Sitzungsberichte

der

philosophisch-philologischen
und der
historischen Classe

der

k. b. Akademie der Wissenschaften

zu München.

Jahrgang 1892.

München
Verlag der K. Akademie
1893.

In Commission bei G. Franz.

Philosophisch-philologische Classe.

Sitzung vom 5. März 1892.

Herr Stumpf hielt einen Vortrag:

„Ueber den Begriff der mathematischen Wahrscheinlichkeit."

Das Anwendungsgebiet der Wahrscheinlichkeitsrechnung ist in beständiger Ausdehnung begriffen. Wenn wir auch nicht mehr mit Pascal durch die an Glücksspielen entwickelten Begriffe den religiösen Glauben stützen oder gar mit Craig das Jahr ausrechnen wollen, in welchem die abnehmende Wahrscheinlichkeit der evangelischen Berichte so klein geworden sein wird, dass Christus wiederkommen muss; wenn wir auch den ausführlichen Theorien geschichtlicher und gerichtlicher Zeugnisse, wie sie Condorcet und Poisson entwickelten, schon wegen ihrer Unanwendbarkeit wenig Interesse mehr entgegenbringen: so überraschen uns doch die Naturwissenschaften von den Tagen des Laplace und Gauss bis zu denen Maxwells und Boltzmanns mit immer neuen weittragenden Verwertungen. Aber auch in die Moral- und Geisteswissenschaften ist die Wahrscheinlichkeitsrechnung wiedereingedrungen, sowol durch ihre Beziehungen zur Statistik, der ja alle Thatsachenwissenschaften ohne Ausnahme unterworfen sind, als durch die exacteren Principien der Hypothesenschätzung, die sie an die Hand gibt. Wenn der

Philosoph des Unbewussten mit der Wahrscheinlichkeit 0,999 999 999 6 auf das Mitwirken geistiger Ursachen bei der Embryo-Entwickelung schloss, so war dies freilich ein Fehlschluss, aber die Form des Schlusses war correct. Auch die Geisteswissenschaft im engeren Sinn, die Psychologie, und selbst die Wissenschaft vom Schönen öffneten ihre Pforten; gerade von hier aus wurde Fechner wieder zu Bereicherungen der mathematischen Theorie geführt. Die modernen Verfechter der Telepathie stützen sich auf mathematische Wahrscheinlichkeit. Endlich sind auch die Grundprobleme der Erkenntnistheorie, das der Induction, des Causalgesetzes, der Aussenwelt, sogar die Frage nach einem letzten gemeinsamen Princip aller Dinge in älterer und neuerer Zeit von verschiedenen Seiten unter den Gesichtspunct der Wahrscheinlichkeitsrechnung gestellt. Grund genug für den Philosophen, dem Werkzeug, mit welchem hier operirt wird, besondere Aufmerksamkeit zu widmen.

Das Interesse der Mathematiker selbst, denen wir die Ausbildung der Lehre verdanken, galt allezeit vorwiegend der Lösung von Aufgaben, die aus triftigen methodischen Gründen hauptsächlich den Glücksspielen entnommen waren, sowie der Entwickelung bestimmter Rechnungsmethoden, die zur Lösung analoger Aufgaben führen konnten. Erst die immer manichfacheren und kühneren Anwendungen der aufgefundenen Principien wurden hie und da, und wiederum weniger für die Fachmänner als für Naturforscher und Philosophen, Veranlassung, den Begriff der mathematischen Wahrscheinlichkeit selbst und die darin etwa in Hinsicht seiner Anwendung enthaltenen Voraussetzungen genauer zu prüfen.

Gegenüber neueren Untersuchungen, welche einer wesentlichen Einschränkung oder Umformung des älteren Wahrscheinlichkeitsbegriffes das Wort reden, möchte ich im Folgenden zeigen, dass er nur etwa einer genaueren Formulirung bedarf, während die Grenzen seiner Anwendung eher weiter als enger gezogen werden müssen. In der Weise der Anwendung allerdings müssen wir um so vorsichtiger sein.

Definitionen sind insofern und insoweit willkürlich, als damit nur gesagt sein soll: „Ich für meine Person verstehe im Folgenden unter diesem Wort diesen Begriff.“ Hierüber wäre denn nicht zu streiten. Aber in der Regel beansprucht man damit zugleich den Sinn eines in der Wissenschaft bereits eingebürgerten und in wichtigen Sätzen angewandten Ausdrucks so wiederzugeben, dass er genau die Merkmale bezeichnet, aus denen die Consequenzen in Wirklichkeit gezogen wurden, während zugleich alle etwaigen Unbestimmtheiten, Unklarheiten und Widersprüche getilgt sind. In Fällen wie dem unsrigen handelt es sich aber ausserdem um einen Ausdruck und Begriff, den die Wissenschaft dem gewöhnlichen Denken entnommen hat und der auch in seinen Consequenzen ausgesprochenermassen nicht zu unerträglichen Abweichungen vom gemeinen Menschenverstand führen darf. Die Wahrscheinlichkeitsrechnung, sagt Laplace, ist nichts anderes als die mathematische Rechtfertigung der gesunden Vernunft (le bon sens réduit en calcul). Feineren Bestimmungen ist die blosse Schätzung auf Grund dieser „gesunden Vernunft“ allerdings nicht gewachsen und eben darum bedürfen wir der Rechnung. Auch wird gerade das Wahrscheinlichkeitsurteil am leichtesten durch Affecte u. dergl. mitbestimmt. Aber wo die Rechnungsergebnisse der natürlichen Schätzung allzusehr widerstreiten, da werden wir immer nachzusehen haben, ob in der Aufstellung der Grundformeln oder der Bedingungen ihrer Anwendung nicht ein Versehen platzgegriffen hat, ein Verstoss gegen die gemeinsamen Kriterien des wissenschaftlichen und des gewöhnlichen Denkens, die logischen Principien der Evidenz.

Aus diesen Gesichtspuncten also ist auch eine Kritik von Definitionen in unserem Fall erlaubt und erforderlich.

I. Allgemeine Fassung des Laplace'schen Wahrscheinlichkeitsbegriffes.

1. In dem berühmten „Philosophischen Versuch über die Wahrscheinlichkeiten"[1]) führt Laplace den Begriff mit folgenden Worten ein: „Die von einem Luft- oder Dampfteilchen beschriebene krumme Linie ist ebenso gesetzlich bestimmt wie die Planetenbahnen, mit dem einzigen Unterschied, dass wir ihr Gesetz nicht kennen. — Die Wahrscheinlichkeit hängt teils von dieser Unwissenheit, teils von unseren Kenntnissen ab. Zuweilen wissen wir, dass sich von drei oder mehr Begebenheiten Eine ereignen wird, und doch ist kein Grund vorhanden, dass wir glauben sollten, die eine werde sich wahrscheinlicher zutragen als die andere .... Die Theorie des Zufalls besteht darin, alle gleichartigen Begebenheiten auf eine gewisse Anzahl möglicher Fälle zurückzuführen, d. h. solcher Fälle, über deren Dasein (existence) wir in gleicher Unwissenheit sind, und dann die Anzahl der Fälle zu bestimmen, welche für die Begebenheit, deren Wahrscheinlichkeit man sucht, günstig sind. Das Verhältnis dieser Zahl zur Anzahl aller möglichen Fälle bildet das Mass dieser Wahrscheinlichkeit, die (das?) nichts anderes ist als ein Bruch, dessen Zähler die Zahl der günstigen und dessen Nenner die Zahl aller möglichen Fälle angibt."

Wir werden zunächst den Begriff der mathematischen Wahrscheinlichkeit, wie er Laplace vorschwebte, mit Beseitigung gewisser Ungenauigkeiten und unnötiger Beschränkungen formuliren, die darin liegenden erkenntnistheoretischen Consequenzen hervorziehen und sie gegen einige weitverbreitete Misverständnisse und Einwendungen verteidigen. Im nächsten Abschnitt besprechen wir Angriffe und Umformungen, welche die Grundlagen selbst betreffen.

---

1) Zuerst erschienen 1814 als Einleitung zur zweiten Auflage der „Théorie analytique des Probabilités".

Es liegt zu Tage, dass sich eine Ungenauigkeit eingeschlichen hat, wenn Laplace die sog. gleichmöglichen Fälle, die in den Begriff der Wahrscheinlichkeit eingehen, als solche definirt, bei denen man keinen Grund hat, die eine für wahrscheinlicher zu halten als die andere. Man darf nicht in die Definition des Wahrscheinlichen den Begriff des Wahrscheinlicheren einführen.[1]) Die Meinung von Laplace ist denn auch vollkommen ausgedrückt, wenn wir sagen: Gleich möglich sind Fälle, in Bezug auf welche wir uns in gleicher Unwissenheit befinden. Und da die Unwissenheit nur dann ihrem Masse nach gleich gesetzt werden kann, wenn wir absolut Nichts darüber wissen, welcher von den unterscheidbaren Fällen eintreten wird, so können wir noch bestimmter diese Erklärung dafür einsetzen.

Soviel allerdings ist richtig, dass gleichmögliche Fälle immer auch gleichwahrscheinlich sind, nämlich jeder $= \frac{1}{N}$ bei N gleichmöglichen Fällen. Aber die gleiche Wahrscheinlichkeit ist erst die Folge der gleichen Möglichkeit. Es haben ja auch gleiche Summen gleichmöglicher Fälle untereinander gleiche Wahrscheinlichkeit. Zuerst also muss die Gleichmöglichkeit erkannt sein.

Der Ausdruck „günstige Fälle" (chances favorables, Bernouilli's casus fertiles seu foecundi gegenüber den casus steriles), auch kurzweg Chancen, bedeutet im Sinne der Definition nicht etwa Umstände oder Bedingungen, welche

1) Seltsamer Weise findet sich die nämliche Wendung noch bis in die neueste Zeit. Wenn man „wahrscheinlicher" hier im Sinne des sog. Philosophisch-Wahrscheinlicheren nehmen wollte, wäre nicht geholfen. Denn es liegt der „philosophischen" und der mathematischen, d. h. der nicht (oder nicht genau) messbaren und der messbaren, Wahrscheinlichkeit doch ein gemeinsamer Begriff zu Grunde. Ueberdies wäre es ja thatsächlich schon eine Messung, wenn wir die eine Begebenheit „nicht wahrscheinlicher" als die andere nennen: wir würden eben die Wahrscheinlichkeiten gleich gross setzen.

dem Ereignis günstig sind, d. h. welche (im Verein mit den sonstigen Bedingungen) es herbeiführen werden, sondern vielmehr diejenigen unter den möglichen Fällen selbst, unter welche wir das bezügliche Ereignis logisch subsumiren müssen; z. B. wenn nach der Wahrscheinlichkeit eines Pasches gefragt ist, die Fälle „1 und 1", „2 und 2" u. s. f. unter den 36, über deren jeden wir uns in Unwissenheit gegenüber den anderen befinden. Diese 36 bilden die sämmtlichen untersten Arten des Gattungsbegriffes „eine mit 2 Würfeln zu werfende Zahlencombination"; und von diesen untersten Arten fallen 6 unter den nächsthöheren Artbegriff „Pasch".[1])

In gewissen später (IV.) zu erwähnenden Fällen werden unter den Chancen in der That reale Umstände verstanden,

---

1) A. Meyer's vielbenützte „Vorlesungen über Wahrscheinlichkeitsrechnung" (deutsch v. Czuber 1879) beginnen folgendermassen: „Jedes Ereignis ist die Folge eines Zusammenwirkens zweier Arten von Umständen; die einen, bekannt oder unbekannt, sind notwendig zu seiner Hervorbringung, während die anderen, stets unbekannten, nur zufällig dazu beitragen. Die Umstände der ersteren Art nennt man Ursachen oder Chancen des Ereignisses; die anderen in ihrer Gesammtheit bilden das was man als Zufall bezeichnet." Später wird freilich bemerkt (S. 8—9), dass das Wort Ursache in der Wahrscheinlichkeitslehre etwas anderes als sonst bedeute: „nicht das, was einen Erfolg oder ein Ereignis herbeiführt" (wie man allerdings nach jener Erklärung denken sollte) „sondern das was einem Ereignis die ihm eigentümliche Wahrscheinlichkeit erteilt. Es sind dies die Chancen des Ereignisses an und für sich". Abgesehen davon, dass dies gleich anfangs hätte gesagt werden müssen, ist die neue Erklärung auch nicht sonderlich deutlich. Was erteilt dem Ereignis seine eigentümliche Wahrscheinlichkeit? Wir drehen uns im Kreise.

Uebrigens legen auch Mathematiker, die die Chancen durch Ursachen definiren, bei der Aufstellung des Summenprincips alsbald doch unseren obigen Begriff zu Grunde. Beim Pasch sollen die günstigen Fälle summirt werden: man meint hier offenbar nicht die Ursachen, die einen Pasch herbeiführen, sondern einfach die Arten des Ereignisses selbst, die Fälle, welche unter jenen Begriff gehören.

und zwar diejenigen, auf denen die Anzahl der möglichen und günstigen Fälle (Chancen im vorigen Sinne) beruht; wie wenn wir sagen, die Chancen eines Ereignisses haben sich verändert. Der Begriff ist, wie man sieht, auf den vorigen zurückzuführen; es ist aber zur Klarheit nötig, die doppelte Bedeutung des Wortes auseinanderzuhalten. Vorläufig haben wir es nur mit der ersten zu thun.

2. Eine Frage von hervorragender Wichtigkeit betrifft das Zeitmoment in der Wahrscheinlichkeitsdefinition.

Es lag in der Natur der concreten Aufgaben, aus denen die Wahrscheinlichkeitsrechnung erwuchs, dass dabei immer von zukünftigen Begebenheiten die Rede war. Die möglichen und die günstigen Fälle betrafen den Ausgang von Spielen. Wir sehen diese Bezugnahme auf Künftiges auch an der Art, wie Laplace den Wahrscheinlichkeitsbegriff einführt. Aber wenn man sie in die Definition hereinnehmen wollte, würde eine unnötige und durch nichts gerechtfertigte Beschränkung entstehen. Wir nennen es offenbar in demselben Sinne $\frac{1}{6}$ wahrscheinlich, dass die Zahl 4 bei einem künftigen Wurf mit dem Würfel erscheinen wird, dass sie bei einem vorhin stattgefundenen erschienen ist, und dass sie gegenwärtig oben liegt. In den letzten Fällen bedeutet der Nachsatz nicht etwa, dass wir diese Zahl erblicken werden, wenn wir hinsehen: denn auch wenn uns jede Möglichkeit benommen ist, jemals den Sachverhalt durch Beobachtung festzustellen, behaupten wir seine Wahrscheinlichkeit, und zwar diese bestimmte. Es gibt also Wahrscheinlichkeits-Aussagen, die sich ausschliesslich auf Vergangenes oder Gegenwärtiges als solches beziehen. Zuletzt haben doch gerade alle vergangenen Thatsachen nur Wahrscheinlichkeit, wenn auch eine graduell unendlich verschiedene und in den meisten Fällen nicht bestimmt messbare. Laplace hat denn auch selbst als die möglichen Fälle diejenigen bezeichnet, über deren Dasein — nicht über deren Eintritt — wir in gleicher Unwissenheit sind. Er hätte vielleicht, um die Beziehung auf eine bestimmte Zeit ganz abzuschneiden, noch genauer sagen können: über deren Wahrheit.

Die obige Beschränkung ist neuerdings durch keinen Geringeren als Lotze[1]) ausdrücklich sanctionirt worden. Er gründet den Wahrscheinlichkeitsbegriff von vornherein nur auf das praktische Bedürfnis, unsere Handlungen mit Bezug auf künftige Ereignisse zu regeln, polemisirt gegen Laplace' Berechnung der Wahrscheinlichkeit einer gemeinschaftlichen Ursache für die nahe übereinstimmenden Richtungen der Planetenbahnen, und will die Wahrscheinlichkeit durchaus nur als Mass des Vertrauens zu dem Eintritt künftiger Ereignisse angesehen wissen.

Dem sonst so scharfen Denker sind hier wie in anderen Puncten der Wahrscheinlichkeitslehre sonderbare Misverständnisse begegnet. So viel ist gewiss, dass der mathematische Wahrscheinlichkeitsbegriff sich ohne die Beschränkung auf das Künftige ganz ebenso scharf definiren lässt, und dass er in diesem Fall nicht blos eine praktische sondern auch eine theoretische Bedeutung hat. Es könnte sich also nur etwa fragen, ob die aus einer solchen allgemeineren Fassung gezogenen Consequenzen auch mit dem gewöhnlichen Menschenverstand und Sprachgebrauch übereinstimmen: und daran kann meines Erachtens kein Zweifel sein, vorausgesetzt, dass

---

1) Logik (1874) S. 414. 432 f. 434. Gleiches lehrt Wundt (Logik I 393) mit der Begründung, dass sich Wahrscheinlichkeit immer auf erwartete Thatsachen beziehe; was freilich nur eine Wiederholung der Behauptung ist.

Oefters begegnet man wol auch der Wendung: alles Vergangene sei fest gegeben, und es sei sinnlos, nach der Wahrscheinlichkeit einer gegebenen Thatsache zu fragen. Aber in Wahrheit ist ja das Vergangene uns niemals gegeben, nicht anders als das Künftige. Aendern lässt es sich freilich nicht; aber die Wahrscheinlichkeitsrechnung will nichts anders machen als es ist, und würde dies auch in Hinsicht des Künftigen nicht zuwegebringen.

die Folgerungen correct gezogen werden. Ob diese Voraussetzung bei Laplace' Hypothese über die Entstehung des Sonnensystems durchaus zutrifft, mag hier unerörtert bleiben: unbestreitbar ist doch wieder, dass zufällige Coincidenzen in einer grossen Anzahl im Allgemeinen auch vom gewöhnlichen Menschen in dem gleichen Sinne unwahrscheinlich genannt werden, mögen sie der fernsten Vergangenheit oder der Gegenwart oder der Zukunft angehören. Berichtet ein alter Schriftsteller, dass er beim Würfeln 1000 Mal nacheinander die Zahl 3 erhalten habe, so werden wir entweder die Ehrlichkeit seines Berichtes oder die seines Spiels oder die gleichmässige Structur des Würfels in Zweifel ziehen, weil die zufällige d. h. nicht durch eine gemeinsame Ursache bedingte Wiederholung äusserst unwahrscheinlich $\left(=\frac{1}{6^{1000}}\right)$ ist, mag sie der vergangenen, gegenwärtigen oder zukünftigen Zeit angehören.

Allerdings wird bei der Wahrscheinlichkeitsberechnung für eine Hypothese u. A. gefragt, wie wahrscheinlich die gegebene Thatsache unter Voraussetzung der Hypothese sei (vgl. u. IV). Das Gegebene wird also hier in Gedanken als ein Mögliches neben anderen gleich möglichen Fällen betrachtet, d. h. als ein Fall, über den wir uns ebenso wie über die anderen in Unwissenheit befänden. Aber nicht als ein Künftiges. „Angenommen — so lautet die Frage — es sei uns nicht gegeben: wie wahrscheinlich wäre sein Vorhandensein auf Grund der einen und der anderen concurrirenden Hypothese?“ Zu dieser Bestimmung dient eine logische Coordination mit den übrigen denkbaren Fällen. Weder in der Fragestellung noch in der Lösung liegt eine Nötigung, den Fall fictiv in die Zukunft oder uns selbst in die Vergangenheit zu verlegen; wenn auch die dadurch rein logische Fragestellung bei derartigen Problemen anschaulicher werden mag.

Es muss auch zugegeben werden, dass dem Ausdruck Wahrscheinlichkeit im ganz populären Gebrauch ein solcher Beigeschmack, eine praktische Beziehung auf Künftiges, auf Lotteriegewinnste u. dgl. anhaftet. Aber die Consequenzen, die aus einer wissenschaftlichen Fixirung dieses Merkmals entstehen würden, widersprechen doch selbst dem gewöhnlichen Gebrauch aufs Bestimmteste.

Die Beschränkung auf Künftiges ist also durch keinerlei sachliche Erwägung gefordert. Wir könnten mit demselben Recht auch eine räumliche Beschränkung, etwa auf die sublunarische Region, in den Wahrscheinlichkeitsbegriff einführen.

Im Grunde folgt übrigens die Irrelevanz der Zeit schon daraus, dass es für die Wahrscheinlichkeitsbestimmung zugestandenermassen einerlei ist, ob wir 6 mal nacheinander oder gleichzeitig würfeln.[1]) Wir können dann offenbar auch 2 von den Würfen schon gemacht haben, 2 eben machen und 2 noch machen wollen: die Wahrscheinlichkeit für das 6 malige Eintreffen einer Seite ist immer die nämliche.

3. Hiemit hängt nun eine weitere Verallgemeinerung zusammen. Es ist unnötig und ungerechtfertigt, nur Ereignissen (événements) eine mathematische Wahrscheinlichkeit zuzuschreiben. Denn ebenso wie wir von einem Ereignis sagen können, dass es eines unter einer bestimmten Zahl von Ereignissen sei, über die wir nur wissen, dass eines von ihnen wirklich ist (sein wird, gewesen ist), aber nicht, welches: ebenso können wir uns auch in Bezug auf jede beliebige sonstige Urteilsmaterie in einem analogen Stande des Wissens und Nichtwissens befinden. Es sei uns gegeben — um an ein gebräuchliches Schema anzuknüpfen —, dass eine

---

1) Früher begegnete auch dies hie und da Zweifeln, pflegt aber jetzt in den Lehrbüchern ausdrücklich betont oder als selbstverständlich vorausgesetzt zu werden. Vgl. A. Meyer S. 10.

Urne zwei weisse oder schwarze Kugeln enthält, so ist die Wahrscheinlichkeit, dass zwei weisse darin seien, $= (\frac{1}{2})^2$, u. s. f. Es handelt sich hier nicht um ein Ereignis sondern um einen Thatbestand; und wenn derselbe von Ewigkeit zu Ewigkeit unverändert existirte, so würde die Wahrscheinlichkeit dieselbe bleiben, solange nur unser Wissen sich nicht verändert. In diesem Sinne hat G. Kirchhoff[1]) die Wahrscheinlichkeit, dass die Coincidenz von 60 hellen Linien im Eisenspectrum mit 60 dunklen Linien im Sonnenspectrum keine zufällige sei, = mindestens $1 - \frac{1}{2^{60}}$ bestimmt und eben darauf (in Verbindung mit unseren Kenntnissen über die Auslöschung solcher hellen Linien) seinen Schluss gestützt, dass Eisen in der Sonne vorhanden sei. Auch dieser Thatsache schrieb er in Folge dessen die erwähnte Wahrscheinlichkeit zu, und es ist dabei offenbar vollkommen gleichgültig, ob man sie sich als ein vorübergehendes Ereignis oder als ewigen Bestand vorstellt. In gleichem Sinne können wir denn auch von dem wahrscheinlichen Durchmesser eines Wasserstoffmoleculs oder (wenn einer auch diesen als veränderlich betrachten will) von dem wahrscheinlichen Werte der chemischen Constanten der als absolut unveränderlich gedachten Uratome reden.

Endlich kann statt eines concreten Thatbestandes auch eine **allgemeine, abstracte** Urteilsmaterie als wahrscheinlich in gleichem Sinne bezeichnet werden. Auch da können wir in der Lage sein, zu wissen, dass eines von m Gesetzen wahr sein muss, ohne das Geringste darüber zu wissen, welches. Wir werden jedem von ihnen die Wahrscheinlichkeit $\frac{1}{m}$ zuerkennen. Nur wird sich in solchen Fällen selten mit Bestimmtheit eine feste endliche Zahl m angeben

---

1) Untersuchungen über das Sonnenspectrum. Abhdl. der Berliner Akad. 1861 S. 79.

lassen. Auch ist die Voraussetzung, dass wir über die einzelnen Möglichkeiten vollkommen gleich unwissend seien, hier in der Praxis selten erfüllt. Der Wahrscheinlichkeitsansatz wird dann mehr oder weniger an Bestimmtheit verlieren, ohne dadurch wissenschaftlich wertlos zu werden. Ist doch auch die Wahrscheinlichkeit $\frac{1}{2}$ für die Coincidenz je zweier Linien im obigen Falle von Kirchhoff ausdrücklich nur approximativ geschätzt und gleichwol Grundlage einer der wertvollsten Entdeckungen.

