Der Satz von Bayes: Eine geschichtsträchtige Idee mit vielfältigen Anwendungsmöglichkeiten

KATRIN WÖLFEL, ERLANGEN-NÜRNBERG

Zusammenfassung: *Der Satz von Bayes ist eines der bekanntesten Theoreme der Wahrscheinlichkeitstheorie. Wird er rein formal hergeleitet, gerät die ursprüngliche Problemstellung der Wahrscheinlichkeit von Ursachen jedoch in den Hintergrund. Bei der Anwendung des Satzes zur Lösung vielfältiger Probleme ist dieser Grundgedanke aber von großer Bedeutung. Im Folgenden soll deshalb die intuitive Idee hinter dem Theorem aus historischer Perspektive erörtert und aufgezeigt werden, wie Fragen verschiedenster Disziplinen dadurch „bayesianisch" gelöst werden können.*

1 Einleitung

Jeder, der sich heute ein wenig mit Wahrscheinlichkeitstheorie beschäftigt, stößt schnell auf den sog. Satz von Bayes, der häufig auch als „Formel" oder „Regel" von Bayes bezeichnet wird und in diesem Sinn im Hochschulbereich auch meist rein formal notiert und hergeleitet wird (vgl. z. B. Koop 2003, S. 1): Unter der Annahme, dass ein Ereignis A zuerst eintritt, würde man die Wahrscheinlichkeit des Eintretens dieses Ereignisses A und eines weiteren Ereignisses B ($pr(A \cap B)$) als Produkt der Wahrscheinlichkeit, dass A eintritt ($pr(A)$), und der Wahrscheinlichkeit, dass B eintritt unter der Bedingung, dass A bereits eingetreten ist ($pr(B|A)$), mit Hilfe der ersten Pfadregel berechnen, d. h.

$$pr(A \cap B) = pr(A)pr(B|A). \tag{1}$$

Nimmt man hingegen an, dass B vor A eintritt, würde sich $pr(A \cap B)$ berechnen lassen als

$$pr(A \cap B) = pr(B)pr(A|B). \tag{2}$$

Da die linken Seiten der Gleichungen (1) und (2) übereinstimmen, muss dies auch für die rechten Seiten gelten (da die zeitliche Präzedenz der Ereignisse A und B für die Berechnung von $pr(A \cap B)$ irrelevant ist):

$$pr(B|A)pr(A) = pr(A|B)pr(B) \tag{3}$$

Löst man (3) nach $pr(B|A)$ auf, ergibt sich die als Satz von Bayes bekannte Formel (die zur Betrachtung endlich vieler Ereignisse noch verallgemeinert werden kann):

$$pr(B|A) = \frac{pr(A|B)pr(B)}{pr(A)} \tag{4}$$

Bei dieser formalen Herleitung des Satzes gerät dessen ursprüngliche Problemstellung, die auch heute noch für vielerlei Anwendungen des Satzes von Bedeutung ist, ebenso wie der damalige stufenweise Beweis dieses Theorems in den Hintergrund. Stone (2013, S. 31) findet für diese Problematik folgende treffende Worte:

> „even though Bayes' rule follows in a few lines of algebra from the rules of probability, no amount of staring at the rules themselves will make Bayes' rule obvious. Perhaps if it were more obvious, Bayes and others would not have had such trouble in discovering it, and we would not expend so much effort in understanding its subtleties."

Neben der Tatsache, dass die Geschichte der Herkunft des Satzes von Bayes für sich selbst interessant ist, da es nicht sein Namensgeber Thomas Bayes (1701–1761) selbst war, der die Bedeutung des Satzes erkannte und ihn deshalb bekannt machte (sondern dessen Freund Richard Price und später unabhängig davon Pierre Simon Laplace), offenbart diese Geschichte die intuitive Idee hinter dem berühmten Satz, die bei jeder modernen Anwendung des Satzes bewusst sein sollte.

2 Grundidee des bayesianischen Ansatzes

Im Zentrum der bayesianischen Idee steht die Frage nach der Wahrscheinlichkeit von Ursachen (engl. probability of causes) bzw. die der „umgekehrten Wahrscheinlichkeit" (engl. inverse probability): Wenn die Folge einer unbekannten Ursache bekannt ist, was kann dann über die Wahrscheinlichkeit verschiedener möglicher Ursachen ausgesagt werden? Gesucht wird also die Wahrscheinlichkeit einer von n möglichen Ursachen $C_i (i \in \{1, \ldots, n\})$ bei Kenntnis der Folge E: $pr(C_i|E)$. Diese Situation der Unsicherheit besteht bei vielen Fragestellungen verschiedenster Disziplinen: Man beobachtet ein bestimmtes Ereignis und möchte auf dessen Ursachen schließen können. Die bayesianische Methode zur Lösung dieses Problems besteht darin, dass man zunächst eine subjektive (möglicherweise sogar willkürliche) Vermutung über die Ursache abgibt („guess“, vgl. Bayes und Price 1763, S. 392) und diese anschließend mit Hilfe immer neuer (objektiver) Informationen aktualisiert, wodurch man eine korrigierte Vermutung bekommt, die stets durch neue Informationen weiter verbessert werden kann. Die Verbindung dieser intuitiven Herangehensweise mit dem dahinter stehenden mathematischen Konzept wird deutlich, wenn man die Entwicklung des Theorems genauer betrachtet.

