Über mathematische Wahrscheinlichkeitsurteile

Letzte Woche haben wir einen kurzen Einblick in das Verständnis von Wahrscheinlichkeitsurteilen im Mittelalter bekommen. Dabei kam heraus, dass Wahrscheinlichkeit eine Art von Ableitung aus bestimmten Aussagen darstellt, die wiederum begründen, warum eine Aussage für Wahr und somit wahrscheinlich zu halten ist. Platt gesagt, ist somit die Wahrscheinlichkeit mit Blick auf das 13. Bis 17. Jhd. ein Für-Wahr-Halten von Aussagen.

Wie gesagt, handelte es sich in der letzten Woche nur um einen kurzen Einblick. Wenn Sie sich erinnern, so habe ich verschiedene Umstände benannt, die für das Verständnis des Wahrscheinlichkeitsbegriffs wichtig sind, wie die Summa Theologiae des Hl. Thomas, den Gnadenstreit und Augustinus Erbsündenlehre.[1] Sofern Sie die Materie tiefer und genauer verstehen wöllten, müssten Sie auch ein Verständnis dieser Texte haben. Das geht aber weit über die Möglichkeiten der Vorlesung hinaus. Schließlich lassen sich mit dem Studium jeder der nötigen Texte ganze Semester verbringen. Verzeihen Sie mir daher bitte, wenn ich auch in dieser Sitzung nur ein Überblick liefere; oder anders gesagt, den Anreiz sich mit den Themen selbstständig näher zu beschäftigen.

Auch wenn man in seinen Darstellungen für gewöhnlich historisch korrekt vorgehen soll, so werde ich dieses Mal eine Ausnahme machen. Ich werde zunächst nicht, wie es sich gehören sollte, bei dem französischen Großvater der mathematischen Wahrscheinlichkeitstheorie beginnen, i.e. Blaise Pascal, sondern bei einem seiner Enkel, nämlich Pierre-Simon Laplace. Der Grund dafür ist, dass Ihnen der Name Laplace vermutlich durch den nach ihm benannten Würfel, den Laplace-Würfel, bekannter sein dürfte. Außerdem ist der Name in der heutigen Zeit mit der Invasion des Bösen für einen jeden Vertreter des freien Willens verbunden, dem so genannten laplaceschen Dämon.

Gehen wir aber zunächst auf den Laplace-Würfel ein. Unter einem solchen lässt sich ein Würfel verstehen, für dessen Flächen gilt, dass wenn er geworfen wird, es für jede Fläche gleichwahrscheinlich ist, dass er auf jener zur Ruhe kommt. Da es sich bei einem Würfel in der Regel um einen regelmäßigen Hexaeder handelt, der somit nur sechs Flächen und insofern nur auf einer dieser Flächen liegen kann, besteht ein Verhältnis von eins zu sechs, dass er auf

[1] Vgl. Augustinus : PL 34 : De genesi ad litteram libri duodecim. lib. X, cap. XIV. n. 24 col. 418 et : Augustinus : PL 32 : De libero arbitrio lib. I. cap. XV n. 32. col. 1238-1239 et cap. XVI n. 34. col. 1239-1240.

einer der Flächen liegen bleibt. Laplace selbst hat die Ermittlung dieses Verhältnisses so bestimmt, dass

„die gesuchte Wahrscheinlichkeit eines Ereignisses durch Zurückführung aller Ereignisse derselben Art auf eine gewisse Anzahl gleich möglicher Fälle, das sind solcher, über deren Existenz wir in gleicher Weise unschlüssig sind, und durch Bestimmung der dem Ereignis günstigen Fälle.“[2]

Angewandt auf den Laplace-Würfel bedeutet dies, dass erst die Möglichkeit abgeschätzt wird, auf welchen Flächen der Würfel zur Ruhe kommen kann. Von diesen möglichen Alternativen können wir bei einem Wurf nicht wissen, welche eintreten wird. Letztlich können wir aber wissen, dass eine der Alternativen eintreten muss, wenn der Würfel geworfen wird. Wir können also sagen, dass wir hinsichtlich unseres Würfels sechs gleich mögliche Fälle des Eintretens eines Ereignisses ausmachen können, jedoch nur einen möglichen Fall, der für die Bestimmung eines Ereignisses günstig ist.

Lassen Sie uns damit ein wenig Schabernack treiben und Laplace Überlegungen vereinfachen indem wir sie Verkomplizieren. Wenn wir aussagen wollen, dass ein bestimmtes Ereignis aufgrund einer Wahrscheinlichkeit eintritt, so können wir dies nur dann, wenn wir wissen, was die notwendigen Bedingungen für den Eintritt eines Ereignisses sind. Das Ereignis selbst, über das wir etwas aussagen wollen, ist in seinem Auftreten durch die notwendigen Bedingungen selbst bedingt. Ohne dass die notwendigen Bedingungen gegeben sind, bleibt also ein Ereignis, wie das Ruhen des Würfels auf einer seiner sechs Flächen aus. Der Ereigniseintritt ist jedoch wiederum nur eine hinreichende Bedingung. Denn nur weil uns die sechs Flächen des Würfels gegeben sind, ist damit in keinster Weise bestimmt, dass er auf einer ganz bestimmten Fläche zur Ruhe kommt. Betrachtet man nämlich die einzelnen Flächen, so muss er nicht auf dieser oder jener Fläche ruhen, sondern kann auch auf einer anderen zur Ruhe kommen. Sagen wir nun ein Wahrscheinlichkeitsurteil aus, so scheint dies einer Aussage über das logische Verhältnis von notwendigen und hinreichenden Bedingungen zu sein. Sprechen wir davon, dass bei einem Laplace-Würfel die Wahrscheinlichkeit für das Eintreten eines Ereignisses für jede Fläche 1/6 ist, so sagen wir auf den ersten Blick ein Implikationsverhältnis aus. Dieses lässt sich durch folgende Implikation aussagen:

„Wenn ein nummerierter Würfel notwendigerweise sechs Flächen hat und geworfen wird, dann kann er entweder auf der Fläche mit der 1 oder auf der Fläche mit der 2 oder auf der

[2] Vgl. Laplace, Philosophischer Versuch über die Wahrschenlichkeit, 4.