4. Allgemein also lässt sich im Sinne von Laplace, in consequenter Ausdehnung seiner Bestimmungen, sagen: Jede beliebige Urteilsmaterie nennen wir $\frac{n}{N}$ wahrscheinlich, wenn wir sie auffassen können als eines von n Gliedern (günstigen Fällen) innerhalb einer Gesamtzahl von N Gliedern (möglichen Fällen), von denen wir wissen, dass eines und nur eines wahr ist, dagegen schlechterdings nicht wissen welches.

Auf diese Weise sind gleichsam die Eierschalen abgestreift, die dem Begriff der mathematischen Wahrscheinlichkeit von seinen Ursprungsbeispielen her noch anhafteten. Wenn wir gleichwol im Folgenden uns ebenfalls vorwiegend an solche Beispiele halten, geschieht es der Anschaulichkeit und Einfachheit halber; denn sie geben immer das beste Schema, auf welches auch complicirtere Verhältnisse reducirt werden können. Besonders zeigt sich dies bei der sog. empirischen Wahrscheinlichkeit (IV).

Unter den neueren Logikern hat namentlich Sigwart mit Recht betont, dass das Wahrscheinlichkeitsurteil auf dem disjunctiven Urteil gründet.[1]) Es ist nicht selbst ein disjunc-

---

1) Logik II (1878) S. 265 f., wo die Wahrscheinlichkeitsschlüsse in interessanter Weise als eine specielle Classe der Schlüsse dargestellt werden, die man aus einer Combination disjunctiver Urteile entwickeln kann.

A. Lange, der die Wahrscheinlichkeitslehre ebenfalls auf das tives Urteil, aber eine Folgerung aus einem solchen in Verbindung mit einer zweiten Prämisse, der Anerkennung völligen Nichtswissens über die einzelnen disjungirten Glieder.

Legen wir diese Auffassung zu Grunde, so folgt, dass der mathematische Wahrscheinlichkeitsbegriff keinerlei Voraussetzungen oder Ueberzeugungen hinsichtlich der objectiven Welt einschliesst[1]), insbesondere auch nicht die der Gültigkeit des Causalgesetzes, mag man es dahin aussprechen, dass jedes Ereignis seine Ursache hat oder dass unter gleichen Umständen immer gleiche Folgen eintreten.

Denken wir uns, es sei nichts Körperliches vorhanden als sechs Atome, die unter sich keine Kräfte ausübten, sondern im leeren Raum, jedes als eine Welt für sich, schwebten, und es sei uns, die wir als reine Intelligenz existirten, nur gegeben, dass eines davon die Kugelform, fünf die tetraedrische Form besässen, so würde die Wahrscheinlichkeit der Kugelform für ein bestimmtes dieser Atome $\frac{1}{6}$, die des Tetraeders $\frac{5}{6}$ sein, und diese Aussage hätte ihren Sinn gleich jeder anderen mathematischen Wahrscheinlichkeit.

Indem Laplace in seiner Einleitung von der Unverbrüchlichkeit des Causalgesetzes und unserem unbedingten Glauben an dasselbe ausgeht, hat er einen vielleicht didaktisch bequemen aber für seine Definition nicht unumgänglichen Ausgangspunct gewählt.

---

disjunctive Urteil gründete (Logische Studien 1877 S. 108 f.), hat doch durch seine Lehre von der integrirenden Bedeutung räumlicher Anschauungsformen für logische Begriffsverhältnisse gerade hier ein nicht nur unwesentliches sondern irreleitendes Moment eingeführt, indem er die (alsbald zu besprechende) Forderung einer physischen Gleichheit der möglichen Fälle daraus ableitete.

Lotze deutet die Beziehung zum disjunctiven Urteil ebenfalls an (S. 414), ohne aber Consequenzen daraus zu ziehen.

1) Auch hierüber glaube ich mich mit Sigwart in Uebereinstimmung zu befinden, vgl. a. a. O. 273.

So wird ja auch in vielen Beispielen aus der wirklichen Welt bei der Wahrscheinlichkeitsbestimmung von Causalverhältnissen völlig abstrahirt, wie in obigem Urnenbeispiel (S. 47 o.). Es würde uns hier auch gar nichts helfen, auf Causalverhältnisse zurückzugehen, da immer dieselben Disjunctionen, dasselbe Zahlenverhältnis möglicher Fälle herauskommt. Oder wenn wir die Wahrscheinlichkeit, dass ein vorliegender regelmässiger Körper, von dem wir nur wissen, dass er nicht über 20 Seitenflächen hat, ein Tetraeder sei, $= \frac{1}{5}$, oder die Wahrscheinlichkeit, dass ein vorliegendes Dreieck absolut genau rechtwinklig sei, $= \frac{1}{\infty}$ bestimmen, so ist auch hier wie in allen rein geometrischen Dingen von Ursache und Wirkung keine Rede. Der Geometrie ist der Causalbegriff so fremd wie der Unterschied von Gut und Böse.

Es versteht sich, dass überall, wo speciell die Wahrscheinlichkeit von Ereignissen zu bestimmen ist und wo unsere Kenntnisse eine Disjunction der möglicherweise vorhandenen Ursachen (d. h. der Combinationen von Bedingungen, von denen eine vorhanden ist, ungewiss welche) gestatten, damit auch eine Disjunction der möglichen Ereignisse selbst gegeben ist, die aus je einer dieser Ursachen notwendig fliessen. Und es versteht sich, dass wir in allen Fällen, wo uns causale Kenntnisse zur Verfügung stehen, die zu einer anderen Disjunction der Ereignisse führen als die directe Betrachtung der Ereignisse selbst, auf die Ursachen zurückgehen müssen. Denn der Wahrscheinlichkeitsbegriff verlangt, dass alle uns gegebenen Kenntnisse über die bezüglichen Ereignisse — und dazu gehören auch die über ihre Entstehung — berücksichtigt werden. Wo uns aber solche Kenntnisse nicht gegeben sind, da wird der Wahrscheinlichkeitsansatz keineswegs unmöglich; wir entnehmen ihn eben der Disjunction der Ereignisse selbst. Wir mögen auch da von der allgemeinen Ueberzeugung durchdrungen sein, dass jedes Ereignis irgendwelche Ursachen hat und dass gleiche Ursachen immer gleiche Wirkungen haben: die Ueberzeugung hat für die mathematische Wahrscheinlichkeit des Ereignisses hier nicht die geringste Bedeutung; und wer sie nicht teilte (wie ja wirklich die Anhänger einer indeterministischen Willensfreiheit ihre Allgemeingültigkeit in Abrede stellen) für den würde gleichwol der Wahrscheinlichkeitsansatz und dessen Sinn der nämliche bleiben.

Ja selbst da, wo wir auf die Ursachen zurückgehen können und müssen, wo uns Kenntnisse darüber zur Verfügung stehen, welche eine Disjunction der möglichen Fälle gestatten: selbst da haben wir in dieser Disjunction doch auch wieder nur eine Disjunction gewisser Ereignisse, die dem fraglichen Ereignis möglicherweise vorausgehen. Statt also dieses Ereignis E selbst als ein Glied (bezw. eins von n Gliedern) innerhalb einer vollständigen Disjunction von N Gliedern aufzufassen, fassen wir das vorangehende Ereignis E_1, woraus jenes hervorgehen muss, wenn E_1 existirt, als ein Glied innerhalb einer solchen Disjunction. Es bleibt also zuletzt in allen Fällen bei einer Disjunction von Gliedern ohne weitere Rücksicht auf ihre Herkunft und Verursachung.

Die Grenze bei dem Rückgang auf die Ursachen ist dann gegeben, wenn die Kenntnisse über Ursachen derart sind, dass sie uns zu keiner anderen Disjunction mehr führen können als die Kenntnis der Wirkungen. Häufig genug ist dies schon beim ersten Schritt der Fall.

Wo es sich speciell um die Wahrscheinlichkeit von Ereignissen handelt, die aus ihren Ursachen erschlossen werden sollen, da schliesst die mathematische Wahrscheinlichkeitsbestimmung nicht blos das allgemeine sondern auch bestimmte concrete Causalgesetze darum ein, weil die Disjunction der möglichen Fälle nur mit Rücksicht darauf als eine vollständige gelten kann. Ist mir nicht gegeben, dass ein Würfel gefallen ist (oder fallen wird), sondern dass er geworfen ist (oder werden wird), so kann ich nur dann behaupten, dass das Fallen einer Würfelseite $\frac{1}{6}$ wahrscheinlich ist, wenn ich voraussetze, dass der geworfene Körper überhaupt zu Boden fällt und nicht etwa in der Luft hängen bleibt oder sich in Nichts auflöst. Dann allein weiss ich, dass einer von den 6 unterschiedenen Fällen eintritt. Aber nicht alle Anwendungen der Wahrscheinlichkeit sind von dieser Art, die Bedingung ist also keine allgemeine. Ferner brauchen wir, auch wo ein solcher Fall vorliegt, keineswegs eine allen Zweifel ausschliessende Sicherheit für das oder die benützten Causalgesetze. Die entgegengesetzten Möglichkeiten müssen dann eben in die Disjunction aufgenommen werden. Wenn wir z. B., um den extremsten Fall zu nehmen, absolut nichts darüber wüssten, ob der geworfene Körper überhaupt fallen wird oder nicht, so würde für jede dieser beiden nach der Vorraussetzung gleichen Möglichkeiten die Wahrscheinlichkeit $\frac{1}{2}$ und somit für jede der unter dem ersten Glied befassten unter sich gleichen Einzelmöglichkeiten $\frac{1}{2} \cdot \frac{1}{6}$ resultiren (ganz ebenso wie wenn wir bei voller Kenntnis der Fallgesetze nichts darüber wissen, ob er factisch geworfen wird oder nicht). Wenn das Fallen des geworfenen Körpers eine kleinere oder grössere angebbare Wahrscheinlichkeit hätte, d. h. also sich unter eine vollständige Disjunction von mehr als zwei Fällen gleicher Unwissenheit ordnete oder mehr günstige Fälle unter sich befasste, so wäre der entsprechend kleinere oder grössere Bruch mit $\frac{1}{6}$ zu multipliciren. Hätte endlich das Fallen des Körpers eine nur abschätzbare nicht genau bestimmbare Wahrscheinlichkeit, so bliebe eben auch das Product nur abschätzbar. Vielfach ist dies wirklich die Sachlage, da nicht alle Naturgesetze die Sicherheit des Fallgesetzes besitzen. Dann hilft auch die festeste Ueberzeugung von dem allgemeinen Causalgesetz nicht weiter.

Ja wir können auch dieses selbst als fraglich ansehen oder ihm nur eine gewisse Wahrscheinlichkeit zuschreiben, ohne dass Wahrscheinlichkeitsbestimmungen der genannten

Art principiell unmöglich würden. Sie sind dann eben in gleicher Weise mit der Wahrscheinlichkeit des Causalgesetzes zu multipliciren.

Läge in jeder Wahrscheinlichkeitsaussage an sich schon die Voraussetzung des Causalgesetzes, so würden freilich diejenigen Philosophen, die das Causalgesetz selbst nur für wahrscheinlich (sei es auch unendlich wahrscheinlich) ansehen, den offenbarsten Cirkel begehen. Von solcher Absurdität also müssen wir sie freisprechen.[1])

1) Die in Deutschland und zumal im Kreise der Kantianer vorherrschende Auffassung formulirt O. Liebmann einmal in Kürze so: „Wer ausgerechnet hat, dass beim Ziehen aus einer verdeckten Urne, welche w weisse und s schwarze Kugeln enthält, die Wahrscheinlichkeit, im einzelnen Falle weiss zu ziehen, $= \frac{w}{w+s}$ ist, der setzt eben schon voraus, dass nicht durch ein Zauberkunststück oder ein Wunder die Anzahl der Kugeln unter der Hand vermehrt oder vermindert werde. Das heisst, er setzt objective Gültigkeit des Causalprincips voraus." (Klimax der Theorien S. 91.)

Ich will hier unentschieden lassen, ob man ein Zauberkunststück notwendig für eine Durchbrechung des Causalgesetzes ansehen müsste. Das Mischungsverhältnis könnte ja auch auf unzweifelhaft natürlichem Wege verändert werden: die Wahrscheinlichkeitsbestimmung würde auch dann ungültig. Schon daraus erhellt, dass das Causalgesetz hier unnötig herangezogen ist. Bei der Wahrscheinlichkeitsbestimmung für ein künftiges Ereignis müssen natürlich auch die Verhältnisse, worauf sie sich gründet, als in jenem künftigen Moment stattfindende angenommen werden (ebenso wie gegenüber einem vergangenen oder gegenwärtigen für den bezüglichen Moment). Die Wahrscheinlichkeit, die wir für die Unveränderlichkeit dieser Verhältnisse haben, ist je nach der Materie und der Entfernung des fraglichen Zeitpunkts äusserst verschieden. Aber der Wert $\frac{w}{w+s}$ drückt auch weiter nichts aus als die Wahrscheinlichkeit, dass eine weisse Kugel gezogen werde, wann und wo dieses Mischungsverhältnis gegeben ist. Diese Wahrscheinlichkeit soll ja nicht für bestimmte Zeit an dieser Urne haften, sondern ist die zeitlose logische Consequenz dieses Mischungsverhältnisses. Und nicht das Causal-

5. Von den Anhängern der Laplace'schen Definition wird vielfach auch die Folgerung als besonders wichtig hervorgehoben, dass die Wahrscheinlichkeit etwas durchaus Subjectives sei. Die wirklichen Dinge und Ereignisse seien nur wahr und falsch.

Indessen sind genau gesprochen doch auch Wahrheit und Falschheit nicht Eigenschaften von Dingen, die ihnen abgesehen von ihrer Beurteilung zukämen, vielmehr haben auch diese Prädicate, wie bereits Aristoteles betonte, nur Sinn mit Bezug auf ein Urteil des Verstandes.

Man replizirt, etwas Wahres könne doch nie falsch werden, während ein Wahrscheinliches seine Wahrscheinlichkeit verändern und auch gleichzeitig für verschiedene Individuen verschiedene Wahrscheinlichkeit besitzen könne. Für einen unendlichen Verstand würde jede Materie nur sicher sein, dagegen würde der Unterschied von Wahr und Falsch auch für ihn nicht hinwegfallen. Insofern scheint also Wahrscheinlichkeit doch in besonderem Sinn subjectiv zu sein.

Man übersieht, dass eine Veränderung und eine Verschiedenheit des Wahrscheinlichkeitsgrades nur möglich ist, wenn die Materie des Urteils sich ändert oder verschieden

---

gesetz sondern das der Identität verlangt, dass ich das im Problem Gegebene nicht auch als nichtgegeben ansehe.

Ein anderer Kantianer interpretirt folgendermassen: „Es beschäftigt uns (bei der Wahrscheinlichkeitsbestimmung) nur der Zusammenhang von Ursachen und Wirkungen. Wir urteilen, dass keine der 6 Würfelseiten durch beständig wirkende Ursachen begünstigt wird und auch weder durch den Werfenden noch durch den Wurf eine Begünstigung möglich ist; wir bezeichnen in diesem und in keinerlei mathematischem Sinne die 6 möglichen Fälle als gleich möglich.“ (A. Elsas, Philos. Monatshefte XXV, 1889, S. 567 f.)

Im Gegenteil! Ich zweifle gar nicht, dass eine der 6 Seiten durch beständig wirkende Ursachen begünstigt ist. Ich weiss nur nicht, welche, und habe nicht den geringsten Anhaltspunct für eine von ihnen. Daher $\frac{1}{6}$.

ist. Die Wahrscheinlichkeit, mit einem gefälschten Würfel eine bestimmte Seite zu werfen, ist für den, der seine Construction kennt, natürlich nicht $= \frac{1}{6}$. Da zur Materie eines Wahrscheinlichkeitsurteils, wenn sie so vollständig ausgesprochen wird, wie sie in der Problemstellung ausgesprochen werden muss, auch alle Daten gehören, aus denen die Wahrscheinlichkeit abgeleitet wird, so hat in der That jede Materie nur einen einzigen bestimmten und unveränderlichen Wahrscheinlichkeitsgrad, den nämlichen für jeden beliebigen Verstand. Auch für einen unendlichen Verstand würde, wenn wir ihn nicht nach der Ausdehnung sondern nur nach der Kraft des Denkens unendlich setzen, das Fallen einer bestimmten Würfelseite ohne weitere Data nur die Wahrscheinlichkeit $\frac{1}{6}$ haben. Die mathematische Wahrscheinlichkeit ist etwas durchaus Festes, eine Function der Urteilsmaterie. Nennen wir nun „objectiv gültig" dasjenige, was von allen Subjecten bei gleicher Urteilsmaterie anerkannt werden muss, so ist das Wahrscheinliche objectiv gültig. Nur dann kann hierüber ein Misverständnis entstehen, wenn man an Stelle der Materie im weiteren und vollständigen Sinn nur die Materie im engeren und unvollständigen Sinne in Betracht zieht, d. h. das Subject des Wahrscheinlichkeitsprädicats mit Hinweglassung der uns gegebenen näheren Bestimmungen; wie es allerdings vielfach nicht blos in der gewöhnlichen Rede sondern bei nachlässiger Formulirung auch im wissenschaftlichen Gebrauche geschieht.

6. Ganz verkehrt endlich erscheint die ebenfalls noch öfters wiederkehrende Wendung, das Wahrscheinliche stehe zwischen Wahrem und Falschem in der Mitte. Der Satz, dass alles entweder wahr oder falsch ist, dass zwischen contradictorischen Gegensätzen kein Drittes liegt, wird auch durch das Wahrscheinliche in keiner Weise gefährdet. Ebensogut könnte man sagen, das Hypothetische („Wenn A ist, ist B") liege in der Mitte zwischen Wahrem und Falschem.

Es ist Eine Einteilung unserer Urteile, wenn wir sie in wahre und falsche, und es ist eine andere Einteilung, wenn wir sie in wahrscheinliche verschiedenen Grades (einschliesslich der unendlich wahrscheinlichen) und in (absolut oder mathematisch) sichere scheiden. Eine behauptete Wahrscheinlichkeit kann wahr oder falsch d. h. auf Grund der gegebenen Materie richtig oder unrichtig bestimmt sein. Ferner kann eine richtig bestimmte Wahrscheinlichkeit sich beziehen auf die Wahrheit oder auf die Falschheit einer Annahme. Endlich wenn sie sich auf die Wahrheit einer Annahme bezieht, d. h. auf das Stattfinden des Angenommenen, so kann dieses Wahrscheinliche selbst (die Materie des Wahrscheinlichkeitsurteils, und zwar sowol im vollständigen als unvollständigen Sinne), z. B. das Fallen der Zahl 4, das einemal wahr, das anderemal falsch sein. Die Wahrscheinlichkeit und das Wahrscheinliche stehen also nicht in der Mitte, sondern sind jedesmal auch nur eines von beiden. Alles Wahrscheinliche ist zugleich und ausserdem entweder wahr oder falsch.

Eine besondere Beziehung ergibt sich dadurch, dass der Begriff des Wahren für die Definition des Wahrscheinlichen in der aus dem Wortlaut derselben ersichtlichen Weise vorausgesetzt wird. Jener ist logisch der frühere (*λόγῳ πρότερον*). Insofern ist die zweite Einteilung nicht unabhängig von der ersten. Aber sie ist keine Untereinteilung derselben, und noch weniger lassen sich beide zu einer dreigliederigen Einteilung coordinirter Glieder verbinden.

7. Die mathematische Wahrscheinlichkeit wird auch als Mass unserer vernünftigen Erwartung bezeichnet. Erwartung ist der Glaube an das künftige Eintreten eines Ereignisses. Insofern deckt sich also die Erklärung nicht mit allen Fällen der mathematischen Wahrscheinlichkeit. Will man sie verallgemeinern, so muss man eben dem Ausdruck „Erwartung“ einen entsprechend allgemeineren Sinn

unterlegen oder „Vermutung“ dafür setzen, worin keine notwendige Beziehung auf die Zukunft liegt.

Abgesehen aber von dieser unnötigen Einschränkung, die sich durch terminologische Bestimmungen aufheben liesse, ist die Erklärung, wie mir scheint, unbedenklich und unschwer zu rechtfertigen. Unsere Erwartungen werden, wie unser Glauben und Fürwahrhalten überhaupt, im Leben nicht durch blos logische Erwägungen sondern durch zahlreiche instinctartige oder sozusagen blindwirkende Mächte (Gewöhnung, Wünsche und Neigungen, überhaupt Affecte aller Art) mitbestimmt. Ebendarum wird die mathematische Wahrscheinlichkeit nur als Mass der vernünftigen Erwartung bezeichnet, d. h. der Erwartung, soweit sie von der blossen Vernunft bestimmt ist oder wäre.

Der Psychologe möchte hier freilich noch die Frage aufwerfen, ob man denn wirklich den Glauben messen kann. Besitzt er eine Intensität wie ein Sinneseindruck oder gar eine lebendige Kraft wie ein Naturprocess? — In der That wird durch die mathematische Wahrscheinlichkeit nicht eine derartige Eigenschaft unseres Urteils gemessen. Was gemessen wird, ist lediglich die Urteilsmaterie und zwar speciell in Hinsicht der Anzahl von Disjunctionsgliedern, von denen wir wissen u. s. f. Dieser Materie kommt die Wahrscheinlichkeit primär als Prädicat zu, dem Urteil nur eben darum und insofern, als eine solche Materie seinen Gegenstand bildet. Also wird hier nur in einem indirecten Sinne von einem mathematischen Mass unseres Urteils gesprochen; und in diesem Sinne ist die Bezeichnung der mathematischen Wahrscheinlichkeit als des Masses der vernünftigen Erwartung nicht nur richtig sondern fällt mit der Definition zusammen.

Ich will hiemit der Frage, ob nicht doch das, was wir Zuversicht und speciell vernünftige Zuversicht nennen, **auch** verschiedener Grade fähig und in diesen seinen Graden durch die mathematische Wahrscheinlichkeit bedingt sei, nicht vorgreifen. Sie kann aber als eine rein psychologische für die gegenwärtige Untersuchung dahingestellt bleiben.

II. Angriffe und Umbildungen.

1. Durch die Schwierigkeit, die Stellung des Wahrscheinlichen zum Wahren und Falschen zu bestimmen, zwischen denen es doch ein Mittleres nicht geben könne, sah sich A. Fick[1]) zu einer neuen Definition geführt, wonach mathematische Wahrscheinlichkeit eine Eigenschaft eines unvollständig ausgedrückten hypothetischen Urteiles ist, nämlich „der als ächter Bruch dargestellte Teil des ganzen Bereichs der Bedingung, an dessen Verwirklichung der im Nachsatz ausgedrückte Erfolg notwendig geknüpft ist". Wir schreiben dem Urteil: „Wenn eine Münze auf den Tisch geworfen wird, so wird die Wappenseite oben liegen" die Wahrscheinlichkeit $\frac{1}{2}$ zu. Aber der Vordersatz ist unvollständig und lautet ergänzt: „Wenn eine Münze auf den Tisch geworfen wird und ihre Schriftseite mit einem Winkel zwischen 0° und 90° auftrifft, so wird die Wappenseite oben liegen." Dazu gehört dann als Ergänzung der Regel selbst die andere: „Wenn ihre Schriftseite mit einem Winkel zwischen 90° und 180° auftrifft, wird die Schriftseite oben liegen." Jede dieser beiden Regeln (die wieder je unendlich viele Einzelregeln umfassen) ist ein unverbrüchliches Naturgesetz. Wahrscheinlichkeit kommt also nur einer Regel zu, welche unvollständig ausgedrückt und darum der Ausnahmen fähig ist; und sie wird gemessen durch das Verhältnis der Zahl der Regeln, in denen mit dem unvollständig ausgedrückten Vordersatz der Nachsatz verbunden ist, zur Gesammtzahl der Regeln, welche denselben Vordersatz mit einem zugehörigen Nachsatz überhaupt verbinden.