3 Ursprung des Theorems bei Thomas Bayes (1701–1791)

Heutiger Namensgeber des Satzes ist der Engländer Thomas Bayes (1701–1761), in dessen Werk der Satz seinen Ursprung zu haben scheint. Bayes war Amateur-Mathematiker. Auf Grund seiner Zugehörigkeit zur presbyterianischen Kirche hatte er keinen Zugang zu englischen Universitäten, weshalb er in Schottland (Edinburgh) Theologie und Mathematik studierte, bevor er ca. 1722 durch seinen ebenfalls geistlichen Vater zum Pastor ernannt wurde und als solcher von da an in London und später in Tunbridge Wells wirkte. Die einzige mathematische Veröffentlichung, die Bayes zu seiner Lebenszeit tätigte, war anonym und stammt aus dem Jahre 1736 („An Introduction to the Doctrine of Fluxions, and a Defence of the Mathematicians Against the Objections of the Author of the Analyst“[1]). Darin verteidigte er Newtons Ideen zur Infinitesimalrechnung und widersetzte sich damit George Berkeley (der der Autor des „Analyst“ war). Zu Beginn der 1740er Jahre erlangte Bayes zunehmend Respekt für seine mathematischen Arbeiten und wurde deshalb 1742 in die „Royal Society“ in London aufgenommen, die zu dieser Zeit eine Vereinigung von Amateur-Mathematikern war und die Bayes die Möglichkeit bot, seine Ideen mit anderen zu diskutieren und sich mathematisch weiter zu entwickeln.[2]

Auf die Frage nach der Wahrscheinlichkeit von Ursachen wurde Bayes durch ein bekanntes Werk von Abraham de Moivre aufmerksam, der „Doctrine of Chances“, von der drei Auflagen in der Zeit zwischen 1718 und 1756 erschienen. Dort findet sich die Idee, dass man die Ordnung des Universums möglicherweise durch Betrachtung verschiedener Naturphänomene erschließen könne (Moivre 1756, S. 252). So versuchte Bayes diesem Problem in den 1740er Jahren mit einem Gedankenexperiment näher zu kommen: Eine Kugel wird auf einen quadratischen Tisch geworfen. Nun stellt sich die Aufgabe, zu erschließen – ohne den Kugelwurf gesehen zu haben –, an welcher Stelle des Tisches die Kugel gelandet ist. Dazu wird eine weitere Kugel auf den Tisch geworfen, deren Position in Bezug zur ersten Kugel man beobachten kann (d. h. man weiß, ob die zweite Kugel rechts oder links bzw. über oder unter der ersten Kugel gelandet ist). Unter der Annahme, dass eine Kugel mit gleicher Wahrscheinlichkeit an jeder Stelle des Tisches auftreffen kann, kann man aus der Lage der zweiten Kugel die der ersten Kugel etwas präzisieren: Liegt die zweite Kugel links (rechts) von der ersten Kugel, ist es wahrscheinlicher, dass die erste Kugel in der rechten (linken) Tischhälfte liegt. Das Werfen weiterer Kugeln ermöglicht es, die Position der ersten Kugel stets weiter auf diese Weise einzugrenzen. Somit kann man also mit Hilfe immer neuer Beobachtungen (Positionen der geworfenen Kugeln) nähere Aussagen über die Vergangenheit (die Position der ersten Kugel) treffen und diesen Aussagen sogar Wahrscheinlichkeiten zuordnen. (Bayes und Price 1763, S. 385–388; McGrayne 2011, S. 7 f.)

Dies sei an einem fiktivem Beispiel erläutert, das zwar nicht schultauglich ist, aber die originäre bayesianische Idee verdeutlichen soll: Man nehme an, Bayes’ Gedankenexperiment führt nach fünf Kugelwürfen (im Anschluss an den Wurf der ursprünglichen Kugel) zu folgendem Ergebnis:

- Erster Wurf: Kugel liegt rechts und oberhalb der ursprünglichen Kugel.
- Zweiter Wurf: Kugel liegt links und oberhalb der ursprünglichen Kugel.
- Dritter Wurf: Kugel liegt links und unterhalb der ursprünglichen Kugel.
- Vierter Wurf: Kugel liegt links und oberhalb der ursprünglichen Kugel.
- Fünfter Wurf: Kugel liegt links und unterhalb der ursprünglichen Kugel.

