

## Richter und Astronomen<sup>1</sup>

---

Die Wahrscheinlichkeitsrechnung – als Verfahren, mit dem die Rationalität von Entscheidungen in Situationen der Ungewißheit begründet werden soll – erblickte in einem exakt eingrenzbaeren Zeitabschnitt zwischen 1650 und 1660 das Licht der Welt. Ian Hacking (1975, [117]), der das „Auftauchen“ (*emergence*) der Wahrscheinlichkeitsrechnung beschreibt, hebt die anfängliche Dualität dieses Werkzeugs hervor, das gleichzeitig präskriptiv und deskriptiv, epistemisch und frequentistisch war. Dabei vergleicht er die Begriffe *Glaubensgrund* und *Zufall*. Überlegungen zu dieser Dualität finden wir bei Condorcet, der zwischen „Glaubensgrund“<sup>2</sup> und „Fazilität“ unterscheidet, bei Cournot, der von „Chance“ und von „Wahrscheinlichkeit“ spricht, und bei Carnap, der die „induktive Wahrscheinlichkeit“ der „statistischen Wahrscheinlichkeit“ gegenüberstellt.

Die Prädominanz der einen oder der anderen Interpretation wird häufig aus historischer Sicht dargestellt. Demnach scheint der Entscheidungsaspekt im 18. Jahrhundert (der „klassischen Ära der Wahrscheinlichkeitsrechnung“) sehr wichtig gewesen zu sein, vor allem im Zusammenhang mit den Verfahren, die sich aus dem Satz von Bayes ergeben: Bayes schlug vor, in die Berechnungen eine unvollständige Information über frühere Ereignisse einzubeziehen, wobei sein Ziel darin bestand, die Wahrscheinlichkeit statistischer Methoden zu messen. Diese Verfahrensweisen wurden in der Folgezeit im 19. Jahrhundert vom frequentistischen Standpunkt aus angefochten. Die Verfech-

---

<sup>1</sup> Dieses Kapitel zur Genese der wahrscheinlichkeitstheoretischen Argumentation und deren Anwendung in den Naturwissenschaften ist etwas schwierig für Leser mit rein geisteswissenschaftlicher Bildung. Für die Lektüre der hiernach folgenden Kapitel ist jedoch kein vollständiges Verständnis dieses Kapitels erforderlich.

<sup>2</sup> Condorcet stellte durch folgende Bemerkung eine Verbindung zwischen Wahrscheinlichkeit und Glauben her: Je größer die Wahrscheinlichkeit eines Ereignisses ist, desto fundierter ist der Grund unseres Glaubens („motif de croire“) daran, daß das Ereignis eintritt. Dale (1991, [360]) übersetzt den Begriff durch „reason for belief“ ins Englische; im Deutschen findet man auch die Übersetzung „Motiv zu glauben“.

ter dieser Richtung machten einen deutlichen Unterschied zwischen Entscheidungen, die sich auf nicht quantifizierbare Urteile stützen (zum Beispiel die Entscheidungen der Geschworenen eines Schwurgerichts), und Entscheidungen, die auf wiederholten Beobachtungen beruhen, etwa auf den Beobachtungen, die von den entstehenden Einrichtungen für Verwaltungsstatistik geliefert wurden. Den „Frequentisten“ kam das Bayessche Verfahren – das eine kleine Anzahl von Beobachtungen mit einer rein mutmaßlichen „*A-priori*-Wahrscheinlichkeit“ verknüpfte, um hieraus auf eine besser gesicherte „*A-posteriori*-Wahrscheinlichkeit“ zu schließen – wie ein Phantasiegebilde vor. Dagegen erfuhren im 20. Jahrhundert die Überlegungen zur Art und Weise, in der Entscheidungen in Situationen der Ungewißheit getroffen werden, eine Neubelebung durch die Arbeiten von Keynes, De Finetti und Savage. Die Diskussionen über den Bayesianismus und seine Bedeutung erlangten erneut eine erstrangige Wichtigkeit. Man führte Forschungsarbeiten auf dem Gebiet der experimentellen Psychologie durch, um festzustellen, ob das menschliche Gehirn tatsächlich entsprechend derartigen Verfahren arbeitet (Gigerenzer und Murray, 1987, [106]).

Das Hin und Her zwischen dem entscheidungsbezogenen und dem beschreibenden Standpunkt zieht sich durch die gesamte Geschichte der wahrscheinlichkeitstheoretischen Formalisierungen, die gegen 1660 begann.<sup>3</sup> Aber die Fragen, die diese neue Sprache zu beantworten suchte, sind aus sehr viel ältere Debatten über die Möglichkeit hervorgegangen, über den Zufall nachzusinnen – entweder indem man ihm die Entscheidung in schwierigen Fällen anvertraute, oder indem man die Beurteilung einer ungewissen Zukunft in die Gegenwart einbezog. Diese Archäologie der Wahrscheinlichkeitsrechnung wurde von Ernest Coumet (1970, [50]) in einem überaus nützlichen Artikel rekonstruiert, in dem er diesen Kalkül und die dadurch implizierten *Äquivalenzen* als Antwort auf ein Problem der *Billigkeit*<sup>4</sup> oder *Fairness* darstellte.

## Aleatorische Verträge und faire Abmachungen

Der Artikel von Coumet trägt die Überschrift „Ist die Theorie des Zufalls zufällig entstanden?“ Er wies in dieser Arbeit nach, wie die Frage der Gerechtigkeit zu den Äquivalenzkonventionen für *Erwartungswerte* geführt hat – zu einem Begriff also, der zeitlich vor dem Begriff der Wahrscheinlichkeit zu liegen scheint. Die Schwierigkeit rührte daher, daß der Zufall – das heißt das ungewisse Auftreten eines zukünftigen Ereignisses – zur damaligen Zeit als ein hei-

<sup>3</sup> Die mathematische Wahrscheinlichkeitsrechnung, die Statistik und die Philosophie der Wahrscheinlichkeit werden oft als „siamesische Drillinge“ bezeichnet.

<sup>4</sup> In der lateinischen Formulierung durch „*aequitas*“ ausgedrückt. Dieses Wort hat u.a. folgende Bedeutungen: Gleichheit, Rechtsgleichheit und Billigkeit (im moralischen Sinne). Aus dem lateinischen Wort leitet sich das französische *équité* (Billigkeit, Angemessenheit, Gerechtigkeit, Rechtlichkeit) und das englische *equitableness* (Fairness, Billigkeit) ab.

liges Attribut aufgefaßt wurde, das die Menschen niemals beherrschen würden – außer wenn sie Gott zurate ziehen. Die thomistische Theologie erkannte in sehr präzise bestimmten Notfällen die Möglichkeit an, eine Frage durch das Ziehen eines Loses zu entscheiden, und unterschied zwischen *konsultatorischen* Losen (*sortes consultatoriae*), *divinatorischen* Losen (*sortes divinatoriae*) und *divisorischen* Losen (*sortes divisoriae*). Divisorische Lose konnten verwendet werden „um zu entscheiden, wem das betreffende Ding zufällt oder was den betreffenden Personen zugesprochen werden soll – zum Beispiel Besitz, Ehren oder Würden“.<sup>5</sup> In jedem anderen Fall stellte der Rekurs auf „Glücksspiele“ eine „schwerwiegende Sünde“ dar.

Würfel und Lose sind bereits im Altertum nicht nur in Glücksspielen verwendet worden; sie waren auch Mittel zur Divination, zur Erforschung des göttlichen Willens, allenfalls der Zukunft. Das Anrufen einer Gottheit, das sich im Ziehen von Losen ausdrückt, um besonders heikle Streitfälle zu entscheiden, schien demnach eine gerechte Lösung gewesen zu sein: sie stützte sich nämlich auf eine Abmachung, die *über* den prozeßführenden Parteien stand und daher von ihnen akzeptiert werden konnte. Diese Art Lösung war bereits vom hl. Augustin erwähnt worden. Für ihn war das Los kein Übel an sich, sondern zeigte dem zweifelnden Menschen den Willen Gottes an:

Nimm zum Beispiel an, daß du eine überflüssige Sache hast. Du solltest sie jemandem geben, der sie nicht hat. Aber du kannst sie nicht zwei Menschen geben. Erscheinen nun zwei Menschen, von denen keiner über den anderen die Oberhand gewinnt – sei es durch Not, sei es durch die Bande der Freundschaft zu dir – ist es dann nicht die gerechteste Lösung für dich, durch das Los entscheiden zu lassen, welcher der beiden das erhalten soll, was du nicht beiden gleichzeitig geben kannst? (Vgl. hl. Augustin, *De la doctrine chrétienne*<sup>6</sup>, zitiert von Coumet, 1970, [50].)

Das Ziehen von Losen war für einen Richter eine Art und Weise, *das* auszuführen, was für ihn auch heute noch eine unabdingbare Pflicht ist – die Verpflichtung nämlich, ein Urteil zu fällen. Ein Richter kann auf der Grundlage eines im Gesetzbuch stehenden Artikels oder aufgrund seiner inneren Überzeugung entscheiden – das einzige, was er *nicht* tun darf, besteht darin, *keine* Entscheidung zu treffen. Man kann sich ohne weiteres vorstellen, daß das Los – als Ausdruck eines über den Menschen stehenden Willens – lange Zeit hindurch in dramatischen Fällen eine Möglichkeit zu bieten vermochte,

<sup>5</sup> Gemäß der Doktrin der *sortes divisoriae* wurde die Aufteilung eines Gutes unter mehreren Personen durch einen Zufallsmechanismus unter bestimmten Bedingungen als legitim erachtet. Das lateinische Wort *sors* (Plural: *sortes*) hat u.a. folgende Bedeutungen: Losstäbchen, Lostäfelchen, Weissagungstäfelchen; Losen, Verlosung, Los, Orakelspruch; Teil, Anteil; Los = Schicksal, Geschick.

<sup>6</sup> Augustinus, Aurelius (354-430): *De doctrina christiana*. Deutsch von P. Sigisbert Mitterer; erschienen 1925 in: Bibliothek der Kirchenväter, Bd. 49, München, Kösel und Pustet.

diese schwere Last erträglicher zu machen. Die Verpflichtung, *Urteile* sogar in denjenigen Fällen zu verkünden, bei denen ein Zweifel bleibt, kann mit der Verpflichtung des *Kodierens* verglichen werden, das heißt mit der Tätigkeit des Statistikers, wenn er bei einer Erhebung die Antworten (oder „Nichtantworten“) in Äquivalenzklassen einteilt. Auch für den Statistiker sind zweifelhafte Fälle eine Last und die Versuchung kann groß sein, dann auf den Zufall zurückzugreifen, um sich dieser Bürde zu entledigen. Der Rückgriff auf den Zufall wird in den sogenannten *Hot-Deck*-Verfahren formalisiert und systematisiert, bei denen im Falle einer „Nichtantwort“ zufällige Antworten zugeordnet werden. Diese Zuordnung erfolgt entsprechend den Gesetzen für bedingte Wahrscheinlichkeiten, die auf der Grundlage der gegebenen Antworten konstruiert werden. Man kann also die *sortes divisoriae*, den ungebändigten Zufall eines abgespannten Kodierers und die *Hot-Deck*-Verfahren als eine ökonomische Art und Weise betrachten, in problematischen Fällen den Kopf aus der Schlinge zu ziehen.