1) Philosophischer Versuch über die Wahrscheinlichkeiten, 1883.

Wahrscheinlichkeit bezieht sich nach dieser Definition niemals auf ein individuelles Ereignis.[1]) Gegenüber einem solchen könne, sagt Fick, in der That nicht von Wahrscheinlichkeit die Rede sein, da jedes individuelle Ereignis im Causalnexus entweder notwendig (und daher wirklich) oder unmöglich sei. Laplace habe dieser Schwierigkeit durch den Recurs auf unsere Unwissenheit entgehen wollen. Aber eben diese Behauptung, dass die Wahrscheinlichkeit teils von unserem Wissen teils von unserer Unwissenheit abhänge, sei geradezu falsch. Das Wahrscheinlichkeitsurteil habe mit irgend einer Unkenntnis absolut nichts zu thun; es drücke einen objectiven Sachverhalt aus. Auch setze jede Wahrscheinlichkeitsbestimmung die objective Geltung des Causalgesetzes voraus.

Erinnern wir uns der vorangeschickten Bemerkungen über Definitionsstreitigkeiten, so werden wir dem philosophischen Naturforscher nicht etwa entgegenhalten, dass Definitionen willkürlich und darum niemals falsch seien. Es bleiben doch die reellen Fragepuncte, wie sie dort bezeichnet sind. Aber eben in dieser Hinsicht kann ich die Schwierigkeiten der alten Lehre, welche Fick findet, nicht als solche anerkennen, und muss andererseits widersprechen, wenn die von ihm gegebene Definition als Ausdruck des gewöhnlichen wie des wissenschaftlichen Bewusstseins gelten soll. Von Wahrscheinlichkeit wäre nach Fick überhaupt nicht mehr die Rede, sobald wir uns nur präcis und vollständig ausdrücken. Und dies hat ja z. B. beim Münzwerfen nicht die geringste Schwierigkeit, obgleich immerhin auch die absolut horizontale Lage und mathematische Beschaffenheit der Unter-

1) Wie die Auffassungen sich entgegenstehen, sieht man daran, dass nach Lotze gerade nur bei einzelnen Thatsachen von mathematischer Wahrscheinlichkeit die Rede wäre. Er verhandelt die Wahrscheinlichkeitsregeln in dem Kapitel „Bestimmung singularer Thatsachen“.

lage unter den Voraussetzungen mitgenannt werden müssten, wie ja auch die Münze selbst von Fick als mathematische Ebene gedacht ist. Aber einer so nachlässigen Formulirung, wie in jenem „unvollständig ausgedrückten hypothetischen Urteil“, wird sich überhaupt Niemand schuldig machen. In anderen Fällen mögen wir ausser Stande sein, den Vordersatz genau zu formuliren, die Bedingungsclassen vollständig anzugeben. Dann ist aber auch eine Abzählung der Bedingungen, unter denen das im Nachsatz erwähnte Ereignis eintritt, und derjenigen unter denen es nicht eintritt, also eine mathematische Wahrscheinlichkeitsbestimmung nach Fick's Vorschriften unmöglich. Somit wären, wenn ich recht verstehe, nur Fälle einer unverantwortlich schlechten Redeweise das Anwendungsgebiet dieser hoch entwickelten Disciplin.[1])

Ueberdies, was nennen wir eigentlich bei jener unvollständig ausgedrückten Regel wahrscheinlich? Die Regel selbst oder das Ereignis, von welchem in ihrem Nachsatz die Rede ist? Die Regel als solche ist nicht wahrscheinlich, sondern sicher falsch. Denn wenn in einem hypothetischen Satze die Folge nicht immer eintritt, ohne dass irgend eine Clausel angedeutet wäre, so ist der Satz so, wie er ausgesprochen ist, falsch. Zum mindesten müsste der Nachsatz lauten: „so kann das Wappen oben liegen“. Und dann wäre die Regel wieder sicher und zweifellos wahr. Was wir wahrscheinlich nennen, ist in der That nicht die Regel

1) Wir müssen natürlich auseinanderhalten einen abgekürzten und einen unvollständigen Ausdruck. Wenn wir die unendlich vielen Bedingungen, welche durch die Winkel zwischen 0^0 und 90^0 gegeben sind, statt sie einzeln aufzuzählen, unter eben diesen Einen Ausdruck zusammenfassen, so kann man dies eine abgekürzte, aber nicht eine unvollständige Ausdrucksweise nennen. Und so ist auch die mathematische Wahrscheinlichkeit, wie schliesslich jedes Rechnungsergebnis, ein abgekürzter Ausdruck. Aber sie ist weder selbst eine unvollständige Ausdrucksweise noch ist sie eine Eigenschaft einer solchen.

sondern das Ereignis, und wir nennen es $\frac{1}{2}$ wahrscheinlich, nicht weil wir wissen, dass es unter der Hälfte der Bedingungen eintritt, unter der anderen fehlt, sondern weil wir nicht wissen, welche von den zwei Bedingungsarten in einem gegebenen Fall zutrifft. Das Vorhandensein der einen Bedingungsart und darum auch das Eintreffen des einen der beiden Ereignisse, also eine individuelle Thatsache, ist es, worauf sich der Wahrscheinlichkeitswert $\frac{1}{2}$ in solchem Falle bezieht und allein beziehen kann. Irgend etwas Anderes kommt überhaupt nicht in Frage; alles Uebrige ist sicher oder zum mindesten nicht blos $\frac{1}{2}$ wahrscheinlich.

Verhält es sich so, dann brauchen wir auch nicht mehr die absolut horizontale Lage u. s. f. als gegeben vorauszusetzen. Solange wir uns darüber ebenso in absoluter Unwissenheit befinden, wie über den Auffallswinkel — und dies ist der Fall, solange in der Problemstellung nichts über diese Puncte erwähnt ist —, ergibt sich der nämliche Wahrscheinlichkeitswert. Wären wir aber über sämmtliche Kräfte und Verhältnisse unterrichtet, so würde die Wahrscheinlichkeit $\frac{1}{2}$ in die Sicherheit eines bestimmten einzelnen Erfolges übergehen. Es scheint mir daher klar, dass nichts anderes als unsere Unkenntnis uns hier veranlassen kann, von Wahrscheinlichkeit überhaupt zu sprechen, dass also auch in dieser Beziehung die ältere Auffassung im Rechte bleibt.

2. Auch den ausführlichen, sehr anregenden Untersuchungen, mit welchen v. Kries die philosophische Litteratur bereichert hat,[1]) liegt der Gedanke zu Grunde, dass die „gleich möglichen Fälle“, auf deren Interpretation alles ankommt, nicht blos Fälle gleicher Unwissenheit sondern bestimmte objective Verhältnisse, einen physischen Spielraum bedeuten müssen. v. Kries gründet den Angriff gegen die

1) Die Principien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung. 1886.

ältere Lehre vorzugsweise auf den Umstand, dass danach eine feste Berechnung der Wahrscheinlichkeit überhaupt nicht möglich sei. Wir folgen schrittweise seinen Erwägungen.

a. Schon dies sei bedenklich, dass die nämliche Wahrscheinlichkeit $\frac{1}{2}$ resultire, wenn wir wissen, dass in einer Urne gleichviele weisse und schwarze Kugeln, und wenn wir nur wissen, dass überhaupt (nur) weisse und schwarze Kugeln darin sind.

Nun ist es an sich nicht wunderbar, wenn verschiedene Probleme zu einem gleichen Wahrscheinlichkeitswert führen. Was Kries befremdlich findet, ist, dass die Berechnung sich im einen Fall auf das objectiv bestehende und uns bekannte Verhältnis der Zahlengleichheit, im anderen dagegen nur auf unsere Unkenntnis gründe. Indessen gründet sich, wenn ich recht sehe, die Wahrscheinlichkeit doch auch im ersten Falle auf unsere Unkenntnis, nur nicht auf unsere Unkenntnis über das Zahlenverhältnis der Kugeln, sondern über ihre Lagerung und über die Richtung der hineingreifenden Hand. Welche von den unendlich vielen Combinationen dieser beiden Bedingungen auch vorhanden sein mag: immer kann ebensowol eine *w* als eine *s* Kugel die Stelle ausfüllen, nach welcher die Hand gerade greift. Wir erhalten also zwei Classen von gleich unendlich vielen Fällen und befinden uns bezüglich jeder Classe gegenüber der anderen in völliger Unwissenheit.[1])

Wird unser Wissen auch in Bezug auf die Richtung der Hand vervollständigt, so bleibt immer noch wegen der Unkenntnis der Lagerung dieselbe Wahrscheinlichkeit. Sind wir aber auch darüber orientirt, so ist von Wahrscheinlichkeit überhaupt nicht mehr die Rede. Nur also solange und

1) Zugleich wieder ein Beispiel, in welchem wie bei der geworfenen Münze die Disjunction der Ursachen nicht weiter führt als die directe Disjunction der Ereignisse. Kries spricht in solchen Fällen von „ursprünglichen Spielräumen".

soweit wir in Unkenntnis sind, solange und soweit findet eine Wahrscheinlichkeit statt, und die Berechnung gründet sich gerade auf das mit dem Wissen verknüpfte Nichtwissen, genauer auf die Anzahl der Fälle, über die wir in der fraglichen Beziehung nichts wissen.

Man könnte wol sagen, dass zwischen den beiden oben gegenübergestellten Wahrscheinlichkeitsurteilen ein Unterschied des Erkenntniswertes stattfinde. Demselben, richtig bestimmten, mathematischen Wahrscheinlichkeitsgrad eines und desselben Ereignisses kann man einen verschiedenen Wert zuschreiben, jenachdem wir uns nur über wenige oder über viele Umstände in absoluter Unkenntnis nach der Seite der *w* wie *s* Kugeln befinden. Denn es ist immer besser, wenigstens in theoretischer Hinsicht, mehr zu wissen als weniger; und unvernünftig wäre überhaupt jede Wahrscheinlichkeitsbestimmung, welche nicht auf so vielen Kenntnissen ruhte, als augenblicklich zu erlangen sind. Aber in Fällen, wo wir vorläufig oder überhaupt nicht weiter kommen als im Fall einer Urne mit unbekanntem Mischungsverhältnis, ist das Wahrscheinlichkeitsurteil darum doch nicht zu verachten. Am wertvollsten wäre die Wahrscheinlichkeitsbestimmung, wenn uns sämmtliche Umstände bekannt sind mit Ausnahme eines einzigen, z. B. der Lagerung der Kugeln (wenn man es nicht vielleicht für noch angenehmer halten möchte, zu wissen, dass die in gleichem Verhältnis vorhandenen *w* und *s* auch räumlich vollkommen regelmässig alterniren, immer eine *w* neben einer *s*, oder dass die *w* alle auf einer Seite, rechts oder links, liegen). Aber diese Wertunterschiede, wenn man sie so nennen will, haben nichts zu thun mit der Definition und Berechnung des Wahrscheinlichkeitsgrades. Er bleibt genau derselbe, ist in allen diesen Fällen gleich scharf definirt und stimmt nicht blos in sich selbst sondern auch in seinen mathematischen Consequenzen (vgl. III und IV) vollkommen mit dem gesunden Menschenverstand überein.

Diejenigen Kenntnisse, wodurch die Problemstellungen in den erwähnten Fällen sich unterscheiden, betreffen eben nicht blos die w oder die s, sondern beide gleichmässig; sie sind darum für die Wahrscheinlichkeit irrelevant.[1])

b. Hieran schliesse ich sogleich die Prüfung eines Bedenkens, auf das Kries mehrfach entscheidendes Gewicht legt (S. 31—34. 59). Wenn wir nur wissen, dass unter n Kugeln sowol weisse als schwarze und nur solche vorhanden sind, aber nicht wissen, in welchem Verhältnis, so sind nach der herkömmlichen Auffassung $n-1$ Fälle gleichmöglich, nämlich die Combinationen $1\,s + (n-1)\,w$, $2\,s + (n-2)\,w$ u. s. f. bis zu $(n-1)\,s + 1\,w$. Jede Combination hat also die Wahrscheinlichkeit $\frac{1}{n-1}$.[2]) Werden auch die Fälle mit-

1) A. Nitzsche bezeichnet den hier besprochenen Unterschied in einem soeben erschienenen Aufsatz (Viertelj. Sch. f. wissensch. Philos. 1892, XVI, S. 20 f.) als „Dimensionen der Wahrscheinlichkeit“. Denn es handle sich bei gleichem Wahrscheinlichkeitsansatz noch um die verschiedene Sicherheit, mit der die möglichen Fälle als gleichmöglich betrachtet werden können. Aber wenn es sich wirklich darum handelte, so würde dies nicht eine Wahrscheinlichkeit von anderer Dimension sein, sondern die vorher bestimmte wäre nicht richtig bestimmt, sie müsste noch mit der zweiten multiplicirt werden und wir hätten dann erst die richtig bestimmte Wahrscheinlichkeit, aber Wahrscheinlichkeit in demselben Sinne.

A. Meinong, auf dessen kritische Bemerkungen zu Kries' Theorie (Göttinger gel. Anzeigen 1890, S. 68 f.) diese Dimensionenlehre zurückgeht, hatte doch nicht von Dimensionen der Wahrscheinlichkeit gesprochen, sondern von Dimensionen des Urteils, dessen Eine Dimension die Wahrscheinlichkeit wäre, dessen andere Dimension er bereits als einen Wertunterschied bezeichnete. Wenn ich auch gegen die daran geknüpften psychologischen Ausführungen und die Anwendung des Dimensionsbegriffes manches einwenden möchte (wir gehen hier nicht auf psychologische Fragen ein), so trifft doch das obige Bedenken diese Fassung nicht.

2) Auch hieraus kann man ohne Weiteres die Wahrscheinlichkeit $\frac{1}{2}$ für den Zug einer w ableiten. Denn unter der Voraussetzung,

gezählt, dass blos *s* oder blos *w* vorhanden, so hat jede Combination die Wahrscheinlichkeit $\frac{1}{n+1}$.

Kries bestreitet nun, dass hier überhaupt ein fester Ansatz möglich sei. „Wir könnten recht wol auch sagen, dass die Füllung des Gefässes mit Kugeln blos einer Sorte und andererseits eine zufällige Durcheinandermischung beider Sorten am ehesten anzunehmen sei; es würde danach, wenn 1000 Kugeln vorhanden sind, den Annahmen, dass 1000, dass 500 oder dass gar keine schwarz sei, grössere Wahrscheinlichkeit zugeschrieben werden müssen, als etwa der, dass 873 schwarz seien. Der Versuch einer vollständigen Zergliederung aller Möglichkeiten würde sich in ein endloses Labyrinth verlieren und notwendig resultatlos bleiben.“ (S. 34.) „Es würden ... Fragen der verschiedensten Art sich aufdrängen, etwa ob die Anordnung in einer absichtlichen Weise, zu diesem oder jenem Zweck hergestellt worden sei u. dgl. ..... Demgemäss kann auch der übliche Wahrscheinlichkeitsansatz nur dann als begründet gelten, wenn entweder jede Anordnung der Kugeln als gleich wahrscheinlich angesehen werden darf, d. h. wenn man diese sehr sorgfältig durcheinandergemischt hat, oder wenn bekannt ist, dass in jedem kleinen Teile des ganzen Gefässes schwarze und weisse Kugeln in sehr annähernd oder genau demselben Zahlenverhältnis *n* zu *m* enthalten sind.“ (S. 59.)

---

dass 1 *w* auf $n-1\,s$ kommt, ist die Wahrscheinlichkeit, dass eine *w* gezogen wird, $= \frac{1}{n} \cdot \frac{1}{n-1}$. Unter der Voraussetzung, dass $2\,w$ auf $(n-2)\,s$, ist sie $= \frac{2}{n} \cdot \frac{1}{n-1}$ u. s. f. Die Wahrscheinlichkeit überhaupt, dass eine *w* gezogen wird, ist die Summe dieser Wahrscheinlichkeiten $= \frac{1+2+3+\cdots+(n-1)}{n\,(n-1)} = \frac{1}{2}$. Ebenso ergibt sich, wenn auch die Fälle, dass blos *s* oder blos *w* vorhanden, mitgezählt werden, die Summe $\frac{1+2+3\cdots+n}{n\,(n+1)} = \frac{1}{2}$.

Ich muss hier gestehen, dass mir die herkömmliche Berechnung irrig erscheint. Kries hat sich ein Verdienst erworben, indem er sie in Zweifel zog. Aber nicht kann ich zugeben, dass ein fester Ansatz überhaupt unmöglich wäre. Bei gegebener Anzahl n und unbekanntem Mischungsverhältnis ist, da jede individuelle Kugel gleichgut s und w sein kann, die Berechnung genau dieselbe, wie wenn ich n mal (nacheinander oder gleichzeitig) mit der Münze werfe. Die Combinationen: n weisse oder n schwarze Kugeln sind weniger wahrscheinlich als die gemischten, und bei diesen wiederum richtet sich die Wahrscheinlichkeit nach der Anzahl der möglichen Permutationen.[1])

Wenn gegeben ist, dass in einem Gefäss zwischen 1 und 1000 Kugeln sind und nach der Wahrscheinlichkeit des Vorhandenseins einer bestimmten Anzahl (ohne Rücksicht auf Farbe) gefragt wird, so ist diese natürlich für jede Anzahl, für 500 wie für 873, identisch $\frac{1}{1000}$. Anders liegt die Sache, wenn wie hier nach der Verteilung zweier Farben innerhalb 1000 Kugeln gefragt ist.

---

1) Auch hieraus folgt weiterhin für das Ziehen einer w die Wahrscheinlichkeit $\frac{1}{2}$. Sei die Anzahl der Kugeln n, so sind $n+1$ Hypothesen möglich: dass n, $n-1, \cdots$ bis dass 0 von den Kugeln weiss seien. Die Wahrscheinlichkeiten dafür sind $= (\frac{1}{2})^n \cdot \overset{1}{\underset{n}{B}}$, $(\frac{1}{2})^n \cdot \overset{2}{\underset{n}{B}}, \cdots$ bis $(\frac{1}{2})^n \cdot \overset{n+1}{\underset{n}{B}}$, wobei $\overset{k}{\underset{n}{B}}$ den k^{ten} Binomialcoefficienten (den Coefficienten des k^{ten} Gliedes) für den Exponenten n bedeutet, der erste und letzte also $= 1$ sind. Unter der ersten Voraussetzung nun ist die gesuchte Wahrscheinlichkeit, dass eine w gezogen werde, $= 1$. Unter der zweiten $= \frac{n-1}{n}$. Unter der dritten $= \frac{n-2}{n}$. U. s. f. Nun hat man die Producte aus den Werten der ersten und zweiten Reihe zu bilden und zu summiren. Dies gibt $(\frac{1}{2})^n \cdot \Sigma \underset{n-1}{B}$. Da aber $\Sigma \underset{n-1}{B} = 2^{n-1}$, so erhalten wir $\frac{2^{n-1}}{2^n} = \frac{1}{2}$.

Also diese, allerdings weittragende, Correctur der älteren nur auf einer Inconsequenz ruhenden Berechnung erscheint notwendig. Im Uebrigen aber vermag ich die Einwendungen nicht triftig zu finden. Sicherlich ist es denkbar, dass die Anordnung absichtlich hergestellt sei, und zwar zu den allerverschiedensten Zwecken. Aber in den Bedingungen des Problems liegt es (da in der Problemstellung alles, was wir wissen sollen, angegeben sein muss), dass wir absolut nichts wissen, ob eine absichtliche Anordnung hergestellt sei oder nicht. Ferner wenn eine solche Absicht herrschte, so kann sie freilich sehr verschieden gedacht werden, aber soviel ist offenbar: zu jeder Absicht, welche den *w* einen bestimmten Vorteil der Anzahl oder Lagerung verschaffen würde, lässt sich eine Absicht denken, welche denselben Vorteil den *s* verschaffte. Der Ordner könnte z. B. alle *w* — wie viel es auch sein mögen — absichtlich auf die rechte Seite gelegt haben, aber ebensogut auch alle *s* u. s. f. Wir brauchen also diesen Möglichkeiten nur ins Auge zu sehen, so verlieren wir uns keineswegs in ein Labyrinth, sondern erkennen mit mathematischer Evidenz, dass die Sache für unser Urteil in Hinsicht der *w* und *s* genau gleich steht, und erhalten darum auch nicht blos ungefähr, sondern genau den Wert $\frac{1}{2}$.

Zu schütteln pflegt man das Gefäss in Fällen, wo es sich um fortgesetzte Ziehungen handelt und wo die gezogenen Kugeln wieder hineingelegt werden: weil sich beim Hineingreifen leicht, auch ohne Absicht, eine gewisse Gleichförmigkeit ausbildet und dadurch in Verbindung mit einer gleichbleibenden Lagerung der Kugeln die vorausgesetzte Unabhängigkeit der Fälle gefährdet würde. (Das Gleichbleiben der Lagerung würde für sich allein noch keine Gefahr sein, wenn die aufeinanderfolgenden Handbewegungen völlig freie Variabilität besässen. Nur weil hierüber Verdacht besteht, muss dieser mögliche constante Factor durch fortwährend veränderte Lagerung der Kugeln unschädlich gemacht werden.) Wenn aber ein Gefäss vor mir steht, in das ich überhaupt noch nicht hineingegriffen habe und von dem ich nichts weiss als was in der Problemstellung angegeben ist, also auch nicht das Geringste über das Zustandekommen der Mischung, so liegt kein Anlass vor, die Kugeln zuerst noch in eine andere Lage zu bringen. Die Wahrscheinlichkeit wird hier weder verändert noch besser begründet, wenn ich die Urne vorher noch so kräftig schüttle.

Ich kann also nur wieder finden, dass der ältere Wahrscheinlichkeitsbegriff, consequent durchgeführt, in bester Uebereinstimmung bleibt sowol mit sich selbst als mit dem gewöhnlichen Menschenverstand.

c. Aber ein noch schwererer Vorwurf bleibt zu erledigen: Die alte Definition soll bei gleicher Problemstellung zu mehreren, ja beliebig vielen verschiedenen Wahrscheinlichkeiten führen. Ich erlaube mir, das erste Beispiel, woran Kries dies erläutert, sogleich auf ein durchsichtigeres allgemeines Schema zu bringen.[1]) Eine Kugel falle auf eine begrenzte Ebene, von der wir nur wissen, dass sie in 5 Teile $a\,b\,c\,d\,e$ zerfällt, während uns über die relative Ausdehnung derselben nichts bekannt ist. Für jeden Teil also Wahrscheinlichkeit $\frac{1}{5}$. Nun wird uns gesagt, dass der Teil a wieder in drei Teile $\alpha\,\beta\,\gamma$ zerfällt. Für jeden dieser Teile

1) Kries fragt: Wie wahrscheinlich ist das Auftreffen eines Meteors auf einen bestimmten Teil der Erdoberfläche, wenn wir nur wissen, dass es verschiedene Länder Namens Dänemark, Spanien u. s. f. gibt, aber nichts über ihre relative Grösse? Dann, wenn wir einige dieser Länder unter einen allgemeineren Begriff (genauer: einen umfassenderen Bezirk), wie Europa, zusammenfassen? — Was bei dieser besonderen Formulirung vorzüglich die unbefangene Erwägung stört und jede der berechneten Wahrscheinlichkeiten immer schon an sich wunderlich erscheinen lässt, ist der Umstand, dass wir thatsächlich so viel mehr über die relative Ausdehnung der Länder und der Erdteile wissen. A. Nitzsche hat denn auch in seiner Antwort auf dieses Argument dessen Spitze, wie mir scheint, nicht genügend erfasst.

also Wahrscheinlichkeit $\frac{1}{15}$. Wir können aber ebensogut diese 3 Teile von vornherein auch als selbständige Teile neben *b c d e* ansehen, und danach würde sich für je einen derselben vielmehr $\frac{1}{7}$ ergeben. Und so können wir überhaupt willkürlich jede beliebige Wahrscheinlichkeit für einen und denselben Teil berechnen. Kries sieht sich daher durch solche Erwägungen zu der positiven Forderung geführt, dass die zu vergleichenden Fälle objectiv gleiche „Spielräume" bilden und als solche uns bekannt sein müssen; also z. B. dass in unserem Falle die Ebene in eine bestimmte Anzahl unter sich gleicher Bezirke geteilt sei.

