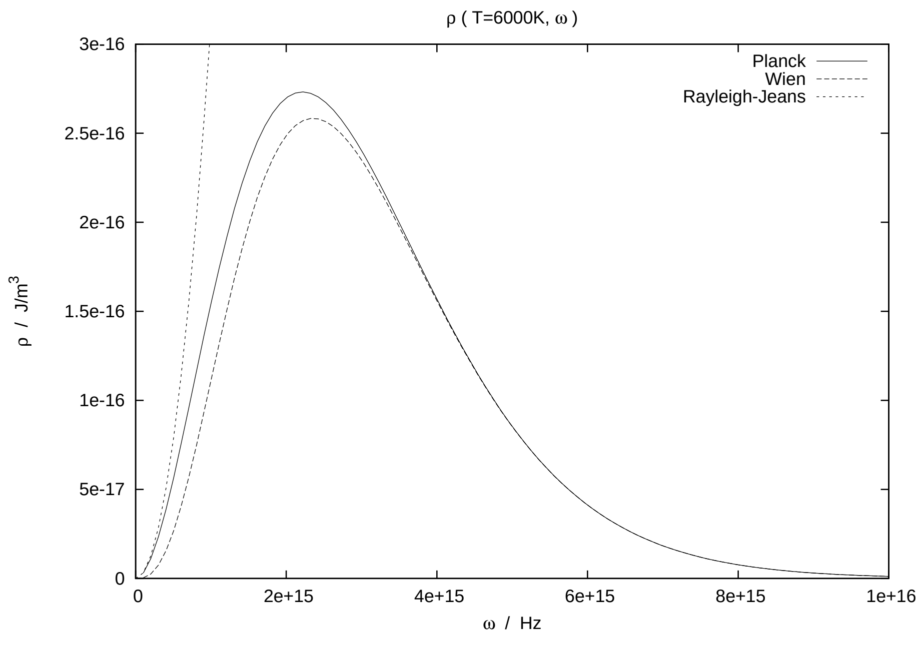

Abb. 1: Verlauf der Strahlungsgesetze

Jeans-Gesetz gute Werte für sehr niedrige Frequenzen, das Wiensche Strahlungsgesetz für sehr hohe Frequenzen.

[5] siehe (Giulini und Straumann, 2000)
[6] Formeln nach (Greiner, 2000, Kap. 2))

Um an empirische Daten zu gelangen, habe ich eine Simulation[7] durchgeführt. Dabei bin ich vom Planckschen Gesetz ausgegangen und habe eine Million realistischer Datenpunkte erzeugt. Für jeden Datenpunkt wurde eine zufällige Temperatur zwischen 5000 und 7000 K und ein ω im Bereich $[0 : 10^{16}\,\mathrm{Hz}]$ mit einem zufälligen Fehler zwischen 0.5 und 5 Prozent als Parameter verwendet. Mittels der Polar-Methode von Marsaglia[8] wurde dann ein normalverteilter Datenpunkt erzeugt.

Bei der Analyse wurde von einem Messfehler von 1% ausgegangen. Die Likelyhood wurde gemäß Gl. (16) berechnet.

Die anfänglichen Priorwahrscheinlichkeiten wahren für alle drei Hypothesen konstant 1/3. In einer realistischeren Bayesianischen Untersuchung würde man Expertenmeinungen oder vorherige Ergebnisse zur Ermittlung einer Priorwahrscheinlichkeit heranziehen.

Ein Beispiel für die Verrechnung eines Datenpunktes gibt der folgende Ausschnitt aus der Ausgabe des Programms:

```
T = 5017 K, omega = (2.3461 +- 0.0458)e+15 Hz
Priors     : 0.33, 0.33, 0.33
Likelyhoods: 0.32, 0.06, 8.18e-106
Products   : 0.10, 0.02, 2.72e-106
Posteriors : 0.83, 0.16, 2.08e-105
```