Außer diesen ziemlich unzweideutigen Fällen, in denen sich ein Richter auf zufällige Entscheidungen als letzten Ausweg verlassen konnte, kam der Zufall auch in verschiedenen anderen Situationen ins Spiel, deren gemeinsames Merkmal darin bestand, daß sie sich auf zukünftige unsichere Gewinne bezogen: Investition von Kapital in den Seehandel, Versicherungen, Leibrenten und Glücksspiele. In allen diesen Fällen wird etwas, dessen man sich in der Gegenwart sicher ist, gegen eine zufällige Bereicherung in der Zukunft eingetauscht. Ist eine derartige „grundlos legitime“ Bereicherung statthaft? Ebenso wie im Falle von Zinsdarlehen oder Wucher diskutierten die Theologen erbittert über die Gerechtigkeit von Vereinbarungen, bei denen Menschen durch zukünftige Ereignisse gebunden werden. So stellte etwa der hl. Franz von Sales Gewinne infrage, die nicht auf die Bemühungen der Vertragspartner zurückzuführen sind:

Wir haben uns also geeinigt, werden Sie mir sagen? Das taugt, um zu zeigen, daß derjenige, der gewinnt, keinem Dritten schadet. Aber daraus folgt nicht, daß auch die Abmachung selbst nicht ebenso unvernünftig ist wie der Einsatz. Denn der Gewinn, der ein Preis für den Fleiß sein sollte, wird ein Preis des Zufalls, der jedoch keinen Preis verdient, da er ganz und gar nicht von uns selbst abhängt. (hl. Franz von Sales, *Introduction à la vie dévote*, 1628, zitiert von Coumet, 1970, [50].)

Während sich diese Mißbilligung vor allem auf Glücksspiele bezog, konnte das für den Seehandel erforderliche Risikokapital dagegen – und zwar gerade aufgrund des Risikos – zu einem Gewinn führen. Für die mittelalterliche Theologie war dieser Umstand sogar eine mögliche Rechtfertigung für Zinsdarlehen, die im Allgemeinen verboten waren und mit Wucher gleichgesetzt wurden (Le Goff, 1962, [175]). Im Übrigen wurden Versicherungsverträge als zulässig angesehen.

Auf diese Weise entstand ein ganzer Arbeitskomplex, der sich auf unterschiedliche Bereiche bezog (*sortes divisoriae*, risikobehaftete Darlehen, Versicherungsverträge, Glücksspiele), aber dazu tendierte, den Begriff einer *fairen Abmachung* zu formalisieren. So konnte man etwa die *sortes divisoriae* aus dem Bereich der Heiligen herausnehmen und als Grundlage für derartige Abmachungen rechtfertigen. Ebenso erwies es sich im Interesse einer soliden Rechtsgrundlage für Verträge, die zukunftsbezogene Risiken enthielten, oder im Interesse der Rechtfertigung eines Glücksspiels als notwendig, daß die Vertragspartner oder die Spieler in den Genuß „gleicher Voraussetzungen“ kamen. Diese Forderung nach Gleichheit, die sich vom Gerechtigkeitsgedanken leiten ließ, eröffnete den Weg zur Konstruktion eines gemeinsamen begrifflichen Rahmens für Aktivitäten, die im Übrigen vollkommen unterschiedlich waren: Würfelspiele, Lebensversicherungen und Gewinne, die man von ungewissen Handelsgeschäften erwartete. Die Glücksspiele sind nur *ein* Beispiel für *aleatorische Verträge*: derartige Verträge „beruhten auf freiwilligen Abmachungen, denen zufolge der Erwerb eines Vermögens eine ungewisse Glückssache war; um legitim zu sein, mußten diese Abmachungen bestimmten Voraussetzungen der Fairness<sup>7</sup> genügen“ (Coumet, 1970, [50]). Das Problem der Fairness tauchte in der Praxis im Zusammenhang mit *Aufteilungen* auf, wenn ein Geschäft oder ein Spiel unterbrochen werden mußte: Wie sind die Gewinne oder die Einsätze aufzuteilen? Hierbei handelt es sich um einen Spezialfall des sogenannten *Teilungsproblems*. Eine ähnliche Frage, die der Chevalier de Méré an Pascal gerichtet hatte, war die Grundlage der Pascalschen Wahrscheinlichkeitsrechnung. Aber diese scheinbar harmlose Frage zur belanglosen Tätigkeit eines Spielers hing tatsächlich mit den alten Debatten über das Problem der Gerechtigkeit von Abmachungen zusammen, auf denen aleatorische Verträge beruhten.

Bei der Einführung seiner neuen Formalisierung ließ sich Pascal natürlich Ausdrücke aus der Sprache der Juristen, denn das war die Sprache, in der die zeitgenössischen Debatten geführt wurden. Aber Pascal schuf auch eine neue Art und Weise, die Rolle eines *über* den besonderen Interessen stehenden Schiedsrichters zu spielen – eine Rolle, die früher den Theologen zugekommen war:

... Die Moralisten, welche die Voraussetzungen zu bestimmen versuchten, die ein Spiel erfüllen muß, um fair zu sein, nahmen eine Position ein, die über den Begehrlichkeiten und Antagonismen der Spieler stand; ein Mathematiker, der eine „faire Verteilung“ berechnen will, nimmt einfach nur eine Haltung ein, die noch rigoroser als die Einstel-

<sup>7</sup> Im Zentrum der juristischen Debatte über die Doktrin der *contrats aléatoires* stand die Frage nach der Billigkeit dieser Kontrakte: die Frage, ob die ungewisse Aussicht auf ein Gut gegen ein sicheres Gut, einen Geldbetrag, billigerweise aufgewogen werden kann; es handelt sich also um die Frage, ob und inwieweit dem Träger eines finanziellen Risikos ein bestimmter Betrag, eine Prämie, rechtmäßig zusteht. (Vgl. Hauser, 1997, [397].)

lung der Moralisten ist: er ist der Schiedsrichter. (Vgl. Coumet, 1970, [50].)

Aber worauf konnte sich diese Objektivierung stützen, die es Pascal ermöglichte, die Position des Schiedsrichters zu definieren – zum Beispiel bei seiner Beantwortung der Frage des Chevalier de Méré (bezüglich einer fairen Aufteilung der Anteile im Falle einer Spielunterbrechung)? Die berechneten *Erwartungswerte* (oder *Nützlichkeitswerte*) müssen so beschaffen sein, daß die Alternativen (Spiel beenden oder fortsetzen) für die Spieler belanglos sind:

... die Regelung dessen, was ihnen gehören sollte, muß in angemessener Weise dem entsprechen, was sie berechtigterweise vom Glück erwarten konnten, so daß es für jeden von ihnen vollkommen gleichgültig ist, ob sie *das* nehmen, was man ihnen zuteilt, oder ob sie das Abenteuer des Spiels fortsetzen ... (Pascal, zitiert von Coumet, 1970, [50].)

Man führte also zunächst Äquivalenzen zwischen denjenigen „Erwartungswerten“<sup>8</sup> ein, die sich unter Umständen unter den Spielern austauschen ließen – weil die Werte gleich groß waren –, um dann den Begriff der Wahrscheinlichkeit einzuführen, indem man den Erwartungswert durch die Höhe des Einsatzes dividierte. Diese Erwartungswerte gehörten zum gemeinsamen Wissen der Spieler und stellten einen Punkt dar, an dem die Kontroversen eingestellt werden konnten: Man hatte eine Regel, auf deren Grundlage sich das Spiel unterbrechen ließ, ohne daß sich dadurch einer der Beteiligten verletzt fühlte. Diese Wesensverwandtschaft der Sorgen von Richtern und Vermessungsingenieuren implizierte für beide Berufsgruppen einen gewissen Abstand zum betreffenden Rechtsstreit oder zur betreffenden Entscheidung; sie mußten sich in die Position von Zuschauern, unvoreingenommenen Schiedsrichtern oder in die Lage von Beratern eines Fürsten versetzen.

Tatsächlich verhielt es sich so, daß man diesen Abstand zum Sachverhalt aus unterschiedlichen Gründen wahrte: entweder um zwischen zwei Spielern zu schlichten, falls das Spiel unterbrochen werden mußte, oder um einen Tatmenschen zu beraten, der mit der Wahl zwischen mehreren Entscheidungen konfrontiert war, deren Konsequenzen von ungewissen zukünftigen Ereignissen abhängen. Es handelte sich aber um dieselben Berechnungen von Erwartungswerten oder Nützlichkeitswerten, an denen sich die Urteile von Richtern oder die Ratschläge von Beratern orientierten. Im erstgenannten Fall strebte man ein Fairnessprinzip an, das die Aufstellung von harmonischen zwischenmenschlichen Beziehungen ermöglichte. Im zweiten Fall ging es um die innere Konsistenz der von Tatmenschen getroffenen Entscheidungen und um die Suche nach einer Rationalität, die unterschiedliche Lebensmomente umfaßte. Diese Umstände förderten die Akzeptanz des neuen Kalküls, der objektivierbar war und sich von *einem* Fall auf einen *anderen* übertragen ließ.

<sup>8</sup> Der Begriff „Erwartungswert“ ist hier im eigentlichen Sinne des Wortes zu interpretieren, das heißt als „zu erwartender Wert“.

Wir müssen jedoch erkennen, daß die Konstruktion derartiger Äquivalenzräume für zukünftige, noch nicht eingetretene, inkompatible und heterogene Ereignissen nur mit Schwierigkeiten zugestanden werden konnte. Man spürt das sogar, wenn man sich das Pascalsche Wett-Argument noch einmal durchliest, das auf dieser Art des Vergleichs aufbaute. Wir stellen diese Schwierigkeit auch bei scheinbaren Paradoxa fest, die man heute gut versteht: zum Beispiel, wenn man die schwierige Wahl zwischen Situation *A* (*garantierter* Gewinn von *einer* Million Euro) und Situation *B* (*Chance von Eins zu Zehn* für einen Gewinn von *zwanzig* Millionen Euro) treffen muß. Die Berechnung der Erwartungswerte führt zur Wahl von *B*, aber nur wenige „rational denkende“ Menschen würden eine solche Wahl treffen. Viele würden dagegen *A* bevorzugen, das heißt einen sicheren, wenn auch geringeren Gewinn. Wir wissen jetzt, daß die Berechnung der Erwartungswerte (was in diesem Fall zur Wahl von *B* verleiten würde) nicht wirklich plausibel ist – es sei denn, man nimmt den Standpunkt der Frequentisten ein: Würde diese (fiktive!) Auswahl viele Male (zehnmal? zwanzigmal?) wiederholt werden, dann würde man offensichtlich *B* wählen. Das dürfte auch der Fall sein, wenn es um die *einmalige Wahl kleiner Beträge* ginge (sicherer Gewinn von zehn Euro oder eine Chance von Eins zu Zehn für einen Gewinn von zweihundert Euro).

Anhand dieser verschiedenen Fälle (Schlichtung, Hilfe bei Entscheidungsfindungen) erkennt man, daß die – häufig nicht sehr intuitive – probabilistische Rationalität ihre Berechtigung in einer allgemeinen, alles überschauenden Sichtweise fand: in der Sichtweise des Richters, der über den prozeßführenden Parteien steht, in der Sichtweise des Bankiers oder des Versicherers, der mit zahlreichen Risiken umgeht, oder auch in der Sichtweise eines isolierten Individuums, das mit Mikro-Entscheidungen konfrontiert ist, die kein großes Engagement bedeuten (Wahl zwischen kleinen Geldbeträgen). Es gab also einen engen Zusammenhang zwischen der Fähigkeit, Äquivalenzräume begrifflich zu erfassen und zu konstruieren, und der Möglichkeit, derartige alles überschauende Positionen einzunehmen. Die Spannung zwischen objektiver und subjektiver Wahrscheinlichkeit läßt sich folgendermaßen in eine Standpunktfrage zurückübersetzen: Eine einzige und kontingente Auswahl oder eine allumfassende und verallgemeinerungsfähige Position?