Das scharfsinnige Argument lehrt ohne Zweifel, dass in den Voraussetzungen irgend eine Absurdität liegen müsse. Aber sie liegt, wie mir scheint, nicht in der Fassung des Wahrscheinlichkeitsbegriffes, nicht in den darin niedergelegten allgemeinen Bedingungen seiner Anwendung, sondern in der besonderen Natur der Voraussetzungen, unter denen er hier angewandt werden soll. Wir sollen zuerst über die Ebene nichts weiter wissen, als dass sie in 5 mit den Buchstaben *a* bis *e* bezeichnete Teile zerfällt. Aber wir wissen ja factisch bei jedem Continuum, dass jeder seiner Teile wieder Teile hat und so ins Unendliche; und dies ist nicht eine Eigenschaft, die wir etwa in Folge gewohnheitsmässiger Associationen oder von Erfahrungen hinzudenken und von der wir darum in der Problemstellung absehen könnten, sondern sie liegt in dem Begriff des Continuums. Wir wussten von vornherein, dass der Teil *a*, und nicht blos dieser, wieder 3 Teile, und nicht blos 3, enthält. Jede Abgrenzung ist willkürlich. Daran ändert auch die Buchstabenbezeichnung nichts, denn auch das wissen wir vorher, dass auf dem mit *a* bezeichneten Teil beliebig viele andere Buchstaben stehen können und ebenso auf jedem anderen (die Grössenvorstellungen sind ja beliebig), sowie dass es für die Wahrscheinlichkeit keinen Unterschied macht, welchem Alphabet sie angehören.

Die Fragestellung enthält also, indem sie eine unmögliche Beschränkung des Wissens verlangt, von Anfang an eine innere Absurdität; kein Wunder, dass sie in den Consequenzen zum Vorschein kommt. Es hat hier, da ich mich in Bezug auf alle beliebig kleinen Teile in absoluter Unwissenheit befinde, keine andere Fragestellung von vorherein einen Sinn als diese: „Welcher mathematische Punct wird getroffen?" — wobei sich für jeden die Wahrscheinlichkeit $\frac{1}{\infty}$ ergibt, vorausgesetzt, dass die auftreffende Kugel ebenfalls eine mathematische Kugel ist. Ausserdem sind so viele Teile der Ebene von vornherein zu zählen, als wie vielmal das von der Kugel bedeckte Flächenelement in der Ebene enthalten ist.

Modificiren wir das Beispiel in folgender Art. Eine begrenzte Ebene sei in 5 Bezirke geteilt, über deren relative Grösse wir nichts wissen (oder sogar wissen, dass sie ungleich gross sind), jeder Bezirk bilde aber den Eingang eines mit einem Buchstaben *a* — *e* bezeichneten Beutels. Wie wahrscheinlich ist es, dass eine Kugel, von der wir lediglich wissen, dass sie von oben her in der Ebene auftrifft und dass sie nicht zu gross ist, um in irgend einen der Beutel zu fallen, sich in einem bestimmten Beutel z. B. *c* finden wird? Hier liegt in der Frage keine Absurdität, weil das Continuum von Teilen in ein Discretum verwandelt ist. Auch hier wissen wir freilich, dass jeder Teil der Ebene wieder Teile enthält, aber es werden nur solche als verschieden gezählt, die zu verschiedenen Beuteln gehören, und dieser sind nicht mehr und nicht weniger als 5. Es ist überhaupt nicht mehr nach Teilen der Ebene sondern nach Beuteln gefragt. Der Sinn der Frage wäre derselbe, wenn die 5 Beutel sich an ganz verschiedenen weitgetrennten Orten befänden und wir nur wüssten, dass in Einem derselben eine Kugel liegt.

Machen wir aus einem Beutel drei, so geht freilich auch in unserem Beispiel die Wahrscheinlichkeit für jeden der übrigen aus $\frac{1}{5}$ in $\frac{1}{7}$ über. Aber dann ist eben das Problem, das Gegebene, auch ein anderes, und da wir absolut nicht wissen, welcher geteilt wurde — es kann der kleinste, der grösste u. s. f. gewesen sein — so ist die Veränderung der Wahrscheinlichkeit ebenso gerechtfertigt, wie wenn wiederum die 5 Beutel an weit getrennten Orten wären und noch 2 dazu kämen, während uns nur gegeben wäre, dass in Einem eine Kugel liegt.

Kries hat aber auch durch dieses Argument, obschon es sein eigentliches Ziel nicht trifft, uns auf eine bemerkenswerte Eigentümlichkeit gewisser Fälle hingewiesen, welche zu Trugschlüssen leiten kann, wie solche ja im Wahrscheinlichkeitsgebiete besonders zahlreich und interessant sind. Dagegen dürften in einem weiteren Beispiel doch nur logisch-mathematische Versehen vorliegen:

d. „Dass Eisen im Sirius vorhanden sei, hat, wenn gar keine Anhaltspuncte für oder gegen vorliegen, die Wahrscheinlichkeit $\frac{1}{2}$. Ebensogross ist die Wahrscheinlichkeit, dass es sich nicht dort findet. Das Nämliche gilt von jedem anderen irdischen Stoff. Dass also alle 68 irdischen Elemente — wenn wir so viele annehmen — sich n i c h t im Sirius finden, hat die Wahrscheinlichkeit $(\frac{1}{2})^{68}$; somit ist die Wahrscheinlichkeit, dass mindestens eines derselben dort sei, $1 - (\frac{1}{2})^{68}$. Diese ungeheure Wahrscheinlichkeit, während wir doch nach der Voraussetzung gar keinen Anhaltspunct für und gegen jedes Element haben, widerspricht schon dem gesunden Menschenverstand. Aber wir kommen auch in einen inneren Widerspruch: denn wenn wir sogleich nach der Wahrscheinlichkeit fragen, dass der Sirius überhaupt irdische Stoffe enthält, so ergibt sich $\frac{1}{2}$. Dieselbe Frage gibt also bei denselben Anhaltspuncten einmal eine ungeheure und einmal nur eine mittlere Wahrscheinlichkeit."

In diese (hier etwas verkürzte) Argumentation hat sich, wenn ich recht sehe, doch ein starker Fehler eingeschlichen. Angenommen, dass die Wahrscheinlichkeit $\frac{1}{2}$ für das Nichtvorhandensein eines bestimmten Elementes richtig angesetzt sei, so ergibt sich $(\frac{1}{2})^{68}$ dafür, dass nicht alle 68 zusammen da sind. Aber das contradictorische Gegenteil hievon, wofür sich dann der Wert $1 - (\frac{1}{2})^{68}$ ergibt, ist nicht, wie Kries sagt: „dass mindestens eines derselben vorhanden sei", sondern: „dass entweder nur ein Teil der 68 (also 67 oder 66 u. s. f.) oder gar keines von ihnen vorhanden sei". Nimmt man dieses zweite Glied der Alternative, welches Kries übersehen, mit hinzu, so dürfte die berechnete Wahrscheinlichkeit dem gewöhnlichen Menschenverstand schon nicht mehr so ferne stehen. Denn wenn man wirklich absolut nichts wüsste, was zu Gunsten einer Gleichheit der Materie in der Welt spräche, auch nichts von der Allwirksamkeit der Schwerkraft, die auch schon einigermassen den Gedanken an stoffliche Einheit begünstigt, so würde man die Annahme, dass sämmtliche irdische Stoffe auf einem beliebig herausgegriffenen Stern vertreten seien, höchst unwahrscheinlich, und das Gegenteil davon, die obige Alternative, entsprechend wahrscheinlich finden (das zweite Glied dieser Alternative vielleicht noch wahrscheinlicher als das erste — wie es denn Aristoteles trotz der schon von Anaxagoras richtig gedeuteten Meteorsteine nur natürlich fand, dass die Sterne aus einem fünften Element beständen — doch ist über die relative Wahrscheinlichkeit der beiden Glieder der Alternative in dem Wahrscheinlichkeitswerte $1 - (\frac{1}{68})^2$ nichts gesagt).

Ausserdem ist der Ansatz $\frac{1}{2}$ für das Vorhandensein eines bestimmten Elementes, wie des Eisens, im Sirius unter Voraussetzung völligen Mangels empirischer Gründe pro und contra viel zu hoch gegriffen. Eisen ist, wenn wir vorerst unsere chemische Definition zu Grunde legen, charakterisirt durch das Verbindungsgewicht 56. Wenn wir nun absolut nichts darüber wissen, innerhalb welcher Grenzen die Verbindungsgewichte der in der Welt vorhandenen Grundstoffe überhaupt variiren können (unsere Kenntnisse sollen sich ja voraussetzungsgemäss auf die Erde beschränken), so ist die Zahl der denkbaren Verbindungsgewichte, also der denkbaren Elemente, offenbar immer noch unbegrenzt. Bedenken wir oder geben wir zu, dass irgend welche Grenzen doch vorhanden sein müssen, dass ferner durch die chemischen Gesetze, mit denen der Begriff des Verbindungsgewichts zusammenhängt, stetige Uebergänge ausgeschlossen sind, so wird die Zahl zwar endlich, aber sie kann immer noch, da wir nicht wissen, wie kleine discrete Abstufungen zwischen den Verbindungsgewichten möglich sind, ungeheuer, ja beliebig gross gesetzt werden. Das Vorhandensein eines Stoffes mit dem Verbindungsgewicht 56 würde also bei solchem Zustand unserer Kenntnisse Ein Fall unter einer ungeheuren Anzahl gleich möglicher Fälle, daher äusserst unwahrscheinlich sein. Noch mehr, wenn wir „Eisen" ausser durch sein Verbindungsgewicht noch durch eine Reihe chemischer und physikalischer Eigenschaften definirt sein lassen, welche erfahrungsgemäss zusammen vorkommen, aber vorläufig nicht auseinander ableitbar sind, von denen wir also nicht wissen, ob einzelne darunter nicht bei einem Element von gleichem Verbindungsgewicht auf dem Sirius einen anderen Wert besitzen könnten. Ein solches Element könnte dann eine so wesentlich verschiedene allotrope Modification darstellen, dass wir Bedenken tragen würden, es noch unter den Begriff des „Eisens" zu subsumiren. Dadurch wird dann die Wahrscheinlichkeit, dass genau diese Verknüpfung von Eigenschaften, die wir mit diesem Worte belegen, sich vorfinde, noch geringer.

Nennen wir diese Wahrscheinlichkeit $\frac{1}{Z}$, wo also Z eine ungeheure Zahl bedeute, so folgt für die Wahrscheinlichkeit, dass sich nicht alle unsere Elemente auf dem Sirius befinden, dass also nur irgend ein Teil davon oder gar keines vorhanden sei, $1 - \left(\frac{1}{Z}\right)^{68}$. Diese Wahrscheinlichkeit ist also jedenfalls noch viel grösser als sie schon nach Kries' Berechnung sein würde.

Legen wir statt der Definition durch das Verbindungsgewicht nur die durch physikalische Eigenschaften zu Grunde, so folgt doch Aehnliches, da auch diese ebenso verschieden gedacht werden können. Gebrauchen wir das Wort „Eisen" überhaupt in einer vageren Bedeutung, entsprechend etwa den Vorstellungen des Altertums, das von bestimmten Schmelzpuncten, Festigkeitscoefficienten u. s. f. ebenso wenig wie von Verbindungsgewichten wusste und wol kaum Anstand genommen hätte, sich die Eigenschaften eines „Elementes", wie überhaupt eines gegebenen Stoffes, innerhalb einer gewissen Breite stetig variabel zu denken: so wächst zwar die Wahrscheinlichkeit für das Vorhandensein eines solchen Stoffes, bleibt aber auch dann noch ausserordentlich gering.

Dass ein genauer Wert sich auch unter den vorhergenannten Voraussetzungen nicht berechnen lässt, liegt nicht — wie man hier vielleicht einwenden möchte — an der Unbestimmtheit des Wahrscheinlichkeitsbegriffes, sondern nur an der Unmöglichkeit, die im Begriffe verlangten und an sich vollkommen genügenden Bedingungen im vorliegenden Falle zu erfüllen, nämlich eine bestimmte Zahl für die „gleichmöglichen Fälle" anzugeben. Kein noch so verclausulirter Wahrscheinlichkeitsbegriff würde darüber hinweghelfen. Man würde fast ebensogut eine mathematische Wahrscheinlichkeit dafür verlangen können, dass Menschen auf dem Sirius wohnen. Aber dies alles hindert nicht, dass wir den Wert $\frac{1}{2}$ für das Eisen ebenso wie für Menschen als unrichtig bestimmt und als ungeheuer viel zu gross bezeichnen müssen.

Drittens endlich ist auch der Ansatz $\frac{1}{2}$ für die Glieder der letzten Alternative (S. 71 u.): „ob der Sirius überhaupt irdische Stoffe enthält oder nicht", nicht richtig aus den Wahrscheinlichkeitsregeln abgeleitet. Die Wahrscheinlichkeit, dass der Sirius überhaupt (= irgendwelche) irdische Elemente enthalte, ist die Summe der Wahrscheinlichkeiten für die einzelnen Elemente; ebenso wie die Wahrscheinlichkeit, beim Würfeln eine gerade Zahl zu werfen, die Summe der Wahrscheinlichkeiten für jede einzelne gerade Zahl ist. Sie ist also, wenn wir 68 irdische Elemente statuiren, $= \frac{68}{Z}$, bleibt somit immer noch ein äusserst kleiner Bruch.

Hieraus sieht man zugleich wieder, wie wenig es angeht, die Wahrscheinlichkeit für ein bestimmtes irdisches Element $= \frac{1}{2}$ zu setzen. Denn es würde ja dann als Wahrscheinlichkeit für irgend ein irdisches Element $\frac{68}{2}$ resultiren.[1])

1) Auch A. Nitzsche's Berechnung kann ich nicht zustimmen, obschon er erkannt hat, dass der Fall „Irgend eines ist vorhanden" dem Fall „Keines ist vorhanden" nicht gleich steht sondern eine Summe vieler Fälle umfasst. Er setzt dagegen die Fälle als gleich, dass von den 68 keines oder eines oder 2 oder 3 . . . oder alle vorhanden seien, und findet so für jeden dieser Fälle, auch den ersten, $\frac{1}{69}$; woraus er dann weiter für das Vorhandensein eines bestimmten Elementes auch wieder $\frac{1}{2}$ ableitet. Aber der erste Fall ist nicht „gleichmöglich" mit einem der folgenden.

e. Wenn zwei Karten verkehrt auf dem Tische liegen, so bestehen nach gewöhnlicher Annahme hinsichtlich ihrer Farbe 4 gleichmögliche Fälle: sie können beide schwarz, beide rot, es kann *a* (z. B. die rechtsliegende) rot und *b* schwarz, und es kann *b* rot und *a* schwarz sein. Indessen, sagt Kries, kann man ebensogut einteilen: keine *s* (schwarz), beide *s*, eine *s*. Es ergibt sich daher für eine und dieselbe Annahme, z. B. „beide *s*“, sowol die Wahrscheinlichkeit $\frac{1}{4}$ als $\frac{1}{3}$.

Aber hier ist die übliche Einteilung doch allein consequent. Die andere unterscheidet sich nur dadurch von ihr, dass sie zwei Fälle unter Einem Ausdruck zusammen nimmt. Dass aber diese beiden als zwei zu zählen sind, darüber lässt uns der alte Wahrscheinlichkeitsbegriff nicht im Zweifel. Wir befinden uns in absoluter Unwissenheit sowol über die Karte *a* als *b*, und jedesmal über zwei Farben. Immer muss als Zahl der möglichen Fälle das Maximum der innerhalb der Problemstellung noch unterscheidbaren Fälle genommen werden. Jede der Eigenschaften, um die es sich handelt, muss in Bezug auf jedes Individuum, um das es sich handelt, als ein besonderer Fall betrachtet werden; und nicht minder muss jede mögliche Permutation als besonderer Fall gelten. Wird ausdrücklich in der Fragestellung die Ordnung (Permutation) als irrelevant bezeichnet, so sind eben nachher die betreffenden Einzelfälle zusammenzuzählen. Aber zunächst müssen sie als gleichmögliche allen übrigen coordinirt werden. Sobald man sich freilich gestattet, mehrere innerhalb der Problemstellung noch unterscheidbare Fälle sogleich unter Einen Ausdruck zu fassen, ist der Willkür Thür und Thor geöffnet. Wir könnten dann auch noch einfacher disjungiren: entweder beide sind *s* oder nicht, und erhielten dann sogar $\frac{1}{2}$ für jedes Glied der Alternative, statt $\frac{1}{4}$ für das erste und $\frac{3}{4}$ für das letzte. Und wären es 50 Karten, so könnten wir ebenso disjungiren: entweder alle 50 sind *s* oder nicht, und erhielten ebenfalls $\frac{1}{2}$, statt $(\frac{1}{2})^{50}$ für das erste und $1 - (\frac{1}{2})^{50}$ für das letzte Glied.[1])

Allerdings ist Poisson, welchen Kries hier citirt, in einen ähnlichen Fehler verfallen. Wenn wir eine der Karten aufdecken und sie schwarz finden, so ist die Wahrscheinlichkeit, dass die andere *s* und dass sie *r*, für den unbefangenen Verstand je $= \frac{1}{2}$. Und mit Recht, denn von den 4 Möglichkeiten sind jetzt 2 hinweggefallen, dass nämlich beide rot und dass nur *a* (dies mag die aufgedeckte sein) rot wäre. Bleiben also 2 Möglichkeiten, ganz so wie wenn von vornherein nur die jetzt übriggebliebene auf dem Tische gelegen hätte. Poisson berechnet aber $\frac{2}{3}$ dafür, dass sie *s*, $\frac{1}{3}$ dafür, dass sie *r* sei. Nach Kries wären auch hier beide Ansätze auf Grund des alten Wahrscheinlichkeitsbegriffes gleichberechtigt.

Aber es liegt wieder nur ein Irrtum auf Seiten Poisson's vor, der sich daraus erklärt, dass er hier eine Regel, nach welcher aus der Wahrscheinlichkeit vergangener Ereignisse diejenige künftiger bestimmt werden soll (Bayes' Princip) in unrichtiger Weise angewandt hat. Ich hoffe anderwärts zu zeigen, dass auch aus diesem complicirteren Gedankengange bei correcter Durchführung dieselbe Wahrscheinlichkeitsbestimmung folgt wie aus der einfachsten Ueberlegung.[2])

---

1) Seltsam rechnet hier Nitzsche. Es ergebe sich, sagt er, auch nach der Kries'schen Einteilung die Wahrscheinlichkeit $\frac{1}{2}$ für eine Farbe, z. B. rot, als durchschnittliche Wahrscheinlichkeit. Es sei nämlich die Wahrscheinlichkeit, dass eine der Karten *r*, für den Fall, dass beide *s*, $= 0$; für den Fall, dass beide *r*, $= 1$; und für den Fall, dass sie gemischt sind, $= \frac{1}{2}$. Somit durchschnittlich $= \frac{1}{2}$. Ich verstehe nicht, wiefern die Wahrscheinlichkeit, dass eine der Karten rot sei, für den Fall, dass eine rot und eine schwarz ist, noch $= \frac{1}{2}$ und nicht vielmehr $= 1$ sein soll.

2) Wenn aus einer Urne mit *w* oder *s* Kugeln von unbekanntem Mischungsverhältnis eine *w* gezogen und nicht wieder hineingelegt wurde, so ergibt sich die Wahrscheinlichkeit, dass beim nächsten Zug *w* erscheine, unabhängig von der Anzahl der Kugeln stets $= \frac{1}{2}$

Ohnedies wird man, wo eine verwickeltere Deduction mit dem evidenten Ergebnis der einfachsten Ueberlegung in Widerspruch tritt, nicht dieser letzteren oder beiden sondern nur jener Deduction mistrauen müssen.

Man findet sich hier auch an einen von D'Alembert gegen die Grundlagen der Wahrscheinlichkeitslehre erhobenen und zeitlebens festgehaltenen Einwurf erinnert. Die Wahrscheinlichkeit, dass bei zweimaligem Werfen einer Münze wenigstens einmal das Wappen erscheint, ist nach der gewöhnlichen Berechnung $\frac{3}{4}$. D'Alembert meinte, da man doch nicht weiter spielt, wenn sogleich beim ersten Mal Wappen erscheint, so seien nur drei Fälle möglich: Schrift Schrift, Schrift Wappen, Wappen —; also die verlangte Wahrscheinlichkeit $\frac{2}{3}$. Aber es ist ja für die logische Disjunction

---

(nach Poisson = $\frac{5}{6}$). Wird sie wieder hineingelegt, so ist diese Wahrscheinlichkeit bei m Kugeln $= \frac{m+1}{2m}$ $\left(\text{nach Poisson } \frac{2m+1}{3m}\right)$. Diese Unterschiede rühren daher, dass wir die Wahrscheinlichkeit der verschiedenen Mischungsverhältnisse nicht als gleich ansehen dürfen (s. o. S. 66).

Wurde n mal eine w gezogen und immer wieder hineingelegt, so ist die Wahrscheinlichkeit für einen $n+1^{\text{ten}}$ weissen Zug =

$$\frac{\Sigma \overset{k}{\underset{m-1}{B}} \cdot (m-r)^n}{m\,\Sigma \overset{k}{\underset{m-1}{B}} \cdot (m-r)^{n-1}}$$

worin k von 1 bis m zunimmt, $r = k - 1$, und $\overset{k}{\underset{m-1}{B}}$ die k^{ten} Binomialcoefficienten (den ersten = 1 mitgezählt) für den Exponenten $m - 1$ bedeutet. Die Formel geht für $n = 1$ in die vorige über.

Die berühmte Formel $\frac{n+1}{n+2}$ ist für Fälle abgeleitet, in denen unendlich viele Hypothesen g l e i c h m ö g l i c h wären. Aber die Mischungsverhältnisse in einer Urne sind niemals alle gleichmöglich; und so stehen denn auch die Folgerungen, die sich aus der Formel für diesen Fall ergeben würden, keineswegs in Einklang mit dem gesunden Menschenverstand.

der Fälle ganz gleichgültig, was der Spieler thun wird. Gegen die Principien der Wahrscheinlichkeitsrechnung folgt daraus ebensowenig etwas, als durch die bekannte Vexirfrage der Schulkinder: „Wie viel Sperlinge sitzen noch auf dem Dach, wenn von 100 einer weggeschossen wurde?" die Grundlagen der Arithmetik ernstlich gefährdet sind. Solche Irrtümer lehren nur, wie nicht blos mathematisch sondern auch philosophisch geschulte Denker in Sachen der Wahrscheinlichkeit Misgriffen ausgesetzt sind, die darum nicht weniger Misgriffe bleiben.[1])

Nach Allem kann ich mich nicht überzeugen, dass die frühere Auffassung, richtig verstanden und durchgeführt, in

---

1) Hat doch selbst Leibniz die Wahrscheinlichkeit, mit zwei Würfeln 12 zu werfen, und die Wahrscheinlichkeit, 11 zu werfen, gleich gross, und die Wahrscheinlichkeit, 7 zu werfen, nur 3 mal grösser als die für 12 gesetzt, während diese Wahrscheinlichkeiten bezw. $= \frac{1}{36}, \frac{2}{36}, \frac{6}{36}$ sind. Vgl. Todhunter, History of the mathematical Theory of Probability 1865 p. 48. Daselbst Ch. XIII (auch p. 253, 256) die Geschichte des von D'Alembert geführten Krieges.