Fläche mit der 3 oder der Fläche mit der 4 oder auf der Fläche mit der 5 oder auf der Fläche mit der 6 zur Ruhe kommen."

Das erscheint im Grunde ganz einfach. Diese Einfachheit täuscht jedoch und bietet das schon angesprochene Einfallstor für das Böse, den von vielen Juristen, Philosophen, Neurowissenschaftlern aber kaum einem Naturwissenschaftler ins Feld geführte Laplaceschen Dämon. Die Idee hinter selbigen sagt mit Laplace eigenen Worten aus:

„Eine Intelligenz, welche für einen gegebenen Augenblick alle in der Natur wirkenden Kräfte sowie die gegenseitige Lage der sie zusammensetzenden Elemente kennte, und überdies umfassend genug wäre, um diese gegebenen Größen der Analyses zu unterwerfen, würde in derselben Formel die Bewegung der größten Weltkörper wie der leichtesten Atome umschließen, nichts würde ihr ungewiß sein und Zukunft wie Vergangenheit würden ihr offen vor Augen liegen."

Mit Blick auf unser Würfelbeispiel bedeutet dies, wenn eine Intelligenz um alle notwendigen Bedingungen weiß, aus der alle möglichen Alternativen ersichtlich sind, wüsste es um die Wahrscheinlichkeit des Eintretens aller Ereignisse. So wie wir wissen, dass beim Wurf eines Würfels, da dieser notwendigerweise sechs Flächen hat, er auf einer der sechs Flächen zur Ruhe komme muss, wüsste diese Intelligenz bei der Kenntnis aller Zustände zu einem bestimmten Zeitpunkt im Universum, was nach diesem Zeitpunkt geschehen wird. Begründet ist diese Annahme dadurch, dass ein jeder Zustand der Gegenwart notwendigerweise ein Resultat der Vergangenheit ist. Alles was also in der Vergangenheit passiert ist, ist die notwenige Bedingung für das hier und jetzt. Nehmen wir daher den Laplaceschen Dämon ernst, so können wir sagen, dass Wahrscheinlichkeitsurteile nicht nur ein logisches Verhältnis zwischen notwendigen und hinreichenden Bedingungen angeben, sondern mit dem klassischen Verständnis von Naturgesetzen gleichzusetzen sind. In diesem Fall muss letztlich etwas passieren, was durch die Wahrscheinlichkeitsaussage ausgesagt wird. Dass ein Würfel aufgrund bestimmter Umstände vielleicht nicht zum Ruhen kommt, wäre somit ausgeschlossen. Hieraus müsste dann gefolgert werden, dass das Eintreten einer der möglichen Alternativen gewiss ist.

Machen wir nun die Annahme des laplaceschen Dämons noch ein wenig stärker und lassen Sie uns die Wahrscheinlichkeitsaussage als eine Art von Naturgesetz ansehen. Sofern wir dies annehmen, ist für uns aus einer Wahrscheinlichkeitsaussage erkennbar, welche Alternativen aufgrund bestimmter Bedingungen eintreten müssen. Diese Annahme verändert jedoch das

Verständnis von unserem Implikationsverhältnis grundlegend. Bemühen wir den Laplace-Würfel unter der Annahme, dass Wahrscheinlichkeiten Naturgesetz aussagen erneut. Die sechs Flächen bleiben weiterhin unsere notwendige Bedingung, damit der Würfel zur Ruhe kommen kann. Theoretisch kann er immer noch auf jeder der sechs Flächen zur Ruhe kommen, wenn er geworfen wird. Die Wahrscheinlichkeit für das Ruhen auf einer Fläche ist also für jede Fläche immer noch 1/6. Gibt die Wahrscheinlichkeit von 1/6 jedoch eine Naturgesetzlichkeit an, weil die notwendigen Bedingungen jeden möglichen Fortgang im Weltganzen festlegen, so können wir nicht mehr davon sprechen, dass unser Laplace-Würfel mit einer Wahrscheinlichkeit von 1/6 auf einer der Flächen zur Ruhe kommen kann. Dass etwas etwas kann, sagt nämlich die Möglichkeit des Eintretens von etwas aus. Unter der Möglichkeit von etwas verstehen wir aber für gewöhnlich, dass etwas aktual sein kann, aber nicht aktual sein muss. Wenn die notwendige Bedingung der sechs Flächen des Laplace-Würfels aber jede mögliche Alternative vorgibt, so *muss* der Würfel, wenn er geworfen wird, auf einer der Flächen zur Ruhe kommen. Dann handelt es sich bei dem Ereignis des Ruhens auf einer Fläche aber nicht mehr um eine hinreichende Bedingung, sondern zwangsläufig um eine notwendige Bedingung. Schließlich gilt für jede Fläche, dass wenn der Würfel zur Ruhe kommt, dies für diese Fläche bereits determiniert war.

Unter dieser Interpretation von Wahrscheinlichkeitsurteilen als Naturgesetzlichkeiten, sind solche Urteile keine Implikationen mehr, sondern Bisubjunktionen. Denn sie geben an, dass bei der Gegebenheit notweniger Bedingungen weitere notwendige Bedingungen ableitbar sind. Die Wahrscheinlichkeitsaussage für den Wurf eines Laplace-Würfels würde dann also lauten:

„Dann und nur dann, wenn ein nummerierter Würfel notwendigerweise sechs Flächen hat und geworfen wird, dann muss er notwendigerweise entweder auf der Fläche mit der 1 oder auf der Fläche mit der 2 oder auf der Fläche mit der 3 oder der Fläche mit der 4 oder auf der Fläche mit der 5 oder auf der Fläche mit der 6 zur Ruhe kommen."

Wie wir hieran leicht sehen, ist jede Alternative, die nicht 1, 2, 3, 4, 5, 6 ist, ausgeschlossen. Auf einer der Flächen muss ein geworfener Würfel also zur Ruhe kommen. Das scheinen wir ganz ohne Erfahrung wissen zu können bzw. ohne dass überhaupt jemand einen Würfel werfen muss. Des Weiteren können wir der Aussage entnehmen, dass wenn der Würfel auf einer der Flächen zur Ruhe gekommen ist, das Eintreten des Ereignisses auf die dafür notwendigen Bedingungen zurückzuführen ist. Anders gesagt, weil die notwendigen

Bedingung der Sechsflächigkeit gegeben war, musste ein geworfener Würfel z.B. auf der 1 oder 2 zur Ruhe kommen.