Demnach mußten die ersten „Probabilisten“ – ausgehend vom Standpunkt eines Juristen oder Beraters eines Entscheidungsträgers – große Schwierigkeiten überwinden, als sie eine „Geometrie“ zur Behandlung *disparater Größen* entwickelten und die Menschen davon überzeugen wollten, diese Geometrie zu akzeptieren. Leibniz wies auf die zögerlichen Versuche und Schwierigkeiten hin, gleichzeitig an den möglichen Gewinn und an die Wahrscheinlichkeit zu denken, und beides zu einem einzigen „Nützlichkeitswert“ zusammenzusetzen, ähnlich wie sich die Oberfläche eines Rechtecks durch die Multiplikation von Länge und Breite ergibt:

Da die Größe der Konsequenz und die des Sukzedenten zwei heterogene Betrachtungen sind (oder Betrachtungen, die nicht miteinander

verglichen werden können), waren die Moralisten, die einen Vergleich gewollt hatten, ziemlich verwirrt, wie es im Falle derjenigen Moralisten offensichtlich ist, die sich mit der Wahrscheinlichkeit befaßt hatten. Die Wahrheit ist, daß hier – wie bei anderen disparaten und heterogenen Schätzungen, die mehr als *eine* Dimension umfassen – die betreffende Größe das zusammengesetzte Ergebnis der einen oder anderen Schätzung ist, ähnlich wie bei einem Rechteck, wo es zwei Betrachtungen gibt, nämlich die der Länge und die der Breite; und bezüglich der Größe der Konsequenz und der Grade der Wahrscheinlichkeit fehlt uns immer noch dieser Teil der Logik, die eine Schätzung dieser Größen bewirken muß. (Leibniz, *Nouveaux essais sur l'entendement humain*.<sup>9</sup>)

Das Bestreben, einen Komparabilitätsraum für heterogene Größen zu konstruieren, ergab sich aus den Debatten der Juristen des 17. Jahrhunderts, die darauf bedacht waren, die Fairness der aleatorischen Verträge zu rechtfertigen. Diese Bestrebungen wurden im 18. Jahrhundert fortgeführt und erweitert, um einen homogenen Raum von Graden der Sicherheit aufzustellen, der seinerseits mit den Erfordernissen von Handlungen und Entscheidungsfindungen verknüpft war.

## Konstruktiver Skeptizismus und Überzeugungsgrad

Die Geschichte der Entfaltung des wahrscheinlichkeitstheoretischen Denkens im Rahmen einer Gesamtheit von praktischen Problemen, die eine Ungewißheit implizierten, wurde von Lorraine Daston in einem 1989 [55] erschienenen Artikel zusammengefaßt. Der Artikel setzt die von Coumet begonnene Arbeit fort und stellt eine andere Wurzel des Werkzeugs vor, das es ermöglichte, die verschiedenen Aspekte der Ungewißheit innerhalb ein und desselben begrifflichen Rahmens zu erfassen: die Debatten über Sicherheit und Wissen, die im Ergebnis von Reformation und Gegenreformation geführt wurden.

Diese Debatten, welche die Glaubensgrundlagen betonten (die Offenbarung für die Protestanten und die Tradition für die Katholiken), löste gegenseitige Denunziationen aus, die allmählich die verschiedenen Bestandteile des Glaubens unterhöhlten und zum Skeptizismus führten. Eine extreme Form des Skeptizismus – der von gelehrten Freigeistern vertretene Pyrrhonismus<sup>10</sup> – leugnete sogar die Evidenz von Empfindungen und mathematischen Beweisen. Zur gleichen Zeit versuchten mehrere Autoren auf halbem Wege zwischen

<sup>9</sup> Im Manuskript vollendet um 1705. Deutsche Übersetzung unter dem Titel *Neue Abhandlungen über den menschlichen Verstand*. Neuausgabe als Band 498 der Philosophischen Bibliothek, Felix Meiner Verlag, Hamburg, 1996. (1. Auflage 1873 als Band 56 der Philosophischen Bibliothek).

<sup>10</sup> Pyrrhon, griechischer Philosoph aus Elis, um 360–270 v. Chr., gilt als Begründer des Skeptizismus.



den dogmatischen Fideisten – die sich auf die Gewißheiten des wahren Glaubens stützten – und den äußerst bissigen Skeptikern, eine Vorstellung von *dem* zu definieren, was „einfach wahrscheinlich“ war, eine Vorstellung vom „Überzeugungsgrad, der ausreicht, einen besonnenen Geschäftsmann zum Handeln zu ermuntern ..., wobei die Überzeugung von einer intuitiven Bewertung der möglichen Pläne und der damit zusammenhängenden Risiken abhängt“ (Daston, 1989, [55]).

Diese „konstruktiven Skeptiker“ (um einen von Popkin 1964 [238] geprägten Ausdruck zu verwenden) betrachteten demnach das Handeln als Grundlage des Wissens (und nicht umgekehrt). Sie waren – anders als die Juristen, von denen Pascal seine Inspirationen erhielt – weniger an *Fairness* interessiert als am *rationalen Glauben*, der die Orientierung für eine Entscheidung vorgab. Doch auch sie nutzten die Doktrin der aleatorischen Verträge, um daraus Beispiele zu entnehmen, die zeigten, daß es mitunter vernünftig war, einen gegenwärtigen sicheren Besitz gegen einen unsicheren zukünftigen Besitz einzutauschen. Diese Erkenntnisphilosophie wies der Wahrscheinlichkeitsrechnung eine klare „epistemische“ Rolle zu, denn sie orientierte sich am Aspekt des *unzulänglichen Wissens* und nicht am Aspekt des Zufalls. Die konstruktiven Skeptiker integrierten jedoch Glücksspiele, riskante Tätigkeiten (Handel, Impfungen) und Entscheidungen von Geschworenen zur möglichen Schuld eines Angeklagten in ein und dasselbe Modell und bereiteten dadurch den Übergang von einem Aspekt zum anderen vor.

Es ist interessant, die philosophische Haltung dieser *konstruktiven Skeptiker*, die zwischen den Fideisten und den radikalen Skeptikern stehen, mit der Position zu vergleichen, die ich in der Einleitung vorgeschlagen habe, das heißt mit der Position einer modernen Wissenssoziologie, die sich sowohl vom wissenschaftlichen Objektivismus – für den „Fakten Fakten sind“ – als auch vom Relativismus unterscheidet, für den Objekte und Niederschriften gänzlich von kontingenten Situationen abhängen. Die beiden historischen Konfigurationen des 17. und des 20. Jahrhunderts unterscheiden sich jedoch radikal voneinander, und sei es nur, weil im ersteren Fall der Pol der Gewißheit durch die Religion, im letzteren Fall aber durch die Wissenschaft verkörpert wird. Von diesem Standpunkt ist die probabilistische Verfahrensweise – welche die Ungewißheit durch Quantifizierung objektiviert – Bestandteil eines Säkularisierungsprozesses. Das ist auch der Grund dafür, warum sich heute sowohl religiöse Menschen als auch (religiöse und nichtreligiöse) Wissenschaftler gleichermaßen unwohl fühlen, wenn sie sich mit Pascals Wette befassen: Beide Gruppen spüren, daß sich in diesem berühmten Text zwei Argumentationsweisen überschneiden, die fortan getrennt voneinander verliefen.

Hinter der Distanz, welche die „konstruktiven Skeptiker“ zur nihilistischen Skepsis der Pyrrhoneer hielten (die man heute als radikale Relativisten betrachten würde), stand die Absicht, Objekte zu erzeugen, auf die man sich beim Handeln stützen konnte. Diese Objekte waren „Überzeugungsgrade“, das heißt *probabilisierte Glaubensakte*. Wie Lorraine Daston sagt: „Die Betonung des Handelns als Grundlage des Glaubens – und nicht umgekehrt –

ist der Schlüssel zur Verteidigung gegen den Skeptizismus. Schriftsteller wie Wilkins machten häufig die Bemerkung, daß auch der überzeugteste Skeptiker sein Abendessen so verspeist, *als ob* die Außenwelt wirklich existiert“. (Wilkins führte das Beispiel eines Kaufmanns an, der die Risiken einer langen Reise in der Hoffnung auf einen höheren Gewinn auf sich nimmt, und empfahl, derartige Handlungsregeln auch in wissenschaftlichen und religiösen Fragen zu befolgen). Die wichtigen Wörter in der Fabel vom skeptischen Abendessensgast sind „so, *als ob*“. Dadurch bezieht man sich nicht auf ein Problem von essentieller Realität (wie es ein Fideist oder ein heutiger Realist tun würde), sondern auf das praktische Verhalten, auf eine Handlungslogik.

Die genannten Autoren konstruierten deswegen einen Rahmen, der eine gemeinsame Konzeptualisierung von Gewißheitsformen gestattete, die zuvor voneinander verschieden waren: die *mathematische* Gewißheit eines Beweises, die *physikalische* Gewißheit der sensorischen Evidenz und die *moralische* Gewißheit von Aussage und Vermutung. Sie ordneten die verschiedenen Formen der Gewißheit auf einer Ordinalskala an und meinten, daß die meisten Dinge nur auf der untersten Ebene, das heißt auf der Aussagenebene, gewiß sind. Auf diese Weise erlangte in Frankreich das Wort „probabilité“<sup>11</sup>, das im Mittelalter „eine amtlich beglaubigte Stellungnahme“ bezeichnete, die Bedeutung von „Grad der Zustimmung entsprechend der Evidenz von Dingen und Zeugen“.

Dann fügten Leibniz und Niklaus Bernoulli, die beide Mathematiker und Juristen waren, drei Ebenen der Gewißheit in ein *Kontinuum* ein, das jeden Grad an Zustimmung einschloß<sup>12</sup> – vom Unglauben bis hin zur vollständigen Überzeugung.

Die drei Ebenen der Gewißheit entsprachen drei sehr verschiedenen Weisen der Bewertung von Wahrscheinlichkeiten: (1) *gleiche Möglichkeiten* auf

<sup>11</sup> Das französische Wort geht, ebenso wie das englische „probability“, auf das lateinische „probabilitas“ zurück, das sich vom Adjektiv „probabilis“ (annehmbar, glaublich, wahrscheinlich, tauglich) ableitet. Dieses Wort hängt seinerseits mit dem lateinischen „probare“ (prüfen, billigen, gutheißen, gelten lassen, anerkennen) zusammen. Das deutsche Wort „Wahrscheinlichkeit“ leitet sich von „wahrscheinlich“ (= „mit ziemlicher Sicherheit“) ab und wurde im 17. Jahrhundert vermutlich nach dem Vorbild des gleichbedeutenden niederländischen „waarschijnlijk“ gebildet. Das niederländische Adjektiv ist wohl eine Lehnübertragung des lateinischen „verisimilis“ (wahrscheinlich), einer Zusammensetzung aus „verus“ (wahr) und „similis“ (ähnlich).