Die Lage der ursprünglichen Kugel ist unbekannt. Doch bieten die fünf anschließenden Kugelwürfe Informationen über die relative Lage der ursprünglichen Kugel. Was kann man also daraus für die Position der ursprünglichen Kugel schließen, wenn man bayesianisch argumentiert? Dazu sei angenommen, dass die ursprüngliche Kugel a-priori genau in der Mitte des Tisches liegt, dass eine Kugel mit gleicher Wahrscheinlichkeit an jeder Stelle des quadratischen Tisches aufkommen kann und die Kugeln jeweils tatsächlich an einem festen Platz landen, nach dem Wurf also nicht mehr weiterrollen.

Die a-priori-Annahme, dass die ursprüngliche Kugel genau in der Mitte des Tisches gelandet ist, kann mit

Hilfe der Informationen aus den anschließenden fünf Kugelwürfen aktualisiert werden: Da die anschließend geworfene Kugel bei vier der fünf Würfe links der ursprünglichen Kugel aufgekommen ist, ist es wahrscheinlicher, dass die ursprüngliche Kugel in der rechten Hälfte des Tisches positioniert ist.

Stellt man sich ein Raster vor, das den Tisch von links nach rechts in sechs gleich große Abschnitte einteilt, würde man deshalb vermuten, dass die Kugel im zweiten Abschnitt von rechts liegt. Außerdem lässt die Beobachtung, dass die Kugel dreimal oberhalb und zweimal unterhalb der ursprünglichen Kugel gelandet ist, darauf schließen, dass diese eher in der unteren Tischhälfte liegt. Betrachtet man wiederum eine Rastereinteilung des Tisches (in sechs horizontale Streifen), würde man die ursprüngliche Kugel deshalb im dritten Abschnitt von unten einordnen. Abb. 1 veranschaulicht die Aktualisierung der a-priori-Vermutung durch neue Informationen bei jedem weiteren Kugelwurf. Übersetzt man Bayes' Ideen in die heutige übliche Formulierung, so entspricht die Wahrscheinlichkeit der ursprünglichen bzw. ersten Vermutung über eine Ursache der sog. a-priori-Wahrscheinlichkeit (engl. prior) $pr_{prior}(C_i)$. Die Wahrscheinlichkeit $pr(E|C_i)$ ermöglicht die Aktualisierung der ursprünglichen Hypothese, da sie die Wahrscheinlichkeit einer tatsächlichen Folge E unter der Annahme, ihr läge eine bestimmte Ursache C_i zu Grunde, beschreibt. Im Englischen wird hierfür der Begriff Likelihood verwendet. Die Wahrscheinlichkeit der damit überarbeiteten Vermutung $pr(C_i|E)$ wird heute mit a-posteriori-Wahrscheinlichkeit (engl. posterior) bezeichnet. Daraus lässt sich die Formel

$$pr_{prior}(C_i) \cdot pr(E|C_i) \sim pr(C_i|E) \tag{5}$$

folgern (~ wird dabei als „proportional zu" gelesen; in der englischsprachigen Literatur wird meist $\propto$ hierfür verwendet). Hierbei muss betont werden, dass in Bayes' Arbeit niemals eine Formel wie (5) erscheint. Stattdessen verwendete er eine geometrische Betrachtung in der damals üblichen Schreibweise nach Newton und formulierte obige Formel nie explizit.[3]

Bayes war sich wohl auch der Bedeutung dieser heute berühmten Aussage nie bewusst, denn er veröffentlichte seinen Aufsatz hierzu nicht, noch erwähnte er darin irgendeine Anwendungsmöglichkeit des Satzes. Dass der Ursprung des Theorems heute bekannt ist, ist Richard Price (1723–1791) zu verdanken, der als Freund Bayes' nach dessen Tod dessen Aufzeichnungen fand, deren Potential erkannte und sie, ergänzt durch eigene Anmerkungen, über die Royal Society veröffentlichte. In „An Essay Towards Solving a Problem in the Doctrine of Chances" verlieh Price Bayes' Idee zudem eine religiöse Rechtfertigung, nämlich „to confirm the argument taken from final causes for the existence of the Deity" (Bayes und Price 1763, S. 374).

4 Der Verdienst von Pierre Simon Laplace (1749–1827)

Obwohl er heute kaum explizit mit dem Satz von Bayes in Verbindung gebracht wird, war es der französische Mathematiker Pierre Simon Laplace (1749–1827), der dem Theorem zu seinem eigentlichen Durchbruch verhalf.[4] Laplace, der heute als einer der berühmtesten Wahrscheinlichkeitstheoretiker gilt, war schon in jungen Jahren professioneller Mathematiker und wurde so bereits 1773 im Alter von 24 Jahren in die Académie des Sciences in Paris aufgenommen (Hahn 2005, S. 41). Als er 1774 sei-