Wir können nun also folgern, wenn eine Wahrscheinlichkeitsaussage eine Naturgesetz in Form einer Bisubjunktion aussagt, dass aus ihr ersichtlich ist, dass das Eintreten irgendeines Ereignisses immer schon feststehend ist. Da die Berechnungen in der Stochastik, die wir alle einmal machen durften, dies zu bestätigen scheinen, ist die Annahme des Laplaceschen Dämons gar nicht so abwegig.

Nichtsdestotrotz, nachdem wir den Dämon stark genug gemacht haben, wollen wir dazu übergehen, ihn zu vernichten und aufzeigen, was die mathematische Wahrscheinlichkeit genauerhin ist.

Führen wir zunächst ein mathematisches Argument an. Die Mathematik ist die Wissenschaft von den Verhältnisses. Führen wir eine mathematische Operation durch, so gibt das Ganze der Operation eine Einheit aus Relationen an. D.h. kommen wir in unseren Operationen zu einem Ergebnis, so lässt bei der Bekanntheit der Operationen rekursiv vom Ende der Operation zum Anfang zurückschreiten. Was durch die Einheit aus Relationen ausgesagt wird, ist immer dasselbe; egal ob ich es vom Ende zum Anfang oder vom Anfang zum Ende betrachte bzw. selbst dann, wenn ich einzelne Operationen regelkonform umforme. Dementsprechend können wir für mathematische Wahrscheinlichkeitsurteile sagen, dass die durch sie ausgesagte Einheit aus Relationen stets gleichbleibend sein muss. Betrachte ich das Verhältnis der Ereignisse bei einem Würfelwurf vor dem Werfen oder nach dem Werfen, macht die Betrachtung keinen Unterschied, da die Verhältnisse doch immer dieselben sind. Wollen wir nun veranschaulichen, dass eine Wahrscheinlichkeitsaussage stets eine Bisubjunktion in Form einer Naturgesetzlichkeit ist, sollten keine Widersprüche auftreten, noch irgendetwas irrational daran erscheinen. Um das Beispiel des Laplace-Würfels zu veranschaulichen, können wir auf eine so genannte Markov-Kette zurückgreifen. Eine Markov-Kette ermöglicht es, dass man verschiedene Zustände und Zustandswechsel anhand von Wahrscheinlichkeiten darstellen kann. Beim Werfen eines Laplace-Würfels können wir von zwei Zuständen sprechen. Zum einen haben wir den Anfangszustand vor dem Wurf. Von diesem Wissen wir, dass sechs Flächen gegeben sind und dass der Würfel auf einer der sechs Flächen zur Ruhe kommen muss. Somit sind sechs mögliche Endzustände gegeben, die auf den Anfangszustand folgen können. Die Wahrscheinlichkeit, dass von dem Anfangszustand auf einen der Endzustände übergegangen wird, ist 1/6. In dieser Angabe ist unser Anfangszustand jedoch noch nicht vollständig bestimmt. Denn aus der Angabe für die Wahrscheinlichkeit des

Eintretens des Endzustandes geht noch nicht die Relation zu den möglichen anderen Alternativen hervor. Da wir bei der Angabe vollständiger Relationen immer auf die Einheit überhaupt kommen müssen, also 1, muss in der Markov-Kette die Wahrscheinlichkeit vermerkt sein, zu der 1/6 in Beziehung steht und durch deren Hinzunahme die Einheit gebildet wird. Diese Wahrscheinlichkeit ist 5/6. Denn wenn eine Alternative realisiert wird, wurden 5 von 6 Alternativen nicht realisiert. Soviel zur Betrachtung unseres Wahrscheinlichkeitsurteils vom Anfang zum Ende. Nehmen wir nun die Endzustände zu unserem Ausgangspunkt, so ist zunächst zu vermuten, da die Relationen ja immer gleichbleiben sind, dass uns auch hier wieder die Wahrscheinlichkeiten von 1/6 und 5/6 begegnen. Das unangenehme bei der Deutung der Bisubjunktion war jedoch, dass wenn ein Ereignis eingetreten ist, es aufgrund der gegebenen Bedingungen mit Notwendigkeit eintreten musste. D.h. sein Eintreten hätte nicht ausbleiben können. In der mathematischen Deutung heißt dies, dass das Eintreten gewiss war. Dementsprechend müsste gesagt werden, dass von jedem der möglichen Endzustände eine Wahrscheinlichkeit von 1, also 100% besteht, dass das Ereignis durch die notwendigen Bedingungen hervorgebracht ist. Auf den ersten Blick erscheint das verständlich. Denn wenn ein Ereignis nur durch bestimmte Bedingungen hervorgebracht werden kann, dann lässt sich das Ereignis vollständig auf diese Bedingungen zurückführen. Wie die Gegenwart notwendigerweise aus unserer Vergangenheit folgt, folgt auch das Eintreten eines Ereignisses vollkommen aus den dafür notwendigen Bedingungen. Hier begehen wir jedoch einen Denkfehler, wie die Markov-Kette zeigt. Die Endzustände sind jeweils mögliche Ereignisse, die mit Blick auf unseren Anfangszustand hervorgehen können. Als solche Endzustände stehen auch sie zueinander in Beziehung und sind daher nicht unabhängig voneinander. Wenn wir nun behaupten, dass das Eintreten eines Ereignisses zu 100% aus dem Anfangszustand hervorgeht und dies für jedes der sechs Ereignisse gültig sein soll, so haben wir sechs Endzustände, die zu 100% aus dem Anfangszustand hervorgehen. In diesem Sinne hätten wir zwei absurde Schlussfolgerungen. Zum einen müssten wir sagen, dass ein jeder der möglichen Endzustände bei einem Würfelwurf eintritt, denn die Endzustände folgen mit Gewissheit aus dem Anfangszustand. Das ein Würfel aber auf allen sechs Flächen zugleich zur Ruhe kommt, ist widersinnig und daher zu verwerfen. Zum anderen stehen, da die Endzustände sich aus den Anfangszuständen ergeben, diese selbst wieder in Beziehung zueinander, sodass sie eine Einheit aus Relationen bilden. Weil wir in der Wahrscheinlichkeitstheorie die Einheit durch eine 1 oder 100% ausdrücken, ist es ausgeschlossen, dass die Summe der Wahrscheinlichkeitswerte der Endzustände über 100%

liegt. In unserem Beispiel liegt sie aber bei 600%; was eine Unmöglichkeit darstellt und insofern falsch sein muss.