<sup>12</sup> Leibniz, 1705, [412]: „... wenn die Natur der Dinge nichts enthält, was für oder gegen ein bestimmtes Faktum spricht, so wird es, sofern es durch das Zeugnis unverdächtiger Leute bestätigt wird (z.B. daß Julius Caesar gelebt hat), mit einem *festen Glauben* aufgenommen. Wenn aber die Zeugnisse dem gewöhnlichen Naturlauf widerstreiten oder untereinander widersprechend sind, so können die Wahrscheinlichkeitsgrade sich bis Unendliche verschieden gestalten und daher stammen alle jene Grade, welche wir *Glauben, Vermutung, Ungewißheit, Mißtrauen* nennen; und daher ist denn strenge Prüfung nötig, um ein richtiges Urteil zu bilden und unsere Zustimmung gemäß den Graden der Wahrscheinlichkeit zu erteilen.“

der Grundlage physikalischer Symmetrie (nur für Glücksspiele geeignet); (2) *Beobachtungshäufigkeiten* von Ereignissen (die eine Sammlung von Statistiken voraussetzten, welche ihrerseits – wie Graunts Sterbetafeln aus dem Jahr 1662 – eine hinreichende zeitliche Stabilität aufwiesen); und schließlich (3) *Grade subjektiver Gewißheit* oder Glaubensgrade (zum Beispiel juristische Praktiken bei der Gewichtung von Indizien und Mutmaßungen). Dieses Konstrukt ist deswegen so überraschend, da wir hier bereits eine Zusammenfassung dessen finden, was in der Folgezeit erneut in Bezug auf die Unterscheidung zwischen den sogenannten *objektiven* Wahrscheinlichkeiten und den *subjektiven* Wahrscheinlichkeiten zusammengetragen wurde: die ersten hängen mit den *Zuständen der Welt* (hier: den ersten beiden Ebenen) zusammen, die letzteren entsprechen dagegen den *Zuständen des Verstandes* (dritte Ebene). Aber diese Unterscheidung gilt nur dann, wenn man die Genese des Problems der Überzeugung außer Acht läßt: Wie gelangt man vom Glauben (dritte Ebene) zu den objektiven Gewißheiten der ersten Ebene und der zweiten Ebene? Dieses Problem, das ein wenig an die von den Verfechtern des „starken Programms“ in der Wissenschaftssoziologie gestellte Frage erinnert, ließ den Philosophen des Jahrhunderts der Aufklärung, den „Trägern des Lichts“, keine Ruhe.

So bekundeten etwa Locke, Hartley und Hume offen eine Assoziations-theorie, bei der sie den Verstand mit einer Addiermaschine verglichen, welche die Häufigkeiten vergangener Ereignisse zusammenzählt und deren mögliche Wiederholung berechnet. Hartley stellte sich vor, daß wiederholt auftretende Empfindungen „Rillen ins Gehirn“ graben und sich dort verfestigen. Hume sprach von der „zusätzlichen Lebhaftigkeit“, die eine wiederholte Erfahrung einem geistigen Bild verleiht und betonte den Begriff der *Gewohnheit* – ein Begriff, der mit dem *Habitus* von Bourdieu vergleichbar ist (Héran, 1987, [130]). Das Ziel dieser Konstrukte besteht darin, die Gewißheiten der dritten Ebene (Glaubensüberzeugungen) mit denen der zweiten Ebene (Häufigkeiten, wiederholte Empfindungen) zu verbinden. Diese Fragen treten erneut bei der Wahl des Gegenstandes der Wissenschaftssoziologie auf, deren Ziel es ist, den Entstehungsprozeß der Wissenschaft und die Herausbildung einer Überzeugung zu beschreiben. Will ein Wissenschaftler von seinesgleichen anerkannt werden, dann muß er die Experimente in eindeutig formulierten Standardprotokollen aufzeichnen, damit sie wiederholt werden können. Ebenso wird vor Gericht ein einzelner Zeuge als rechtsunwirksam angesehen. Historiker und Journalisten müssen ihre Quellen vergleichen und überprüfen. Wiederholungen sind Indizien für Objektivität und können als Beweis geltend gemacht werden.

Die „klassische Wahrscheinlichkeitsrechnung“ des 18. Jahrhunderts ist in gewisser Weise eine Vorläuferin der „großen Trennung“, die zwischen dem objektiven wissenschaftlichen Wissen – das die vom Menschen unabhängigen Dinge beschreibt – und den Glaubensüberzeugungen stattgefunden hat, die für primitive Gesellschaften und für vorwissenschaftliches Denken charakteristisch sind. Tatsächlich fügten die betreffenden Gelehrten die Wahrscheinlichkeiten des Würfelspiels, die Regelmäßigkeiten von demographischen

Statistiken und die persönlichen Überzeugungen von Richtern in einen einzigen Rahmen ein. Aber durch dieses Integrationsbestreben bereiteten sie den Boden für die große Trennung vor, indem sie versuchten, das Territorium des Objektivierbaren immer weiter auszudehnen, und zwar auch auf Probleme wie zum Beispiel Schwurgerichtsentscheidungen, die von Condorcet und Poisson diskutiert wurden.

Die Trennung wurde vor allem durch mehrere Debatten vorbereitet, in denen Zweifel bezüglich des rationalen Charakters von Verhaltensweisen zur Sprache kamen, die einzig und allein von der Berechnung der Erwartungswerte diktiert waren. Ein berühmtes Beispiel hierfür war das „St. Petersburger Paradoxon“<sup>13</sup> (Jorland, 1987, [140]). Daniel Bernoulli hatte seine interessante Diskussion des Paradoxons in den „Petersburger Commentarien für 1730/31“, die 1738 erschienen, veröffentlicht. Von dieser Publikation hat das Problem seinen Namen, und dort wird es folgendermaßen formuliert<sup>14</sup>:

Mein sehr verehrter Oheim, der berühmte Nicolaus Bernoulli, Professor beider Rechte an der Akademie zu Basel, legte einmal dem bekannten Monmort fünf Probleme vor, die man in dem Buche *Analyse sur les jeux de hazard* von Herrn De Montmort findet. Das letzte dieser Probleme lautet folgendermaßen: Peter wirft eine Münze in die Höhe und zwar so lange, bis sie nach dem Niederfallen die Kopfseite zeigt; geschieht dies nach dem ersten Wurf, so soll er dem Paul 1 Dukaten geben; wenn aber erst nach dem zweiten: 2, nach dem dritten: 4, nach dem vierten: 8, und so fort in der Weise, daß nach jedem Wurf die Anzahl der Dukaten verdoppelt wird. Man fragt: Welchen Wert hat die Gewinnhoffnung für Paul?

Der Gewinnerwartungswert von Paul ist demnach

$$\frac{1}{2} + 2 \left( \frac{1}{2} \right)^2 + 2^2 \left( \frac{1}{2} \right)^3 + \dots + 2^{n-1} \left( \frac{1}{2} \right)^n + \dots$$

Dieser Wert ist also unendlich groß. Demnach läge es entsprechend der Theorie der Erwartungswerte in Pauls Interesse, in diesem Spiel eine *beliebige Summe* einzusetzen, da der „wahrscheinlich erwartete Gewinn“ stets größer als diese Summe ist. Aber im Gegensatz zur Theorie zeigt der gesunde Menschenverstand, daß niemand mehr als ein paar Dukaten einsetzen würde. Dieser Sachverhalt verunsicherte die zeitgenössischen Gelehrten sehr. Die heute übliche Fassung dieses Problems erwuchs aus Diskussionen zwischen Niklaus Bernoulli, De Monmort und Gabriel Cramer in den Jahren 1713 bis 1728.

Das Problem führte auch zu einer höchst lebhaften Diskussion zwischen Niklaus und Daniel Bernoulli, die von Jorland und Daston analysiert wurde. Es

<sup>13</sup> Auch als „Petersburger Paradoxon“ oder als „Petersburger Problem“ bezeichnet.

<sup>14</sup> Zitiert nach Daniel Bernoulli, *Versuch einer neuen Theorie der Wertbestimmung von Glücksfällen* (Specimen theoriae novae de mensura sortis), (übersetzt von A. Pringsheim), Leipzig 1896.

ist das Verdienst dieser Diskussion, die verschiedenen möglichen Bedeutungen der Berechnung von Wahrscheinlichkeiten aufzuzeigen. Ohne auf Einzelheiten einzugehen, stellen wir hier nur fest, daß der Gegensatz der Standpunkte der beiden Bernoullis bedeutsam ist. Für Daniel kommt die klassische Art der Berechnung von Erwartungswerten in der Person eines unparteiischen Richters zum Ausdruck, der die individuellen Merkmale des Spielers nicht kennt, während es sich für den Spieler weniger um *Fairness* als um *Vorsicht* handelt. Dem „mathematischen“ Erwartungswert setzt Daniel also einen „moralischen“ Erwartungswert entgegen, der sich durch das Produkt der Wahrscheinlichkeit des Ereignisses mit dessen „Nützlichkeit“ (im Sinne der Wirtschaftstheorie) ausdrückt.

Daniel entnahm seine Argumentation der Geschäftswelt, während der Jurist Niklaus den Einwand machte, daß dieser „moralische Erwartungswert“ nicht „der Gleichheit und der Gerechtigkeit“ entspricht.

Hierauf entgegnete Daniel, daß seine Überlegung „in vollkommener Übereinstimmung mit der Erfahrung stehe“. In der Tat stützte sich Niklaus auf die Gleichheitsbedeutung aleatorischer Verträge, während Daniel eine Art kommerzieller Vorsicht verteidigte. Der besonnene Kaufmann stand dem unparteiischen Richter gegenüber. Bezüglich des im vorhergehenden Abschnitt „Aleatorische Verträge und faire Abmachungen“ geschilderten einfacheren Paradoxons finden wir einerseits den Richter in der Position, die Lage von oben überschauen zu können (zur Not sehen wir sogar einen utopischen Spieler von unbegrenztem Reichtum, der das Spiel unendlich oft unter dem Einsatz großer Summen spielen könnte), und andererseits den „normalen“ Spieler, der mit einem umsichtigen Kaufmann von begrenztem Glück vergleichbar ist und der es sich nicht leisten kann, eine große Summe auf einen riesigen, wenn auch sehr unwahrscheinlichen Gewinn zu setzen.

Der Nachteil der hier erwähnten Paradoxa bestand darin, daß sie mit fiktiven Problemen zu tun hatten und intellektuelle Spiele zu sein schienen. Ganz anders verlief die Debatte über die Pockenimpfung, welche die Gelehrten zur gleichen Zeit gegeneinander aufbrachte (Benzécri, 1982, [12]). Die Präventivmaßnahme der allgemeinen Impfung senkte das Auftreten der Krankheit erheblich, führte aber unglücklicherweise auch in *einem* von dreihundert Fällen zum Tod der geimpften Personen im Jahr der Impfung. Die Bilanz war jedoch positiv und Daniel Bernoulli rechnete aus, daß die Lebenserwartung von geimpften Personen trotz dieser fatalen Vorfälle drei Jahre höher ist, als die Lebenserwartung ungeimpfter Personen. Vom Standpunkt der öffentlichen Gesundheit konnte diese Maßnahme demnach obligatorisch gemacht oder zumindest mit Nachdruck empfohlen werden. Begreiflicherweise waren jedoch einige (häufig als „Familienväter“ beschriebene) Personen in Bezug auf sich und ihre Kinder mehr als zurückhaltend. Aus diesem Beispiel ist ersichtlich, daß der frequentistische Standpunkt mit einer makrosozialen Position (das heißt mit der Position des Staates oder einer Allgemeinheit) zusammenhing, während es sich beim epistemischen Standpunkt um die Position einer Person handelte, die für sich selbst Entscheidungen treffen mußte. Das Problem kam

dann im 19. Jahrhundert erneut in anderen Debatten auf, in denen es um die Anwendung der statistischen Methode in der Medizin ging.