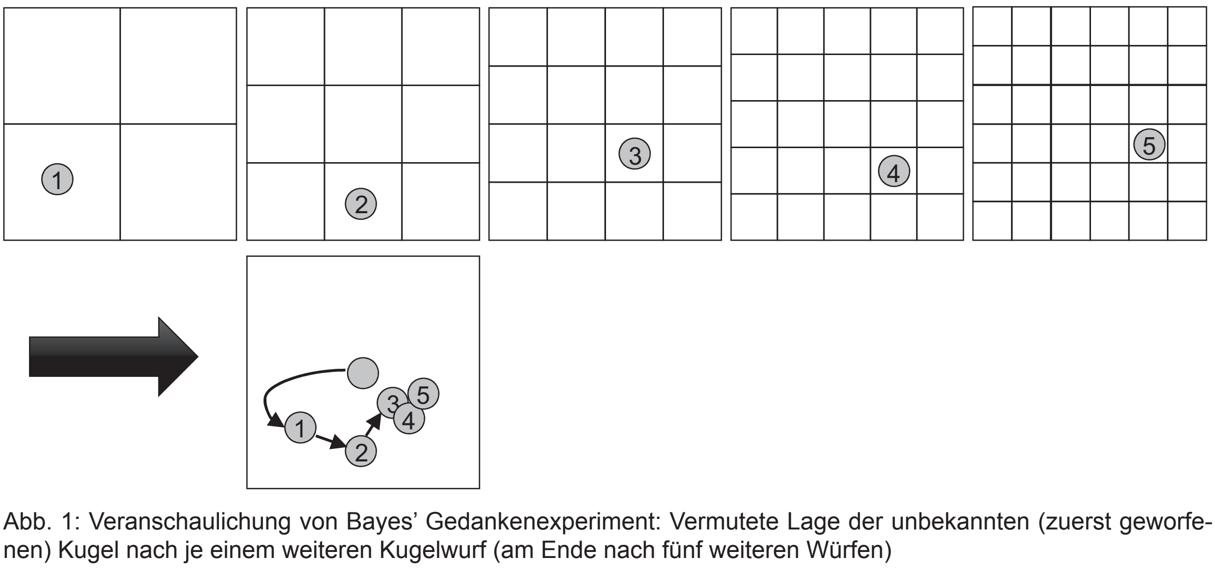

Abb. 1: Veranschaulichung von Bayes' Gedankenexperiment: Vermutete Lage der unbekannten (zuerst geworfenen) Kugel nach je einem weiteren Kugelwurf (am Ende nach fünf weiteren Würfen)

ne „Mémoire sur la Probabilité des Causes par les Événements“ veröffentlichte, wusste er höchstwahrscheinlich nichts von der Arbeit des Engländers Thomas Bayes’ (Gillispie et al. 1997, S. 16). Er schrieb darin das Prinzip nieder, das als „erster Versuch“ der heutigen Regel von Bayes angesehen werden kann:

> „Si un événement peut être produit par un nombre *n* de causes différentes, les probabilités de l’existence de ces causes prises de l’événement sont entre elles comme les probabilités de l’événement prises de ces causes, et la probabilité de l’existence de chacune d’elles est égale à la probabilité de l’événement prise de cette cause, divisée par la somme de toutes les probabilités de l’événement prises de chacune de ces causes“ (Laplace 1774, S. 29).

Dies kann (sinngemäß) wie folgt übersetzt werden:

> „Wenn ein Ereignis *n* verschiedene Ursachen haben kann, verhalten sich die Wahrscheinlichkeiten der Ursachen unter der Voraussetzung des Ereignisses untereinander wie die Wahrscheinlichkeiten des Ereignisses unter Voraussetzung dieser Ursachen, und die Wahrscheinlichkeit jeder von ihnen ist gleich der Wahrscheinlichkeit des Ereignisses unter Voraussetzung der Ursache geteilt durch die Summe aller Wahrscheinlichkeiten des Ereignisses unter Voraussetzung jeder dieser Ursachen“

In heutiger mathematischer Formulierung ergibt das genau folgende Aussage:

$$pr(C_i|E) = \frac{pr(E|C_i)}{\sum_{r=1}^{n} pr(E|C_r)} \tag{6}$$

Im Unterschied zur heute bekannten (allgemeineren) Formulierung unterstellte Laplace hier (implizit) die Annahme, dass jede der *n* Ursachen gleich wahrscheinlich ist (andernfalls ist (6) keine wahre Aussage). Außerdem findet sich an dieser Stelle noch kein Beweis dieser Aussage. Bemerkenswert ist weiterhin, dass es Laplace war, der hier als erster den Begriff der Wahrscheinlichkeit von Ursachen (franz. probabilité des causes) verwendete – im Aufsatz von Bayes findet man diesen Begriff nicht. In Bayes’ Arbeit wird die Idee der Wahrscheinlichkeit von Ursachen auch hauptsächlich durch den von Price ergänzten Anhang deutlich (Bayes und Price 1763, S. 405 f.; Dale 1982, S. 29), wo dieser kurze Anwendungsbeispiele des Satzes ergänzt. Doch gerade dieser Begriff der Wahrscheinlichkeit von Ursachen war es, der in späteren Zeiten in viel Zwiespalt mündete (vgl. Abschnitt 5).