(Markov-Vollständig!)

Was ist aber falsch an unserer Überlegung? An unserer Deutung eines Wahrscheinlichkeitsurteils als Bisubjunktion scheint doch zunächst nichts falsch. Denn wenn eine notwendige Bedingung gegeben ist und diese determiniert, welche Alternativen es gibt, so treten die Alternativen auch mit Gewissheit ein. Hier liegt nun jedoch der Fehler. Nur weil eine notwendige Bedingung vorliegt, heißt dies nicht, dass sich irgendeine der durch sie angegebenen Alternativen aktualisieren wird, noch dass wenn eine Alternative eintritt, ihr Eintreten gewiss war. Ein Umstand, den viele Philosophen, Juristen und andere Wissenschaftler, die gern mit Wahrscheinlichkeiten arbeiten gern übersehen. Die notwendige Bedingung gibt nur einen festen Rahmen von Alternativen vor, die eintreten können, welche der Alternativen realisiert wird, bleibt zufällig. Daher gibt ein Wahrscheinlichkeitsurteil streng genommen ein Naturgesetz im alten Sinne an. D.h. es beschreibt, wenn wir es auf die Natur anwenden, den Rahmen innerhalb dessen etwas geschieht, also sich bewegt, oder wenn es rein mathematisch gedeutet wird, so gibt es die Bedingungen für mögliche relationale Bestimmungen vor. In keinster Weise wird also etwas darüber ausgesagt, dass ein Ereignis auch wirklich eintritt, noch dass der Eintritt eines Ereignisses vorherbestimmt sei. Die Annahme, dass sich aus Wahrscheinlichkeitsurteilen ein Determinismus ableiten lasse, ist somit hinfällig.

Dieses Verständnis von Wahrscheinlichkeitsurteilen, dass Wahrscheinlichkeitsurteile nur einen bestimmten Rahmen von Alternativen aussagen, wird nun ganz ihrer ursprünglichen Geschichte gerecht. Diese beginnt im Wesentlichen im 16. Jhd. Damals wie heute neigten die Menschen zum Glückspiel. Vor allem Würfel und Kartenspiele erfreuten sich damals großer Beliebtheit. Da man den Spaß am Spiel oftmals mit Geldeinsätzen koppelte, überlegte so manch ein Spieler, wie er seinen Gewinn maximieren konnte bzw. wie er abschätzen kann, ob er noch gewinnen kann oder auf dem besten Wege ist zu verlieren. Einige der exzessivsten Spieler waren Mathematiker, wie z.B. der Gelehrte Gerolamo Cardano, der eines der ersten mathematischen Werke verfasst hat, die sich mit Würfelspielen auseinandersetzt[3] oder aber Galileo Galilei. Dabei geht es vor allem darum herauszufinden, wenn man zwei Würfel wirft,

[3] Cardano, Gerolamo: Liber de Ludo Aleae.

welche Summe an Augenzahlen häufiger zu erwarten sind, als andere bzw. wie oft es vorkommen kann, dass zwei Würfel dieselbe Augenzahl aufweisen. Alles zum Zweck um erfolgreich auf das Auftreten bestimmter Augenzahlen wetten zu können.

Vor ein wirklich großes Problem stellte die Mathematiker jedoch die Berechnung der Gewinnausschüttung, wenn man ein Spiel z.B. nur nach erfolgreichen 10 Spielrunden gewinnen konnte und das Spiel aber unvollendet blieb. Der Gewinn bzw. Jackpot des Spiels soll schließlich an den Gewinner des ganzen Spiels vollständig ausgezahlt werden. Da aber der Fall eintreten kann, dass wenn drei Leute spielen, einer 8 Runden gewonnen hat und die anderen bloß 6 Runden und das Spiel aus irgendwelchen Umständen abgebrochen werden muss, kann der Gewinn nicht vollständig an irgendeinen der Spieler ausgezahlt werden. Weil man die Mühen des Spiels aber auf sich genommen hat und auch den Gewinn wollte, suchte man nach einem Maßstab, nach dem sich der Jackpot aufteilen ließ. Die Lösung des Problems ist zugleich die Sternstunde der mathematischen Wahrscheinlichkeitstheorie.

Die Lösung für das Problem lieferten erstmals der Mathematiker Blaise Pascal und Pierre Fermat im Rahmen eines Briefwechsels.[4] In diesem vereinfacht Pascal das Problem zunächst auf eine überschaubare Anzahl an Runde, um zu gewinnen und einen Einsatz, mit dem sich gut rechnen lässt. So geht er davon aus, dass wir zwei Spieler haben, die drei Runden gewinnen müssen, um einen Einsatz von 64 Geldeinheiten zu gewinnen. Jeder der Spieler hat bereits zwei Spiele gewonnen, sodass jeder von ihnen nur noch einmal gewinnen muss. Soll der Gewinn nun verhältnismäßig aufgeteilt werden, so bekommt jeder der beiden Spieler 32 Geldeinheiten. Doch warum eigentlich? Aus den Spielregeln wissen wir, dass der gewinnt, der drei Runden gewonnen hat. Jedem der Spieler fehlt nun noch eine Runde, damit er gewinnen kann. Insgesamt haben wir zwei mögliche Ausgänge des Spiels, für jeden Spieler ist nur ein Ereignis gegeben, bei dem er gewinnen kann. Gemessen an den Möglichkeiten des Gewinns hat jeder der beiden Spieler von den zwei möglichen Gewinnereignissen einen Anteil daran. Man hat also für jeden Spieler ein Gewinnverhältnis von 1 zu 2 oder etwas moderner gesagt, jeder der Spieler hat eine 50%-ige Wahrscheinlichkeit zu gewinnen. Die Wahrscheinlichkeit auf den Gewinn angewandt ergibt die Hälfte des Einsatzes für die beiden Spieler. Pascal kam nun auf die Idee, dass man von den letzten Zuständen des Spiels immer weiter im Spielverlauf zurückgehen kann, um jede mögliche Gewinnverteilung zu ermitteln bzw. die Wahrscheinlichkeit für den Gewinn eines Spieles ab einem beliebigen Punkt im Spiel zu ermitteln. So lässt sich nun z.B. annehmen, dass einer der Spieler nur ein Spiel