Die Diskussion führte dazu, die Verwendung von Erwartungswerten bei der rationalen Betrachtung menschlicher Entscheidungen in Zweifel zu ziehen und bereitete den Weg für die spätere Trennung der wahrscheinlichkeitstheoretischen Betrachtungsweisen vor. Im 19. Jahrhundert wurde eine vorläufige Grenze errichtet, welche die mit den „Zuständen des Verstandes“ verknüpfte Seite der Wahrscheinlichkeitsrechnung zurückwies und eine Beschränkung auf die „Zustände der Welt“ und insbesondere auf die frequentistische Richtung beinhaltete. Als etwa Auguste Comte die Wahrscheinlichkeitsrechnung im Allgemeinen attackierte und sie anklagte, daß sie außerstande sei, die Komplexität der menschlichen Verhaltensweisen zu berücksichtigen, griff er im Grunde genommen die epistemische Interpretation dieses Kalküls an, die den Überlegungen der klassischen Probabilisten Nahrung gegeben hatte.

Es gab mehrere Typen von Beispielen, mit denen man diese Art von Kritik untermauern konnte. In einigen dieser Beispiele – wie etwa im St. Petersburger Paradoxon oder beim Pockenimpfungsproblem – wurde die Nützlichkeitsfunktion infrage gestellt, die mit der ihrerseits nur wenig diskutierten objektiven Wahrscheinlichkeit assoziiert war (geometrisch<sup>15</sup> für das Spiel „Kopf oder Zahl“ und frequentistisch im Falle der Impfung). In anderen Fällen dagegen – wie etwa bei den auf Abstimmung beruhenden Entscheidungen von Geschworenen in Strafsachen, in denen eine Beurteilung der Schuld unter Berücksichtigung von Indizien und Vermutungen erfolgte, aber auf keinem vollständigen Beweis aufbaute – bezweifelte man sogar die Möglichkeit der Abschätzung der (subjektiven) Wahrscheinlichkeit einer „Ursache“ (der Schuld des Angeklagten), weil man wußte, daß bestimmte „Effekte“ (Indizien oder unsichere Zeugenaussagen) eine Rolle gespielt hatten. Die Frage nach der *Ursachenwahrscheinlichkeit* (oder *inversen Wahrscheinlichkeit*) spielte im 18. Jahrhundert bei Bayes und Laplace eine wichtige Rolle. Im 19. Jahrhundert hatte man diese Frage weitgehend verdrängt, bevor sie im 20. Jahrhundert im Rahmen der Spieltheorie, der Entscheidungstheorie und in der Kognitionswissenschaft wieder zum Leben erweckt wurde.

## Der Bayessche Ansatz

Das Problem der Gerechtigkeit, definiert durch die Gleichheit von Erwartungswerten, beherrschte die Arbeiten der Begründer der Wahrscheinlichkeitsrechnung: Die „Logik von Port Royal“<sup>16</sup> machte zum Beispiel umfassenden Gebrauch davon. Der frequentistische Standpunkt, der sich implizit in den

<sup>15</sup> Hiermit ist die zugrunde liegende geometrische Folge der Partialsummen gemeint.

<sup>16</sup> Das 1662 anonym in Paris erschienene Werk *La Logique ou l'Art de penser* war außerordentlich einflußreich. Ein Nachdruck wurde von Baron Freytag von Löringhoff, Stuttgart 1965, herausgegeben. Die lateinische Fassung erschien unter dem Titel *Ars cogitandi*. Als Verfasser des Buches gelten Antoine Arnauld und Pierre

Glücksspielen wiederfindet, führte Jakob Bernoulli (1654–1705) zum ersten Beweis des Gesetzes der großen Zahlen, das 1713 nach seinem Tod veröffentlicht wurde: Die Häufigkeit des Auftretens eines Ereignisses, das eine gegebene Wahrscheinlichkeit hat, strebt gegen diese Wahrscheinlichkeit, wenn man die Anzahl der Versuche erhöht<sup>17</sup> (Meusnier, 1987, [196]). Danach vervollständigte Abraham de Moivre (1667–1754) im Jahre 1738 den Beweis: er berechnete die Wahrscheinlichkeit dafür, daß diese Häufigkeit in einem beliebig kleinen Intervall auftritt, wenn man eine hinreichend große Anzahl von Ziehungen durchführt. Nebenbei gab er die erste präzise Formulierung der späteren „Normalverteilung“<sup>18</sup>, indem er sich auf Stirlings asymptotische Formel<sup>19</sup> für  $n!$  ( $n$  Fakultät) stützte ( $n! = 1 \times 2 \times 3 \times \dots \times n$ ).

Die vorhergehenden Ergebnisse gestatteten es, erwartete Wirkungen aus einer bekannten Ursache (Füllung einer Urne) – das heißt die Wahrscheinlichkeiten der Beobachtung von Häufigkeiten – abzuleiten. Aber das inverse Problem wurde oft in allen denjenigen Fällen gestellt, in denen es wünschenswert war, auf der Grundlage der beobachteten Ereignisse etwas über die (unbekannten) Ursachen zu sagen. Erst Bayes (1702–1761) und dann Laplace (1749–1827) machten sich an die Beantwortung dieser Frage – und ihre Antworten führten zu weiteren Debatten.

Das auf diese Weise von Bayes und Laplace gestellte Problem der „inversen Wahrscheinlichkeit“ spielt – bis zum heutigen Tage – eine entscheidende Rolle in der Geschichte der Anwendung der Wahrscheinlichkeitsrechnung und der Statistik in den verschiedenen Wissenschaften. In der Tat liegt dieses Problem an der Nahtstelle zwischen dem objektivistischen Standpunkt (die Wissenschaft *entschleiert* eine verborgene Realität – den Inhalt der Urne – oder nähert sich dieser Realität immer mehr an) und dem konstruktivistischen Standpunkt (die Wissenschaft *konstruiert* Objekte und Modelle mit Hilfe von Kriterien und Werkzeugen, und verleiht dadurch diesen Objekten und Modellen eine relative Stabilität). Nun ist aber die Methode, bei der man von

---

Nicole, beide prominente Vertreter der Jansenisten, einer katholischen Reformbewegung im Umkreis des Klosters Port Royal bei Paris, die in scharfem Konflikt mit der römischen Kirche lag. Auch Blaise Pascal stand dieser Bewegung nahe und war womöglich ein Mitverfasser des Buches.

<sup>17</sup> Formaler ausgedrückt beinhaltet das Gesetz der großen Zahlen: Die Häufigkeit  $h$  eines Ereignisses mit Wahrscheinlichkeit  $p$  in  $n$  unabhängigen Versuchen unterscheidet sich von  $p$  mit beliebig großer Wahrscheinlichkeit nur um beliebig wenig, sobald  $n$  hinreichend groß ist. Man drückt dasselbe auch folgendermaßen aus: Die Häufigkeit  $h$  konvergiert fast sicher für  $n \rightarrow \infty$  gegen  $p$ . Oder auch: Für  $n \rightarrow \infty$  ist  $h$  eine konsistente Schätzung für  $p$ .

<sup>18</sup> Die Bezeichnungsweise ist sowohl im Französischen als auch im Deutschen uneinheitlich. Der Autor verwendet hier „loi normale“, was im Deutschen auch durch „Normalverteilungsgesetz“ wiedergegeben wird.

<sup>19</sup> Stirlingsche Näherungsformel für die Fakultät:  $n! \approx (\frac{n}{e})^n \sqrt{2\pi n}$ , wobei  $e$  die Basis der natürlichen Logarithmen bezeichnet.

registrierten Ereignissen auf Ursachen schließt, beider Interpretationen fähig: Ursachen werden entweder aufgedeckt oder konstruiert.

Die Bayesschen Formalisierungen zielten darauf ab, *Glaubensgründe* unter Berücksichtigung vorheriger Erfahrungen so abzuschätzen, daß eine konkrete Situation beurteilt und eine Entscheidung getroffen werden kann. Der von Bayes 1764 unter dem Titel *An Essay Towards Solving a Problem in the Doctrine of Chances*<sup>20</sup> veröffentlichte Text beginnt bei der Formulierung des zum Problem von Bernoulli und de Moivre inversen Problems mit Worten, die heute seltsam klingen (die „Chance einer Wahrscheinlichkeit“), deren Zweideutigkeit jedoch beim Verständnis dessen helfen kann, wonach damals gesucht wurde:

*Gegeben* ist die Anzahl Male, die ein unbekanntes Ereignis eingetreten und ausgeblieben ist. *Gesucht* ist die Chance, daß die Wahrscheinlichkeit seines Eintretens bei einem einzelnen Versuch irgendwo zwischen zwei angebbaren Graden der Wahrscheinlichkeit liegt (Bayes, 1764, neu herausgegeben von Pearson und Kendall, 1970, [221]).

Die Wörter „Wahrscheinlichkeit“ und „Chance“ werden anschließend durch Begriffe aus der Theorie der Erwartungswerte definiert, das heißt durch ein Verhältnis von Werten, die zu schätzen sind, damit eine Wette berechtigt ist.

Die Wahrscheinlichkeit eines Ereignisses ist das Verhältnis zwischen dem Werte, welcher einer an das Eintreten des Ereignisses geknüpften Erwartung zu geben ist, und dem Werte des in diesem Falle erwarteten Gewinns. Unter Chance verstehe ich dasselbe wie Wahrscheinlichkeit. (Bayes, op. cit.)

Demnach wird den Wörtern *Chance* und *Wahrscheinlichkeit* per Deklaration ein und dieselbe Bedeutung zugewiesen, was aber nicht zur Klärung des ersten Satzes beiträgt. Jedoch zeigt die nachfolgende Überlegung, daß das Wort „Chance“ im Sinne von „Glaubensgrund“ mit Blick auf eine Entscheidung verwendet wird, während das Wort „Wahrscheinlichkeit“ eine objektive Bedeutung hat, ähnlich wie der Begriff „Füllung der Urne“. Es handelt sich demnach um die „Wahrscheinlichkeit“, daß die Füllung dieser unbekannten Urne (also das Verhältnis zwischen den Anzahlen der schwarzen und der weißen Kugeln) innerhalb eines gegebenen Intervalls liegt. Das heißt es geht um eine „Ursachenwahrscheinlichkeit“, die sich – ebenso wie die „Wahrscheinlichkeit“ der Schuld eines Angeklagten – nur durch einen „Grad der Sicherheit“ interpretieren läßt, der für eine Entscheidung notwendig ist. Die einzige Art und Weise, dieser „Ursachenwahrscheinlichkeit“ eine formalisierte Bedeutung zu geben, wäre die Annahme, daß die Urne ihrerseits aus einer großen Anzahl von Urnen verschiedener Füllungen gezogen wurde. Aber das führt auf

<sup>20</sup> In deutscher Sprache erschienen unter dem Titel *Versuch zur Lösung eines Problems der Wahrscheinlichkeitsrechnung* in „Ostwalds Klassikern der exakten Wissenschaften“ (Leipzig 1908).



die Frage der Verteilung dieser Zusammensetzungen, das heißt auf eine „*A-priori*-Wahrscheinlichkeit“ zurück, und genau das ist der Punkt, an dem der Bayessche Ansatz am meisten kritisiert wird.