Ein erster Beweisversuch von (6) befindet sich in der „Mémoire sur les Probabilités“ (Laplace 1778–1781, S. 415–417), der vollständige Beweis folgt in der späteren „Mémoire sur les Approximations des Formules qui sont Fonctions de Très Grands Nombres (Suite)“ (Laplace 1783–1786, S. 300 f.). Zu dieser Zeit hatte Laplace wohl auch bereits Bayes’ Arbeit kennengelernt, da Bayes in einem Vorwort zur „Mémoire sur les Probabilités“ von Condorcet erwähnt wird (Gillispie et al. 1997, S. 16, 78), und Laplace in sein Werk offensichtlich die ursprünglichen bayesianischen Ideen integrierte (Gillispie et al. 1997, S. 72). Möglicherweise war der Aufenthalt von Richard Price in Paris im Jahre 1781 die Ursache der Verbreitung von Bayes’ Arbeit unter den französischen Wissenschaftlern (McGrayne 2011, S. 23). Laplace’ Beweis des Bayes-Theorems unter Annahme gleicher Wahrscheinlichkeiten für alle möglichen Ursachen, ist wie folgt strukturiert: Ausgehend davon, dass sich die a-posteriori-Wahrscheinlichkeit $pr(C_i|E)$ gemäß der bereits bewiesenen (grundlegenden) Überlegungen der Wahrscheinlichkeitstheorie berechnen lässt als

$$pr(C_i|E) = \frac{pr(C_i \cap E)}{pr(E)}, \tag{7}$$

können Zähler und Nenner des Bruches auf der rechten Seite von Gleichung (7) spezifiziert werden. Dazu nimmt Laplace – wie bereits erwähnt – an, dass die a-priori-Wahrscheinlichkeit jeder der möglichen *n* Ursachen $1/n$ ist (Laplace 1783–1786, S. 301). Deshalb kann $pr(C_i \cap E)$ durch

$$pr(C_i \cap E) = \tfrac{1}{n} pr(E|C_i) \tag{8}$$

und $pr(E)$ durch

$$pr(E) = \sum_{r=1}^{n} \tfrac{1}{n} pr(E|C_r) \tag{9}$$

ausgedrückt werden. Werden (8) und (9) in (7) eingesetzt, folgt unmittelbar (6). (Laplace 1783–1786, S. 300 f.).

Nach der Französischen Revolution (1789/99), während der die (Natur-)Wissenschaften sehr große Anerkennung erlangten, publizierte Laplace 1812 sein bekanntes Werk „Théorie Analytique des Probabilités“, das eine Zusammenfassung seiner Hauptbeiträge zur Wahrscheinlichkeitstheorie darstellt. Auch der Satz von Bayes in der Fassung von Laplace (1774) bzw. Laplace (1783–1786) ist darin enthalten (Laplace 1812, S. 177), jedoch nicht der Beweis des Allgemeinfalls. Erst 1814 in seinem „Essai Philosophique sur les Probabilités“ verallgemeinerte er ihn zu der uns heute bekannten Form, bei der die Ursachen a-priori nicht gleichwahrscheinlich sein müssen:

> „si ces diverses causes considérées à priori, sont inégalement probables; il faut au lieu de la probabilité de l’événement, résultante de chaque cause, employer le produit de cette probabilité, par celle de la cause elle-même“ (Laplace 1814, S. 18)

Sein wiederum nur in Worten formuliertes sechstes Prinzip lässt sich daher wie folgt übersetzen und darstellen:

> „wenn diese verschiedenen a-priori betrachteten Ursachen nicht gleichwahrscheinlich sind, muss man anstelle der Wahrscheinlichkeit des Ereignisses, die sich bei jeder Ursache ergibt, das Produkt dieser Wahrscheinlichkeit mit der Wahrscheinlichkeit der Ursache selbst verwenden“

$$pr(C_i|E) = \frac{pr(E|C_i)pr_{prior}(C_i)}{\sum_{r=1}^{n} pr(E|C_r)pr_{prior}(C_r)} \tag{10}$$

Das ist das uns heute als „Satz von Bayes“ bekannte Theorem (vgl. (4)). Im Zentrum steht die Berechnung der a-posteriori Wahrscheinlichkeit $pr(C_i|E)$, d. h. der Wahrscheinlichkeit einer bestimmten Ursache C_i, nachdem eine tatsächlich eingetretene Folge E bekannt ist. Zu deren Berechnung benötigt man eine erste Einschätzung der Wahrscheinlichkeit von C_i, die sog. a-priori-Wahrscheinlichkeit $pr_{prior}(C_i)$, die Wahrscheinlichkeit des Eintretens von Folge E, sollte C_i tatsächlich Ursache davon sein ($pr(E|C_i)$), und die Gesamtwahrscheinlichkeit einer derartigen Folge E ($pr(E)$), die sich als Summe der Eintrittswahrscheinlichkeiten von E unter allen möglichen, jeweils verschiedenen Ursachen C_r ($r = 1, \ldots, n$) ergibt