Das Ergebnis der Analyse ist in Abb. 2 dargestellt. Es entspricht genau

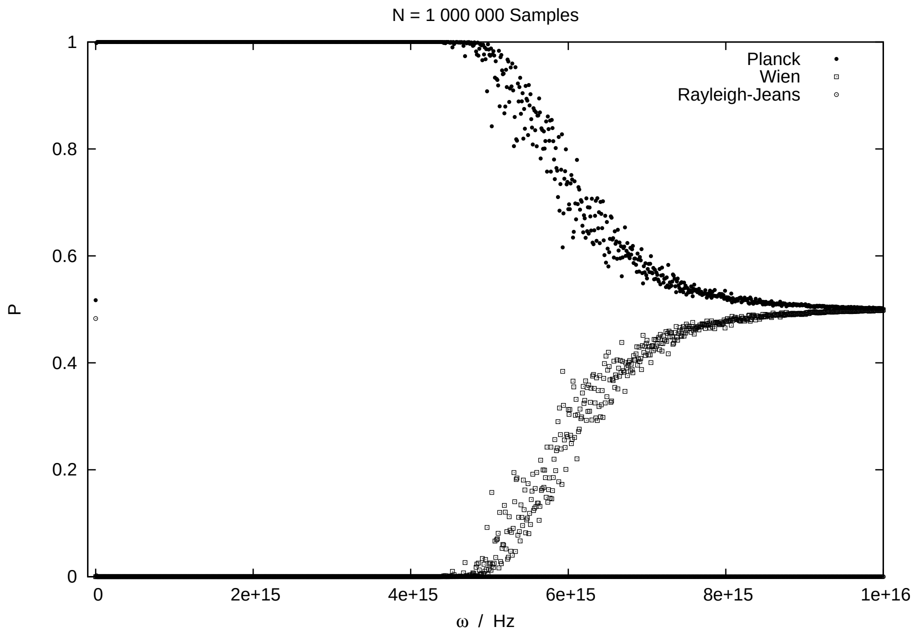

Abb. 2: Posteriorwahrscheinlichkeiten der Strahlungsgesetze

der Erwartung: Im ganz niedrigen Temperaturbereich hat das Rayleigh-Jeans-

[7] Programmcodes und Daten unter
`http://www.physik.fu-berlin.de/~goerz/studies/programme/bayes.zip`

[8] `http://de.wikipedia.org/wiki/Polar-Methode`

Gesetz einen nennenswerten Bestätigungsgrad, ebenso wie die Plancksche Formel. Bis zu einer Frequenz von ca. $4.5 \cdot 10^{15}$ Hz ist dann das Plancksche Strahlungsgesetz mit faktisch 1 bestätigt, darüber hinaus nähern sich Wien und Planck so nah an, dass sie die Daten gleichermaßen gut unterstützen: Zwischen beiden Theorien kann nicht mehr unterschieden werden, sie pendeln sich beide auf einer Posteriorwahrscheinlichkeit von 0.5 ein.

Die Streuung der Punkte im Übergangsbereich ist auf die beschränkte Anzahl von Daten zurückzuführen, je größer die Anzahl der Samples, desto geringer fällt sie aus.

Wollte man eine Gesamtbestätigungsgrad für die verschiedenen Theorien ermitteln, würde man über die Kurven integrieren. Für das Plancksche Gesetz ergäbe sich dabei der größte Bestätigungsgrad.

5 Kritik am Bayesianismus

Nachdem wir nun gesehen haben, dass die Bayesische Formel sehr gut in der Lage ist, den Bestätigungsgrad einer Theorie auf der Basis empirischer Daten zu ermitteln, ist es an der Zeit, ihn in seine Schranke zu weisen.

Ist der empirische Bestätigungsgrad, oder der subjektive Überzeugungsgrad von der Wahrheit einer Theorie der Kern der Wissenschaft? Meiner Ansicht nach nicht. Selbstverständlich sind auch statistische Analysen und Studien „wissenschaftlich". Wenn beispielsweise allein aufgrund statistischer Analyse nachgewiesen wird, dass mit einer Wahrscheinlichkeit von mehr als 90% menschliche Aktivitäten zur mehr als 70% zur globalen Erderwärmung beitragen, so wäre es sowohl irrational als auch unwissenschaftlich, einen solchen Zusammenhang zu bestreiten; eine solche Studie ist aber wissenschaftlich noch *unbefriedigend*. Befriedigend wird die Theorie erst dann, wenn z.B. erklärt werden kann, welche Prozesse, der Ausstoß welcher Mengen von Treibhausgasen etwa, in welcher Art und Weise und durch welche Interaktionen zu diesem Ergebnis führt (was in dem Zusammenhang des Klimawandels ja auch geschehen ist).

So widerspreche ich denn auch der gängigen Meinung, das definierende Ziel von Wissenschaft sei die Vorhersage von Ereignissen. Ziel ist vielmehr das *Verständnis* von Prozessen, welches dann Vorhersagbarkeit nach sich zieht.