[4] Oeuvres de Fermat. Tome Deucième. S. 290f.

gewonnen hat und der andere bereits zwei. Folgt man Pascals Methode, so muss nun rückläufig festgestellt werden, wie viele Möglichkeiten einem jeden Spieler zugesprochen werden können, damit er das Spiel gewinnt. Zwei der Möglichkeiten sind uns bereits bekannt, wenn nämlich beide zwei Spiele gewonnen haben und entweder der eine oder der andere die nächste Runde gewinnt. Eine weitere Möglichkeit des Gewinns besteht, wenn der Spieler, der nur einmal gewonnen hat, zweimal gewinnt. Einen weiteren Anteil am Gewinn hätte derselbe Spieler aber auch genau dann, wenn er einmal gewinnt und dann einmal verliert. Pascals Gedanke ist nun, dass man den einzelnen Spielern die Anzahl an Aussagen zuweist, die die Möglichkeit zu gewinnen aussagen. Denn daraus ergibt sich ein Wert aus dem ersichtlich ist, wie sehr ein Spieler hoffen kann, zu gewinnen. Insgesamt lassen sich vier Aussagen über mögliche Gewinne bilden. Der Spieler, der bereits zwei Runden gewonnen hat, hat letztlich drei Möglichkeiten, um zu gewinnen. Der Spieler, der nur ein Spiel gewonnen hat, hat indes nur eine Möglichkeit den Gewinn zu erzielen. Die Chancen, dass der Spieler gewinnt, der nur einmal gewonnen hat, stehen also 1 zu 4 bzw. 25%, während die Chancen des anderen Spielers 3 zu 4 bzw. 75% stehen. Nach diesen Wahrscheinlichkeiten wäre also auch der Gewinn bei einer Unterbrechung auszuzahlen, was bei einem Jackpot von 64 Geldeinheiten 48 Geldeinheiten für den einen und 16 Geldeinheiten für den anderen macht.

Was offenbart nun aber Pascals Lösung hinsichtlich des Charakters von Wahrscheinlichkeitsurteilen? Wir können zum einen sehen, dass die Bildung von Wahrscheinlichkeitsurteilen nicht im luftleeren Raum stattfindet. Um in unserem Beispiel ein Wahrscheinlichkeitsurteil formulieren zu können, müssen wir um die Spielregeln wissen. Denn nur allein in Bezug zu den Spielregeln können wir aussagen, welche Relationen und somit Alternativen in den einzelnen Phasen des Spiels bestehen können. Sofern wir nämlich nicht wüssten, dass zwei Spieler spielen und nur der den Sieg davon trägt, der drei Runden gewonnen hat, wüssten wir nicht, nach welchen Regeln man bei den Verhaltensweisen der Spieler fort- bzw. zurückschreiten muss. Es würden also die Regeln fehlen, die für uns erkennbar machen, dass sich die Verhaltensweisen der Spieler aufeinander beziehen, sodass eine relationale Einheit ausgesagt werden könnte. Da die relationale Einheit in unserem Beispiel das Spiel selbst bildet, wüssten wir also ohne Spielregeln auch nicht, dass überhaupt ein Spiel gespielt wird. Diese Erkenntnis mag nun trivial sein, vor allem, wenn wir uns den Laplace-Würfel noch einmal vergegenwärtigen. Denn schließlich ist es auch bei diesem so, dass wir aufgrund der Definition des Würfels Wahrscheinlichkeitsaussagen folgern können und sofern wir dies wöllten, könnten wir aufgrund der Definition sogar für mehrere Würfelwürfe hintereinander die Wahrscheinlichkeit für das Ruhen des Würfels auf bestimmten Flächen aussagen. Auch beim Würfel wäre dies ohne einen vorher bestimmten Rahmen nicht möglich. Besonders ist unsere Erkenntnis in bestimmter Hinsicht jedoch schon. Denn wenn wir versuchen würden Wahrscheinlichkeitsurteile als Aussagen zu klassifizieren, so müssen sie als abgeleitet Aussagen verstanden werden. In ihrem Bestehen und somit in ihrem Wahrheitswert sind sie immer durch andere Aussagen bedingt. Würden diese Aussagen wegfallen, wäre die Bildung eines Wahrscheinlichkeitsurteils überhaupt nicht möglich. Wüssten wir z.B. nicht um die Eigenschaften eines Würfels, dann könnten wir auch keine Aussage über seinen Wurf noch über die Möglichkeiten seines Ruhens formulieren.

Was heißt das jetzt aber für den Wahrheitswert von Wahrscheinlichkeitsurteilen? Augenscheinlich können Wahrscheinlichkeitsurteile doch niemals wahr oder falsch sein, da sie etwas aussagen, dass irgendwie dazwischen liegt. Wie wir aber festgestellt haben, ist ein Wahrscheinlichkeitsurteil immer durch ihm vorgelagerte Aussagen bedingt. Das Urteil besteht mithin nur mit ständigem Blick auf die vorauszusetzenden Aussagen. Aufgrund dieser Eigenschaft lässt sich auch in bestimmter Weise von der Wahrheit einer Wahrscheinlichkeitsaussage sprechen. Nur besteht diese Wahrheit nicht in der Referenz einer Aussage auf ein Ding in der Welt. Denn dies würde bedeuten, dass das Ausgesagte einer Aussage mit den Eigenschaften eines Dings in der Welt übereinstimmt. Vielmehr ergibt sich die Wahrheit der Wahrscheinlichkeitsaussage aus ihrer Kohärenz mit den vorauszusetzenden Aussagen. Oder etwas leichter verständlich formuliert: Eine Wahrscheinlichkeitsaussage ist dann wahr, wenn sie sich tatsächlich aus den vorauszusetzenden Aussagen ableiten lässt. Die Wahrheit lässt sich somit als Ableitbarkeit definieren, während die Falschheit schlichtweg dann besteht, wenn eine Wahrscheinlichkeitsaussage nicht ableitbar ist. Somit bekommen wir nun ein sehr gutes Verständnis für mathematische Wahrscheinlichkeitsurteile. Denn letztlich lassen sie sich als bedingte Aussagen verstehen, die hinsichtlich bestimmter Ableitbarkeitskriterien wahr oder falsch sein können.