Ein anderes von D'Alembert's Bedenken (das. p. 268 Nr. 484 und p. 277 Nr. 505) wurde später auch von Fries, ebenfalls Mathematiker und Philosoph zugleich, in seiner Logik (2. Aufl. 1819 S. 457) wiederholt. „Wenn nach dem wahrscheinlichen Lebensalter eines neugeborenen Kindes gefragt wird, so sagt Halley: man müsse es bei uns auf $21^1/_2$ Jahre setzen, denn von 10000 Menschen sterben 5000 vor diesem Alter und 5000 nach ihm, der Einzelne ist also vorher mit überwiegender Wahrscheinlichkeit unter den Lebenden, nachher unter den Todten. Deparcieux hingegen sagt, dies Alter sei 29 Jahre, denn 10000 Menschen leben zusammen 290000 Jahre, wo also im Durchschnitt auf einen 29 Jahre kommen. Wer hat nun Recht? Das lässt sich im Allgemeinen gar nicht bestimmen." — In Wahrheit sind durchschnittliches Lebensalter und wahrscheinliches Lebensalter zwei ganz verschiedene Begriffe, und wer sie zusammenwirft, begeht einfach eine Verwechslung. Vgl. Todhunter l. c.

In seinem „Versuch einer Kritik der Principien der Wahrscheinlichkeitsrechnung" (1842) ist denn auch Fries selbst nicht mehr auf dieses Bedenken zurückgekommen sondern hält beide Begriffe auseinander, S. 162.

irgend einem Puncte zu unannehmbaren Folgen führte, und dass besondere Voraussetzungen hinsichtlich des objectiven Thatbestandes die unumgängliche Bedingung für den mathematischen Wahrscheinlichkeitsansatz bildeten. „Objectiv gleiche Spielräume" müssen allerdings in dem Sinne vorliegen, dass das Princip der Disjunction aus der Materie genommen sein muss, wie bei jeder guten Einteilung. Wenn der Fall „2 mal rot" für mich besonderes Interesse hat, dürfen wir nicht um deswillen einteilen: „2 mal rot" und „nicht 2 mal rot". Aber man kann es ja gerade den Vertretern der alten Wahrscheinlichkeitslehre im Allgemeinen weniger als ihren Gegnern zum Vorwurf machen, dass sie gegen diese Vorschrift gesündigt hätten.[1])

III. Der Wahrscheinlichkeitsbegriff im Theorem Jacob Bernoulli's.

Wie ist es nun möglich, aus einem Wahrscheinlichkeitsansatz, wenn er nur unsere Unkenntnis in Bezug auf eine

1) Auch die Bedingungen, welche Kries als „indifferente" und „vergleichbare" Spielräume bezeichnet, kann man in gewissem Sinn zugeben. Indessen hat es keinen Zweck, im Ausdruck Uebereinstimmung zu suchen, wo die Behauptungen doch inhaltlich auseinandergehen. Um so wertvoller ist mir die wirkliche Uebereinstimmung in mehreren wesentlichen Puncten: dass der Wahrscheinlichkeitsbegriff (gegenüber Fick) auf das wirkliche Stattfinden eines einzelnen Ereignisses, und nicht blos eines Ereignisses sondern auch einer Thatsache, auf einen bestehenden Sachverhalt z. B. räumliche Collocationen („Configurations-Spielräume"), ja auf Naturconstanten angewandt werden kann; ferner, dass Wahrscheinlichkeitsbestimmungen nicht auf dem Causalprincip beruhen. Kries setzt diesem ein ebenbürtiges „Princip der freien Spielräume" zur Seite, welches unsere Erwartungen in solchen Fällen statt des Causalprincips regle. Doch möchte ich glauben, dass wir hier nicht ein eigenes logisches Princip nötig haben, sondern dass der Wahrscheinlichkeitsansatz als eine Ableitung aus einer vollständigen Disjunction auf den gewöhnlichen logischen Axiomen ruht.

Anzahl coordinirter möglicher Fälle ausdrückt, mit Bernoulli einen annähernd sicheren Schluss zu ziehen auf die Häufigkeit eines Ereignisses in einer grossen Anzahl künftiger Fälle?

Es liesse sich eine Maschine herstellen, die einen Würfel so drehte, dass ganz regelmässig eine Zahl nach der anderen obenauf käme. Sind aber überhaupt Schwankungen und Unregelmässigkeiten der Folge möglich, warum muss dann doch unter vielen Fällen eine bestimmte Zahl annähernd ihrer Chance gemäss auftreten? Man fühlt sich versucht, an einen geheimnisvollen physischen Process der Ausgleichung zu denken, durch welchen allmälig, wie sich ja Laplace auch einmal ausdrückt, Ordnung in das Chaos der Erscheinungen kommt. Und doch widerspricht eine solche Annahme der vorausgesetzten Unabhängigkeit der einzelnen Fälle, ohne welche auch jene vorgängigen Chancen gar nicht behauptet werden könnten.

Prevost und Lhuilier bezeichneten es geradezu als eine notwendige Fiction der Wahrscheinlichkeitsrechnung, dass die gleichmöglichen Fälle der Reihe nach darankommen, wie in der obigen Maschine.[1]) Aber wir werden uns kaum zu dieser verzweifelten Ausflucht entschliessen.

J. St. Mill ergriff in den ersten Auflagen seiner Logik den Ausweg, die Wahrscheinlichkeit kurzweg durch die relative Häufigkeit des Vorkommens innerhalb einer grossen Anzahl von Fällen zu d e f i n i r e n. Später ist er davon zurückgekommen[2]), während J. Venn denselben Standpunct

---

1) Mémoires de l'Academie de Berlin 1796. Cl. de phil. spec. p. 4. „Lorsqu' en vertu d'une certaine determination des causes plusieurs événements nous paraissent également possibles, nous feignons que tous ces événements ont lieu successivement tour à tour et sans répétition." („Hypothèse stochastique.")

2) Allerdings nicht ganz und consequent. Was in der Gomperz'schen Uebersetzung II (Ges. Werke III) 260—1 steht, kann ich mit der Zurücknahme auf S. 254 nicht zusammenreimen.

noch energischer verfocht.[1]) Dieses Durchhauen des Knotens beruht aber vor allem auf mangelnder Einsicht in das Bernoulli'sche Gesetz selbst. Es sei zunächst gestattet, die diesem Gesetz zu Grunde liegenden Gedanken, absehend von der Beweisführung, die ja in jedem Lehrbuch zu finden ist, so darzulegen, dass der Wahrscheinlichkeitsbegriff darin möglichst rein und scharf hervortritt. Wir mögen dabei der Kürze halber nur von Ereignissen sprechen und an die gewöhnlichen schematischen Beispiele denken, aber im Gedächtnis behalten, dass Alles ebenso auf beliebige Urteilsmaterien anzuwenden wäre, welche die in I angegebenen Merkmale besitzen.

1. Bernoulli's Theorem.

Werfen wir zweimal mit einer Münze (a bedeute Wappen, b Schrift), so kann aa, bb, ab, ba resultiren. Kommt es uns nur auf die Anzahl der in einem solchen „zusammengesetzten Ereignis" vorkommenden einfachen Ereignisse an, nicht auf ihre Anordnung, so hat die Verteilung $2a$ die Wahrscheinlichkeit $\frac{1}{2^2}$, $2b$ ebenso, dagegen $1a\ 1b$ wegen der möglichen Permutation $\frac{2}{2^2}$. Werfen wir 3 mal, so ergibt sich für $3a$ und $3b$ je $\frac{1}{2^3}$, dagegen für $2a\ 1b$ und $2b\ 1a$ je $\frac{3}{2^3}$. U. s. w. Die Verhältnisse sind in leicht ersichtlicher

1) Logic of Chance 1866. Venn's gewaltige Misverständnisse zeigen sich besonders im Ch. IX. Wer den Satz aussprechen kann (p. 213): „If the chances (gegen das 10 malige Treffen von „Wappen") be 1023 to 1..., it undoubtedly will happen once in 1024 times", der hat keinen Begriff von dem Bernoulli'schen Theorem. Das Vorkommen in grossen Anzahlen gemäss den Chancen wird ihm absolut gewiss (weil er eben die Chancen durch das Vorkommen definirt), während ihm zugleich die Bestimmung der Chancen selbst in allen Fällen höchst fraglich wird, da die Erfahrungen niemals dazu hinreichen (p. 218).

Weise durch die Entwickelung des Binoms $a + b$ ausgedrückt, dessen Exponent die Zahl der Fälle (x) ist. Die Entwickelung beruht ja auf demselben Verfahren der Combination und Permutation. Denjenigen zusammengesetzten Ereignissen, worin die beiden einfachen Ereignisse möglichst gleichmässig — das heisst ihrem Chancenverhältnis entsprechend — verbunden sind, entspricht immer eine grössere Anzahl möglicher Fälle als denen, worin nur je ein einfaches Ereignis vertreten ist; mit andern Worten eine grössere Wahrscheinlichkeit. Die absolute Wahrscheinlichkeit jedes einzelnen zusammengesetzten Ereignisses, einer jeden Verteilungsart, wird zwar immer geringer mit wachsender Zahl der Fälle x, aber die der gleichmässig gemischten wird relativ zu den aus blosser Wiederholung Eines Ereignisses bestehenden immer grösser.

Für den Fall ungleicher Chancen der einfachen Ereignisse mögen Beispiele den Gang der Verhältnisse erläutern. a habe die Wahrscheinlichkeit $\frac{4}{5}$, b die $\frac{1}{5}$ (d. h. also: es seien 5 noch einfachere gleichmögliche Ereignisse $\alpha\, \beta\, \gamma\, \delta\, \varepsilon$ unterscheidbar, wovon 4 unter den Begriff a, 1 unter b fallen), und die Zahl der einfachen Ereignisse a oder b, welche sich zu einem zusammengesetzten verbinden, sei 5: so erhalten wir als gemeinsamen Nenner der Wahrscheinlichkeitsbrüche 5^5, als Zähler aber für $5\,a$: 1024, für $5\,b$: 1, für $4\,a\;1\,b$ (mit 5 Permutationen): 1280; für $3\,a\;2\,b$ (mit 10 Permutationen): 640, für $2\,a\;3\,b$ (mit 10 Permutationen): 160. Die wahrscheinlichste Verteilung, d. h. die, auf welche die zahlreichsten möglichen zusammengesetzten Ereignisse entfallen, ist also $4\,a\;1\,b$, wiederum entsprechend den Chancen der einfachen Ereignisse.

Bei 6 Fällen erhalten wir als Zähler der Wahrscheinlichkeitsbrüche (mit dem gemeinsamen Nenner 5^6) für die Verteilung $6\,a$: 4096, für $6\,b$: 1, für $5\,a\;1\,b$: 6144, für $4\,a\;2\,b$: 3840, für $3\,a\;3\,b$: 1280, für $2\,a\;4\,b$: 240, für $1\,a$

$5\,b$: 24. Hier kann keine Verteilung mit dem Verhältnis der Chancen der einfachen Ereignisse $4:1$ genau zusammenfallen. Aber diejenige hat die grösste Wahrscheinlichkeit, welche diesem Verhältnis am nächsten steht: $5\,a\;1\,b$.

Ueberhaupt, setzen wir die Zahl der Fälle $x = r\,(m + n)$, wo m und n die Chancen von a und b (die Zähler ihrer Wahrscheinlichkeiten) bedeuten, so ist am wahrscheinlichsten die Verteilung, welche a in der Anzahl $r\,m$, b in der Anzahl $r\,n$ enthält, bezw. (wenn r keine ganze Zahl) welche die grösste Annäherung an $r\,m$ und $r\,n$ enthält. So ist im letzten Beispiel $r\,m = 4\frac{4}{5}$, $r\,n = 1\frac{1}{5}$, welchen Werten die Verteilung $5\,a\;1\,b$ am nächsten steht.

Das nun, worauf es Bernoulli am meisten ankam und was den eigentlichen Kern seiner Entdeckung bildet, ist der Gang, den die Zahlenwerte für die verschiedenen Verteilungen nehmen, wenn x immer grösser genommen wird. Die absolute Wahrscheinlichkeit jeder einzelnen Verteilung wird natürlich immer geringer, aber die Abnahme erfolgt am langsamsten bei derjenigen, die dem Verhältnis der Chancen m und n (möglichst) entspricht, immer schneller dagegen, je weiter sich eine Verteilungsart davon entfernt. Schreiben wir die Verteilungsarten auf die Abscissenaxe und tragen die entsprechenden Wahrscheinlichkeiten als Ordinaten auf, so kann man bei sehr zahlreichen x durch Verbindung der Ordinaten eine Curve ziehen, welche mit wachsenden x zwar ihrer absoluten Höhe nach abnimmt, aber in der Maximumgegend immer steiler und nach aussen immer flacher wird. Wenn man daher die Wahrscheinlichkeiten von der grössten an nach beiden Seiten bis zu einer bestimmten Grenze summirt, also die Wahrscheinlichkeit berechnet, dass irgend eine der darin eingeschlossenen Verteilungsarten stattfinde, so kann man durch Vermehrung der Fälle x das Uebergewicht dieser Wahrscheinlichkeitssumme über die Summe aller übrigen Wahrscheinlichkeiten immer mehr steigern und das Verhältnis beider beliebig der Einheit (Sicherheit) annähern. Geometrisch: das durch die Ordinaten zweier bestimmter Abscissenpuncte rechts und links vom Maximum abgegrenzte Flächenstück beträgt einen immer grösseren Teil der ganzen Curvenfläche. Man kann für einen gewünschten Grad der Annäherung an die Sicherheit die erforderliche Zahl x von Fällen berechnen und umgekehrt. So fand Bernoulli für $m = \frac{3}{5}$, um eine Wahrscheinlichkeit von mindestens $\frac{1000}{1001}$ zu erhalten, dass die Anzahl der a zwischen den Grenzen $\frac{31}{50}\,x$ und $\frac{29}{50}\,x$ liege, $x = 25550$.

2. Die in Bernoulli's Theorem ausgedrückte Wahrscheinlichkeit fällt unter die im I. Abschnitt gegebene Definition.

Sie ist nach dem Vorangehenden das Zahlenverhältnis bestimmter Verteilungsarten (immer einschliesslich der Permutationen) zu den sämmtlichen Verteilungsarten in einem zusammengesetzten Ereignisse. Man könnte den Ausdruck Wahrscheinlichkeit, der hier wegen der populären Nebenbedeutungen manches Misverständnis verursacht, auch ganz bei Seite lassen und das Theorem als einen Satz der Combinations- und Permutationslehre ausdrücken. Ja man könnte in jedem Fall ohne Hilfe der Algebra und der höheren Analysis durch mechanische Abzählung der möglichen Combinationen und Permutationen jenes Verhältnis feststellen, wenn man sich die Mühe nehmen wollte. Freilich würde uns bei grösseren x bald Zeit und Kraft vergehen. Zur Ersparung dieser Mühe dienen zunächst die gewöhnlichen Formeln der Combinations- und Permutationsrechnung, bei grossen x aber kunstvoll erdachte Näherungsformeln (zumal die Stirling'sche: $1 \cdot 2 \cdot 3 \cdots n = e^{-n}\, n^n \sqrt{2\pi n}$). Dass man auf diesem Wege das gewünschte Zahlenverhältnis nicht absolut genau erhält, bedeutet eine Unvollkommenheit unserer vollkommensten Rechnungsmethoden; an sich ist es aber ein durchaus bestimmtes und würde von einem noch besseren Rechenmeister, als selbst Stirling, Euler und Laplace es waren, ebenso genau angegeben werden, wie das der Verteilung $2\,a\;1\,b$ bei den Chancen $\frac{1}{2}$.

Und nicht blos ist die Bernoulli'sche Wahrscheinlichkeit (wenn der Ausdruck gestattet ist) eine Wahrscheinlichkeit in demselben Sinne wie die der einfachen Ereignisse, sondern ist in und mit dieser implicite gegeben. Behaupten, dass das einzelne Ereignis a die Wahrscheinlichkeit $\frac{4}{5}$ und b $\frac{1}{5}$ habe, heisst zugleich behaupten, dass die Verteilung $5\,a\;1\,b$ die Wahrscheinlichkeit $\frac{6144}{15625}$ habe u. s. f. Und beides heisst nichts weiter, als dass das bezügliche einfache oder zusammengesetzte Ereignis der durch den Zähler angegebenen Anzahl von Fällen unter der durch den Nenner angegebenen Anzahl von Fällen entspricht, über welche wir in gleicher Weise disjunctiv absolut unwissend sind. Poinsot hat auch hier Recht: man liest aus den Formeln der Mathematik nichts heraus, was man nicht hineingelegt hat.

Was aber hier besonders wieder betont werden muss, ist die Irrelevanz aller zeitlichen Bestimmungen. Selbst Lehrbücher, die den rein mathematischen Charakter des Theorems ausdrücklich betonen, pflegen es in der Form einer Prophezeiung auszusprechen. Cournot z. B. sagt, gerade um jenen Charakter zu erläutern: es bedeute nichts anderes als dass unter allen Combinationen oder Hypothesen, die man machen kann, die den Chancen von a und b entsprechenden häufiger stattfinden werden als die übrigen. Es leuchtet aber ein, dass auch hier der Sinn und die Berechtigung unseres Wahrscheinlichkeitsurteils ungeändert bleibt, wenn nach der wahrscheinlichsten Zusammensetzung von x vergangenen Ereignissen gefragt wird, von denen wir nur wissen, dass sie von der Art a und b sein können und dass diese einzeln die Wahrscheinlichkeit m und n besitzen; oder wenn gar nicht von Ereignissen sondern von unveränderlich stattfindenden Thatsachen oder allgemeinen Gesetzen die Rede ist. Es können auch die Ereignisse, wenn es sich um solche handelt, teilweise vergangen, teilweise gegenwärtig, teilweise zukünftig sein; sie können, statt aufeinanderzufolgen, gleichzeitig sein oder teils das eine teils das andere: die Zeit hat im Allgemeinen mit der Wahrscheinlichkeit einer Materie so wenig zu schaffen wie mit der Wahrheit von $2 \times 2 = 4$.

Auch brauchen die Glieder der Reihe, wenn sie aufeinanderfolgen, keineswegs ununterbrochen aufeinanderzufolgen, ebensowenig wie sie räumlich in einer bestimmten Beziehung stehen müssen. Wenn wir 191 Würfe aufschreiben, die ich eben jetzt durch einen Würfel erhielt, dazu 57, die in 14 Tagen von einem Freund in New-York gemacht werden, ferner 326, die vor 10 Jahren irgendwo gemacht und aufgeschrieben wurden, und so ganz willkürlich weiter, bis die Summe der im wörtlichsten Sinn zusammengewürfelten Würfe eine Million ausmacht, so findet das Bernoulli'sche Gesetz darauf ebenso Anwendung, wie wenn ich jetzt und hier ununterbrochen fortwürfle. Die einzige Bedingung muss gewahrt sein, dass wir in Bezug auf keinen Fall irgend etwas weiter wissen, was die Wahrscheinlichkeit des einfachen Ereignisses verändern könnte.

Hieraus erhellt nun recht deutlich, wie haltlos die versteckte Supposition irgend eines causalen Einflusses der einzelnen Ereignisse aufeinander ist, dem zufolge die späteren gezwungen würden, nach und nach Ordnung in die Reihe zu bringen, nachdem sich die früheren allzusehr haben gehen lassen. Weder mit Causalverhältnissen noch mit physischen Thatsachen hat das Bernoulli'sche Theorem an sich das Geringste zu thun. Der Begriff constanter objectiver Bedingungen, der Begriff realer Bedingungen überhaupt ist darin nicht mehr enthalten als in dem nächsten besten arithmetischen Satze.

Die zu Anfang erhobenen Fragen sind hiemit noch nicht erledigt; aber es ist unerlässlich zu ihrer Erledigung, dass man diesen rein mathematischen Charakter des Satzes durchschaue und festhalte.

3. Die in Bernoulli's Theorem ausgedrückte Wahrscheinlichkeit setzt eine physische Deutung der „gleichmöglichen" Fälle nicht voraus.

Dieser Punct, an sich nur eine weitere Durchführung des vorangehenden, bedarf doch wegen erschwerender Umstände einer besonderen Betrachtung. Wenn wir wissen, dass in einer Urne w und s Kugeln in gleicher Anzahl vorhanden sind (= Fall A), so entspricht es nicht blos der Rechnung sondern auch dem gesunden Menschenverstand, dass bei sehr vielen Ziehungen (mit Wiederhineinlegen und Schütteln) ungefähr dasselbe Verhältnis sich herausstellt. Wenn wir aber nur wissen, dass überhaupt w und s Kugeln darin sind (= Fall B), so entspricht es dem gesunden Menschenverstand keineswegs, bei sehr vielen Ziehungen mit fast völliger Sicherheit eine Verteilung 1 : 1 zu erwarten. Ebenso wird es uns nichts weniger als vernünftig erscheinen, beim Werfen mit einem unbekannten 6 seitigen Körper von vielleicht ganz unregelmässiger Gestalt mit hoher Wahrscheinlichkeit zu erwarten, dass jede Seite ungefähr gleich oft darankomme. Niemand würde eine solche Folgerung anders als lächerlich finden, Keiner darauf einen Einsatz wagen. Und doch haben wir die Wahrscheinlichkeit der einfachen Ereignisse auch in solchen Fällen als correct mit $\frac{1}{2}$ bezw. $\frac{1}{6}$ bestimmt und als Ausdruck unseres natürlichen Bewusstseins angesehen. Und wiederum folgt allem Anschein nach aus dieser Bestimmung mit mathematischer Notwendigkeit alles Uebrige.

Müssen wir also doch die Grundbestimmungen revidiren und die Bedingung stellen, dass die „gleichmöglichen Fälle" nicht blos ein gleiches Nichtwissen bedeuten sondern eine physische Gleichheit, von der wir Kenntnis haben, im

Sinne des Falles A? — Dem steht Alles entgegen was unter I und II verhandelt wurde.

Oder sollen wir einen doppelten Wahrscheinlichkeitsbegriff statuiren, einen, der auch den Fall B umfasst aber nur für die Wahrscheinlichkeit einfacher Ereignisse gilt, und einen anderen, der nur A umfasst aber auch für die Verteilung gilt? — Unmöglich, da die Wahrscheinlichkeiten der Verteilung aus denselben Principien folgen wie die der einfachen Ereignisse (Summirung und Multiplicirung der Wahrscheinlichkeiten, Combination und Permutation) und überhaupt zwischen einfachen und zusammengesetzten Ereignissen nur ein relativer Unterschied ist (denken wir z. B. an das Werfen eines Pasches).

Oder soll man sagen, das Bernoulli'sche Gesetz lehre nur das relative Uebergewicht bestimmter Verteilungen über andere, und relativ sei ja immerhin auch im Falle B die Verteilung 1 : 1 für w und s die wahrscheinlichste, wenn auch jede einzelne Verteilung absolut genommen weniger wahrscheinlich sei als im Falle A?[1]) — Unmöglich, da dann die Summe aller Wahrscheinlichkeiten bei B nicht mehr $= 1$ wäre. Gerade auch die absoluten Wahrscheinlichkeiten ergeben sich nach der Rechnung für den Fall A und B als die nämlichen.

Ich muss gestehen, dass mich diese Schwierigkeit einige Zeit peinigte, da jede Thüre verschlossen schien. Jedenfalls dürfte hier das Hauptmotiv für die Anhänger der physischen Deutungen des Wahrscheinlichkeitsbegriffes liegen. Gleichwol ist die Lösung einfach genug.