$$pr(E) = \sum_{r=1}^{n} pr(E|C_r)pr_{prior}(C_r).$$

Laplace formulierte die Regel nicht nur in moderner Form und erbrachte deren Beweis, er zeigte auch zahlreiche Anwendungsbeispiele dafür auf, die über die Naturwissenschaften hinaus v. a. in den Bereich der Sozialwissenschaften (Demographie und Justizwesen) reichten. Durch Laplace wurde die eigentliche Bedeutung und Nützlichkeit des Satzes von Bayes erschlossen. Dementsprechend wäre es wohl auch verdient gewesen, Laplace als Namensgeber des Satzes zu würdigen. Allerdings gab dieser öffentlich in seinem Aufsatz von 1814 zu, dass es Thomas Bayes war, der die Grundidee dazu bereits vor ihm hatte (Laplace 1814, S. 186). Doch selbst daran gibt es Zweifel. So schreibt Stigler (1983) von einem Buch mit dem Titel „Observations on Man“ von David Hartley aus dem Jahr 1749, in dem ein „erfinderischer Freund“ erwähnt wird, der die Lösung des „inversen Problems“ gefunden habe. Stiglers Suche dieses nicht beim Namen genannten Freundes ergibt, dass neben Bayes auch Nicholas Saunderson dafür in Frage kommen und damit auch der Urheber des Satzes von Bayes sein könnte.

5 Das bayesianische Konzept in der Kritik

In den Jahrhunderten nach Laplace wurde die bayesianische Idee zu einem Streitpunkt, an dem sich die Geister vieler Wissenschaftler schieden.[5] Das liegt v. a. daran, dass bald zwei Ansätze unterschieden wurden, die – obwohl von Laplace ursprünglich als vereinbar und äquivalent beschrieben (McGrayne 2011, S. 36) – als streng gegensätzlich interpretiert wurden: Dem bayesianischen Konzept steht das „frequentistische“ Konzept (engl. frequentist, abgeleitet von „frequency“ (Häufigkeit)) gegenüber. Gemäß dem frequentistischen Ansatz werden Wahrscheinlichkeiten grundsätzlich aus der Häufigkeit, mit der ein bestimmtes Ereignis in einer großen Grundgesamtheit auftritt, abgeleitet. Eine Wahrscheinlichkeit $pr(E|C_i)$ wird demnach bestimmt, indem eine Stichprobe betrachtet wird und mithilfe der asymptotischen Theorie (in deren Mittelpunkt Theoreme wie der zentrale Grenzwertsatz und das Gesetz der großen Zahlen stehen) von der Häufigkeit des Auftretens des Ereignisses E unter der Bedingung C_i in der Stichprobe die allgemeine Wahrscheinlichkeit gefolgert wird. Im frequentistischen Ansatz werden bedingte Wahrscheinlichkeiten stets auf diese Art betrachtet, während der bayesianschen Sichtweise der Wahrscheinlichkeiten von Ursachen $pr(C_i|E)$ viel Skepsis entgegen gebracht wird. Frequentisten stoßen sich v. a. an der subjektiven a-priori-Vermutung, die Bayesianer verwenden und die der für gewöhnlich objektiven mathematischen Herangehensweise zu widersprechen scheint. So kann man sich beispielsweise im oben beschriebenen fiktiven Beispiel zu Bayes' Gedankenexperiment fragen, warum man ohne weitere Überlegung zunächst davon ausgeht, dass die Kugel genau in der Mitte des Tisches gelandet ist. Wäre es nicht genauso gut möglich, anzunehmen, dass sie z. B. am rechten Tischrand liegt und würde das die folgende bayesianische Argumentation nicht entscheidend beeinflussen? Bayesianer entgegnen dem, dass zum Einen a-priori-Wahrscheinlichkeiten häufig durch relative Häufigkeiten (d. h. frequentistisch) gewählt werden können (vgl. Schlussbeispiel). Zum Anderen bietet die Wahl einer beliebigen a-priori-Wahrscheinlichkeit (für den Fall, dass keinerlei begründete Vermutung möglich ist) einen Startpunkt zur Bestimmung der Wahrscheinlichkeit einer Ursache, was rein frequentistisch schlicht unmöglich wäre. Zudem mindert eine genügend umfangreiche Aktualisierung der a-priori-Wahrscheinlichkeit – zur besseren Vorstellung sei wieder auf das Kugelbeispiel verwiesen – mit frequentistischen Mitteln (zur Bestimmung von Wahrscheinlichkeiten der Art $pr(E|C_i)$) die Gewichtung der anfänglichen Vermutung in der zu bestimmenden a-posteriori-Wahrscheinlichkeit. Hier wird der von Laplace postulierte ergänzende Charakter der beiden so häufig als gegensätzlich beschriebenen Ansätze deutlich.