Dieses Verständnis mutet jedoch hinsichtlich unseres Alltags bzw. auch so mancher Wissenschaft als sehr seltsam an. Werden wir z.B. gefragt, für wie wahrscheinlich wir es halten, dass Morgen gutes Wetter ist, werden wir uns in unserem Wahrscheinlichkeitsurteil wohl kaum auf eine Menge von Aussagen stützen, aus der wir unser Urteil abgeleitet haben. Vielleicht verhält es sich aber auch einfach so, dass wir überhaupt kein Wahrscheinlichkeitsurteil fällen, sondern letztlich nur abschätzen, was für ein Wetter wir am nächsten Tag gern hätten. Wir beantworten also schlichtweg eine andere Frage, wenn wir auf die ursprüngliche Frage keine Antwort wissen. Ein Umstand, der durch psychologische Untersuchungen bestätigt ist.

Wenn es nun aber stimmt, dass Wahrscheinlichkeitsurteile abgeleitete Aussagen sind, dann müsste uns hinsichtlich des pascal‘schen Beispiels und des Laplace-Würfels etwas nicht Uninteressantes auffallen. In beiden Fällen wissen wir um die Rahmenbedingungen, aus denen sich Wahrscheinlichkeitsurteile ableiten lassen. In beiden Fällen können wir, ohne auf Erfahrung zurückgreifen zu müssen, sagen, wie die Verhältnisse der Ereignisse zueinander sind. Doch was sollen das für Ereignisse sein? Sind es Ereignisse, die sich tatsächlich in unserer Welt vollziehen können? Eine Wahrscheinlichkeitsaussage, die nur aus Spielregeln oder der Definition eines geometrischen Körpers abgeleitet ist, kann sich doch überhaupt nicht auf unsere Welt beziehen. Denn damit ein Würfel geworfen werden kann, bedarf es doch außer dem Würfel noch viel mehr, z.B. eine Welt in der sich der Würfel befindet und ein Ding, dass Kraft auf den Würfel ausübt, sodass er in der Welt zur Ruhe kommen kann. Auch die Wahrscheinlichkeitsaussagen im pascal’sche Spiel lassen diese Umstände außer Acht, weshalb das Spiel, wie der Würfel allein eine Vorstellung in unseren Gedanken sein kann. Ohne Weiteres sind die Überlegungen daher überhaupt nicht auf die Welt anwendbar. Das hört sich für uns, die wir gewohnt sind, dass uns Wahrscheinlichkeitsurteile über das Wetter oder den Ausgang einer Wahl vorgelegt werden, sicherlich seltsam an. Schließlich scheinen uns Wahrscheinlichkeitsurteile ständig etwas über die Welt zu erzählen. In dieser primitiven Form, wie uns die Urteile durch Pascal und Laplace vorliegen, stimmt das aber nicht. Dessen sind sich die Autoren auch bewusst. Da für beide die Operationen der Mathematik zunächst im Kopf stattfinden und gleichfalls den Bedingungen unterliegen, die unser Verstand ihnen vorgibt. So gelten die Wahrscheinlichkeiten des Laplace-Würfels doch nur, wenn er ein Objekt der euklidischen Geometrie ist. Euklids Geometrie beschreibt ideale Körper, die allein kraft dadurch existieren, dass wir sie nach den Grundsätzen seiner Geometrie denken. Da diese Körper über keine Masse verfügen und auch keinen Kräften ausgesetzt sind, lässt sich leicht davon sprechen, dass wenn ein Würfel geworfen wird, er auf jeder seiner Flächen gleichwahrscheinlich zur Ruhe kommen kann. In der wirklichen Welt unterliegen würfelförmige Objekte aber weitaus mehr Bedingungen. Werden diese Bedingungen nicht auch in die Bildung eines Wahrscheinlichkeitsurteils einbezogen, besteht nicht einmal die Möglichkeit, dass diese Urteile uns etwas über die Welt verraten. Denn sie würden etwas über die Welt der euklidischen Geometrie oder irgendeiner anderen Welt aussagen.

Dieser Umstand, dass die Bedingungen der Welt immer einzubeziehen sind, damit über die Welt auch etwas ausgesagt werden kann, ist seit der Idealisierung der Mechanik Sir Isaac Newtons als das Werkzeug zur Voraussage für Körperbewegungen leider immer mehr in Vergessenheit geraten bzw. wird auch explizit abgelehnt. Warum aber seit der Newtonschen Mechanik? In seiner Schrift Philosophiae Naturalis Principia Mathematica, beschreibt Newton sein Vorgehen als Induktion. Dabei abstrahiert man von den konkreten Gegenständen der Erfahrung, um etwas Allgemeingültiges schlussfolgern zu können. Er betont dabei jedoch, dass seine Mechanik ein gedankliches Konstrukt ist, die beschreibt, wie sich natürliche Körper bewegen müssten, wenn sie sich wie die Bestandteile des gedanklichen Konstrukts verhalten würden. Nach Newton bleiben die Aussagen seiner Mechanik also reine Aussagen über Gedankenoperationen und nicht über die Welt selbst. Da die ihm nachfolgenden Gelehrten jedoch feststellten, dass man anhand seiner Mechanik sehr gute Prognosen über die Bewegung von Körpern machen konnte, setzte man das gedankliche Konstrukt mit einer Beschreibung über die Beschaffenheit der Welt gleich. Dementsprechend ging man in der Physik, aber auch damaligen Randbereichen der Mathematik, wie z.B. der Ökonomie, immer mehr dazu über, dass man Gedankenkonstrukte, die durch Induktion gewonnen wurden, mit Aussagen über die Beschaffenheit der Welt gleichsetzte. Daher kam man besonders durch den Einfluss des Ökonomen und Philosophen John Stuart Mill[5] in der Wahrscheinlichkeitstheorie zu der Annahme, dass auch Wahrscheinlichkeitsaussagen, die unter Bedingungen idealisierter Welten formuliert worden sind, eine exakte Beschreibung unserer Welt sind. Schließlich kann man ja auch scheinbar mit der Theorie eines Laplace-Würfels sehr gut voraussagen, wie die Wahrscheinlichkeiten für das Eintreten des Ruhens eines Würfels auf einer bestimmten Fläche sind.