Durch die Schaffung des Begriffes der „bedingten Wahrscheinlichkeit“ führt das Verfahren die Irreversibilität der Zeit in die Formulierung ein („*A dann B*“) und das ist die Ursache für den Doppelcharakter dieses Begriffes (Clero, 1986, [48]). In der Tat läßt sich die Überlegung auf der Grundlage der folgenden doppelten Gleichheit aufbauen:

$$P(A \text{ und } B) = P(A \text{ falls } B) \times P(B) = P(B \text{ falls } A) \times P(A)$$

und hieraus folgt

$$P(A \text{ falls } B) = P(B \text{ falls } A) \times \frac{P(A)}{P(B)}.$$

Übertragen auf das Problem der Wahrscheinlichkeit einer Ursache  $H_i$  (in einer Menge von  $n$  sich gegenseitig ausschließenden Ursachen) für ein Ereignis  $E$  läßt sich dieser Sachverhalt unter Verwendung einer moderneren Schreibweise folgendermaßen wiedergeben:

$$P(H_i|E) = \frac{P(E \cap H_i)}{P(E)} = \frac{P(E|H_i) \cdot P(H_i)}{\sum_{j=1}^n P(E|H_j) \cdot P(H_j)}.$$

Diese Formel wurde 1774 von Laplace in einem langen Satz ausgedrückt, der heute schwer lesbar ist. Die Schwierigkeit läßt sich mit der andersartigen Schwierigkeit der vorhergehenden mathematischen Formel vergleichen.

Läßt sich ein Ereignis durch eine Anzahl  $n$  von verschiedenen Ursachen erzeugen, dann verhalten sich die Wahrscheinlichkeiten der Existenz dieser Ursachen für das betreffende Ereignis wie die Wahrscheinlichkeiten des Ereignisses für diese Ursachen, und die Wahrscheinlichkeit für die Existenz einer jeden Ursache ist gleich der Wahrscheinlichkeit des Ereignisses bei Vorliegen dieser Ursache, dividiert durch die Summe aller Wahrscheinlichkeiten des Ereignisses bei Vorliegen aller dieser Ursachen. (Laplace, 1774, *Mémoire sur la probabilité des causes par les événements*. In: Oeuvres complètes, t. VIII, S. 28.)

Aber der entscheidende Punkt im Beweis von Bayes besteht darin, daß die Symmetrie der doppelten Gleichheit, welche die bedingten Wahrscheinlichkeiten  $P(A \text{ falls } B)$  und  $P(B \text{ falls } A)$  definiert, für ihn nicht existierte und daß beide Gleichheiten gesondert und unabhängig voneinander bewiesen werden (Stigler, 1986, [267]), und zwar mit Hilfe von zwei verschiedenen „Propositionen“. Diese Überlegungen stützen sich auf *Zunahmen* der Gewinnerwartungswerte, die durch das Auftreten eines Anfangsereignisses  $A$  eingeführt wurden. Jeder Beweis ist die Schilderung einer Folge von hypothetischen Ereignissen und deren Konsequenzen in Bezug auf Gewinne. Aber

diese Schilderungen können nur dann zu einer Schlußfolgerung führen, wenn man den Ursachen *A-priori*-Wahrscheinlichkeiten zuordnet, das heißt in diesem Fall die Wahrscheinlichkeit  $P(H_i)$  – und das sogar tut, *bevor* man irgendein partielles Wissen erlangt hat. Die Hypothesen der „gleichbleibenden“ *A-priori*-Wahrscheinlichkeit derartiger Ursachen sind häufig angefochten worden, indem man mit Hilfe von Beispielen zeigte, daß diese „gleichbleibenden“ Wahrscheinlichkeiten auf verschiedene Weisen gewählt werden können und daher reine Konventionen sind.

Die Spannung und die Fruchtbarkeit des Bayesschen Ansatzes ist auf den Doppelcharakter des Anfangsausdrucks zurückzuführen: die ursprüngliche doppelte Gleichheit ist *formal* symmetrisch, aber *logisch* asymmetrisch, da die Zeit eine Rolle spielt und da die Ereignisse *bekannt* sind, während die Ursachen aus ihnen *geschlußfolgert* werden. Dieser Doppelcharakter war übrigens ein inhärenter Bestandteil des Apparates, den Bayes zur Untermauerung seiner Überlegungen ersonnen hatte. In der Tat erwies sich das Beispiel der Urne als unzulänglich, denn es war schwierig, eine Reihe von Ereignissen zu konstruieren, bei denen man *zuerst* Urnen und *dann* Kugeln ziehen mußte: die Asymmetrie war zu stark. Bayes schlug deswegen vor, nacheinander zwei Kugeln auf einem quadratischen Billardtisch derart in Bewegung zu setzen, daß die Wahrscheinlichkeitsdichten ihrer Haltepunkte auf dem grünen Tuch gleichmäßig verteilt sind. Die Kugel *A* wird zuerst gestoßen und nachdem sie zum Stillstand gekommen ist, wird eine vertikale Gerade durch ihren Haltepunkt gelegt, wodurch das Quadrat in zwei Rechtecke *P* und *Q* zerlegt wird. Danach wird die Kugel *B* gestoßen und die Wahrscheinlichkeit dessen untersucht, daß sie im Rechteck *Q* zum Stehen kommt. Man hat demnach eine Folge von zwei Ereignissen und kann die damit zusammenhängenden bedingten Wahrscheinlichkeiten berechnen, indem man sowohl deren symmetrischen Charakter (geometrisch) als auch ihren asymmetrischen Charakter (zeitlich) aufrecht erhält.

Die Methode der Bayesschen Inferenz kann aus der Sicht des vorliegenden Buches vom Standpunkt der *Konstruktion von Äquivalenzklassen, Taxonomien* und *Kodierungen* interpretiert werden. Tatsächlich impliziert der aus dieser Sichtweise postulierte Kausalitätsbegriff, daß ähnliche Ursachen zu ähnlichen Folgen führen können. Mit anderen Worten: man kann sich Kategorien von Ursachen und Wirkungen vorstellen und zukünftige Ereignisse können mit vergangenen Beobachtungen derart verknüpft werden, daß die Ungewißheit der zukünftigen Ereignisse dadurch umschrieben wird. Diese Herangehensweise steht im Gegensatz zur Einstellung des nominalistischen Skeptizismus, der in der Philosophie des 18. Jahrhunderts (zum Beispiel bei Berkeley) sehr verbreitet war. In dieser Philosophie kann nichts mit etwas anderem verglichen werden und es ist keine allgemeine Darstellung möglich (Clero, 1986, [48]). In dieser Hinsicht schloß sie sich an die Philosophie von Hume an, der das menschliche Wissen als Produkt einer Akkumulation von „Bleistiftstrichen“ beschrieb, die „Rillen in das Gehirn“ eingraben, denen jede weitere Erfahrung eine „zusätzliche Lebhaftigkeit“ verleiht. Die Vorstellung von der

mit neuem Wissen zusammenhängenden Zusätzlichkeit ist im Text von Bayes außerordentlich gegenwärtig.

Dieser Umstand ist auch für die *Indizienwissenschaften* typisch, die Ginzburg (1980) den *Galileischen Wissenschaften* gegenüberstellt. Während die letzteren mit Hilfe mathematischer und statistischer Methoden eine große Masse von Informationen gleichzeitig verarbeiten, um daraus allgemeine Gesetze abzuleiten, gehen die ersteren von einzelnen Merkmalen aus, um „Geschichten“ zu konstruieren oder um einzelne Fälle allgemeinen Familien von Ursachen zuzuordnen. Das ist die Vorgehensweise des Historikers, des Polizisten und des Arztes, der auf der Grundlage von Symptomen eine Diagnose stellt.

Folglich kann die Kodierung von „Todesursachen“ – durch Verarbeitung der medizinischen Meldepflicht auf Totenscheinen – entsprechend der *Internationalen Klassifikation der Krankheiten, Verletzungen und Todesursachen*<sup>21</sup> als Bayessches Verfahren beschrieben werden (Fagot-Largeault, 1989, [90]): sowohl die einzelnen Symptome (Ereignisse) als auch der bereits bekannte Verbreitungsgrad einer Krankheit (*A-priori*-Wahrscheinlichkeit) werden berücksichtigt. In diesem Fall, der eine Feinanalyse der Entstehung und Entwicklung der ICD-Nomenklatur darstellt, tritt die Kodierung der Todesursachen als Konvention auf. Tatsächlich darf eine Ursache, deren *statistische Beschreibung* als *aufschlußreich* beurteilt wird, in der Folge der dem Tod vorangehenden Ereignisse nicht zu früh auftreten (andernfalls wäre ihr Einfluß indirekt, schwach und verwässert), aber diese Ursache darf auch nicht zu eng mit Tod verknüpft sein (für den sie ein *Synonym* wäre: das Herz schlägt nicht mehr). Vielmehr muß es sich um eine Kategorie handeln, die zwischen den beiden genannten Kategorien steht und den Effekt hat, die Wahrscheinlichkeit des Todes signifikant zu steigern, ohne ihn jedoch zur Gewißheit zu machen. Dabei wird angenommen, daß man *auf die betreffende Ursache* durch Prävention oder durch geeignete therapeutische Maßnahmen *einwirken* kann, um diese Wahrscheinlichkeit zu senken. Die Analyse der medizinischen Diagnose und der Kodierung eines Totenscheins ist in doppelter Hinsicht für den Bayesschen Ansatz typisch (vgl. Fagot-Largeault, 1989, [90]): einerseits durch die Konventionen, welche die Kausalität durch eine Kategorisierung der Ereignisse definieren, deren Kenntnis wiederum Wahrscheinlichkeitsschwankungen impliziert – andererseits durch die Verarbeitung des *Kodierungsmoments*, das als *Entscheidung* angesehen wird, die sowohl einzelne Ereignisse als auch *A-priori*-Wahrscheinlichkeiten berücksichtigt.

Der Bayessche Ansatz war also eine wesentliche Episode in der Geschichte der Wahrscheinlichkeitsrechnung, denn dieser Ansatz war der Ursprung des Begriffes der *inversen Wahrscheinlichkeit* oder *Ursachenwahrscheinlichkeit*. Dieser Begriff hat keine Bedeutung in einer rein axiomatischen Theorie der Wahrscheinlichkeitsrechnung (etwa in der Theorie von Kolmogorow), erweist sich aber als sinnvoll in der Analyse so manchen Entscheidungsverhaltens und

<sup>21</sup> International Classification of Diseases, Injuries and Causes of Death (ICD).

bei der praktischen Konstruktion von Äquivalenzklassen – einer Schlüsselphase der statistischen Tätigkeit. Das Moment der Kodierung und seine speziellen Randbedingungen geraten mitunter in Vergessenheit, wenn man statistische Tabellen verarbeitet und interpretiert – ebenso wie der Bayesianismus lange Zeit in der statistischen Denkweise verdrängt worden ist.