Denken wir uns statt Einer Urne im Falle B soviele Urnen als Ziehungen stattfinden sollen, z. B. 20000, und jedesmal werde aus einer anderen gezogen. Wir mögen

---

1) Dies würde etwa einem Gedanken entsprechen, den wir oben S. 64 von Nitsche (so, und nicht Nitzsche, zu lesen) hörten.

wissen, dass in jeder Urne *w* oder *s* oder eine Mischung aus beiden Sorten sich befinde, aber nicht den geringsten Anhaltspunct haben, irgend eine Absicht oder sonstige constante Ursache zu vermuten, welche eine gewisse Gleichmässigkeit in der Füllung der Urnen hätte bedingen können. Hier wird es auch dem gesunden Menschenverstand nicht zuwiderlaufen, als Endergebnis der Ziehungen mit grosser Wahrscheinlichkeit eine annähernd gleiche Verteilung von *w* und *s* zu erwarten. Ebenso, wenn unter sonst gleichen Voraussetzungen unendlich viele Urnen vorlägen und jedesmal blindlings ohne bestimmte Regel der Abfolge in irgend eine Urne gegriffen würde.

So aber oder in aequivalenter Weise müssen wir die Frage stellen, wenn die wahrscheinlichste Verteilung bei Unkenntnis des Mischungsverhältnisses berechnet werden soll. Bei jeder Wahrscheinlichkeitsbestimmung müssen diejenigen Umstände, über welche wir uns disjunctiv in völliger Unwissenheit befinden, als unbeschränkt variabel (bezw. bei gleichzeitiger Vielheit der Fälle als unbeschränkt verschieden) vorausgesetzt werden dürfen. Dies liegt im Begriffe der Wahrscheinlichkeit, wie er in I definirt wurde: denn bei der geringsten Vermutung über Constanz eines solchen Umstandes hätten wir nicht mehr gleichmögliche Fälle im Sinne gleicher d. h. absoluter Unwissenheit.

Ist uns das Mischungsverhältnis der *s* und *w* gegeben, erstreckt sich also unser Nichtwissen nur auf die Richtung des Greifens, die Lage der Kugeln in der Urne u. s. w., so werden diese Umstände als unbeschränkt variabel und nur das Mischungsverhältnis als constant gedacht. Erstreckt sich aber unsere Kenntnis nur darauf, dass überhaupt *s* und *w* in der Urne sind, während wir über das Mischungsverhältnis ebensowenig wissen, wie über die sonstigen Umstände, so muss eben auch dieses variabel gedacht werden. Wenn ich hier frage: wie wahrscheinlich ist es, dass ich 20 mal *s* ziehe, so ist die Wahrscheinlichkeit nur dann $= (\frac{1}{2})^{20}$, wenn ich aus 20 Urnen ziehe, bei deren jeder dieselben Bedingungen meines Wissens und Nichtwissens zutreffen, unter denen also das Mischungsverhältnis ebenso verschieden sein kann, wie die Richtung meines Armes und die Lage der Kugeln in dem Fall, wo ich das Mischungsverhältnis kenne und wo es constant bleibt. Natürlich kann auch die nämliche Urne benützt werden, wenn mir nur auf irgend eine Weise jeder Verdacht genommen wird, dass das Mischungsverhältnis in den einzelnen Fällen das nämliche sei. Ebenso wie im Falle *A* auch 20000 verschiedene Urnen mit gleichem Mischungsverhältnis benützt werden können.

Der Constanz des Mischungsverhältnisses im Falle *B* würde im Falle *A* die besondere Bedingung entsprechen, dass immerfort die *w* und *s* Kugeln sich in genau derselben Lage in der Urne befänden und die Hand immer genau an dieselbe Stelle in der Urne griffe. Wir hätten dann ebenfalls für das einfache Ereignis die Wahrscheinlichkeit $\frac{1}{2}$, und dennoch würde jetzt auch hier kein Vernünftiger eine hohe Wette eingehen, dass in 20 oder gar 20000 Fällen annähernd ebensoviele *s* als *w* gezogen werden.

Es ist ebenso beim Werfen eines Körpers. Es entspricht vollkommen dem gesunden Menschenverstand, dass unter 24000 6seitigen Körpern, bei deren Bildung keinerlei Absicht oder sonstige gemeinsame Ursache zu vermuten ist, und deren einzelne Flächen principlos, blindlings mit *a* bis *f* bezeichnet wurden, jeder Buchstabe annähernd 4000 mal am Boden aufliege. Die Wahrscheinlichkeitsaussage würde denn auch nichts weiter sein, als das Ergebnis der (indirecten) Abzählung derjenigen unter allen Combinationen und Permutationen, welche unter den Begriff „annähernd 4000 auf 24000“ fallen, sobald er mathematisch genauer definirt wird (z. B. „zwischen 3950 und 4050 auf 24000“). Unter diesen Begriff, in diese Grenzen, fallen ungeheuer viel mehr mögliche Combinationen mit Permutationen als unter den entgegengesetzten Begriff, ausserhalb dieser Grenzen. Und das ist, was wir als Wahrscheinlichkeit von entsprechend hohem Grade bezeichnen.

IV. Die empirisch bestimmte Wahrscheinlichkeit.

Wie aus gegebenen Chancen zweier Ereignisse (allgemeiner: mehrerer Disjunctionsglieder) auf ihre Verteilung in grosser Anzahl, so kann auch aus der gegebenen Verteilung auf die Chancen geschlossen werden; und eine solche Verteilung kann nicht blos als Prämisse eines Rechenexempels sondern auch als Thatsache der Beobachtung gegeben sein. Man nennt in solchem Fall die daraus bestimmte Wahrscheinlichkeit eine aposteriorische oder empirische gegenüber der apriorischen, die nicht aus einer beobachteten Verteilung sondern aus der unmittelbaren Disjunction der gleichmöglichen Fälle hergeleitet wird. Wir nehmen die Ausdrücke vorläufig, ohne zu fragen, ob sie hier in gleichem Sinn wie sonst in der Erkenntnislehre gebraucht werden. Die Bernoulli'sche Wahrscheinlichkeit fällt, wie wir sahen, in jeder Beziehung unter den Begriff der apriorischen Wahrscheinlichkeit. Die Frage ist, ob sich auch die empirische irgendwie unter einen gemeinsamen Begriff bringen lässt oder ob es sich hier um eine Wahrscheinlichkeit in fundamental anderem Sinn handelt; ferner worin im ersteren Fall die unterscheidenden secundären Merkmale bestehen; endlich inwiefern die eine Wahrscheinlichkeit durch die andere corrigirt oder verändert werden kann. Denn es ist bekannt, dass die Wirklichkeit von der auf Grund der apriorischen Wahrscheinlichkeit erwarteten Verteilung oft erheblich abweicht.

Die Uebereinstimmung mit der Wirklichkeit kann hier allerdings von vornherein nicht den Sinn haben wie bei der Addition von Aepfeln, wo nur Ein individuelles Resultat erwartet werden kann. Es konnte ja aus den Chancen mit hoher Wahrscheinlichkeit nicht auf eine individuelle sondern nur auf eine innerhalb gewisser Grenzen liegende Verteilungsart geschlossen werden. In diesem Sinne haben Buffon, de Morgan, Quetelet, R. Wolf, Jevons beim Münzwerfen und ähnlichen Spielen die Probe gemacht und eine gute Uebereinstimmung gefunden.[1]) Das heisst: den Grenzen, innerhalb deren das Resultat von dem Verhältnis der Chancen abwich, entsprach rechnungsgemäss eine hohe Wahrscheinlichkeitssumme. Aber bei anderen Materien, wie in dem classischen Beispiel der Knaben- und Mädchengeburten, entfernt sich die wirkliche Verteilung doch erheblich weiter von dem Chancenverhältnis 1 : 1, welches man vor der statistischen Erfahrung ansetzen musste.

Natürlich schliessen wir nun hier: die Chancen sind eben andere als wir vorausgesetzt hatten, die Wirklichkeit belehrt uns eines Besseren. Aber was heisst dies, wenn „Chancen" überhaupt nichts weiter bedeutet, als einen gewissen Stand unserer Kenntnis und Unkenntnis? Diesen hatten wir nicht vorausgesetzt, sondern er war wirklich kein anderer. „Die falsche Annahme, die wir gemacht hatten, — so wird nun der Anhänger der objectiven Theorie argumentiren — betraf nicht unseren Wissensstand sondern die Ursachen der Ereignisse. Von ihnen hatten wir vorausgesetzt, dass sie

1) Jevons, der immer 10 Münzen auf einmal warf, trieb es bis zu 20480 Fällen und erhielt darunter 10353 mal Wappen, und zwar in der ersten Hälfte der Fälle 5222, in der zweiten 5131 (Principles of Science 1887 p. 208).

Noch mehr Fälle erhielt Rud. Wolf, der mit zwei (nicht etwa zu dem Zwecke besonders sorgfältig construirten, aber auch nicht ganz schlechten) Würfeln 100000 Würfe machte. S. über die Ergebnisse: „Mitteilungen der naturforschenden Gesellschaft in Bern" 1849—53, und Wolf's Handbuch der Mathematik u. s. f. 1870 I S. 65. Von den übrigen erwähnten Versuchen berichtet Jevons a. a. O.

nach beiden Seiten hin gleich lägen. Der Begriff der Chance und somit auch der gleichmöglicher Fälle bedeutet also doch einen objectiven Zustand, ein Verhältnis realer Bedingungen. Die aposteriorische Wahrscheinlichkeit lässt sich nicht mit der alten Definition vereinigen."

Hier liegt also eine neue Schwierigkeit. Zu deren Lösung müssen wir etwas weiter ausholen.

1. Wird aus einer Urne, von der wir wissen, dass sie 1 *w* 6 *s* Kugeln enthält, eine *w* gezogen, so finden wir keinen merklichen Anlass zur Verwunderung. Wissen wir, dass 1 *w* 100 *s* darin, so stiehlt sich vielleicht schon ein leiser Ausruf von den Lippen; und kommen gar auf 1 *w* eine Million *s*, so wird wenigstens der gewöhnliche Mann, wenn beim ersten und einzigen Zug *w* erscheint, kopfschüttelnd behaupten, dass es nicht mit rechten Dingen zugehe.

Kritiker betonen nun vielfach und energisch, dass jedes, auch das unwahrscheinlichste, Ereignis möglich und dass jeder andere von den 1000001 möglichen Fällen, über den man sich nicht im Geringsten verwundert hätte, genau ebenso unwahrscheinlich war wie dieser.

Und gewiss liegt eine logische Absurdität weder in der Thatsache selbst, wenn sie erfolgte (denn etwas logisch Unmögliches kann auch nicht wirklich sein) noch in der Annahme, dass hier der reine Zufall waltete, d. h. dass die *w* Kugel durch keinerlei besondere Umstände oder Bedingungen, wie z. B. durch Taschenspielerkünste, begünstigt war.[1]) Gleichwol hat der gemeine Mann Recht, wenn er

---

1) Unter einer begünstigenden Bedingung verstehen wir hier eine solche, die in den meisten Fällen diese eine Art von Ereignissen zur Folge hat. Auch eine gute Taschenspielerei kann ja einmal mislingen.

Bei zusammengesetzten Ereignissen bedeutet dementsprechend, um dies im Voraus zu bemerken, eine begünstigende Bedingung für eine bestimmte Verteilung solche Umstände, aus denen sich eine der Zufallshypothese immer ungläubiger gegenübersteht, je unwahrscheinlicher im Voraus der Fall war. Sein Verhalten ist logisch dadurch gerechtfertigt, dass das Ziehen einer *w* und das einer *s* Kugel, wenn die Kugeln nur eben durch ihre Farben definirt werden, als je Eine Classe von Ereignissen für unser Bewusstsein definirt ist. Wir fragen nicht danach, welches schwarze Individuum herauskommt. Diese Classe von Ereignissen also, das Ziehen einer *s*, einerlei welcher, hat nicht die Wahrscheinlichkeit $\frac{1}{1000001}$ sondern $\frac{1000000}{1000001}$. Wären die *s* numerirt und die Frage von vornherein so gestellt: „Wird die Eine *w* oder wird die *s* Nr. 31457 gezogen?" dann freilich ständen sich die Wahrscheinlichkeiten gleich.

„Aber irgend etwas muss doch immer eintreten; und dies Eine ist eben factisch eingetreten. Was soll es angesichts der Wirklichkeit heissen, das Geschehene sei weniger wahrscheinlich?" — So hört man beständig wieder fragen. Als wenn es sich um die Wahrscheinlichkeit des geschehenen

---

ähnliche Verteilung wie die beobachtete mit grosser Wahrscheinlichkeit würde erschliessen lassen. Besteht das zusammengesetzte Ereignis in vielmaliger Wiederholung oder vielfachem gleichzeitigen Auftreten eines von zwei möglichen Ereignissen, so fällt der Begriff einer begünstigenden Bedingung zusammen mit dem einer constanten bezw. gemeinschaftlichen Bedingung der einzelnen Fälle.

Der Ausdruck „gemeinschaftliche Bedingung" wird hier auch nicht misverstanden werden. Wenn ich mit der Hand einen Haufen von Münzen in die Luft werfe, so ist die Handbewegung eine gemeinschaftliche Bedingung für die Aufwärtsbewegung der Münzen. Aber für die Seiten, mit denen die verschiedenen Münzen zuletzt auf dem Boden liegen, ist in gewöhnlichen Fällen keine gemeinschaftliche Bedingung vorhanden. Hiefür sind die ursprünglichen Lagen in der Hand, das verschiedene Gewicht, die verschiedene Grösse und Form u. s. w. massgebend, und diese Bedingungen hängen alle unter sich nicht zusammen.

Ereignisses handelte, von welchem wir Kenntnis haben, und nicht vielmehr um die Wahrscheinlichkeit einer bestimmten Art seines Zustandekommens, von dem wir keine Kenntnis haben. Was wir hier in so hohem Grade unwahrscheinlich nennen, ist die specielle Annahme seines zufälligen Zustandekommens.[1])

2. Das hier zu Grunde liegende Princip ist das der Bildung und Wahrscheinlichkeitsbewertung von Hypothesen, d. h. von wahrscheinlichen Annahmen, aus welchen eine gegebene Thatsache oder ein Gesetz oder überhaupt irgend eine gegebene Materie sich folgern oder erklären lässt. Nennen wir die Wahrscheinlichkeit des in einer Hypothese ausgesprochenen Sachverhalts in sich selbst,

1) Dies zu betonen, schien mir nicht überflüssig, nachdem selbst Fick, „um einem weitverbreiteten Misverständnis des Wesens der Wahrscheinlichkeit wirksam zu begegnen, wonach gewisse individuelle Fälle besonders unwahrscheinlich oder gar unmöglich sind", nachdrücklich hervorhebt (a. a. O. 25): wenn beim Roulette die Kugel 6 mal nacheinander auf Rot falle, so sei die Verwunderung nicht im mindesten mehr gerechtfertigt, als wenn etwa *s r r s r s* herauskomme, da jede von beiden Combinationen die Wahrscheinlichkeit $\frac{1}{64}$ habe. Es sei durchaus nicht erstaunlicher, 100 mal nacheinander Rot kommen zu sehen als irgend eine individuelle bestimmte Reihenfolge von *r* und *s*, die man vielleicht soeben wirklich beobachtet habe.

Wir haben hier bereits ein zusammengesetztes Ereignis, doch fallen diese, wie im Text alsbald auszuführen ist, unter denselben Gesichtspunct. Das 100 malige *r* ist in sich selbst nicht erstaunlicher als jede andere Reihenfolge. Aber wir treten an ein solches Spiel mit der Voraussetzung heran, dass es ein Zufallsspiel sei, d. h. hier: dass keine gemeinschaftliche oder constante Bedingung der verschiedenen Fälle vorhanden sei. Das Eintreffen von 100 *r* unter dieser Voraussetzung ist in der That im höchsten Masse erstaunlich, ja unglaublich, — unmöglich allerdings nicht.

Der Abbate Galiani hatte also nicht so Unrecht, wenn er ausrief: „Beim Blut des Bacchus! Die Würfel sind gefälscht!" (wie sie es denn auch wirklich waren), als ein Neapolitaner mit drei Würfeln 5 mal nacheinander 3×6 warf.

abgesehen von dem Gegebenen, das daraus erklärt werden soll, die vorgängige Wahrscheinlichkeit (m) der Hypothese, sodann die Wahrscheinlichkeit, mit welcher das Gegebene daraus folgt, ihren Erklärungswert (n), ferner ihre Wahrscheinlichkeit, abgesehen von den übrigen denkbaren Hypothesen, abstracte Wahrscheinlichkeit (p), endlich ihre Wahrscheinlichkeit im Verhältnis zu allen übrigen denkbaren Hypothesen concrete Wahrscheinlichkeit (P)[1]) — so können wir dieses Bayes-Laplace'sche Princip in seiner allgemeinsten Form so aussprechen:

Die abstracte Wahrscheinlichkeit p einer Hypothese ist das Product aus ihrer vorgängigen Wahrscheinlichkeit m in ihren Erklärungswert n. Ist so die abstracte Wahrscheinlichkeit aller denkbaren Hypothesen bestimmt, $p_1 = m_1 n_1$, $p_2 = m_2 n_2$, ... so wird die concrete Wahrscheinlichkeit einer einzelnen gegeben durch $P_n = \frac{p_n}{p_1 + p_2 + \cdots p_n \cdots + p_r}$.

Ist die vorgängige Wahrscheinlichkeit aller Hypothesen die gleiche, so kann sie als gleicher Factor in allen Gliedern des Zählers und Nenners von vornherein ausser Betracht bleiben.

Die Darstellungen der Wahrscheinlichkeitsrechnung pflegen bei der Formulirung dieses Princips statt von Hypothesen von Ursachen zu sprechen, aber hinzuzufügen, dass das Wort Ursache in der Wahrscheinlichkeitsrechnung etwas Anderes als gewöhnlich bedeute.[2]) Eine umständliche und

1) Die Ausdrücke „abstracte und concrete Wahrscheinlichkeit“ in obigem Sinn sind ungewöhnlich (Poisson verwendet sie in ganz anderer Bedeutung). Aber es fehlt hier überhaupt an einer allgemein gebräuchlichen guten Bezeichnung.

2) Poisson, Lehrbuch der Wahrscheinlichkeit deutsch S. 50 (§ 27). A. Meyer, Vorlesungen über Wahrscheinlichkeit 1879 S. 8 (vgl. o. S. 42). J. Bertrand, Calcul des probabilités 1889 p. 142—3. („Les causes sont pour nous des accidents qui ont accompagné ou précédé un

ganz überflüssige Verkehrung des Sprachgebrauchs. Denn wenn man zusieht, was unter Ursache hier verstanden wird, so ist es eben eine Voraussetzung, eine Hypothese. Sie kann sich auf Ursachen im gewöhnlichen Sinn beziehen, aber auch auf Anderes. In den typischen Beispielen, wie bei dem unbekannten Mischungsverhältnis der Kugeln in der Urne, dessen Wahrscheinlichkeit aus einer beobachteten Reihe von Zügen (einem zusammengesetzten Ereignis) erschlossen werden soll, handelt es sich wirklich nicht um Ursachen oder auch nur um reale Bedingungen im engsten Sinn, da für das Ziehen jeder wirklich gezogenen Kugel immer nur diese Kugel und nicht die übrigen reale Bedingung ist. (In einem weiteren Sinn mag man immerhin die Gesammtheit der Kugeln als reale Bedingung für die Gesammtheit der Ziehungen bezeichnen.) Es handelt sich hier nur um einen Thatbestand, dem gemäss wir, wenn er uns gegeben wäre, für das fragliche Ereignis selbst die Wahrscheinlichkeit n bestimmen würden. In solchen Fällen spricht man wol auch von der Wahrscheinlichkeit einer Wahrscheinlichkeit, wobei unter der letzteren Wahrscheinlichkeit eben dieser hypothetische Thatbestand (das Mischungsverhältnis) zu verstehen ist, aus dem die Wahrscheinlichkeit n folgt. In anderen Fällen, wo obiges Princip Anwendung findet, handelt es sich aber in der That um die realen Bedingungen eines beobachteten Ereignisses; wie wenn wir fragen, ob ein Brand durch Blitzschlag oder durch Brandlegung u. s. f. entstanden sei.[1]) Wieder in anderen Fällen aber betrifft unser Schluss

événement observé. Le mot n'implique pas qu'au sens philosophique l'événement soit un effet produit par la cause.")

1) In solchen Fällen würden m und n bedeuten: Wahrscheinlichkeit für die Existenz der Ursache und für die Hervorbringung der Wirkung durch dieselbe, und würde die Trennung insofern auf einer metaphysisch ungenauen Ansicht beruhen, als eine Ursache nur dann Ursache ist wenn sie wirkt. Aber man kann sich methodo-

umgekehrt die Wirkungen eines beobachteten Ereignisses: wenn nämlich bei lückenhafter Kenntnis der Umstände oder der Wirkungsgesetze die Erwartung zwischen einer Mehrheit von Wirkungen geteilt ist. Ferner kann die Hypothese sich auf die möglichen Arten eines Ereignisses beziehen, wenn wir nur wissen, dass es innerhalb gewisser Grenzen liegt; z. B. wenn mit 3 Würfeln eine Zahl über 16 geworfen ist, sind die Hypothesen möglich, dass die Zahl 17 und dass sie 18 sei.[2]) Endlich kann es sich auch statt um zeitlich ver-

---

logisch häufig mit Nutzen die vollständige Summe der realen Bedingungen, die im eigentlichen Sinne Ursache ist, in zwei Teile zerlegt denken, oder auch gelegentlich in ein allgemeineres und ein specielleres Merkmal (logische Teile), wie z. B.: „Es hat geblitzt, vielleicht hat der Blitz in das Haus geschlagen.“ Oft können wir die Bedingungen zweckmässig in noch viel mehr Teile zerlegen und die Wahrscheinlichkeit p ist dann eben das Product aus den sämmtlichen Teilwahrscheinlichkeiten. Oft ist eine Zerlegung aber auch unnötig oder vergeblich, indem wir die Bedingungen sogleich so vollständig aussprechen, dass die Wirkung sicher daraus hervorgehen musste; dann fällt eben die Wahrscheinlichkeit p der Ursache mit der ihrer Existenz zusammen. Aber auch dann kommt es, wie das Princip weiter sagt, noch auf andere mögliche Ursachen an, aus denen dieselbe Wirkung hervorgehen musste, wenn sie da waren. Denn es kann ja dieselbe Wirkung naturgesetzlich durch verschiedene Combinationen von Bedingungen entstehen.

2) Mit diesem und einem analogen Beispiel erläutert Bertrand a. a. O. (p. 143, 144) den Begriff der „Ursachen“; 17 und 18 seien die möglichen Ursachen des Ereignisses. In diesem Fall hat nun freilich der Begriff mit dem der Causalität absolut nichts mehr zu thun; es handelt sich, wie Bertrand richtig bemerkt, nur um die „manières diverses dont l'événement a pu se présenter“. Es überschreitet aber doch die Grenze des erlaubten Misbrauchs, wenn ich so sagen soll, in solchen Fällen auch noch von einem Hervorbringen des Ereignisses durch die Ursache (qu'elle produise l'événement) zu reden. A. Meyer wird durch diesen Uebelstand zu einem ganz directen Widerspruch mit sich selbst geführt. S. 165 gibt er als ausdrückliche Definition der „Ursache“ in der Wahrscheinlichkeitsrechnung „die constanten Umstände, welche an der Hervorbringung

laufende Ereignisse um bestehende Collocationen oder Eigenschaften oder selbst um allgemeine Wahrheiten handeln (wie denn jedes Naturgesetz eine solche Hypothese ist).

In allen diesen Fällen ist obiges Princip anwendbar, soweit überall eine vollständige Disjunction der möglichen Hypothesen aufgefunden und die Werte m und n numerisch bestimmt werden können. Ausserdem eben in weniger exacter Weise als Leitfaden von Schätzungen.