Ein Zeitgenossen John Stuart Mills, der Mathematiker und Logiker John Venn, hat aber zu Recht gezeigt, dass die Bedingungen der Welt unbedingt zu berücksichtigen sind. Andernfalls gelangen wir nämlich zu Schlussfolgerungen, die absurd sind bzw. vor dem Hintergrund anderer Wissenschaften schlichtweg unverständlich bleiben. Venns Paradebeispiel sind die irrigen Wahrscheinlichkeitsaussagen über einen Münzwurf. In einer idealisierten Welt würden wir folgern, wenn die Münze geworfen wird, so ist es gleichwahrscheinlich, dass die Münze entweder Kopf oder Zahl angezeigt. In dieser Überlegung spielt der Anfangszustand, aus dem heraus die Münze geworfen wird, keine Rolle. Es ist schlichtweg egal, ob die Münze in meiner Hand, bevor ich sie werfe, auf Kopf oder Zahl liegt. Werfe ich immer mit derselben

[5] John Stuart Mill: A System of Logic.

Kraft, so würde die Wahrscheinlichkeit jeweils bei 50% für jede der Seiten liegen. Vor dem Hintergrund der Mechanik ist eine solche Wahrscheinlichkeitsaussage aber nicht zu verstehen. Wird ein Körper nämlich immer mit derselben Intensität an Kraft bewegt, ohne dass er in seiner Bewegung durch den Eindruck einer anderen Kraft beeinflusst wird, so muss er zwangsläufig dieselbe Bewegung immer wieder vollziehen. Liegt meine Münze daher vor dem Wurf immer auf der Seite mit der Zahl und zeigt sie beim ersten Wurf Kopf an, so ist bei einer Wiederholung des Wurfes ausgeschlossen, dass ein anderes Ergebnis herauskommen könnte. Denn die Münze wird nach den Gesetzen der Mechanik die gleiche Bewegung erneut vollziehen, sodass sie zwangsläufig wieder Kopf anzeigen wird. Sind die Gesetze der Mechanik in unserer Welt gültig, ist die Annahme einer 50% Wahrscheinlichkeit für das Auftreten beider Ereignisse unverständlich. Ignorieren wir die Mechanik und betrachten die Münze einfach von ihrem Material her, so können wir vor dem Hintergrund der Feststoffgeometrie sagen, dass solch eine Münze nicht überall gleichschwer ist, weil ihr Material an einigen Stellen verunreinigt ist oder sie schlichtweg so alt ist, dass sie an einigen Stellen eine mehr oder weniger dicke Patina angesetzt hat. Dementsprechend kann es sein, dass eine Seite der Münze schwerer ist als die ihr gegenüberliegende Seite. Da Körper erfahrungsgemäß mit der schwereren Seite zu Boden gehen, ist wiederum die Annahme einer 50% Wahrscheinlichkeit unverständlich.

Nun sind die angeführten Fachbereiche nur zwei Gebiete, die sich mit den Bedingungen beschäftigen, die unserer Welt zugrunde liegen. Wir könnten viele weitere anführen, die uns aufzeigen, dass Wahrscheinlichkeitsurteile, die diese Gebiete nicht mit einbeziehen, nichts über unsere Welt aussagen können, weil sie nach wissenschaftlichen Maßstäben schlichtweg unverständlich sind. Sofern man also möchte, dass Wahrscheinlichkeitsaussagen etwas über die Welt aussagen, müssten alle uns bekannten Bedingungen der Welt mit einbezogen und mathematisiert werden. Um eine präzise und gültige Wahrscheinlichkeitsaussage über einen Münzwurf zu formulieren, müsste sie also aus dem Ganzen der naturwissenschaftlichen Erkenntnisse abgeleitet werden. Denn allein dann ist ihr zumindest der Anspruch inhärent, dass sie etwas über die Welt aussagt, sofern sich die Körper in der Welt so verhalten, wie sie durch die anderen Wissenschaften beschrieben werden. Betrachten wir diesen Anspruch an eine Wahrscheinlichkeitsaussage genauer, so können wir feststellen, dass wir als Menschen vermutlich nicht dazu in der Lage sind, eine solche Aussage zu formulieren. Schließlich setzt dies voraus, dass der Mensch, der die Aussage formuliert, Kenntnis von allen naturwissenschaftlichen Fachbereichen und Erkenntnis hat und diese zudem widerspruchsfrei miteinander verknüpfen kann. Gerade der letztere Umstand erweist sich jedoch als schwierig, da die Erkenntnisse bzw. Grundsätze einiger Wissenschaftszweige miteinander unvereinbar zu sein scheinen, so wie es z.B. die Grundsätze der Thermodynamik und der Mechanik sind. Nehmen wir dieses Problem für die Bildung von mathematischen Wahrscheinlichkeitsaussagen über die Welt ernst, so ergäbe sich nach den Maßstäben der Wissenschaften des 17. Jhd. ein gar schauerliches Urteil über unsere heutigen Sozial- oder Wirtschaftswissenschaften oder über den morgigen Wetterbericht, die meistens die Bedingungen der Welt ignorieren. So würde z.B. ein Blaise Pascal oder Isaac Newton urteilen, dass die Wahrscheinlichkeitsaussagen dieser Wissenschaften denselben Stellenwert wie die Aussagen der okkulten Wissenschaften haben. Der Wetterbericht, wie die Wahrscheinlichkeitsaussage eines Soziologen oder Politikwissenschaftlers wäre also gleichzusetzen mit dem Urteil einer Wahrsagerin oder eines Hühnerknochenlegers, das die Zukunft voraussagen soll.

Berücksichtigen wir die Bedingungen, die für die Bildung von mathematischen Wahrscheinlichkeitsurteilen über die Welt nötig sind, erscheint deren Formulierung durch die modernen Wissenschaften schlichtweg als sinnlos und man stellt sich die Frage, inwiefern überhaupt ein Sinn oder Nutzen aufgefunden werden kann.