## Der „goldene Mittelweg“: Mittelwerte und kleinste Quadrate

Im Rahmen der Techniken, die heute zur Konstruktion und Stabilisierung der sozialen Welt beitragen, spielt die Statistik eine doppelte Rolle. Einerseits bestimmt sie Objekte, indem sie für diese Objekte mittels standardisierter Definitionen Äquivalenzen festlegt. Dadurch wird es möglich, die Objekte zu messen: Mit Hilfe der Sprache der Wahrscheinlichkeitstheorie präzisiert man den Vertrauensgrad<sup>22</sup>, der sich diesen Messungen zuordnen läßt. Andererseits liefert die Statistik Formen, und zwar sowohl zur Beschreibung der *Relationen* zwischen den so konstruierten Objekten als auch zur Prüfung der Konsistenz dieser Relationen. Diese beiden Aspekte, das heißt die Konstruktion von Objekten und die Analyse der Relationen zwischen diesen Objekten, scheinen eng miteinander zusammenzuhängen. Dennoch gehen beide Aspekte aus zwei deutlich voneinander verschiedenen Traditionen hervor, die erst zu Beginn des 20. Jahrhunderts konvergierten. Mitunter konnte am Ende ein und derselbe Formalismus auf ganz verschiedene Fragen angewendet werden, aber die einfache Übertragung dieses geistigen Werkzeugs von einem Bereich auf einen anderen dauerte ein volles Jahrhundert und bei diesem Prozeß waren aufwendige begriffliche Übersetzungen erforderlich.

Ein Beispiel hierfür ist die als *Methode der kleinsten Quadrate* bezeichnete Anpassungsmethode (Armatte, 1991, [5]). Diese Methode war 1805 von Legendre in Antwort auf eine Frage formuliert worden, die während des gesamten 18. Jahrhunderts von Astronomen und Geodäten immer wieder gestellt wurde: Wie kann man die unter verschiedenen Voraussetzungen gemachten Beobachtungen kombinieren, um bestmögliche Schätzungen einer Reihe von astronomischen und terrestrischen Größen zu erhalten, die ihrerseits durch lineare Relationen verknüpft sind? Diese Größen waren mit unvollkommenen Instrumenten unter unterschiedlichen Bedingungen gemessen worden, zum Beispiel in verschiedenen historischen Epochen oder an mehreren Punkten der Erde. Wie ließ sich diese Fülle von Messungen unter Berücksichtigung des Umstandes am besten nutzen, daß sie die theoretisch vorgegebene Relation niemals vollständig bestätigten, sondern das Vorhandensein einer kleinen (auch als *Fehler*, *Residuum* oder *Rest* bezeichneten) *Abweichung* an derjenigen Stelle gestatteten, wo eigentlich der Wert Null auftreten mußte. Mit anderen Worten: die zwei oder drei unbekannten Größen traten als Lösungen

<sup>22</sup> Auch „Konfidenzgrad“ genannt.

eines Systems auf, das zu viele Gleichungen hatte (soviele Gleichungen wie Beobachtungspunkte). Man muß also diese Gleichungen optimal kombinieren, um eine Schätzung der gesuchten Größen zu erhalten. Das war das Problem, das Legendre im Jahr 1805 mit Hilfe einer Methode löste, bei der die *Summe der Quadrate* dieser Abweichungen *minimiert* wird (Stigler, 1986, [267]).

Es handelte sich also darum, die Messung von Objekten mit größtmöglicher Präzision durchzuführen, indem man die unterschiedlichen Beobachtungen ein und derselben Größe bestmöglich miteinander kombiniert. Dagegen war das in den 1890er Jahren von den englischen Eugenikern Galton und Pearson – den Erfindern der *Regression* und der *Korrelation* – gestellte und gelöste Problem gänzlich andersartig: Wie sind die Relationen und wechselseitigen Beziehungen zwischen Objekten zu beschreiben, die weder voneinander unabhängig noch vollständig voneinander abhängig sind? Derartige Fälle treten bei Problemen der Vererbung auf. Die Anpassung *einer* Variablen an eine *andere* Variable mit Hilfe eines linearen Regressionsmodells führte nichtsdestoweniger zu einem System von Gleichungen und zu einem Lösungsansatz, der *formal analog* zum Verfahren von Legendre war. Aber die Bedeutungsinhalte der mathematischen Konstruktion wichen in beiden Fällen derart voneinander ab, daß die Übertragung des Formalismus von Legendre – die dieser 1810 auf der Grundlage seiner auf Gauß und Laplace zurückgehenden probabilistischen Interpretation schuf – nicht wirklich vor 1930 stattfand.

Die gegen 1810 von Laplace und Gauß durchgeführte Synthese ergab sich ihrerseits aus der Vereinigung zweier ganz unterschiedlicher Traditionen. Einerseits waren die Astronomen und die Physiker daran gewöhnt, empirisch ungenaue Beobachtungen zu kombinieren – zum Beispiel durch Mittelwertberechnungen (das heißt durch den „Mittelweg“, den man einschlagen muß), um die Werte von Naturgrößen „so gut wie möglich“ abzuschätzen. Andererseits hatten die probabilistisch orientierten Mathematiker und Philosophen an der Frage des *Grades der Sicherheit* gearbeitet, der sich einem Wissen oder einem Glauben zuordnen läßt. Die Philosophen gelangten auf diese Weise an einen Punkt, an dem sie den von den Mathematikern verwendeten Ausdruck „so gut wie möglich“ infrage stellten: Wie soll man den Vertrauensgrad von etwas abschätzen, das eine Schätzung verdient? Vor Gauß und Laplace hatte niemand eine Antwort auf diese Frage gegeben.

Die erstgenannte Tradition, die sich mit der Messung astronomischer und terrestrischer Größen befaßte, blickte bereits auf eine lange Geschichte zurück (Stigler, 1986, [267]). Dieses Problem versprach bedeutende ökonomische und militärische Anwendungen. Deswegen hat im gesamten 18. Jahrhundert das Streben nach Perfektionierung der Techniken zur Berechnung von Schiffspeditionen (Längen- und Breitenbestimmung) zahlreiche Forschungsarbeiten stimuliert. Seit 1700 war die Berechnung der *Breite* (auf der Grundlage der Höhe der Fixsterne) ziemlich einfach. Dagegen machte die *Längen*berechnung beträchtliche Schwierigkeiten. In England wurde 1714 eine Kommission gegründet, um diese Frage zu untersuchen und Forschungen zu subventionieren, die zur Problemlösung beitragen (zwischen dem Gründungsdatum und 1815

gab man mehr als 100000 englische Pfund zu diesem Zweck aus). Zwei Techniken wurden damals entwickelt: die Präzision der Uhren, die an Bord von Schiffen die Greenwich-Zeit anzeigen, und die Aufstellung von Tabellen, die eine detaillierte Beschreibung der Mondpositionen lieferten.

Im zweitgenannten Fall besteht das Problem darin, daß sich der Mond gegenüber der Erde nicht immer unter ein und demselben Winkel zeigt und daß leichte Schwankungen der Mondrotation (die „Librationen“) die Berechnung der Mondposition außerordentlich komplizieren. Der deutsche Astronom Tobias Mayer (1723–1762) veröffentlichte hierzu eine geistreiche Lösung<sup>23</sup>, indem er die Beobachtungen in geeigneter Weise miteinander kombinierte. Berechnungen hatten ihn dazu geführt, zu verschiedenen Zeitpunkten die Position eines gewissen Mondkraters präzise zu beobachten und diese Beobachtungen führten zur Messung dreier unterschiedlicher astronomischer Größen, die miteinander durch eine Gleichung der sphärischen Trigonometrie verknüpft waren. Da er diese Beobachtungen insgesamt siebenundzwanzigmal gemacht hatte, mußte er ein überbestimmtes System von siebenundzwanzig Gleichungen in drei Unbekannten lösen.

Mayer verfügte über keine Regel zur Minimierung der Fehler zwischen den Erwartungswerten und den durch zufällige Näherung berechneten Werten. Deswegen führte er eine gut durchdachte Umgruppierung seiner siebenundzwanzig Gleichungen in drei Gruppen zu je neun Gleichungen durch und addierte dann gesondert jede der drei Gruppen. Auf diese Weise erhielt er schließlich ein System von drei Gleichungen in drei Unbekannten, und diese Gleichungen lieferten ihm die gesuchten Abschätzungen. Die Richtigkeit der Methode ist auf die scharfsinnige Auswahl dreier Teilwolken von Punkten zurückzuführen, die durch ihre jeweiligen Schwerpunkte ersetzt wurden, so daß der größtmögliche Anteil der ursprünglichen Informationen der siebenundzwanzig Beobachtungen erhalten blieb. Die Tatsache, daß Mayer die Messungen selbst durchgeführt hatte und mit ihnen gründlich vertraut war, verlieh ihm die Kühnheit, die Gleichungen umzugruppieren, und gab ihm die erforderliche Intuition, diese Umgruppierung auf einfallsreiche Weise vorzunehmen. Aber diese empirische Lösung stützte sich auf kein allgemeines Kriterium und konnte deswegen kaum auf andere Situationen übertragen werden. Es handelte sich um eine *Ad-hoc*-Lösung, wie sie für einen Handwerker typisch ist.

Ein allgemeines Kriterium dafür, eine Anpassung zu optimieren, wurde wenig später im Jahre 1755 von Roger Joseph Boscovich<sup>24</sup> in Bezug auf ein anderes Problem vorgeschlagen, das ebenfalls viele Gelehrte des 18. Jahrhun-

<sup>23</sup> Tobias Mayer, *Abhandlungen über die Umwälzung des Mondes um seine Axe*. In: Kosmographische Nachrichten und Sammlungen, von den Mitgliedern der Kosmographischen Gesellschaft zusammengetragen, 1(1748), S. 52–148.

<sup>24</sup> Ursprünglich: Rudjer Josip Bošković (1711–1787). Kroatischer Jesuit, der seit 1740 als Professor für Mathematik am Collegium Romanum in Rom lehrte und 1764 Professor für Mathematik in Pavia wurde. Sein italianisierter Name ist Ruggiero Guiseppe Boscovich.

derts in Unruhe versetzt hatte: das Problem der Erdgestalt. Man vermutete, daß die Erde keine vollkommene Kugel ist, sondern an den Polen leicht abgeplattet<sup>25</sup>, am Äquator dagegen verbreitert ist (einige Gelehrte vertraten übrigens die entgegengesetzte These). Die Überprüfung dieses Problems machte es erforderlich, die Länge eines Meridianbogens an ganz unterschiedlichen Breiten zu messen. Die Messungen wurden in Paris, Rom, Quito, Lappland und am Kap der Guten Hoffnung durchgeführt. In diesem Fall erwies es sich als notwendig, ein System von fünf Gleichungen in zwei Unbekannten zu lösen.

Boscovich argumentierte ganz anders als Mayer – möglicherweise weil er eine kleinere Anzahl von Daten zur Verfügung hatte. Er erfand eine geometrische Technik zur Minimierung der *Summe der absoluten Werte* der Reste, das heißt der Abweichungen zwischen den beobachteten Werten und den angepaßten Werten. Als allgemeines Kriterium ließ sich diese Technik jedoch nur sehr schwer handhaben und die „geometrische“ Lösung war nur aufgrund der kleinen Anzahl von Beobachtungen und unbekannten Größen möglich (Stigler, 1986, [267]). Laplace hatte versucht, die Summe der absoluten Werte mathematisch zu behandeln, mußte aber wegen der Komplexität der damit verbundenen Berechnungen von seinem Vorhaben Abstand nehmen.