Dies lässt sich auch in der diskutierten Fallstudie wiederfinden. Planck selbst führte seine quantenmechanische Strahlungsformel aus empirischen Gründen als Modifikation der hergeleiteten Formel von Wien ein, die theoretische Herleitung gelang ihm erst deutlich später:

> Das war eine rein formale Annahme, und ich dachte mir nicht viel dabei, sondern eben nur das, dass ich unter allen Umständen, koste es was es wolle, ein positives Resultat herbeiführen musste.[9]

Die Tatsache, dass ihr ein hoher Bestätigungsgrad gegeben war, hielt ihn nicht davon ab, in seiner Vorgehensweise einen unbefriedigenden Kunstgriff zu sehen, und an der Existenz gequantelter Größen zu zweifeln:

> [...] indessen scheinen mir augenblicklich verschiedenartige Anzeichen darauf hinzudeuten, daß man trotz der bisherigen Erfolge der

[9] (Planck, 1969)

> atomistischen Theorie sich schließlich doch einmal zu einer Aufgabe derselben und zur Annahme einer continuierlichen Materie wird entschließen müssen.[10]

Sein Zweifel war deshalb plausibel, weil die Quantentheorie noch nicht ausgereift genug war, *Modelle des Verstehens* zu liefern.

Ich behaupte also, dass der Kern einer wissenschaftlichen Theorie gar nicht (oder nicht ausschließlich) in ihrem empirischen Gehalt liegt. Wie Glymour in Bezug auf den Bayesianismus feststellt:

> Angenommen, wir können die Konsequenzen einer Theorie in solche Sätze unterteilen, die Ausdruck tatsächlicher möglicher Beobachtungen bzw. einfacher Generalisierungen solcher Beobachtungen sind, und solche theoretischer Natur[...] Eine Theorie ist nie besser bestätigt als das Kollektiv ihrer Beobachtungskonsequenzen.[11]

Was ist nun also die Rolle, die dem Bayesianismus zuzuschreiben ist? Sicherlich kann er als Modell betrachtet werden, dass wertvolle Beiträge liefern kann, ebenso wie nichtprobabilistische Bestätigungstheorien oder die klassisch-deduktive Logik in bestimmten Bereichen sinnvolle Modelle sein können. In gewissem Sinne stellt der Bayesianismus ein logisches Kalkül dar. Wie (Horwich, 1998) erläutert, bietet er dabei einige Stärken, wie die Auflösung einiger Paradoxien. Allerdings sollte man sich davor hüten, ihn zum Dogma zu erheben, denn auch er stößt schnell an die Grenzen seiner Möglichkeiten und vernachlässigt, wie ich ausgeführt habe, einiges von dem was Wissenschaft im Kern ausmacht.

Literatur

[Christensen u. a.] Christensen, Ronald ; Hanson, Timothy ; Johnson, Wesley O.: *Bayesian Ideas and Data Analysis*. – URL `http://www.stat.ucdavis.edu/~johnson/st145/`

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

[Giulini und Straumann 2000] Giulini, Domenico ; Straumann, Norbert: "...ich dachte mir nicht viel dabei..." – Plancks ungerader Weg zur Strahlungsformel. In: *Physik Journal* 12 (2000), S. 37–41

[Glymour 1998] Glymour, Clark: Why I Am Not A Bayesian. In: Curd, Martin (Hrsg.) ; Cover, J.A. (Hrsg.): *Philosophy of Science, the Central Issues*. W.W. Norton & Company, 1998

[Gregory 2005] Gregory, Phil: *Bayesian Logical Data Analysis for the Physical Sciences*. Cambridge University Press, 2005

[Greiner 2000] Greiner, Walter: *Quantum Mechanics – An Introduction*. 4. ed. Springer, 2000

[10] (Planck, 1958, Band I, Dokument 4, S. 162–163)

[11] (Glymour, 1998, S. 598), eigene Übersetzung

[Horwich 1998] HORWICH, Paul: Wittgensteinian Bayesianism. In: CURD, Martin (Hrsg.) ; COVER, J.A. (Hrsg.): *Philosophy of Science, the Central Issues*. W.W. Norton & Company, 1998

[Papoulis 1984] PAPOULIS, Athanasios: *Probability, Random Variables and Stochastic Processes*. 2. ed. McGraw-Hill, 1984

[Planck 1958] PLANCK, Max: *Physikalische Abhandlungen und Vorträge*. Vieweg, 1958

[Planck 1969] PLANCK, Max: Brief an Robert Williams Wood von 1931. In: HERMANN, A. (Hrsg.): *Frühgeschichte der Quantentheorie*. Physik Verlag, 1969, S. 31

[Salmon 1998] SALMON, Wesley C.: Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes. In: CURD, Martin (Hrsg.) ; COVER, J.A. (Hrsg.): *Philosophy of Science, the Central Issues*. W.W. Norton & Company, 1998