Dass unsere Wahrscheinlichkeitsurteile, die wir nur auf der Grundlage ganz begrenzter Wissensgebiete bilden, jedoch nicht ganz sinnlos sind, lässt sich durch die Art und Weise ihres Gebrauchs aufzeigen. Wir konnten zwar zeigen, dass wir nicht wirklich dazu in der Lage sind, dass wir Wahrscheinlichkeitsurteile bilden können, die etwas über die Welt aussagen, das heißt aber nicht, dass wir durch sie nicht etwas über uns in Bezug zur Welt aussagen könnten. Wie wir bereits gesagt hatten, ist die Mathematik die Wissenschaft von den Relationen. Relationen müssen jedoch erst bestimmt werden, da sie von sich aus nicht in der Welt zu finden sind. Veranschaulichen lässt sich dies sehr gut an der Größe eines Gebäudes. Wir können zwar sagen, dass ein Gebäude groß ist, aber wenn wir keinen Maßstab bestimmen, nach dem wir wiederum die Größe des Gebäudes bestimmen, bliebe die Größe des Gebäudes unbestimmt. Um es sinngemäß mit Leonard Euler zu sagen, brauchen wir zur Bestimmung von Größe immer schon die Kenntnis einer anderen Größe, damit wir sagen können, dass unser Gebäude größer oder kleiner ist, als die uns bekannte Größe bzw. ob die uns bekannte Größe größer oder kleiner ist als das Gebäude, dessen Größe wir bestimmen wollen. Gleiches gilt nun auch für die Relationen, die durch Wahrscheinlichkeitsurteile ausgesagt werden. Die jeweiligen Relationen müssen erst kraft unseres Verstandes gesetzt werden. Da die Grundlage für diese Festsetzung bestimmte Aussagen einer Theorie oder über die Empirie sind, werden aus logischer Sicht also Relationen zwischen uns bekannten Aussagen gesetzt. In Anbetracht der Theorien oder Aussagen über die Empirie, zeigen uns nun mathematische Wahrscheinlichkeitsurteile, wie wir Struktur ins Chaos der Welt bringen. Versuchen wir nämlich bestimmte Prozesse durch mathematische Wahrscheinlichkeitsurteile auszudrücken, so geben wir mit der Hilfe der Mathematik eine Vielzahl von Denkoperationen an, die zwischen verschiedensten Elementen von Aussagen Relationen festlegen, sodass letztlich eine für uns verstehbare und nachvollziehbare Aussage herauskommt.

Doch was ist genau hin damit gemeint? Machen wir z.B. Beobachtungen über das Sozialverhalten von Menschen über einen Tag hinweg, so könnten wir Unmengen an empirischen Aussagen formulieren. Wir hätten also eine wahre Flut an empirischen Daten. Der einzige Zusammenhang, der zunächst in unserem Denken für dies Daten besteht, ist, dass wir sie Minute für Minute gesammelt haben. In dieser Weise besteht also lediglich ein zeitlicher Zusammenhang zwischen den empirischen Daten. In welcher anderen Weise die Daten aufeinander bezogen sein können bzw. wie vielleicht die Bestandteile eines Ereignisses das Auftreten eines anderen Ereignisses bedingt, geht daraus nicht hervor. Insofern herrscht zwischen den beobachteten Ereignissen, wenn wir uns die zeitliche Abfolge einmal wegdenken, zunächst Chaos in unserem Verstand. Erst indem wir die Bestandteile der Ereignisse wechselseitig aufeinander beziehen, ergibt sich für uns eine bestimmte Ordnung. Indem wir z.B. festlegen, dass zwei Bestandteile eines Ereignisses zusammengehören und nur durch diese Zusammengehörigkeit ein anderes Ereignisses hat zu einem anderen Zeitpunkt entstehen können, ergibt sich für uns eine relationale Einheit und somit eine bestimmte vorstellbare Struktur in unserem Verstand. In diesem Sinne geben Wahrscheinlichkeitsurteile Möglichkeiten vorstellbarer Strukturen an. Bezogen auf unsere Erfahrung heißt das, dass wenn ein Wahrscheinlichkeitsurteil aufgrund einer bestimmten Theorie und bestimmter empirischer Daten formuliert wird, dass damit eine mögliche Ordnung für unseren Verstand ausgesagt wird, nach der dieser z.B. ein Chaos an Eindrücken durch die Festlegung bestimmter Relationen strukturieren kann, um somit irgendwie in der Welt zurecht zukommen. Wohlwissend, dass die Welt der durch den Verstand geschaffenen Ordnung auch nicht entsprechen könnte. Insofern gebrauchen wir Wahrscheinlichkeitsurteile, um unserer Erfahrung oder einfach bestimmten Daten eine verstehbare Struktur zu geben, mit dem Zweck, dass wir nach dieser Struktur unser Handeln ausrichten.

Abschließend ergibt sich nun ein recht klares Bild von Wahrscheinlichkeitsurteilen bzw. von der Wahrscheinlichkeit selbst. Die mathematische Wahrscheinlichkeit gibt streng genommen an, welche Relationen bestehen müssen und können, wenn man sie aus bestimmten vorausgesetzten Aussagen ableitet. Dabei gibt sie zugleich an, welche Denkoperationen der Festlegung der Relationen zugrundeliegen müssen, damit die Wahrscheinlichkeitsaussage verstehbar ist. Denn die Wahrscheinlichkeit besteht ja aufgrund der Denkoperationen, nach denen eine bestimmte Struktur und somit Einheit gebildet wird. Daraus ergibt sich nun für den Umgang mit Wahrscheinlichkeitsaussagen, wie z.B. beim Wetterbericht interessantes. Werden uns nur irgendwelche Zahlen mit Prozentzeichen hingeschmettert, so bleibt für uns unverständlich, was diese Zahlen aussagen sollen. Weder wissen wir durch die Zahl nämlich, welche Denkoperationen notwendig sind, um zu der Zahl zu kommen, noch welche Daten oder Theorien für deren Bildung vorausgesetzt sind, damit eine verstehbare Aussage zustande kommt. Gleiches gilt indes für die Verwendung von Wahrscheinlichkeitsurteilen in der Jurisprudenz, wenn man z.B. auf Statistiken zurückgreift, ohne dass erklärt wird, welcher Algorithmus der Berechnung der aus ihr gefolgerten Wahrscheinlichkeiten zugrundeliegt. Was dadurch ausgesagt werden soll, wenn man auf solche Wahrscheinlichkeiten verweist, bleibt schlichtweg dunkel und unverständlich. Denn für den Verstand kann sich durch eine Zahl allein keine Ordnung ergeben.