Wo und inwieweit reale Bedingungen oder Folgen den Gegenstand unseres Wahrscheinlichkeitsschlusses bilden, liegen diesem nicht blos das allgemeine Causalgesetz sondern auch alle bereits erworbenen Kenntnisse über bestimmte Causalzusammenhänge zu Grunde. In anderen Fällen seiner Anwendung aber und in der allgemeinen Fassung des Princips ist von Causalbeziehungen nicht die Rede. Es handelt sich zunächst und allgemein auch hier nur um logische Zusammenhänge.

Wir sehen so auch an diesem wichtigen Princip alles bestätigt, was über den Begriff der Wahrscheinlichkeit im I. Abschnitt gesagt wurde. Es ist ja auch nur eine notwendige und durch sich verständliche Folge des allgemeinen Wahrscheinlichkeitsbegriffes, keiner neuen und speciellen Voraussetzung bedürftig, und die Producte $m\,n$ sind nach der Regel der zusammengesetzten Wahrscheinlichkeit gebildet, die wiederum nur eine unmittelbare Anwendung des Wahrscheinlichkeitsbegriffes ist.

In Fällen nun wie dem unter 1. besprochenen betrifft unser Schluss das Vorhandensein begünstigender Bedingungen für das beobachtete einzelne Ereignis, und es sind, ganz allgemein gesprochen, zwei Hypothesen möglich: dass irgend

---

eines Ereignisses teilnehmen", nachdem er S. 9 ebenso ausdrücklich erklärt hat, Ursache bedeute in der Wahrscheinlichkeitsrechnung nicht das was ein Ereignis herbeiführt.

eine begünstigende Bedingung und dass keine solche vorhanden war. An sich ist das Letzte vielleicht nach den uns bekannten Umständen wahrscheinlicher. Aber die Wahrscheinlichkeit des Ereignisses unter dieser Voraussetzung ist so klein, dass auch $p_1 = m_1 n_1$ ein äusserst kleiner Bruch wird, viel kleiner als p_2. Dadurch wird auch die concrete Wahrscheinlichkeit der ersten Hypothese viel kleiner als die der zweiten, und diese kann unter Umständen bis auf einen verschwindend kleinen Unterschied der Sicherheit nahekommen.

Freilich wird die Berechnung nur ausnahmsweise einmal in allen Puncten genau in Zahlen durchzuführen sein. Aber auch wo wir uns mit einem „ziemlich, sehr, ungeheuer gross oder klein" u. dgl. behelfen müssen, bleibt der Gang der Ueberlegung derselbe, und man kann ihn, wenn man will, durch Einsetzen von Zahlenwerten, die schätzungsweise diesen Begriffen entsprechen, in die mathematische Form bringen.[1])

---

1) Vgl. wiederum Kirchhoff's Ueberlegung, die eine classische Anwendung dieses Princips auf wirkliche Verhältnisse ist; allerdings schon auf ein zusammengesetztes Ereignis — aber die Grenze der einfachen und zusammengesetzten ist ja überhaupt nicht scharf zu ziehen.

Die berechnete Wahrscheinlichkeit $1 - (\frac{1}{60})^2$ gilt auch hier zunächst nur für das Vorhandensein irgendeiner begünstigenden Bedingung. Aber in Erwägung aller Umstände des Falles ist Kirchhoff berechtigt, sie zugleich als Wahrscheinlichkeit dieser bestimmten Bedingung, des Vorhandenseins von Eisen in der Sonne, anzusehen. Fänden sich dieselben 60 Linien neben anderen auch im Spectrum eines anderen Stoffes, oder hätte man keine Ahnung von dem Zustandekommen dunkler Linien im Spectrum, so wäre die specielle Form der Hypothese weniger wahrscheinlich als die allgemeine. Bei Laplace' Hypothese über die Entstehung des Sonnensystems kann die von ihm und Späteren aus der übereinstimmenden Umlaufsrichtung berechnete Wahrscheinlichkeit nicht ohne Weiteres ganz auch für die specielle Form der Hypothese in Anspruch genommen werden, wie denn auch Laplace selbst nur erschliesst „une cause commune, Auch ist der (besonders von Poisson urgirte) Begriff: „merkwürdige oder ausgezeichnete Fälle", mit anderen Worten die Grenze, wo die Hypothese einer begünstigenden Bedingung über die Zufallshypothese das Uebergewicht erlangt, nicht scharf zu fixiren; weder im Allgemeinen noch im einzelnen Fall. Aber es gibt eben doch Fälle, wo das Uebergewicht ausser allem Zweifel steht.

3. Betrachten wir jetzt die zusammengesetzten Ereignisse, die Verteilungen mehrerer einfachen Ereignisse in einer grossen Anzahl. Auch hier ist jede Verteilung logisch wie physisch möglich. Aber wenn man bestimmte Chancen der einfachen Ereignisse a und b zu Grunde legt, so ist nicht jede Verteilung gleich wahrscheinlich. Vielmehr ist die diesen Chancen entsprechende die wahrscheinlichste, und lässt sich nach dem Bernoulli'schen Gesetz für eine gegebene Gesammtzahl von Beobachtungen immer eine bestimmte Wahrscheinlichkeit dafür berechnen, dass die Verteilung sich nicht mehr als um einen bestimmten Betrag von dieser wahrscheinlichsten unterscheide. Hatte nun vor der Beobachtung des zusammengesetzten Ereignisses jedes der einfachen die Wahrscheinlichkeit $\frac{1}{2}$, und weicht dann die beobachtete Verteilung mehr von dem Verhältnis 1 : 1 ab, als jene apriori bestimmbaren Grenzen betragen, so können wir auch dann noch die ursprüngliche Hypothese festhalten. Sie bleibt logisch möglich wie jede andere. Aber sie wird immer unwahrscheinlicher, je mehr bei wachsender Anzahl der Beobachtungen die wirkliche Verteilung von jener (sich dann immer mehr verengenden) Zone abrückt. Mit anderen Worten es wird immer unwahrscheinlicher, dass diese Abweichung

---

qui a dirigé tous ces mouvements dans le sens de la rotation du soleil" (Oeuv. VII, 262). Die specielle Form verdankt ihre hervorragende Wahrscheinlichkeit nur der Verbindung mit noch anderen Thatsachen und Erwägungen.

zufällig sei, und es tritt immer mehr die Hypothese in den Vordergrund, dass die Chancen annähernd der wirklich vorgefundenen Verteilung entsprechen.[1])

Die Schlussweise ordnet sich unter die obige Regel. Alle für das fragliche Ereignis denkbaren Wahrscheinlichkeiten werden als Hypothesen betrachtet, und zwar beim Mangel aller Anhaltspuncte als vorgängig gleichmögliche Hypothesen, aus deren verschiedenem Erklärungswert gegenüber der beobachteten Verteilung dann ihre abstracte und concrete Wahrscheinlichkeit bestimmt wird.

Wahrscheinlichkeiten hypothetisch setzen, kann aber nur den Sinn haben, dass man gewisse reale Verhältnisse hypothetisch setzt, auf Grund deren man, wenn sie uns gegeben wären, dem Ereignis eben diese Wahrscheinlichkeiten zuschreiben würde; z. B. die Mischungsverhältnisse 1 : 1, 2 : 1, 3 : 1 u. s. f., auf Grund deren man dem Zug einer weissen Kugel aus einer Urne die Wahrscheinlichkeiten $\frac{1}{2}$, $\frac{2}{3}$, $\frac{3}{4}$ u. s. f. zuschreiben würde. Das Ergebnis des Schlusses bedeutet also, dass die realen Verhältnisse analoge seien wie beim Ziehen von Kugeln aus einer Urne, worin sie annähernd im Verhältnis der beobachteten Verteilung gemischt wären.

Mischungsverhältnisse haben allerdings nicht vorgängig gleiche Möglichkeit. Aber dies verschlägt hier nichts, wo sie nur benützt werden, um den S i n n der verschiedenen Hypothesen und der resultirenden empirischen Wahrscheinlichkeit zu veranschaulichen.

Hiemit ist denn auch, wie mir scheint, im Princip alles

1) Mathematisch kann man den Schluss als eine Umkehrung des Bernoulli'schen Lehrsatzes ansehen. S. A. Meyer S. 169—170 und 226 f. Ueber die Modalitäten der Anwendung in der Statistik: Lexis, Einleitung in die Theorie der Bevölkerungsstatistik 1875 (Abschnitt V); Hildebrands und Conrads Jahrbücher für Nationalökonomie XXVII S. 209 f.; Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft 1877 S. 23 f.

gesagt, was zur Lösung der eingangs erwähnten Schwierigkeiten beizubringen möglich und nötig ist. Aber führen wir es nach dieser Richtung näher aus.

Consequent muss man sagen: die apriorische Wahrscheinlichkeit, z. B. $\frac{1}{2}$ für Knaben- und Mädchengeburten, war richtig bestimmt. Denn sie war aus dem Stand unserer Kenntnis und Unkenntnis richtig abgeleitet. Auch die aposteriorische Wahrscheinlichkeit ist richtig bestimmt; sie fliesst aus dem unanfechtbaren Princip der Hypothesenschätzung. Also nicht sowol eine Correctur als vielmehr eine Veränderung der Wahrscheinlichkeit liegt vor.

Ein solche ist nur möglich, wenn die Kenntnisse sich verändert haben. Worin liegt nun die Veränderung unserer Kenntnisse, die in solchen Fällen durch die Ergebnisse der Statistik herbeigeführt wird? — Zweifellos bezieht sie sich auf einen realen Thatbestand und zwar in irgend einer Weise auf die realen Bedingungen des einen und anderen Ereignisses. Hierin stimmen beide Theorien, die objective und die subjective, überein, wenn auch die eine ein reales Verhältnis dieser Bedingungen selbst, die andere ein bestimmtes Wissen und Nichtwissen über sie zur Grundlage des Wahrscheinlichkeitsbegriffes macht. Die Schwierigkeit liegt nur darin, anzugeben, was wir eigentlich jetzt mehr oder anderes über die realen Bedingungen wissen. Denn im Grunde, liesse sich sagen, wissen wir doch über die Ursachen des Geschlechts nach wie vor gar nichts.

Auch die Schwierigkeit ist also in Wahrheit beiden Theorien gemein. Und keine wird sich anders helfen können als entweder durch eine ganz abstracte Formulirung (u. S. 106) oder durch eine von der Art, wie sie vorhin versucht wurde: Reduction auf das Urnenschema oder ein ähnliches. Man hat dann allerdings nur ein Gleichnis, aber ein in dem Vergleichungspunct zutreffendes, und damit doch eine wirkliche Erkenntnis, die uns einstweilen die adaequate Definition des Sachverhaltes ersetzen kann.[1]) Dass es sich hiebei um eine wirkliche Bereicherung unserer Kenntnis handelt, sieht man daran, dass auch eine solche blos abstracte oder symbolische Kenntnis uns in Verbindung mit den wachsenden sonstigen Erfahrungen an Menschen und Thieren der wirklichen Erkenntnis der Ursachen näher bringen kann. Die neueren Hypothesen haben darin immerhin einen wichtigen Anhaltspunct mehr als alle früheren.

Dass es aber auch hier nicht ein objectives Verhältnis realer Bedingungen selbst, abgesehen von aller Beteiligung eines Nichtwissens, ist, worauf sich der aposteriorische Wahr-

---

1) Unter besonderen Umständen kann immerhin das Gleichnis der Sache näher stehen. Wenn uns z. B. ein seiner Gestalt nach völlig unbekannter starrer 5seitiger Körper gegeben ist, an welchem, wie wir wissen, 3 Seiten mit a, 2 mit b beschrieben sind, so ist die apriorische Wahrscheinlichkeit, dass er mit einer a-Seite am Boden aufliege, $= \frac{3}{5}$. Verhalten sich nun in 50000 Würfen die a-Würfe zu den b-Würfen wie 13:7, so können wir sagen: die Bedingungen für a und b sind, alles zusammengenommen (die Grösse und Gestalt der einzelnen Flächen, einschliesslich ihrer etwaigen Krümmungen, die so complicirt sein können, dass Niemand die apriorische Wahrscheinlichkeit daraus hätte bestimmen können, u. s. f.) analoge wie für 13 mit a und 7 mit b bezeichnete Seiten eines Ikosaeders. Hier läge das Gleichnis wenigstens in demselben Gebiete wie der Gegenstand selbst.

Nach den Erörterungen unter III, 3 können wir uns auch statt eines solchen Ikosaeders 50000 ihrer Gestalt nach völlig unbekannte starre 20seitige Körper denken, woran je 13 Seiten mit a, 7 mit b bezeichnet wären, ohne irgend ein sonstiges Princip dieser Bezeichnung. Es ist also auch bei der Interpretation der aposteriorischen Wahrscheinlichkeit durch Rückgang auf ein Schema nicht notwendig, irgend eine physische Gleichheit der „gleichmöglichen" Fälle in diesem Schema anzunehmen. Aber natürlich wird man, wenn es sich schon einmal um die Wahl eines anschaulichen Gleichnisses handelt, die anschaulichere Formulirung vorziehen, also das Ikosaeder im obigen Fall oder als allgemeinst-anwendbares Schema die Urne mit bekanntem Mischungsverhältnis.

scheinlichkeitsbruch bezieht, das ist nicht blos an dem Schema sondern auch an der Sache klar. Denn natürlich kann mit dem Verhältnis der Bedingungen nicht ein Uebergewicht der Bedingungen für Knabengeburten in allen einzelnen Fällen gemeint sein, welches nicht blos die Geburt von Mädchen sondern ebendamit auch die von Knaben überhaupt verhindern würde. Aber auch wenn man von einem „durchschnittlichen Uebergewicht der Bedingungen" im Betrag von etwa 106 gegen 100 redet, so ist der Ausdruck misverständlich. Denn man wird den Sachverhalt bei der Geschlechtsdetermination und allen derartigen Vorgängen, die abwechselnd zu zwei verschiedenen Ergebnissen führen, nicht so auffassen dürfen, wenigstens nicht so auffassen müssen, als ob dabei zwei (vorläufig unbekannte) Kräfte einander gegenüberständen, von deren relativer Intensität der jeweilige Ausgang abhinge. Wenn z. B. der Reifezustand des Eies oder der Spermatozoen massgebend wäre, würde sich dies nicht auf jene Formel bringen lassen. Endlich auch wenn wir „Bedingungen" in einem entsprechend weiteren Sinn nehmen wollten, so kann durch 106 : 100 jedenfalls nicht ein durchschnittliches Mass eines solchen Uebergewichtes von Bedingungen bezeichnet sein: darüber sagen die statistischen Zahlen vollends nichts. Nicht einmal das Vorhandensein eines solchen Uebergewichts von Bedingungen in durchschnittlich 106 unter 206 Fällen: auch dies wäre schon eine specielle, nicht notwendige, Auffassung. Das Uebergewicht der Zahlen bedeutet vielmehr weiter nichts, als dass durchschnittlich unter 206 Fällen die Bedingungen für eine männliche Geburt 106 mal überhaupt vorhanden, 100 mal aber überhaupt nicht vorhanden sind.

Und wenn wir nun dieses Zahlenverhältnis als ein Chancenverhältnis bezeichnen, so kann es auch da nicht ein Massverhältnis von gegeneinanderstehenden realen Bedingungen bedeuten, sondern nur ein Zahlenverhältnis gleichmöglicher Fälle, nämlich von 106 günstigen gegen 100 ungünstige (oder umgekehrt) unter 206 gleichmöglichen Fällen von Bedingungscombinationen. Diese 206 möglichen und 106 günstigen Combinationen müssen wir als in sich selbst unterschiedene annehmen, obschon wir sie nicht näher namhaft machen können. Wir müssen schliessen: dass die Summe von Bedingungen, welche in jedem Fall unmittelbar zur Geschlechtsanlage beim Menschen führt, worin sie auch bestehen möge, unter den Verhältnissen, worauf die statistischen Erhebungen sich beziehen, etwa 206 mögliche Combinationen zulässt, von denen 100 unter den Begriff „weiblich", 106 unter den Begriff „männlich" fallen (das weibliche oder männliche Geschlecht herbeiführen), und dass wir uns über jede dieser 206 möglichen Bedingungscombinationen, selbst wenn wir die Bedingungen ihrer Natur nach adaequat angeben könnten, disjunctiv (bezüglich einer jeden gegenüber jeder anderen) in absoluter Unwissenheit befinden würden.

Dass ein solches Verhältnis günstiger und möglicher Fälle anzunehmen sei, sagen uns, um dies zu wiederholen, nicht unmittelbar die statistischen Zahlen. Es ist eine Hypothese, die nach dem oben beschriebenen Gedankengang selbst nur als wahrscheinlich erschlossen wird, neben der an und für sich viele oder unzählige andere Hypothesen logisch möglich bleiben. Damit ist aber, wenn mich nicht alles Vorangehende täuscht, integrirend das Zugeständnis verknüpft, dass in der Definition der so als wahrscheinlich erschlossenen aposteriorischen Wahrscheinlichkeit der Begriff günstiger und möglicher Fälle in keinem anderen Sinn verstanden werden kann als in dem bereits unter I erläuterten.[1])

1) Auch Lexis scheint mir darauf hinauszukommen, obschon er, anknüpfend an Cournot's Unterscheidung einer objectiven und subjectiven Wahrscheinlichkeit, bestrebt ist, die empirische Wahrscheinlichkeit von der blos auf ungenügendes Wissen gegründeten Wahr-

Auch die aposteriorische Wahrscheinlichkeit also involvirt ein Nichtwissen, und sogar in doppelter Beziehung:

---

scheinlichkeit, die nur zur Regelung von Spielen und Wetten diene, zu unterscheiden (s. die letzte der obenerwähnten Schriften S. 13 f.). Wenn die empirische Wahrscheinlichkeit angenähert constant ist und gegen einen festen Grenzwert convergirt, so bedeutet sie nach ihm die „objective, physische Möglichkeit" des Ereignisses. Ich vermisse eine Definition dieser physischen Möglichkeit. Lexis selbst bemerkt, dass in der wirklichen Welt sowol für das Eintreten als für die Verhinderung eines Ereignisses unzählige Ursachen entscheidend seien, die aber selbstverständlich weder gleich möglich noch absolut unabhängig von einander seien, und sieht sich so doch schliesslich zu der Formulirung geführt, dass diese unberechenbare Manichfaltigkeit der Umstände, die das Ereignis hervorrufen oder verhindern, „eine genügende Analogie des absoluten Zufallsspieles mit E günstigen gegen $Z-E$ ungünstige Chancen darbiete" (S. 18).

Die angenäherte Constanz des statistischen Zahlenverhältnisses und das Convergiren gegen einen festen Grenzwert sind Umstände, die die Wahrscheinlichkeit dieser empirischen Wahrscheinlichkeit erhöhen, also unser Zutrauen zur Richtigkeit des Schlusses rechtfertigen, auf Grund dessen wir die Analogie mit einem Zufallsspiel von den angegebenen Chancen behaupten. Aber auf die Definition der empirischen Wahrscheinlichkeit selbst haben diese Umstände, wie mir scheint, keinen Einfluss.

In der zweiten der obenerwähnten Schriften, die sich speciell mit dem Geschlechtsverhältnis der Geborenen beschäftigt, führt Lexis als die einfachste und bequemste Vorstellung zur Deutung der in diesem Verhältnis gegebenen Wahrscheinlichkeit die ein, „dass schon die sehr zahlreichen unbefruchteten Keime in den weiblichen Ovarien für das eine oder andere Geschlecht praedestinirt seien; und zwar dass bei allen weiblichen Individuen — um zunächst eine streng schematische Annahme zu machen — die männlichen Keime die weiblichen in demselben Verhältnis überwiegen" (S. 242). „Die Analogie mit der Urne ist dann einleuchtend", fügt er selbst hinzu. In der That kann man sich ja auch auf diese Art in dem besonderen Falle den Begriff veranschaulichen, wenn es sich nur darum handelt. Das Bild ist als solches brauchbar, auch wenn es vom wirklichen Sachverhalt weit entfernt ist. Es stellt dann ebenso ein Schema des wir wissen nichts über die 206 möglichen Fälle im Vergleich mit einander, und wir wissen ausserdem auch nichts über die wirkliche Beschaffenheit dieser Fälle, die wir nur ganz abstract oder aber mit Hilfe eines Gleichnisses definiren können. Die letztere Unwissenheit ist das charakteristische Merkmal der aposteriorischen Wahrscheinlichkeit, die erstere ist ihr mit der apriorischen gemein.

Ungleiche Wahrscheinlichkeiten der „einfachen Ereignisse" a und b beruhen immer auf einer Zerlegung in noch einfachere Fälle; z. B. wenn $a = \frac{4}{5}$, $b = \frac{1}{5}$, so heisst dies, es sind 5 noch einfachere Fälle $\alpha \beta \gamma \delta \varepsilon$ möglich, wovon α bis δ unter den Begriff a fallen. Aber bei der apriorischen Wahrscheinlichkeit können wir diese einfachsten Fälle in concreter Weise angeben, bei der aposteriorischen nicht.

Wenn einem und demselben Ereignis zuerst in Ermangelung empirischer Anhaltspuncte die Wahrscheinlichkeit $\frac{1}{2}$, später aber eine andere Wahrscheinlichkeit zugeschrieben wird, so hat sich auch die Materie im engeren Sinn (S. 55) geändert: vorher waren die Ereignisse a und b selbst die disjungirten Fälle, jetzt sind es die Combinationen von Bedingungen, auf deren Wechsel der wechselnde Eintritt von a und b beruht. Der begriffliche Unterschied der apriorischen von der aposteriorischen Wahrscheinlichkeit liegt aber nicht hierin: denn auch die apriorische wird vielfach aus einer Disjunction der Bedingungen abgeleitet. Der Unterschied liegt darin, dass wir bei der apriorischen die disjungirten Fälle in sich selbst kennen, während wir sie bei der aposteriorischen nicht kennen und nur ihre Anzahl sowie die der günstigen Fälle aus äusseren Anhaltspuncten erschliessen. Dennoch muss selbstverständlich die empirische

Begriffes dar wie das Urnengleichnis selbst. Aber es ist dann auch an diesem Schema klar, wie genau sich der Begriff hinsichtlich der Bedeutung der „möglichen und günstigen Fälle" mit dem der apriorischen Wahrscheinlichkeit deckt.

Wahrscheinlichkeit, wo immer eine solche vorliegt, an die Stelle der apriorischen treten, da sich unsere Kenntnisse doch vermehrt haben und Wahrscheinlichkeitsbestimmungen sich stets auf sämmtliche verfügbaren Kenntnisse, die einen Unterschied machen können, insbesondere auch auf solche über die Ursachen, stützen müssen.

Hier bietet sich auch ein Seitenblick auf den Begriff des Apriorischen und Aposteriorischen, wie er sonst in der Erkenntnistheorie auftritt, wenn wir Vernunftgesetze und Erfahrungs- oder Naturgesetze unterscheiden. Man kann die letzteren am Ende auch nicht anders definiren denn als Notwendigkeiten, die Denknotwendigkeiten (Vernunftgesetze) sein würden, wenn wir die adaequaten Begriffe hätten.[1]) Insoweit besteht Analogie zu dem gleichnamigen Unterschied der Wahrscheinlichkeiten. Aber ein Erfahrungsgesetz kann niemals in ein apriorisches übergehen[2]), während allerdings bei der aposteriorischen Wahrscheinlichkeit der weitere Verlauf der Forschung uns zu einer wirklichen Kenntnis jener gleichmöglichen Bedingungscombinationen $\alpha \beta \gamma \ldots$ und damit zu einer apriorischen Wahrscheinlichkeitsbestimmung führen kann. Es besteht auch nicht etwa ein fester Unterschied unter den Materien des Wahrscheinlichkeitsurteils derart, dass bei den einen die (direct) apriorische Bestimmung endgültig, bei den anderen aber provisorisch und durch die aposteriorische zu ersetzen wäre: denn auch selbst beim Würfeln ist die apriorische Bestimmung, genau genommen, keine endgültige, da sich bei jedem noch so guten Würfel aus einer ungeheuren Anzahl von Fällen ein etwas ver-

1) Vgl. Abhandl. d. k. b. Akad. I. Cl. XIX. Bd. II. Abth. S. 494 f.