Die Lösung durch Minimierung der *Summe der Quadrate* der Abweichungen scheint zuerst von Gauß bereits 1795 verwendet worden zu sein (zumindest behauptete er das), aber er gab keine explizite Formulierung dafür an. Unabhängig von Gauß konzipierte, formulierte und veröffentlichte Legendre diese Lösung im Jahre 1805, was einen lebhaften Prioritätsstreit zwischen beiden zur Folge hatte (Plackett, 1972, [232]).<sup>26</sup> Gauß behauptete, dieses Kriterium – die *Methode der kleinsten Quadrate* – bereits 1795 benutzt zu haben, äußerte aber später während der Kontroverse, ihm sei das Kriterium so trivial erschienen, daß er es weder für nützlich befunden hätte, es zu veröffentlichen, noch ihm einen Namen für die Nachwelt zu geben. Für Gauß war das Kriterium nur ein Rechenmittel; das Wesentliche für ihn war das damit erzielte Forschungsergebnis. Dagegen nutzten Legendre im Jahre 1805, vor allem aber Gauß selbst im Jahre 1809 und Laplace im Jahre 1810 sehr spezielle Eigenschaften dieser Methode. Insbesondere verwendeten Laplace und Gauß die Beziehungen, die unter den Gesetzen der Wahrscheinlichkeitsrechnung als „Gaußsches Fehlergesetz“ (der zukünftigen „Normalverteilung“) etabliert wurden.

Wir müssen jetzt in unserer Darstellung noch einmal zurückgehen, um kurz die andere Tradition zu verfolgen, die zur Synthese von Gauß-Laplace

<sup>25</sup> Unter der Voraussetzung, daß die Erde wie eine homogene, mit gleichförmiger Winkelgeschwindigkeit rotierende Flüssigkeit behandelt werden kann, hatte Newton in den *Principia* (1687) gezeigt, daß die Erde ein abgeplattetes Rotationsellipsoid ist, wobei der Radius am Äquator um ca. 1/230 länger ist als der Radius am Pol. Die Abplattung der Erde, das heißt der Längenunterschied zwischen der Achse der Erdkugel und des Erdellipsoids, beträgt ca. 42 km.

<sup>26</sup> Dieser Streit ist nicht nur von anekdotischem Interesse, denn er zeigt, wie sich ein wissenschaftliches Werkzeug verfestigt, wie es übertragbar wird und sich in einen anderen Kontext transportieren läßt.

führte. Es geht um die Tradition der Philosophen, die – ausgehend von probabilistischen Beschreibungen – über den Grad der Sicherheit des Wissens arbeiteten. Um ein Wahrscheinlichkeitsgesetz der statistischen Erwartung zu formulieren, muß man sich zunächst über die entsprechenden Gesetze für elementare Beobachtungsfehler verständigen. Danach müssen diese „elementaren Gesetze“ mathematisch „kombiniert“ werden, um daraus ein Gesetz für statistische Berechnungen abzuleiten. Für die Verteilungen der Elementarfehler sind verschiedene Formen vorgeschlagen worden. Simpson (1757) versuchte es mit einer Linearform, die zu einem gleichschenkligen Dreieck führt:  $-a|x| + b$ . Laplace schlug 1774 zunächst eine exponentielle Form  $[\frac{m}{2}e^{-m|x|}]$  und 1777 einen Logarithmus  $[\frac{1}{2a} \log(\frac{a}{|x|})]$  vor. Laplace kam während seiner Arbeit zur theoretischen Fehlerverteilung einer empirischen Verteilung darauf, das Problem der *inversen Wahrscheinlichkeit* oder *Ursachenwahrscheinlichkeit* aus einer Sicht zu betrachten, die der Auffassung von Bayes nahe stand.

Nach der von Gauß und Laplace im Jahre 1810 durchgeführten Synthese der empiristischen und der probabilistischen Auffassung setzte sich die Gaußsche Formel  $e^{-x^2}$  aufgrund ihrer mathematischen Eigenschaften und ihrer guten Übereinstimmung mit den Beobachtungen fast vollständig durch. Die Frage der Verteilung der *Elementarfehler* hatte im Übrigen einen Teil ihrer Bedeutung verloren, nachdem Laplace 1810 den *Zentralen Grenzwertsatz* bewiesen hatte. Dieser Satz zeigt, daß sogar dann, wenn die Wahrscheinlichkeitsverteilung der Fehler keine Normalverteilung ist, die Verteilung der *Mittelwerte* der Fehler gegen eine solche Verteilung strebt, falls die Anzahl der Beobachtungen unbegrenzt wächst.<sup>27</sup> Dieser Umstand verlieh der Gaußschen Form einen entscheidenden Vorteil, auf dem – seit Quetelet und seinem *Durchschnittsmenschen* – die gesamte Statistik des 19. Jahrhunderts beruhte.

Die Ergebnisse von Gauß und Laplace führten also zu einer außerordentlich fundierten Synthese, auf der die Experimentalwissenschaften des 19. Jahrhunderts aufbauten. Diese Synthese vereinigte in sich einerseits die empirischen Arbeiten, die zur Methode der kleinsten Quadrate führten, und andererseits die wahrscheinlichkeitstheoretischen Formalismen, die im Normalverteilungsgesetz und dessen zahlreichen mathematischen Eigenschaften gipfelten. Jedoch sollte es ein ganzes Jahrhundert dauern, bis diese Techniken in den Sozialwissenschaften und insbesondere in den Wirtschaftswissenschaften eingesetzt und formalisiert wurden. Die Gründe hierfür werden wir nachfolgend präzisieren. Eine der möglichen Hypothesen zur Erklärung dieser Verschiebung besteht darin, daß es noch keine Datenaufzeichnungsverfahren gab, die ihrerseits mit der Schaffung moderner Staaten und der Konstruktion der entsprechenden institutionellen Äquivalenzräume zusammenhängen – das heißt mit der Konstruktion von Äquivalenzklassen im institutionellen Bereich und

<sup>27</sup> Genauer gesagt beinhaltet der Zentrale Grenzwertsatz, daß die Verteilungsfunktion einer Summe  $X = \sum_{n=1}^N X_n$  von unabhängigen oder hinreichend schwach korrelierten Zufallsvariablen  $X_n$  nach geeigneter Normierung unter ziemlich allgemeinen Voraussetzungen für  $N \rightarrow \infty$  gegen die Normalverteilung strebt.



in den Bereichen der Rechtsprechung, der Gesetzgebung oder der Gewohnheitsrechte.

## Messungsanpassungen als Grundlage für Übereinkünfte

Wurde ein Richter im 17. Jahrhundert ersucht, einen Konflikt beizulegen, der spätere und somit noch unbekannte Ereignisse implizierte, dann hat er bei seiner Entscheidung gefordert, daß die Prozeßparteien eine Übereinkunft über äquivalente Erwartungswerte erzielen. Daher erschien die Wahrscheinlichkeit, das heißt das Verhältnis zwischen Erwartungswert und Einsatz, als ein Maß, auf dessen Grundlage eine Übereinkunft zwischen Personen aufgebaut werden konnte. Im 18. Jahrhundert versuchten dann die Vermessungsingenieure, Astronomen und Physiker, die sich mit einer Vielfalt unterschiedlicher Naturbeobachtungen konfrontiert sahen, diese Beobachtungen in Messungen zusammenzufassen, die von anderen Wissenschaftlern weiterverwendet werden konnten. Zu diesem Zweck konstruierte man Verfahren (Berechnung von Mittelwerten, Anpassung durch kleinste Quadrate), deren Optimalitätseigenschaften es ermöglichten, Übereinkünfte zu treffen. Und schließlich waren es zu Beginn des 19. Jahrhunderts die Präfekten, die mit dem Stock in der Hand die Wege ihres Departements entlang gingen, um Besonderheiten zu erkunden und zu registrieren. Mit ihren noch unbeholfenen Beobachtungen trugen die Präfekten zum Aufbau der künftigen neuen Administration bei. Die Aufgabe dieser Administration bestand darin, der ganzen Nation Messungen zu liefern, die mit hinreichender Übereinstimmung akzeptiert werden und damit eine sichere Grundlage für öffentliche Debatten und für Entscheidungsfindungen darstellen.

In den betrachteten vier Fällen – das heißt in Bezug auf Richter, Astronomen, Vermessungsingenieure und Präfekten – haben wir es mit Abwandlungen des Wortes *messen* zu tun, das in scheinbar unterschiedlichen Bedeutungen verwendet wird. Die Richter trafen ihre Entscheidungen mit *Maß* (das heißt maßvoll), die Astronomen und die Vermessungsingenieure optimierten ihre *Messungen* und die Präfekten setzten die (auf Messungen beruhenden) *Maßnahmen* ihrer Minister um.<sup>28</sup> Die Nebeneinanderstellung dieser unterschiedlichen Bedeutungen erscheint jedoch nicht als Zufall, sobald man hinter den Messungen (oder den auf ihnen beruhenden Maßnahmen) die Absicht erkennt, eine *Übereinkunft* zwischen prozeßführenden Parteien, unter Wissenschaftlern oder zwischen Bürgern zu erzielen. Demnach haben die scheinbar heteroklitischen Personen und Situationen – die in diesem Buch als Vorläufer im Stammbaum der modernen Verfahren der statistischen Objektivierung auftreten – eine Gemeinsamkeit: sie verbinden die Konstruktion der Objektivität mit der Konstruktion von „Intersubjektivitätstechnologien“ (Hacking, 1991, [120]), das heißt mit *Übereinkunftsformeln*.

<sup>28</sup> Das im französischen Original verwendete Wort *mesure* läßt sich im Deutschen je nach Zusammenhang durch *Messung*, *Maß* oder *Maßnahme* wiedergeben.

Um 1820 existierten jedoch noch keine vereinheitlichten statistischen Verfahren zur Bereitstellung derartiger Übereinkunftsformeln, die heute aufgrund ihrer Solidität allgegenwärtig sind. Die spätere Geschichte läßt sich als eine Folge von Synthesen schildern, die zwischen *a priori* unterschiedlichen Traditionen erfolgten. Ein erstes Beispiel hierfür ist die Laplace-Gauß-Synthese. Diese Synthese erfolgte durch die Verschmelzung der wahrscheinlichkeitstheoretischen Formalisierung der Binomialgesetze von Jakob Bernoulli und Abraham de Moivre mit der Legendreschen Anpassung durch die Methode der kleinsten Quadrate. Weitere Synthesen sollten folgen. In den 1830er Jahren verglich Quetelet die von den Bureaus für Verwaltungsstatistik beobachteten Regelmäßigkeiten mit den Regelmäßigkeiten der astronomischen Messungen und leitete daraus statistische Gesetze ab, die unabhängig von den Individuen waren. Zwischen 1880 und 1900 verknüpften Galton und Pearson die auf Darwin zurückgehenden Fragen der Vererbung mit der von Quetelet festgestellten Normalverteilung der Merkmale der menschlichen Spezies und mit den Anpassungstechniken, die ihren Ursprung in der Theorie der Meßfehler hatten.<sup>29</sup>

---

<sup>29</sup> Am Anfang von Kapitel 9 unterbreiten wir in einer schematischen Darstellung einen Vorschlag für einen Stammbaum der Ökonometrie.

Die Politik der großen Zahlen  
Eine Geschichte der statistischen Denkweise  
Desrosières, A.  
2005, XIII, 434 S., Softcover  
ISBN: 978-3-540-20655-2