Die Ökonometrie (empirische Wirtschaftsforschung) stellt ein typisches Beispiel für eine ganze Wissenschaftsdisziplin dar, in der die Trennung zwischen klassischer (d. h. frequentistischer) und bayesianischer Ökonometrie sehr ausgeprägt ist. Während die klassische Regressionsanalyse vielen geläufig ist, bedarf das bayesianische Konzept noch zunehmender Akzeptanz. Qin (1996) gibt eine knappe Übersicht über die Entwicklung der bayesianischen Ökonometrie, aus der hervorgeht, dass das Potenzial dieses Konzeptes erst in den 1950ern entdeckt wurde (vgl. z. B. Marschak 1954), in den 1960ern in erste Versuche umgesetzt wurde (vgl. z. B. Raiffa und Schlaifer 1961), bevor 1971 das erste Lehrbuch zur bayesianischen Ökonometrie von A. Zellner herausgegeben wurde („An Introduction to Bayesian Inference in Econometrics“). Die grundlegendste Anwendung des Satzes von Bayes in der Ökonometrie besteht darin, die Parameter eines Regressionsmodells zu schätzen, doch kann die bayesianische Idee auch in komplexeren Zusammenhängen, wie z. B. der bayesianischen Modell-Mittelung, eingesetzt werden (vgl. Wölfel 2014 für eine knappe Schilderung dieses Konzepts). Auch hier können frequentistische und bayesianische Ökonometrie stets als gegenseitige Ergänzung betrachtet werden (anstelle der üblichen gegenpoligen Sichtweise).

6 Schlussbeispiel und Fazit

Um die Idee des Satzes von Bayes vor dem historischen Hintergrund schultauglich zu veranschaulichen, sei mit einer typischen Fragestellung abgeschlossen:

> Von einem bestimmten Drogentest ist bekannt, dass er im Durchschnitt bei 95 von 100 Personen, die tatsächlich Drogen konsumiert haben, ein richtiges (d. h. positives) Testergebnis liefert. Wendet man ihn hingegen auf Personen an, die nicht unter Drogeneinfluss stehen, zeigt er mit 99-prozentiger Wahrscheinlichkeit ein richtiges (d. h. negatives) Testergebnis. In Deutschland konsumiert jeder hundertste Jugendliche regelmäßig Drogen. Wie hoch ist die Wahrscheinlichkeit, dass ein Jugendlicher, der positiv auf Drogen getestet wurde, tatsächlich unter Drogeneinfluss steht?

In diesem Szenario lassen sich zwei (sich gegenseitig ausschließende) Ursachen identifizieren: „der Jugendliche hat Drogen genommen“ (C_1) und „der Jugendliche hat keine Drogen genommen“ (C_2). Die Folge E, die hier betrachtet wird, ist ein positives Ergebnis des Drogentests. Um den Satz von Bayes (Gleichung (10)) anzuwenden, müssen die a-priori-Wahrscheinlichkeiten der Ursachen und die Likelihoods spezifiziert werden. Für die a-priori-Wahrscheinlichkeiten der Ursachen gilt $pr_{prior}(C_1) = 0{,}01$ und $pr_{prior}(C_2) = 0{,}99$; für die Likelihoods gilt $pr(E|C_1) = 0{,}95$ und $pr(E|C_2) = 1 - 0{,}99 = 0{,}01$. Für die gesuchte Wahrscheinlichkeit $pr(C_1|E)$ ergibt sich deshalb gemäß (10):

$$pr(C_1|E) = \frac{pr(E|C_1)pr_{prior}(C_1)}{pr(E|C_1)pr_{prior}(C_1) + pr(E|C_2)pr_{prior}(C_2)}$$

$$= \frac{0{,}95 \cdot 0{,}01}{0{,}95 \cdot 0{,}01 + 0{,}01 \cdot 0{,}99} \approx 0{,}4897$$

Das bedeutet, dass ein positiv auf Drogen getesteter Jugendlicher nur mit einer Wahrscheinlichkeit von ca. 49 Prozent tatsächlich unter Drogeneinfluss steht – ein im ersten Moment wohl verblüffendes Ergebnis, wenn man von der zunächst hoch erscheinenden Zuverlässigkeit des Drogentests (95 bzw. 99 Prozent) ausgeht.

Das Bayes-Theorem, dessen Ursprung in der Arbeit von Thomas Bayes zu liegen scheint, dessen moderne Form mit vielerlei Anwendungsbeispielen jedoch auf Pierre Simon Laplace zurückgeht, löst so eine Problemstellung, die in verschiedensten Bereichen und Disziplinen auftritt: Wie kann man von Beobachtungen in der Realität auf deren Ursachen schließen (wenn verschiedene Ursachen als möglich erachtet werden)? Die Betrachtung der Wurzeln des Theorems zeigt die intuitive Lösungsmethode dieses „inversen Problems“: Ausgehend von einer anfänglichen a-priori-Vermutung wird diese anhand immer neuer Beobachtungen des Eintretens bestimmter Ereignisse unter bestimmten Bedingungen aktualisiert und so schließlich eine a-posteriori Wahrscheinlichkeit ermittelt. Diese ursprüngliche Idee bleibt bei der rein formalen Anwendung der Formel verborgen, ermöglicht aber einen interessanten Zugang zur bayesianischen Perspektive, die eine wichtige Ergänzung des frequentistischen Ansatzes darstellt.