2) So laufen z. B. alle Beweisversuche für das Trägheitsgesetz auf Erschleichungen hinaus. Mathematische Sätze werden allerdings häufig zuerst inductiv gefunden und dann apriori bewiesen. Aber sie wären in solchen Fällen an sich auch apriori auffindbar gewesen, und wir bezeichnen sie darum nicht als Erfahrungsgesetze.

ändertes Chancenverhältnis herausstellen wird. Der Unterschied ist nur ein gradueller, obschon als solcher beträchtlich genug, um im Grossen und Ganzen zweierlei Materien auseinanderzuhalten.

4. Wir haben die Veränderung der Chancen definirt, welche stattfindet oder stattfinden kann, wenn an die Stelle der apriorischen Wahrscheinlichkeit eines Ereignisses die aposteriorische tritt. Leicht ergibt sich nun auch, in welchem Sinn wir von der Veränderung der Chancen von einer Reihe beobachteter Verteilungen zu einer anderen Reihe sprechen; wenn sich z. B. das Verhältnis der Geburten in einem bestimmten Lande als ein wesentlich anderes herausstellte. Wir verstehen dann unter Chancen nicht den Zähler des Wahrscheinlichkeitsbruches selbst sondern jene realen Umstände, in Folge deren die Disjunction der möglichen Fälle von Bedingungscombinationen eine andere sein muss. (Vgl. oben S. 42.) Nennen wir sie die realen Chancen, so wird nun wol kein Misverständnis sich daran knüpfen. Es soll nicht ein allgemeinerer Begriff von Chancen damit eingeführt sein, der die in I definirten und die soeben definirten als Arten unter sich befasste, sondern nur eben ein kurzer Ausdruck für „die realen Umstände, aus denen wir das Chancenverhältnis (direct oder indirect) bestimmen". Es liesse sich dafür auch irgend ein ganz neuer einfacher Ausdruck setzen. Man wolle also hier auch nicht eine Cirkeldefinition finden. Reale Chancen in diesem Sinne nun meinen wir, wenn von Veränderung der Chancen von einer Beobachtungsreihe zur anderen die Rede ist. Wir meinen das Analogon eines veränderten Mischungsverhältnisses in der Urne.

Irgendwelche Chancenveränderungen in diesem Sinne muss man geradezu immer und überall, bei Ereignissen aller Art, innerhalb der beobachteten Reihen voraussetzen, da in keinem Beobachtungsgebiete absolute Constanz der Bedingungen herrscht. Nur in Hinsicht der letzten unbeobachtbaren Elemente pflegen wir absolute Constanz anzunehmen. Das Mischungsverhältnis in der Urne der Natur schwankt von Reihe zu Reihe, ja von Fall zu Fall. Der Statistik sind natürlich nur die Schwankungen von Reihe zu Reihe zugänglich. Sie sind bei manchen Materien in langen Zeitläuften und weitem Raumbezirk verschwindend gering, bei manchen sehr gross. Tritt fast genau dasselbe Verhältnis, wie in einer Gesammtreihe, auch in einzelnen Fractionen auf, die wir immer kleiner nehmen können, so werden wir auch keine erhebliche Veränderung der realen Chancen von einer zur anderen annehmen und die im Ganzen gefundene Verteilung mit um so grösserem Zutrauen als den wahren Wert der empirischen Wahrscheinlichkeit betrachten. Bei grosser Verschiedenheit dagegen wird man eben die Fractionen nach Ort, Zeit u. s. f. auseinanderhalten und die Wahrscheinlichkeiten gesondert bestimmen.

Woran es liegt, dass bei manchen Arten von Ereignissen die realen Chancen in so hohem Masse constant, bei anderen so veränderlich sind, diese Frage geht nicht mehr die Wahrscheinlichkeitslehre sondern die Naturphilosophie an, die aber schwerlich eine allgemeinere Antwort darauf geben kann, so interessante Untersuchungen auch im Einzelnen daraus erwachsen. Die mechanischen Bedingungen für die Stabilität eines individuellen Bedingungscomplexes ebenso wie für die Reproduction gleichartiger Bedingungscomplexe durch den Naturlauf können sehr verschieden sein.

Auch in der Veränderung der Chancen kann eine Constanz liegen, sie können um einen mittleren Stand oscilliren (für welchen Fall Poisson den Bernoulli'schen Satz zum „Gesetz der grossen Zahlen“ umgestaltet hat) oder nach einer bestimmten Richtung fortschreiten. Unter den letzteren Fall ordnen sich neben zahlreichen anderen physischen und psychischen Dispositionen (wie z. B. ein Diener mich um so öfter betrügen wird, je häufiger es bereits mit Erfolg geschah) die Erscheinungen der Uebung und der Ermüdung, sowie die generellen organischen Veränderungen im Lauf der Jahrhunderte oder Jahrtausende, wodurch ebenfalls die realen Chancen bestimmter Leistungen in bestimmter Richtung verändert werden. Es können ferner anfänglich sehr veränderliche Chancen immer constanter werden oder umgekehrt. Wenn ich aus einer rohen Kartoffel einen beliebigen 5 seitigen Körper schneide und die Seiten durch Zeichen unterscheide, so wird immer mehr Gleichmässigkeit in der Verteilung der Würfe eintreten (nicht in den Wurfzahlen der verschiedenen Seiten natürlich, sondern in dem Verhältnis dieser Zahlen), weil durch Austrocknen die Form des Körpers und damit die realen Chancen unveränderlicher werden. Aehnliches wieder im Grossen (vgl. Fechner's Princip der „Tendenz zur Stabilität").

Die Classificirung der Ereignisreihen unter solchen Gesichtspuncten wäre ebenfalls von Interesse, aber ohne weiteren Gewinn für unsere begrifflichen Fragen. Nur das Eine: Wenn der Entwickelungsgedanke in irgend einer Form ganz allgemeine Gültigkeit besitzt, so gibt es strenggenommen überhaupt keine blos oscillirenden, sondern nur nach bestimmter Richtung fortschreitende und dabei eventuell oscillirende reale Chancen. Jedenfalls ist bei organischen Erscheinungsreihen die verlangte „Unabhängigkeit" der einzelnen aufeinanderfolgenden Ereignisse von einander und von gemeinschaftlich zu Grunde liegenden veränderlichen Dispositionen keine absolute. Die Abhängigkeit kann nur graduell so gering sein, dass sie vernachlässigt werden darf. Doch würde, auch wo sie grösser ist, die empirische Wahrscheinlichkeit nicht ohne Weiteres unbestimmbar oder bedeutungslos werden; wir müssten nur zugleich für die Veränderung der realen Chancen selbst nach Richtung und Grösse einen Wahrscheinlichkeitswert ermitteln können.

5. Nach dem Vorstehenden beantwortet sich schliesslich

auch die Frage, inwiefern die empirische Wahrscheinlichkeit, gleich der apriorischen, als Mass unserer vernünftigen Erwartung gelten kann.

Es kann sich bei beobachteten Verteilungen überhaupt um zweierlei Erwartungen handeln: um eine Erwartung in Bezug auf das Gleichbleiben der Verteilungsart, aus welcher sie abgeleitet ist, in einer künftigen Reihe von Fällen, oder um eine Erwartung in Bezug auf das Eintreten eines individuellen Falles.

Die vernünftige Erwartung in der ersten Hinsicht ist natürlich nicht gemessen durch die empirische Wahrscheinlichkeit des fraglichen Ereignisses, sondern durch die Wahrscheinlichkeit, die wir für das Gleichbleiben der realen Chancen besitzen. Diese Wahrscheinlichkeit kann nur ausnahmsweise in Zahlen ausgedrückt werden; sie kann sehr gross, aber auch sehr klein sein, je nach den Erfahrungen über die Einflüsse, denen die realen Chancen gerade bei der fraglichen Art von Ereignissen ausgesetzt sind.

Es bedarf übrigens kaum der Bemerkung, dass das Nämliche, was von künftigen, auch von jeder beliebigen vergangenen oder gegenwärtigen Reihe gilt. Wo immer und inwieweit immer sich für unbeobachtete Reihen die realen Chancen als gleich mit denen innerhalb der beobachteten annehmen lassen, da ist auch die gleiche Verteilungsart mit entsprechender Wahrscheinlichkeit anzunehmen. Wahrscheinlichkeit hat auch hier mit Zeitunterschieden principiell nichts zu thun.

Setzen wir einmal einen bestimmten empirischen Wahrscheinlichkeitswert für eine bestimmte Summe von Ereignissen der bezüglichen Art gültig — mögen sie der Gegenwart oder der fernsten Vergangenheit oder Zukunft, mögen sie auch räumlich benachbarten oder fernen Regionen angehören —, so lässt sich dann wieder auf Grund des Bernoulli'schen Satzes die Wahrscheinlichkeit berechnen, dass die Verteilung sich von der beobachteten nicht über einen bestimmten Betrag entferne. Dies ist dann wieder, ganz so wie bei der apriorischen Wahrscheinlichkeit, eine blos logische Consequenz, ein identischer Satz. Nur so ist es zu verstehen, wenn beispielsweise ausgerechnet wird, dass man im Jahre 1784 nahe 4 gegen 1 wetten konnte, dass in Paris in den nächsten 100 Jahren die Zahl der Knabengeburten alljährlich die der Mädchengeburten übertreffen werde (A. Meyer S. 226). Schöne Exempel, aber ohne jede reelle Bedeutung, solange nicht die Wahrscheinlichkeit der obigen Voraussetzung ebenfalls mathematisch bestimmt und mit der so berechneten multiplicirt werden kann. Wenn man in solchen Fällen die nächsten 100 Jahre nimmt, so ist die Gültigkeit des empirischen Wahrscheinlichkeitswertes für die ganze Reihe wol wahrscheinlicher, als wenn man 100 Jahre aus dem vierzigsten Jahrtausend nimmt. Aber didaktisch wäre das Letztere zweckmässiger, da es über die Notwendigkeit jener Voraussetzung und die problematische Bedeutung der Berechnung keine Täuschung aufkommen lässt.

Gegenüber dem individuellen Fall handelt es sich um die Frage, ob der gefundene Wahrscheinlichkeitswert selbst als Mass einer vernünftigen Erwartung des einzelnen Falles gelten kann. Indem wir diese Frage bejahen, widersprechen wir einer heute sehr verbreiteten Lehre.[1]) Man ist geneigt,

1) Die hyperkritische Doctrin hat wol von Fries ihren Ausgang genommen. „Sagten wir, du hast die Wahrscheinlichkeit $\frac{1}{6}$, mit einem Würfel gerade die 4 zu treffen, so heisst das, im Durchschnitt wird unter 6 Würfen immer einmal die 4 getroffen werden. Aber nun gerade für diesen einen Wurf? Ja da weiss ich gar nicht, welcher von den 6 möglichen Fällen eintreffen wird. Dies will wol beachtet sein. Denn gehen wir nun zur Bestimmung der Wahrscheinlichkeit aposteriori über, so finden wir bei ihr einzig den Satz von objectiver Gültigkeit, dass, je länger wir die Beobachtungen in einer bestimmten Sphäre fortsetzen, wir die Ereignisse im Durchschnitt um so genauer in den Verhältnissen finden werden, in welchen die Zahlen der für

die Anwendung einer empirischen Wahrscheinlichkeit auf den individuellen Fall für eine grosse Lächerlichkeit zu halten. Dennoch erscheint sie nicht nur als die einzig mögliche Consequenz, sondern auch wieder als Ausdruck des gesunden Menschenverstandes, ganz so wie bei der apriorischen Wahrscheinlichkeit. Halten wir uns zunächst wieder an Spiele, bei denen ja ebenfalls aposteriorische Wahrscheinlichkeit gegeben sein kann. Wenn in einer Urne, worin, wie uns bekannt, nur w und s Kugeln, bei einer Million von Zügen (mit Hineinlegen und Schütteln) ungefähr 200000 mal w, 800000 mal s gezogen wurden, so werden wir doch ebenso und in demselben Sinne des Wortes die Wahrscheinlichkeit für das Ziehen einer w im einzelnen Fall $= \frac{1}{5}$ setzen und

---

sie stattfindenden gleichmöglichen Fälle stehen. Alles übrige, einzelne Fälle oder auch allgemeine Uebersichten treffende ist nur von subjectiver Bedeutung." (Versuch einer Kritik der Principien der Wahrscheinlichkeitsrechnung 1842 S. 129.)

Fries war meines Wissens der einzige deutsche Philosoph, der in der ersten Hälfte dieses Jahrhunderts der Wahrscheinlichkeitsrechnung eingehendere Beachtung schenkte, und dies darf ihm nicht vergessen werden. Aber seine Kritik ist wenig gelungen, und hier z. B. jeder Satz ungenau und schief. Verkehrt ist es schon, die Wahrscheinlichkeit $\frac{1}{6}$ durch das Bernoulli'sche Gesetz zu definiren, welches vielmehr seinerseits den Wahrscheinlichkeitsbegriff schon voraussetzt. Es behauptet ja auch nur eine Wahrscheinlichkeit und zwar in demselben Sinn wie die Wahrscheinlichkeit $\frac{1}{6}$ selbst, nur grösser. Ebenso verkehrt der zweite Satz. Keine Wahrscheinlichkeitsangabe, auch die Bernoulli's nicht, will sagen, was geschehen wird. Gerade weil ich gar nicht weiss, welcher von den 6 möglichen Fällen eintreffen wird, gerade darum nenne ich den Wurf der Zahl vier $\frac{1}{6}$ wahrscheinlich. Was dann endlich über die Wahrscheinlichkeit aposteriori folgt, wirft alles durcheinander. Bernoulli's Gesetz, das Fries hier offenbar wieder im Sinne hat, ist keine Wahrscheinlichkeitsbestimmung aposteriori, es lautet nicht so wie es hier ausgesprochen wird, und die darin behauptete Wahrscheinlichkeit ist genau so subjectiv und so objectiv, wie „alles Uebrige", was als mathematisch Wahrscheinliches behauptet wird.

diese Wahrscheinlichkeit in gleicher Weise unserer Erwartung zu Grunde legen, als wenn uns von vornherein das Verhältnis der in der Urne befindlichen Kugeln als $2\,w : 8\,s$ unter je 10 Kugeln gegeben ist. Wenn mit einem regulären Körper, über den mir nur bekannt ist, dass seine Seiten mit je einem verschiedenen Buchstaben bezeichnet sind, unter einer Million von Würfen immer nur 6 Buchstaben und jeder fast genau in der gleichen Anzahl zum Vorschein kommen, so werden wir ihn für einen Würfel erklären und die Wahrscheinlichkeit $\frac{1}{6}$ ganz in demselben Sinn und mit demselben Recht auf einen einzelnen Fall anwenden, wie wenn wir ihn im Voraus als Würfel gekannt hätten. Die Voraussetzung müssen wir allerdings auch hier machen, dass die realen Chancen dieselben bei dem neuen Fall seien, wie in der beobachteten Reihe. In den erwähnten Beispielen sind wir dessen so gut wie sicher; in anderen weniger.

Es ist also principiell kein Unterschied zwischen der Anwendung auf einen einzelnen Fall und auf eine Reihe von Fällen. Insoweit wir überhaupt die empirische Wahrscheinlichkeit unserer Erwartung zu Grunde legen dürfen, dürfen wir's ebensogut gegenüber einem neuen Einzelfall wie gegenüber einer neuen Reihe. Wir dürfen es für die Reihe doch schliesslich nur eben weil und wenn wir's für ihre einzelnen Fälle dürfen. Und nicht blos ist es gestattet, die empirische Wahrscheinlichkeit auf Einzelfälle zu beziehen, sondern ich wüsste nicht, auf was anderes man sie überhaupt beziehen sollte. Denn dass die Verteilung in einer künftigen Million von Fällen bei der Kugelziehung wieder $2\,w : 8\,s$ sein wird, dafür haben wir doch nicht die Wahrscheinlichkeit $\frac{1}{5}$ oder $\frac{4}{5}$. Wofür also? Doch auch nicht dafür, dass die Werte $\frac{1}{5}$ oder $\frac{4}{5}$ den realen Chancen wirklich entsprechen. Diese Wahrscheinlichkeit ist wieder viel grösser, unter Umständen fast $= 1$. Was bedeuten also jene Werte selbst? Welchem Subject kommen sie als Prädicate zu? —

Der Einzelfall kann übrigens auch der beobachteten Reihe selbst angehören. Dann gilt die (richtig abgeleitete) empirische Wahrscheinlichkeit ohne Weiteres und bedingungslos. Und im Anschluss daran ist auch hier wieder zu erinnern, dass die empirische Wahrscheinlichkeit principiell nicht blos der vernünftigen Erwartung künftiger Einzelfälle, geschweige denn nur des nächstkommenden, sondern ebenso der vernünftigen Beurteilung eines vergangenen oder gegenwärtigen Einzelfalles zu Grunde gelegt werden kann, soweit überall die nämlichen Voraussetzungen zutreffen. Der einzige und zwar graduelle Unterschied ist, dass diese Voraussetzungen selbst im Allgemeinen immer weniger wahrscheinlich werden, je weiter der Fall von den beobachteten räumlich und zeitlich abliegt. Aber die Schnelligkeit dieser Abnahme hängt wieder von der Materie ab.

Der Widerstand gegen die Anwendung auf Einzelfälle wurzelt ausser in blossen Misverständnissen auch in einigen mehr sachlichen Motiven. Gewisse Misgriffe in der Anwendung machen solchen Eindruck, dass man die Anwendung überhaupt für einen Misgriff hält.

Jede technische Behandlung des Einzelfalls auf Grund statistischer Wahrscheinlichkeiten hält sich nur an gewisse Eigenschaften, die diesem Fall mit einer Classe von Fällen gemeinsam sind. Will einer sein Leben versichern, so kommt es auf sein Lebensalter und auf etwaige lebensgefährliche Gebrechen an, damit ist die Personalbeschreibung zu Ende. Gleichwol wäre es verkehrt, wollte er selbst seine Vermutungen über das Alter, das er etwa erreichen könnte, nur auf die Sterblichkeitstafeln gründen. Er wird vielfach ausschlaggebendere Momente in seinen Lebensgewohnheiten, in den Erfahrungen über die Festigkeit seiner guten Vorsätze u. s. f. finden. Geradezu eine Thorheit wäre die Anwendung statistischer Wahrscheinlichkeit auf den Einzelfall dann, wenn man in der Lage ist, ein völlig genügendes Urteil aus der

Untersuchung des Einzelfalles selbst zu gewinnen; wie wenn der Richter aus einer Statistik von Zeugenaussagen in seinem Lande oder der Historiker aus einer Statistik der Quellen einer gewissen Kategorie die Zuverlässigkeit einer vorliegenden beurteilen wollte. Neben solchen Extremen, die ja auch nicht wirklich vorkommen, nimmt sich der vorige Fall schon etwas weniger bedenklich aus, und wenn wir der Untersuchung eines Einzelfalls noch weniger Indicien entnehmen können als dort oder noch weniger Kenntnisse darüber haben, so tritt die statistische Betrachtung auch noch mehr in ihre Rechte.

Misgriffe anderer Art haben die Deutung empirischer Wahrscheinlichkeiten auf Einzelfälle noch schlimmer discreditirt. Laplace erzählt von einem Manne, der in dem Monat, wo er Vater werden sollte, die vorhergeborenen Knaben und Mädchen abzählte und aus dem ungewöhnlich hohen Ueberschuss der Knaben die für ihn betrübende Folgerung zog, dass ihm ein Mädchen zu Teil würde. Auch das Verfahren von Lotteriespielern erwähnt er, eine länger nicht herausgekommene Nummer mit Einsätzen zu bedecken, in der Erwartung, dass sie nun um so sicherer darankommen werde. Man hat einfach nicht die ungeheure Zahl von Möglichkeiten im Auge, innerhalb deren solche Verteilungen nicht zu den auffallenden gehören.[1]) Ueberdies ist es zweierlei: einer empirischen Wahrscheinlichkeit entsprechend den Eintritt eines einzelnen Falles erwarten, und: bei A b w e i c h u n g e n von der bisherigen Verteilung eine C o m p e n s a t i o n dieser

1) „Das häufigere Herauskommen einer Nummer" — sagt Laplace in dem Abschnitt „Ueber die Täuschungen bei der Bestimmung der Wahrscheinlichkeiten" — „ist nur eine Anomalie des Zufalls; ich habe mehrere dergleichen berechnet, aber stets gefunden, dass sie in Grenzen eingeschlossen waren, welche eine gleiche Möglichkeit des Herauskommens aller Nummern ohne Unwahrscheinlichkeit anzunehmen gestatten".

Abweichungen erwarten. Zu diesem letzteren Schluss ist man, wie mir scheint, niemals berechtigt, es sei denn etwa, dass man für die Unveränderlichkeit der realen Chancen irgendwelche starken deductiven Gründe habe. Relativ kleine Abweichungen entsprechen ohnedies der Erfahrung und widersprechen nicht der Rechnung. Sie finden sich gerade auch in den Beobachtungen, aus denen der empirische Wahrscheinlichkeitswert abgeleitet wurde. Grosse Abweichungen aber, die die früher beobachteten wesentlich überschreiten, würden einfach auf die Vermutung führen, dass die unbekannten massgebenden Umstände, die bisher als wesentlich constant vorausgesetzten realen Chancen, eine augenblickliche Schwankung in der bezüglichen Richtung erlitten haben. Ist dies der Fall, so vermag uns beim Mangel weiterer thatsächlicher Anhaltspuncte keine blosse Wahrscheinlichkeitsrechnung auch nur den leisesten Anhaltspunct darüber zu geben, wann und ob überhaupt dieser Schwankung eine solche in umgekehrter Richtung folgen werde.

„So wahr im Allgemeinen, so trügerisch im Einzelnen" soll Gibbon die Gesetze der Wahrscheinlichkeit genannt haben. Besser würde man sagen: so wahr auch im Einzelnen, aber voll der Fallstricke für Jeden, der nicht auf seiner Hut ist. Die empirische Wahrscheinlichkeit unterscheidet sich auch hierin nicht von der apriorischen.

Untersuchungen wie die vorstehenden werden Solchen, die lieber rechnen, ohne zu fragen womit, überflüssig oder allzu umständlich erscheinen. Allein in Principienfragen, an denen doch zuletzt auch der Sinn und Wert der Rechnungsergebnisse hängt, ist die Gefahr zu grosser Umständlichkeit geringer zu achten als die entgegengesetzte. In der That hätte manche Frage noch tieferes Eingehen verlangt, auch abgesehen von der psychologischen Seite, die wir ganz ausser Betracht liessen. Doch glaube ich soviel behaupten zu dürfen, dass der von Laplace hingestellte Wahrscheinlichkeitsbegriff einer Correctur im Sinne objectiver Voraussetzungen nicht bedarf, dass umgekehrt Einschränkungen, die sich bei Laplace noch in der Fassung des Begriffes oder in der Weise seiner Einführung finden, mit den wesentlichen Elementen in keiner Verbindung stehen, und dass die hierüber im ersten Abschnitt gegebenen Ausführungen für die empirische Wahrscheinlichkeit ebenso zutreffen wie für die apriorische, mit der sie unter Einen allgemeinen Begriff fällt.

Zeitschrift/Journal: Sitzungsberichte der philosophisch-philologische und historische Klasse der Bayerischen Akademie der Wissenschaften München

Jahr/Year: 1892

Band/Volume: 1892

Autor(en)/Author(s): Stumpf Carl

Artikel/Article: Ueber den Begriff der mathematischen Wahrscheinlichkeit 37-120