Anmerkungen

1 Für genauere Erläuterungen zu dieser Veröffentlichung vgl. Dale 2003, S. 182–257.

2 Für eine genauere Analyse des Lebens von Thomas Bayes vgl. Dale 2003, S. 37–100.

3 Bayes selbst formulierte diese Regel als „Proposition 10“ in seinem Aufsatz, vgl. Bayes und Price 1763, S. 394; Dale 1982, S. 27. Wiederum sei für eine ausführliche Untersuchung des Aufsatzes von Bayes und Price auf Dale 2003, S. 258–369, verwiesen.

4 Für eine ausführliche Dokumentation über das Leben Laplace’, seiner Ansichten und seiner Rolle in der Gesellschaft vgl. Hahn 2005.

5 Für eine übersichtliche Diskussion hierzu vgl. McGrayne 2011, S. 34–58.

Literatur

Bayes, Thomas; Price, Richard (1763): An Essay Towards Solving a Problem in the Doctrine of Chances. In: *Philosophical Transaction of the Royal Society of London* (H.53), S. 370–518.

Dale, Andrew I. (1982): Bayes or Laplace? An Examination of the Origin and Early Applications of Bayes’ Theorem. In: *Archive for History of Exact Sciences* 27(1), S. 23–47.

Dale, Andrew I. (2003): Most Honourable Remembrance – The Life and Work of Thomas Bayes. New York (u. a.): Springer.

Gillispie, Charles Coulston; Fox, Robert; Grattan-Guinness, Ivor (1997): Pierre-Simon Laplace – 1749–1827 – A Life in Exact Science. Princeton, NJ: Princeton University Press.

Hahn, Roger (2005): Pierre-Simon Laplace – 1749–1827 – A Determined Scientist. Cambridge, MA (u. a.): Harvard University Press.

Koop, Gary (2003): Bayesian Econometrics. Chichester: Wiley.

Laplace, Pierre Simon (1774): Mémoire sur la Probabilité des Causes par les Événements. In: Œuvres Complètes de Laplace (1878–1912). Bd. 8. Paris: Gauthier-Villars et Fils, S. 27–65.

Laplace, Pierre Simon (1778–1781): Mémoire sur les Probabilités. In: Œuvres Complètes de Laplace (1878–1912). Bd. 9. Paris: Gauthier-Villars et Fils, S. 383–485.

Laplace, Pierre Simon (1783–1786): Mémoire sur les Approximations des Formules qui sont Fonctions de Très Grands Nombres (Suite). In: Œuvres Complètes de Laplace (1878–1912). Bd. 10. Paris: Gauthier-Villars et Fils, S. 296–338.

Laplace, Pierre Simon (1812): Théorie Analytique des Probabilité. Paris: Courcier.

Laplace, Pierre Simon (1814): Essai Philosophique sur les Probabilités. Paris: Courcier.

Marschak, Jacob (1954): Probability in the Social Sciences. *Cowles Foundation Paper* (H.82), S. 166–215.

McGrayne, Sharon Bertsch (2011): The Theory that Would Not Die – How Bayes’ Rule Cracked the Enigma Code, Hunted Down Russian Submarines & Emerged Triumphant from Two Centuries of Controversy. New Haven/London: Yale University Press.

Moivre, Abraham de (1756): The Doctrine of Chances – Or, a Method of Calculating the Probabilities of Events in Play. 3. Aufl., London: A. Millar.

Qin, Duo (1996): Bayesian Econometrics: The First Twenty Years. In: *Econometric Theory* 12(3), S. 500–516.

Raiffa, Howard; Schlaifer, Robert (1961): Applied Statistical Decision Theory. Boston: Division of Research, Graduate School of Business Administration, Harvard University.

Stigler, Stephen M. (1983): Who Discovered Bayes’s Theorem? In: *The American Statistician* 37(4), S. 290–296.

Stone, James V. (2013): Bayes’ Rule: A Tutorial Introduction to Bayesian Analysis. S. l.: Sebtel Press.

Wölfel, Katrin (2014): Bayesianische Modell-Mittelung – Eine Möglichkeit zur Modellierung von Modellunsicherheit. In: *Wirtschaftswissenschaftliches Studium* 43(3), S. 159–161.

Zellner, Arnold (1971): An Introduction to Bayesian Inference in Econometrics. Malabar, FL: Krieger.

Anschrift der Verfasserin

Katrin Wölfel
Institut für Wirtschaftswissenschaft
(Institute of Economics)
Friedrich-Alexander-Universität Erlangen-Nürnberg
Kochstr. 4 (17)
91054 Erlangen
katrin.woelfel@fau.de